Unique venom proteins from *Solenopsis invicta* × *Solenopsis richteri* hybrid fire ants

Steven M. Valles a,*, Jason B. Oliver b, Karla M. Adesso b, Omaththage P. Perera c

a Center for Medical, Agricultural and Veterinary Entomology, United States Department of Agriculture, Agricultural Research Service, 1600, SW 23rd Drive, Gainesville, FL, USA
b Tennessee State University, College of Agriculture, Otis L. Floyd Nursery Research Center, McMinnville, TN, 37110, USA
c Southern Insect Management Research Unit, United States Department of Agriculture, Agricultural Research Service, 141 Experiment Station Road, Stoneville, MS, USA

**A R T I C L E   I N F O**

Handling Editor: Glenn King

Keywords:
Fire ant
Solenopsis venom protein 2
*Solenopsis invicta* × *S. richteri* hybrid Speciation Formicidae

**A B S T R A C T**

The *Solenopsis* venom protein 2 transcript was amplified, sequenced, probed, and analyzed from *Solenopsis invicta* × *Solenopsis richteri* hybrid ant colonies (hybrids) collected from across Tennessee to determine the extent of introgression of each parent allele (*Solenopsis invicta* venom protein 2 [Soli2] and *Solenopsis richteri* venom protein 2 [Solr2]). Chemotaxonomic analyses of venom alkaloids and cuticular hydrocarbons were used to categorize hybrid colonies and their relative relatedness to each parent species. Hybrid colonies were chosen randomly from each chemotaxonomic hybridization category, including “very near *S. richteri*,” “near *S. richteri*,” “near *S. invicta*,” and “very near *S. invicta*.” Lateral flow immunoassays for detection of the Soli2 and Solr2 venom proteins were largely in agreement with the chemotaxonomic analyses for the very near *S. richteri* (100% Solr2) and very near *S. invicta* (80% Soli2, 20% Soli2 + Solr2 detected in the sample) groups, while Soli2 and Solr2 were reported in 60% and 40% in the near *S. invicta* and near *S. richteri* chemotaxonomic groups. Analysis of transcripts from the hybrid colonies revealed a sequence with 100% identity to Soli2 (GenBank Accession L09560) and three unique sequences, which we identify as *Solenopsis* hybrid venom protein 2 (Solh2; GenBank Accession MT150127), Solenopsis hybrid truncated venom protein 2 (Solh2Tr97), GenBank Accession MT150129), and Solenopsis richteri venom protein 2, D to A change at position 69 (Solr2A69; GenBank Accession MT150128). The predicted open reading frame for Solh2 and Solh2Tr97 revealed sequences unique to hybrid ants, with Solh2Tr97 an alternatively spliced form. A third unique sequence, Solr2A69, is likely the correct sequence for Solr2, which appears to have been published previously with a sequencing error (GenBank Accession P35776).

1. Introduction

*Solenopsis invicta* Buren and *Solenopsis richteri* Forel are invasive fire ant species that were introduced into the southern United States in the early part of the twentieth century (Tschinkel, 2006). During this time, *S. invicta* has greatly expanded its range. It is currently found from Florida to Virginia and west into California (Callcott and Collins, 1996). Conversely, *S. richteri* is confined to a relatively small contiguous area in northeastern Mississippi (Streett et al., 2006), northwestern Alabama (Bertagnolli et al., 2007), and southwestern Tennessee (Oliver et al., 2009). At the U.S. population boundaries between *S. invicta* and *S. richteri*, a fertile hybrid imported fire ant (*S. invicta* × *S. richteri*) was detected in 1985 (Vander Meer et al., 1985). The *S. invicta* × *S. richteri* hybrid (henceforth referred to as hybrid) is currently distributed in Alabama, Arkansas, Georgia, Mississippi, and Tennessee (Gardner et al., 2008; Oliver et al., 2009; Streett et al., 2006).

