

Article

Evidence for Two Soybean Looper Strains in the United States with Limited Capacity for Cross-Hybridization

Rodney N. Nagoshi ^{1,*}, Jeffrey A. Davis ², Robert L. Meagher ¹, Fred R. Musser ³, Graham P. Head ⁴, Hector Portillo ⁵ and Henry Teran ⁶

¹ Center for Medical, Agricultural and Veterinary Entomology, United States Department of Agriculture, Agricultural Research Service, Gainesville, FL 32608, USA; rob.meagher@usda.gov

² Department of Entomology, LSU Agricultural Center, 404 Life Science Building, Baton Rouge, LA 70803, USA; jeffdavis@agcenter.lsu.edu

³ Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Mississippi State, MS 39762, USA; fmusser@entomology.msstate.edu

⁴ Bayer Crop Science US, Chesterfield, MO 63017, USA; graham.head@bayer.com

⁵ FMC Agricultural Solutions, Stine Research Center, Newark, DE 19711, USA; hector.portillo@fmc.com

⁶ Corteva Agriscience™, Carr #3 Km 156.5, Salinas, PR 00751, USA; henry.teransantofimio@corteva.com

* Correspondence: rodney.nagoshi@usda.gov

Abstract: The noctuid moth soybean looper (SBL), *Chrysodeixis includens* (Walker) is an economically important pest of soybeans (*Glycine max* (L.) Merr.) in the southeastern United States. It has characteristics that are of particular concern for pest mitigation that include a broad host range, the capacity for annual long-distance flight, and resistance in some populations to important pesticides such as pyrethroids and chitin synthesis inhibitor. The biology of SBL in the United States resembles that of the fellow noctuid fall armyworm (FAW), *Spodoptera frugiperda* (J.E. Smith), a major pest of corn and several other crops. FAW exhibits a population structure in that it can be divided into two groups (host strains) that differ in their host preferences but are broadly sympatric and exhibit incomplete reproductive isolation. In this paper, strategies used to characterize the FAW strains were applied to SBL to assess the likelihood of population structure in the United States. Evidence is presented for two SBL strains that were defined phylogenetically and display differences in the proportions of a small set of genetic markers. The populations exhibit evidence of reproductive barriers sufficient to allow persistent asymmetry in the distribution of mitochondrial haplotypes. The identified molecular markers will facilitate studies characterizing the behaviors of these two populations, with relevance to pest mitigation and efforts to prevent further dispersal of the resistance traits.

Keywords: soybean looper; migration; haplotypes; fall armyworm; strains; genetic structure

Citation: Nagoshi, R.N.; Davis, J.A.; Meagher, R.L.; Musser, F.R.; Head, G.P.; Portillo, H.; Teran, H. Evidence for Two Soybean Looper Strains in the United States with Limited Capacity for Cross-Hybridization. *Genes* **2023**, *14*, 1509. <https://doi.org/10.3390/genes14071509>

Academic Editor: Nico M. Van Straalen

Received: 16 June 2023

Revised: 19 July 2023

Accepted: 21 July 2023

Published: 24 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Phytophagous insects can be classified into three categories based on feeding behavior: monophagous insects feed on a single plant species, oligophagous insects feed on several species within closely related families, and polyphagous insects have the ability to feed across higher phylogenetic groupings. Polyphagous agricultural pests capable of using a multiplicity of plant species present added challenges for mitigation strategies that include an increased number of habitats that can support pest populations. This makes extrapolating the source of migratory populations more difficult and complicates projections of future infestations. Fortunately, a more extreme phytophagy is relatively rare, perhaps because of unfavorable competition with more specialized, better-adapted species [1]. This would restrict the host range over time, with one possible outcome being a form of oligophagy where a few plant species are adapted to as primary hosts and are

preferred by the pest, while it retains some capacity to use a wider range of secondary host species [2].

The soybean looper (SBL), *Chrysodeixis includens* (Walker), and fall armyworm (FAW), *Spodoptera frugiperda* (J.E. Smith), are noctuid moth species that appear to exhibit this type of oligophagous feeding behavior. SBL has a broad host range that includes over 70 plant species but exhibits the most economic damage as a defoliator of soybean, while also causing significant yield losses in cotton, bean, and tomato crops [3]. FAW has a similarly broad and partially overlapping host range that is in part due to the existence of two populations that differ in host usage [4,5]. The C-strain is preferentially found on corn and sorghum while the R-strain predominates in millet and pasture grasses [6,7]. This may be an example of speciation coincident with substantial gene flow similar to that described in the Lepidoptera species complex *Heliconius*, where divergence in some parts of the genome occurred contemporaneously with persistent admixture [8].