In the United States, the economic impact attributed to invasive fire ant species was estimated to exceed $6 billion (U.S.) annually in 2003 (Pereira, 2003). Accounting for inflation, the 2021 figure is estimated at $7.5 billion. More importantly, both species possess a potent multi-component venom that poses a significant human health risk (Hoffman, 1995; Stafford et al., 1989). Indeed, 5% of those stung require medical attention and 2% exhibit serious allergic responses, including anaphylaxis (deShazo et al., 1990; Rhoades et al., 1989; Stafford et al., 1989). The venom is comprised of >95% alkaloids and 0.01% proteins (Fox, 2016) and is extraordinarily complex (dos Santos Pinto et al., 2012, dos Santos Pinto et al. (2012) identified 46 proteins from the venom of *Solenopsis invicta*, which they organized into four different groups based...
on their potential role, including true venom components, housekeeping proteins, body muscle proteins, and proteins involved in chemical communication. However, four venom proteins, Solenopsis venom protein 1, venom protein 2, venom protein 3, venom protein 4, have been studied extensively and remain the best characterized of the protein components (Hoffman, 2010; Ross et al., 1987). Solenopsis venom protein 2 is the most antigenic, eliciting a strong immune response in mammals, and comprises the largest proportion (~67%) of these four proteins (Hoffman, 1995). Each of the invasive fire ant species, S. invicta and S. richteri, expresses an orthologue of Solenopsis venom protein 2 (i.e., Soli2 [GenBank Accession P35775] and Solr2 [P35776], respectively) with 80% identity between the protein sequences (Hoffman, 2010; Valles et al., 2016). Because the hybrid continues to expand its range, we were curious about the expression of Solenopsis venom protein 2 among the hybridized ants. Our objective was to examine which venom protein 2 (Soli2 or Solr2) was dominant in hybridized ant colonies. While Soli2 and Solr2 were expressed in hybrid colonies, two additional Solenopsis venom protein 2 forms were discovered that are unique to the hybrid lineage.

2. Methods

2.1. Ant collections

Solenopsis species worker ants were collected from colonies along roadsides in infested Tennessee counties south of the Federal Imported Fire Ant Quarantine area (Tennessee Department of Agriculture, 2018) from late July to early October in 2015 and 2016. The global position coordinate for each sample was determined (LandMark Spatial Systems, LLC, Starkville, MS). Worker ants were sampled by digging into the ant mound and then allowing ants to crawl up a wooden board (90 by 3.5 by 2.0 cm). Ants were dislodged from the board into an 18.9 L bucket treated with Insect-a-Slip Insect Barrier (BioQuip Products Rancho Dominguez, CA) along the upper, inner lip to prevent escape. Worker ants were transferred from the bucket with featherweight forceps to 20 ml scintillation vials (Wheaton Scientific, Millville, NJ). One vial received ~50 live ant workers for use in hydrocarbon and alkaloid assessments. A corresponding sample of live worker ants was collected and stored at ~80 °C for molecular and protein analyses. The parental species, S. invicta and S. richteri, were included as sequencing controls for Soli2 and Solr2, respectively. The Soli2 nucleotide sequences obtained from parental S. invicta and hybrids were identical to the sequence in GenBank (Accession L09560). The Solr2 nucleotide sequence was never reported previously. The nucleotide sequence for Solr2 we obtained from S. richteri collected in Henderson county, TN was deposited in GenBank with accession number MT150125.

2.2. Cuticular hydrocarbon and venom alkaloid analyses

Cuticular hydrocarbon and venom alkaloid profiles were used to identify Solenopsis invicta x S. richteri hybridized nests using procedures developed by Vander Meer et al. (1985) and Ross et al. (1987). All ant colony samples were analyzed using a Shimadzu QP-2010 Gas Chromatograph - Mass Spectrometer (GC-MS) (Shimadzu, Kyoto, Japan). To prepare ants for GC-MS analysis, 50 live ant workers were placed in n-hexane Chromasolv® (Sigma-Aldrich, Saint Louis, MO). Each sample vial with live ants received just enough hexane to cover the live ants (approximately 5 ml). Extractions proceeded for 24 h, after which, the hexane was removed to a clean glass vial by pipette (to eliminate any debris present in the sample) and allowed to evaporate in a fume hood to about 0.5 ml to concentrate the solvent sample. The concentration sample was separated on a DB-1 capillary column (30 m x 0.25 mm; i.d., 0.25 μm film thickness; Agilent Technologies, Santa Clara, CA). The GC temperature was programmed as in Menzel and Nebecker (2008). The initial temperature was 50 °C, held for 1 min, increased to 240 °C at a rate of 20 °C/min, and held for 29.5 min. The split ratio was 1:10, injection temperature 250 °C, and transfer line temperature 270 °C. The mass spectrometer was operated at 70 eV in the electron impact mode. Cuticular hydrocarbon index (IHC) and venom alkaloid index (IALK) values were calculated (Ross et al., 1987) to determine the degree of hybrid relatedness to the two parent species. The alkaloid index was defined as IALK = (P15:1 × (2.36), where P15:1 is the proportion of C15:1 alkaloid. The hydrocarbon peaks obtained were used to calculate the cuticular hydrocarbon index (IHC), which was defined as IHC = [PSH – (PSR × 0.116)]/[(PSH – PSR × 0.116) + (PSR – PSr × 0.072)], where PSr and PSr are the proportional hydrocarbon peak areas found within the retention times assigned to S. invicta and S. richteri, respectively (Ross et al., 1987). The Hybrid Index has been defined as the mean of IHC and IALK ratios where a colony with a Hybrid Index ≤0.06 considered S. richteri, between >0.06 and <0.85 as hybrids, and ≥0.85 as S. invicta (Ross et al., 1987). A Hybrid Index value was calculated for the 20 colonies used in this study and colonies were separated into the following categories: “very near S. richteri” (hybrid index = 0.061 to 0.099), “near S. richteri” (0.1–0.2), “near S. invicta” (0.4–0.5), and “very near S. invicta” (0.65–0.75). Five nests from each of these hybrid categories were chosen for subsequent analyses (Table 1).

2.3. Lateral flow immunoassay

Worker ants from nests identified as hybrids from alkaloid and hydrocarbon analyses (Table 1) were evaluated by lateral flow immunoassay (LFA) for the presence of venom proteins, Soli2 (P35775) and Solr2 (P35776) (Valles et al., 2018). The multiplexed LFA utilizes monoclonal antibodies (mAb) specific for each venom protein and can discriminate S. invicta and S. richteri based on the presence of the specific venom protein 2. Assays were conducted by placing 5 worker ants into a 1.7 ml microcentrifuge tube containing 25 μL of LFA homogenization buffer (50 mM sodium phosphate buffer, pH 7.1 ± 0.1% sodium dodecyl sulfate). The ants were thoroughly homogenized (~1 min) with a plastic pestle and the tube was pulsed in a centrifuge to force the contents of the tube to the bottom. An LFA strip was inserted into the tube. After 1 min, an additional 25 μL of homogenization buffer was added to the tube. The reaction proceeded for an additional 9 min, after which the strip was removed from the tube and scored. Positive responses for each venom protein could be determined by their positions on the strip (Valles et al., 2018).

2.4. Molecular analysis

The Soli2 and Solr2 transcripts were amplified by reverse transcription polymerase chain reaction (RT-PCR). Total RNA was extracted from a pooled group of ten worker ants from hybrid colonies (9, 37, 83, 94, 162, 193, 212, 274, 316, 344, 377, 420) using the Trizol method according to the manufacturer’s directions. The RNA (10–50 ng) was treated with DNaseAse and used as template for RT-PCR.

Two-step RT-PCR was employed to amplify the Solenopsis venom protein 2 transcript. First, 0.5 μL (50–100 ng) of the DNAse-treated total RNA was mixed with 10 mM dNTPs, 1 μM reverse oligonucleotide primer CS5 (5′-TTATATGCAAAATATTTTATGGTAAACCAACAC) and heated to 65 °C for 5 min. First strand buffer and Superscript Reverse Transcriptase (RT, Life Technologies) were added and the reaction incubated at 55 °C for 30 min before inactivating the RT reaction by heating to 70 °C for 15 min. PCR followed using the cDNA as template and oligonucleotide primer combination: 4CS (5′-GTTAAATGAAAAGTACTGACATACAACACTCCTCTCTCTA) and 5CS under the following cycling conditions: 94 °C for 2 min, 35 cycles of: 94 °C for 15 s, 55 °C for 15 s, and 68 °C for 1.5 min, 1 cycle of 68 °C for 5 min. These oligonucleotide primers were specific to the conserved 5′ and 3′ termini of the mature Soli2 and Solr2 transcripts. Amplicons were separated on 1% Agarose gels and visualized with SYBR safe (Life Technologies). Amplicons were ligated into the pCR4-TOPO vector (Life Technologies) and transformed into TOP10 E. coli competent cells (Life Technologies).
Colonies were picked and inoculated in Luria broth with ampicillin (75 μg/ml), incubated at 37 °C while shaking at 225 rpm, and positively determined for the presence of the insert by PCR with the original oligonucleotide primers.

Insert-positive plasmid DNA was purified using the QIAprep Mini-prep Spin kit (Qiagen) and clones (a minimum of 6 per colony; three per isoform) were sequenced by the Sanger method. Consensus sequences for each clone were used to search databases to identify genomic regions for each clone were used to search databases to identify genomic regions.