There is also evidence for at least two genetically distinct populations of SBL although the data are more suggestive than conclusive. Collections in Brazil were characterized via inter-simple sequence repeat (ISSR) analysis, in which geographical populations were compared in terms of microsatellite-based polymorphisms [9]. Principal coordinate analysis revealed low but significant genetic differentiation between two groups that were sympatric, showed no geographical specificity, and were associated with a high gene flow and low genetic diversity. It was suggested that the two ISSR groups may be the result of a recent recolonization of SBL into Brazil [9]. This would presumably be a transient differentiation that will disappear as the populations continue to cross-hybridize. A subsequent study of Brazilian SBL using a combination of mitochondrial haplotypes and nuclear SNPs came to similar conclusions of overall low genetic diversity with little genetic structure [10]. However, genetic differences were detected between specimens collected from soybean and cotton in one location, suggesting a possible host-driven divergence of populations, but this difference was not observed in collections from a second location. A limitation of both studies is that sampling was from a single year, so the persistence of the observations is not known.

In this study, we tested for genetically differentiated groups in the North American SBL population. This was carried out by applying a methodology developed and successfully used to detect and characterize the FAW strains in wild populations. The strategy is based on two premises. The first is that while moderate gene flow resulting from incomplete reproductive isolation may not allow population-specific mitochondrial haplotypes, differences in the relative proportions of haplotypes might form and be detectable. The second is that the Z-chromosome may play a disproportionate role in speciation [11,12] with supporting evidence documented in several Lepidoptera species [13,14] including FAW [15]. Particularly relevant to this study are reports that polymorphisms in the Z-linked *Triosephosphate isomerase* (*Tpi*) gene can differentiate closely related moth populations such as the FAW host strains [16], the sibling *Helicoverpa* species of *H. armigera* and *H. zea* [17], and the E and Z pheromone strains of *Ostrinia nubilalis* [18]. The *Tpi* product is a housekeeping function and so is unlikely to drive population divergence, leaving its physical linkage to other Z-chromosome functions as the most likely explanation for its effectiveness as a genetic marker.

Based on these observations, SBL specimens were collected from multiple locations in the U.S. over multiple years and characterized for a segment of the mitochondrial *Cytochrome oxidase subunit I* (*COI*) gene previously used to differentiate geographical FAW populations [19], and a highly variable intron segment from the SBL *Tpi* gene, the analog of which in FAW produced a phylogenetic description of the two host strains [17,20]. Evidence was found for two SBL populations with restricted capacity to cross-hybridize.

2. Materials and Methods

2.1. Specimen Collections

Specimens were collected from soybean in field plots either through pheromone trapping (adult males) or as larvae picked directly from soybean (Table 1A). In addition, larvae collected from soybean were reared to produce laboratory colonies, and specimens from these collections were obtained after the first (F1) or second (F2) generation (Table 1B). The collected specimens were stored in ethanol and initially identified as SBL according to morphological criteria before molecular analysis. The gender of the larval and colony specimens was not identified.

Table 1. Source information for SBL collections. All collections from soybean.

Collection	State	Year	Type	Collector
A. Field collections				
AL2017	Alabama	2017	Pheromone trap	R. Meagher
FL2017	Florida	2017	Pheromone trap	R. Meagher
FL2019	Florida	2019	Pheromone trap	R. Meagher
MS2019a	Mississippi	2019	Pheromone trap	R. Meagher
TX2017	Texas	2017	Pheromone trap	R. Parker
KS2017	Kansas	2017	Pheromone trap	B. McCornack
PR2018	Puerto Rico	2018	Larvae, field	H. Portillo
PR2019	Puerto Rico	2019	Larvae, field	H. Teran
B. Colony collections				
AL2016	Alabama	2016	F1 generation	J. Davis
GA2016	Georgia	2016	F1 generation	J. Davis
LA2016	Louisiana	2016	F1 generation	J. Davis
MS2016	Mississippi	2016	F1 generation	J. Davis
NC2016	N. Carolina	2016	F1 generation	J. Davis
SC2016	S. Carolina	2016	F1 generation	J. Davis
AL2019	Alabama	2019	F1-2 generation	F. Musser, B. Catchot
LA2019	Louisiana	2019	F1-2 generation	F. Musser, B. Catchot
MS2019b	Mississippi	2019	F1-2 generation	F. Musser, B. Catchot
PR2020	Puerto Rico	2020	F1-2 generation	F. Musser, B. Catchot
SC2019	S. Carolina	2019	F1-2 generation	F. Musser, B. Catchot
TN2019	Tennessee	2019	F1-2 generation	F. Musser, B. Catchot
VA2019	Virginia	2019	F1-2 generation	F. Musser, B. Catchot

2.2. DNA Preparation

A mix of genomic and mitochondrial DNA from individual specimens was purified from individuals as described in [21], and stored in TE (10 mM Tris (pH 8), and 1 mM EDTA) at $-20\text{ }^{\circ}\text{C}$. The procedure involved using Genomic Lysis Buffer and Zymo-Spin III columns (Zymo Research, Orange, CA, USA), following the manufacturer's instructions. DNA preparation was increased to a final volume of 100 μL with distilled water. Genomic DNA preparations were stored at $-20\text{ }^{\circ}\text{C}$ and analyzed as needed.