Rapid amplification of cDNA ends (5′ RACE) was conducted on RNA isolated from colonies 9, 37, and 316 to confirm the size of the truncated isoform. Insert-positive plasmid DNA was purified using the QIAprep Mini-prep Spin kit (Qiagen) and clones (a minimum of 6 per colony; three per isoform) were sequenced by the Sanger method. Consensus sequences for each clone were used to search databases to identify genomic regions.

2.5. Western blotting

To determine whether the truncated isoform, Solh2, identified by molecular analysis was being translated into protein, Western blotting was conducted. The abdomen of worker ants (n = 5) from colonies 9 and 37, which exhibited the truncated isoform, was extirpated and individually homogenized in 40 μl of lysis buffer (50 mM Tris-HCl, pH 8, 4% sodium dodecyl sulfate). The sample was centrifuged for 1 min at 21,000×g. The supernatant was removed by pipette (~40 μl) and added to 20 μl of Laemmli (1970) sample buffer (60 mM Tris-HCl, pH 6.8, 25% glycerol, 2% sodium dodecyl sulfate, 14.4 mM 2-mercaptoethanol, 0.1% Bromophenol blue). The sample was heated in a block heater for 5 min at 95 °C. The sample was divided in half (~30 μl) and applied to two separate 4–20% gradient SDS-PAGE gels.

SDS-PAGE-separated proteins were electroblotted onto PVDF membranes and blocked in TBS (tris buffered saline; 20 mM Tris-HCl, 500 mM NaCl, pH 7.5) + 1% BSA (bovine serum albumin) for 1 h. One blot was probed with 0.625 μg/ml of mAb 3H6B9 and the other with mAb 6C9D7 for 1 h at room temperature with shaking (40 rpm). The 3H6B9 mAb preparation was specific for the fourteen amino acid sequence (IEAQRVLRKDIAEC) specific to a region of Solh2 near the N-terminal side of the protein. The 6C9D7 mAb preparation was specific for a fifteen amino acid sequence (NPDPAVVKEKNSKMC) toward the carboxyl end (relative to IEAQRVLRKDIAEC) of the Solh2 protein (Valles et al., 2018).

2.6. Phylogenetic analysis

The predicted open reading frame of the venom protein 2 sequences was translated and analyzed by the Maximum Likelihood method based on the JTT matrix-based model with 10,000 bootstrap replicates (Jones et al., 1992). A phylogenetic tree was generated to illustrate relationships between the venom proteins using the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT matrix. A phylogenetic tree was generated to illustrate relationships between the venom proteins using the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT matrix-based model with 10,000 bootstrap replicates (Jones et al., 1992). A phylogenetic tree was generated to illustrate relationships between the venom proteins using the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT matrix-based model with 10,000 bootstrap replicates (Jones et al., 1992). A phylogenetic tree was generated to illustrate relationships between the venom proteins using the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT matrix-based model with 10,000 bootstrap replicates (Jones et al., 1992).

3. Results

Collection data and the respective hybrid index classifications of hybrid colonies are summarized in Table 1. Based on the hybrid index, colonies very near, or near S. richteri were collected from the central and western portions of Tennessee, and those very near, or near S. invicta from the central and eastern parts of the state (Fig. 1). As expected, all 20 hybrid colonies produced an LFA response for either Soli2 or Sol2 venom protein (Table 2). One colony, 212, produced a positive LFA response for both Soli2 and Sol2 venom proteins. A Soli2 response was reported in 7/10 (70%) of colonies classified by hybrid index as near, or very near S. richteri and Sol2 in 8/10 (80%) of colonies classified as near, or very near S. invicta (Fig. 2, Table 2). Overall, the LFA reported a response for Sol2 in 45% (9/20), Soli2 in 50% (10/20), and both Soli2 and Sol2 in 5% (1/20) of the hybrid colonies.

RT-PCR amplification of the Solenopsis venom protein 2 transcript from hybrid colonies revealed two amplicons (596 and 496 nucleotides [nts]) from colonies 9, 37, and 316, and a single amplicon (596 nts) from the remaining colonies (Fig. 3). The amplicons were not extended by 5′ RACE. Sequence analysis of the amplicons from these hybrid colonies,
S. richteri, and S. invicta parents revealed one nucleotide sequence with 100% identity to Soli2 (L09560) and three unique sequences, which we identify as Solenopsis hybrid venom protein 2 (Solh2; GenBank Accession MT150127), Solenopsis hybrid truncated venom protein 2 (Solh2Tr97; GenBank Accession MT150129), and Solenopsis richteri A69 venom protein 2 (Solr2A69; GenBank Accession MT150128; Fig. 4). The nucleotide sequence for Solr2 was never published, so we sequenced it from Solenopsis richteri, which has been deposited in the GenBank database (MT150125). For clarity, Solr2 P35776 refers to the originally reported amino acid sequence for Solr2 (Hoffman, 1997), Solr2 MT150125 refers to the Solr2 predicted amino acid sequence from Solenopsis richteri ants obtained in this study, and Solr2A69 refers to the Solr2 amino acid sequence from hybrid ants obtained in this study. The nucleotide sequence for Solr2A69 was identical (100%) to the Solr2 MT150125 sequence from hybrids. However, the amino acid sequence for Solr2 P35776 (Hoffman, 1997) is different from the Solr2A69 and Solr2 MT150125 sequences by a single residue. Specifically, at amino acid position 69, Solr2A69 and Solr2 MT150125 possess an alanine, while Solr2 P35776 has an aspartic acid residue (Fig. 4). While this discrepancy is possibly the result of a sequencing error in the original work (Hoffman, 1997), we have to assume that the original sequence was correct. Solr2A69 was sequenced 26 times from 7 different hybrid colonies and all exhibited sequences identical to the current parent (i.e., Solr2, which we also sequenced MT150125). Not a single nucleotide sequence would have predicted the translated amino acid sequence published in the original version, P35776. Because this disparity cannot

Table 2
Lateral flow immunoassay (LFA) responses of Solenopsis invicta x Solenopsis richteri hybrid worker ant colonies and the Solenopsis venom protein 2 form detected in each colony (indicated by gray box). LFA responses were determined for five colonies of each hybrid classification and the venom protein isoform was determined in three, which are designated by bold font.

<table>
<thead>
<tr>
<th>Nest designation</th>
<th>Hybrid index classification</th>
<th>LFA response</th>
<th>Venom protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Soli2&lt;sup&gt;P35776&lt;/sup&gt;</td>
</tr>
<tr>
<td>9</td>
<td>Very near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Very near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>Very near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>Very near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>Very near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>Near S. richteri</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>274</td>
<td>Near S. richteri</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>316</td>
<td>Near S. richteri</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>326</td>
<td>Near S. richteri</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>157</td>
<td>Near S. invicta</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>193</td>
<td>Near S. invicta</td>
<td>S. richteri</td>
<td></td>
</tr>
<tr>
<td>344</td>
<td>Near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>367</td>
<td>Near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>377</td>
<td>Near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>162</td>
<td>Very near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>212</td>
<td>Very near S. invicta</td>
<td>S. hybrid</td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>Very near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>337</td>
<td>Very near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>Very near S. invicta</td>
<td>S. invicta</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Map of Tennessee and separate county borders. Black dots show locations where S. invicta x S. richteri hybrid colonies were collected. The number near the dot is the colony designation. Chemotaxonomic (i.e., Hybrid Index) classifications (i.e., very near S. richteri, near S. richteri, near S. invicta, and very near S. invicta) for the colonies are illustrated by colors as shown in the key.
be resolved, we treat the different sequence obtained in our study as distinct from Solr2 P35776. In other words, the Solr2\textsuperscript{A69} amino acid sequence is treated as a unique sequence, distinct from the Solr2 P35776 sequence by a single residue.

The Solh2 (MT150127) sequence was detected in three hybrid ant colonies (i.e., 344, 377, and 420). Nucleotide and amino acid identities between Solh2 and Sol2 were 94% and 85%, respectively, and identities between Solh2 and Solr2 were 93 and 80% (Table 3). Thus, Solh2 is a new Solenopsis venom protein 2 orthologue apparently unique to hybrid ants.

The Solh2\textsuperscript{Tr97} (MT150129) sequence was also unique to hybrid ants. Solh2\textsuperscript{Tr97} is a truncated version of Solh2 (Fig. 4b) that exhibited a 100 nt deletion (nts 98–197) at the 5′ end of the transcript when compared with the other non-truncated venom protein 2 transcripts (Fig. 4b).

Database searches identified a genomic region containing two genes (hereafter G1 and G2 to denote the origin from S. invicta genome assemblies) that could produce transcripts with significant amino acid identity to venom protein 2 proteins Solh2 (Solh2 G1: 4,937,317–4,940,792 bp) and Solr2\textsuperscript{A69}/Solh2\textsuperscript{Tr97} (Solh2 G2: 4,926,521–4,933,021 bp), which are found on the complementary strand of the S. invicta chromosome 10 (BioProject PRJNA 442367; genome assembly GCA_009650705.1). These genes were not annotated in the genome assembly and transcripts were not available in the databases. In addition, predicted polypeptide XP_025991263.1 in the unplaced genomic Si.gnH scaffold 00053 (877,402–887,074 bp) of the genome assembly GCF_000188075.2 of S. invicta (BioProject PRJNA 268798) showed high amino acid identity to Soli2 G2 and Solr2\textsuperscript{A69}. Examination of these genome assemblies indicated that Solenopsis venom protein 2 genes have five exons (Fig. 4a), except Solh2\textsuperscript{Tr97} which was alternatively spliced. The Solh2\textsuperscript{Tr97} transcript resulted from alternative splicing of exon 1 to exon 3, skipping the 100 nt exon 2 (Fig. 4b). Omission of exon 2 generated a stop site 15 codons downstream of the original methionine start site, but another start site 23 nt downstream of the original start codon apparently initiated translation of the truncated protein (Fig. 4b and c).

Western blotting of protein homogenates of ant abdomens from colonies 9 and 37 using mAb 6C9D7 revealed a large band with an empirically determined mean molecular mass of 13,851 ± 696 Da and a smaller band of 11,190 ± 1144 Da (Fig. 5). These mass values are similar to the predicted mass of Solr2\textsuperscript{A69} (i.e., 15,453 Da) and Solh2\textsuperscript{Tr97} (10,854 Da), which are the expressed transcripts identified by RT-PCR in colonies 9 and 37. Monoclonal antibody 6C9D7 recognizes the sequence NPAPAVVKENSKMC, which is identical in both Solr2\textsuperscript{A69} and Solh2\textsuperscript{Tr97}. Only a single large band is identified in S. richteri and S. invicta ants. When the same ant homogenate was probed with mAb 3H6B9, which recognizes the sequence IEAQQRLKDKIAEC near the N-terminus of the protein, only the single, larger band (i.e., Solr2\textsuperscript{A69}) was reported in the hybrids (colonies 9 and 37), S. richteri and S. invicta. The sequence IEAQQRLKDKIAEC is absent in the alternatively spliced/truncated Solh2\textsuperscript{Tr97}.

Heterozygous expression of venom protein 2 was observed in 6/12 (50%) of the hybrid colonies. All venom protein 2 forms were well represented, Solr2 (42%), Solr2\textsuperscript{A69} (58%), Solh2 (25%), and Solh2\textsuperscript{Tr97} (25%). In areas outside the hybrid zone, Soli2 and Solr2 are the only forms described in S. invicta and S. richteri, respectively (Dove and Hoffman, 1988; Hoffman et al., 1988; Schmidt et al., 1993). Nucleotide and translated ORF sequence identities ranged from 0.79 to 1, and 0.54 to 1, respectively. The Solh2 amino acid sequence exhibited greater identity with Soli2 (0.85) than Solr2\textsuperscript{A69} (0.8) (Table 3).

In gene phylogeny, 99% of the bootstrap replicates grouped Soli2 sequenced from hybrid ants (i.e., Soli2H) and Soli2 P35775 with the gene identified from chromosome 10 (Soli2 G1) of the S. invicta genome assembly GCA_009650705.1 (BioProject PRJNA 421367) and these three genes were in turn grouped with Solh2 in 70% of the bootstrap replicates. Therefore, Soli2, Soli2H, and Solh2 transcripts were most likely products of the Soli2 G1 found in the genome assembly GCA_009650705.1. Solr2, Solr2\textsuperscript{A69}, and Solh2\textsuperscript{Tr97} protein sequences were grouped together in 69% of the bootstrap replicates, which in turn was grouped in 96% of the bootstrap replicates with the sequences derived from a second gene (Solli2 G2 and XP_025991263.1) found in both genome assemblies GCF_000188075.2 and GCA_009650705.1. This suggests that Solr2, Solr2\textsuperscript{A69}, and Solh2\textsuperscript{Tr97} are most likely products of a gene orthologous to Soli2 G2 in S. invicta. Solh2 and Solh2\textsuperscript{Tr97} amino acid sequences indicate that their most common ancestors are Soli2 and Solr2, respectively (Fig. 6). Soli2 G1 was not annotated in the genome assembly GCF_000188075.2 (BioProject PRJNA 268798), but Soli2 G2 was present in both genome assemblies. Genome assembly errors may be the reason for the discrepancies present in the sequence of this genomic region. Similarly, polymorphisms observed between transcripts derived from the genome and those obtained by sequencing of cloned cDNA may also be due to pooling of DNA from a large number of individuals for genome sequencing.
Despite sympatry, *S. invicta* and *S. richteri* do not form a hybrid in their native range of South America (Cohen and Privman, 2019), but for unknown reasons, a viable hybrid has formed in the introduced U.S. range (Shoemaker et al., 1996). The hybrid was first detected in the U.S. in 1985 where the two species were sympatric (Vander Meer et al., 1985). By 2006 in Tennessee, surveys showed that the hybrid had expanded its range to include the central and eastern parts of the state (Oliver et al., 2009). Indeed, 55% of statewide samples in Tennessee were classified as hybrids (Oliver et al., 2009). More recently, Pandey et al. (2019) reported that hybrids were found nearly exclusively in the eastern part of the state, predominated in the central portion, and were more common in west Tennessee than the survey from a decade earlier (Oliver et al., 2009). Thus, the hybrid zone is obviously stable and expanding.

While the *Solenopsis* fire ant venom contains a complex mixture of proteins (dos Santos Pinto et al., 2012), *Solenopsis* species venom protein 2 is a principal venom component and exhibits sufficient sequence differences to discriminate *Solenopsis* congeners, including *S. invicta*, *S. richteri*, *S. geminata*, *S. saevissima*, *S. xyloni*, and *S. aurea* (Hoffman, 2010; Valles et al., 2016). Indeed, lateral flow immunoassays have been successfully developed to detect and discriminate *S. richteri* from *S. invicta* ants in the field using antibodies created toward this venom protein (Valles et al., 2016, 2018). Each parent species was known to express only one form of this venom protein (i.e., Soli2 [P35775] in

![Diagram](image-url)
S. invicta and Solr2 (P35776) in S. richteri) (Hoffman, 2010). As anticipated, hybrid colonies were shown to express a parent venom protein transcript (i.e., Soli2 or Solr2 \( \approx \) Solr2\textsuperscript{A69}). However, two new additional venom protein 2 forms were identified and found to be unique to the hybrid, Solh2 and Solh2\textsuperscript{T97}.

Interestingly, the four variants of Solenopsis venom protein 2 (Soli2, Solr2\textsuperscript{A69}, Solh2, and Solh2\textsuperscript{T97}) apparently originate from two distinct genomic locations (Genes 1 and 2) and their similar identities and genomic architectures suggest that these are duplicated genes. Gene duplication has been observed in the venoms of snakes (Giorgianni et al., 2020) and spiders (Haney et al., 2019) providing a mechanism for increasing the diversity of these proteins. Alternative splicing, which also increases protein diversity (Graveley, 2001) was observed in the truncated Solh2\textsuperscript{T97}. In the rattlesnake, Crotalus atrox, a 30-gene complex was traced to a single conserved ancestral venom gene through diversification by alternative splicing (Giorgianni et al., 2020).

Gene 1 encodes Solh2 and Soli2, while Gene 2 encodes Solr2\textsuperscript{A69} and the alternatively spliced variant, Solh2\textsuperscript{T97}. Solr2\textsuperscript{A69} and Soli2 originate from the parent species, S. richteri and S. invicta, respectively, while Solh2 and Solh2\textsuperscript{T97} appear unique to hybrids. The Soli2 transcript obtained from hybrid ant colonies (i.e., Soli2H, MT150126) exhibited 100% nucleotide and predicted amino acid identities with the original sequence for Soli2 (P35775) (Schmidt et al., 1993). Solr2\textsuperscript{A69} differs from the original Solr2 sequence (P35776) by a single amino acid residue at position 69 (aspartic acid replaced with alanine). A nucleotide sequence for the original Solr2 accession was never reported or archived—only an amino acid sequence (i.e., P35776) is available for comparison. Therefore, there is no way to compare the nucleotide sequence of Solr2\textsuperscript{A69} with the original Solr2, which, presumably, predicted P35776. We are confident that no sequencing error occurred in our studies because the Solr2\textsuperscript{A69} transcript was sequenced 26 times from 7 different hybrid colonies and all sequences were identical to the parent (i.e., Solr2, which we also sequenced, MT150125). Not a single transcript sequence would have the predicted the originally published P35776 amino acid sequence. However, we must assume that the original sequence was correct. It is possible that the sequence has mutated since it was first published. Unfortunately, the discrepancy cannot be resolved.

Phylogenetic analysis showed that Solh2 assorts more closely with Soli2 than Solr2, which was anticipated based on the shared genomic origins of gene 1. LFA analysis further supported the close association of Soli2 and Solr2. While the majority (10/12) of the hybrid colonies expressed at least one parent venom protein transcript (i.e., Soli2 or Solr2), colonies 377 and 420 were apparently homozygous for Solh2. Colonies 377 and 420 produced a positive LFA response for S. invicta indicating antibody cross reactivity between Soli2 and Solh2. Lastly, a small conserved sequence near the N-terminal region (i.e., aa 21–27) of Solh2 exhibits 100% identity with Soli2. Indeed, this small region exhibits a unique sequence in each parent species protein (NKEKLII for S. invicta/Soli2 and IEAQRVL for S. richteri/Solr2 = Solr2\textsuperscript{A69}) and forms the basis for discriminating S. invicta from S. richteri (Valles et al., 2016). The Solenopsis venom protein 2 variants, Solh2 and Solh2\textsuperscript{T97} appear to be unique to the hybrid form. Neither sequence has been reported in either parental species or was either sequence detected in the S. invicta genome. These hybrid-specific genes (Solh2 and Solh2\textsuperscript{T97}) that are expressed and translated, and the ever-increasing range expansion of the viable hybrid provide evidence for the possibility that speciation of the S. invicta x S. richteri hybrid is occurring. The hybrids do appear to be better adapted to the colder climate of Tennessee giving them a distinct genetic advantage over either parent (Pandey et al., 2019). However, only a single venom protein was examined, so it is not possible to state this represents a speciation event. Still, our findings would suggest some differentiation is occurring between the parent species and hybrid, at least regarding Solenopsis venom protein 2.

It is possible that the Solh2 transcript and its alternatively spliced transcript Solh2\textsuperscript{T97} found in hybrids are produced by a gene that is a crossover of corresponding S. invicta and S. richteri genes. The cross-over notion is supported by unique sequence regions from each parental species and represented in Solh2. While a large proportion of the parent proteins exhibit identity (Fig. 4), there are regions unique to each protein. As mentioned, the short sequences near the S' end (beginning at amino acid residue 21) are clearly unique to Soli2 (NKEKLII) and Solr2 (IEAQRVL). This short sequence region provides the clearest difference between Solenopsis invicta and S. richteri. In fact, the sequence affords the ability to discriminate the different species (Valles et al., 2018). Conversely, nine amino acid residues were present in Solh2 that were distinct from either parent, including 42 A, 65 F, 77C, 78 L, 92 K, 95 M, 109 R, 115I, and 121 A. Interestingly, despite gene duplication, alternative splicing, multiple loci, and sequence diversity observed in the hybrid form, the Sol2 sequence has not changed and Solr2 may, or may not have, changed at a single amino acid residue. This stability suggests that Soli2 and Solr2 venom proteins are essential and perhaps have functions beyond competitor deterrence or prey incapacitation. Although Solenopsis species venom proteins are known to elicit strong immune responses in mammals, the exact function of these proteins is unknown (Srisong et al., 2018).

Differential expression of the different venom protein 2 genes in hybrids and bona fide species may be due to epigenetic regulation of these genes at the genome level (e.g. methylation). In order to properly
assess these expression differences, studies on areas such as genome organization and epigenetic gene regulation, which are beyond the scope of this study, are needed. Targeted sequencing of the venom protein 2 loci in S. invicta, S. richteri, and S. invicta x S. richteri hybrids may be needed to fully resolve the issue of differential gene expression in hybrids.

Ethical statement

The authors adhered to Elsevier’s publishing ethics policy and ethical guidelines for journal publication.

Author contribution

Steven M. Valles: Conceptualization, Investigation, Methodology, Visualization, Writing – original draft. Jason B. Oliver: Investigation, Writing – review & editing, Investigation. Karla M. Addesso: Investigation, Methodology, Writing – review & editing. Omaththage P. Perera: Investigation, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank C.A. Strong (USDA-ARS, Gainesville, FL) and Calvin Pierce (USDA-ARS, Stoneville, MS) for technical assistance, and Joshua Basham (TSU) for the ant photographs in Fig. 5. The use of trade, firm, or corporation names in this publication are for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

References