2.3. Isolation of the DNA Segments by PCR

Segments of the SBL *COI* gene sequence were identified by aligning the well-characterized FAW *COI* sequences with the SBL mitochondria genome sequence (GenBank LR797832). The primers used to amplify the sCOIB segment were sC891f (5'-TACAC-GAGCTTATTTTACTTC-3'), and sC1457r (5'-ATATCATTCAATAGAAGAGG-3'). PCR amplification was performed in a 30 μL reaction mix containing 3 μL of 10 \times manufacturer's reaction buffer, 1 μL of 10 mM dNTP, 0.5 μL of a 20 μM primer mix, 1 μL of the

DNA template (between 0.05–0.5 µg), and 0.5 units of Taq DNA polymerase (New England Biolabs, Beverly, MA, USA). The thermocycling program was 94 °C (1 min), followed by 33 cycles of 92 °C (30 s), 56 °C (30 s), 72 °C (30 s), and a final segment of 72 °C for 3 min. Briefly, 96 PCR amplifications were performed at the same time using either 0.2 mL tube strips or 96 well-microtiter plates.

The sTpi1140 segment was amplified by the primer pair sTpi891f/sTpi1140r using the same protocol as that used for sCOIB. All primers used for PCR and DNA sequencing were synthesized by Integrated DNA Technologies (Coralville, IA, USA). These included sC891f (5'-TACACGAGCTTATTTTACTTC-3'), sC1457r (5'-ATATCATTCAATAGAA-GAGG-3'), sTpi412f (5'-ATGGCCTGAAAGTCATTGCCTG-3'), and sTpi1140r (5'-GCAG-ACACATTCTTAGCCAGCC-3').

For fragment isolations, 6 µL of 6× gel loading buffer was added to the amplification reaction and the entire sample run on 1.8% agarose horizontal gel containing GelGreen (Biotium, Hayward, CA, USA) in 0.5× Tris-borate buffer (TBE, 45 mM Tris base, 45 mM boric acid, and 1 mM EDTA; pH 8.0). Fragments were visualized on a Blue LED transilluminator (Thermo Fisher Scientific, Waltham, MA, USA) and cut out from the gel. Fragment isolation was performed using Zymo-Spin I columns (Zymo Research, Orange, CA, USA) in accordance with the manufacturer's instructions.

2.4. Characterization of Haplotypes

Each specimen was anticipated to have a single *COI* haplotype that was maternally inherited. The sCOIB haplotype classes used in this study are defined by SNPs sC1035 and sC1272 that are each polymorphic for C/T and produce four observed combinations (C₁₀₃₅C₁₂₇₂, C₁₀₃₅T₁₂₇₂, T₁₀₃₅C₁₂₇₂, and T₁₀₃₅T₁₂₇₂), with C₁₀₃₅T₁₂₇₂ and T₁₀₃₅C₁₂₇₂ being the predominant forms. To simplify the analysis, the sCOIB mitochondrial haplotype categories in this paper were defined solely by sC1035 polymorphisms and designated as either C1035_{COI} or T1035_{COI} if a C or T was present, respectively.

The SBL *Tpi* haplotypes are more complicated as the location of the *Tpi* gene on the Z-chromosome means that two gene copies are present in males (Z/Z) and so heterozygous allele combinations are possible, while females (Z/W) carry only a single copy. Because DNA sequencing is performed on PCR fragments directly produced from the specimen, heterozygosity for the *Tpi* segment will result in the simultaneous sequencing of two different alleles. This will produce chromatographs with overlapping curves at the sequence mismatch. Particularly disruptive are heterozygous indels (insertions and deletions) because these can shift the DNA sequencing frame of the two DNA strands, resulting in continued mismatch downstream of the indel.

2.5. DNA Sequence Analysis, Statistics, and Data Availability

The isolated fragments were analyzed via DNA sequencing using the appropriate primers (Azenta, South Plainfield, NJ, USA). DNA comparisons and alignments were performed using the Geneious Pro 10.1.2 (<https://www.geneious.com>, accessed on 15 June 2023) [22]. All sCOIB haplotypes described in this study were deposited in GenBank, (accession nos. OR128567-OR128591) as were the collection of sTpi1140 sequences (OR124064-OR124621). Statistical analyses were conducted using GraphPad Prism version 7.00 for Mac (GraphPad Software, La Jolla, CA, USA). The generation of graphs was carried out using Excel (version 16.75.2) and Powerpoint (version 16.72, Microsoft, Redmond, WA, USA). Phylogenetic trees were produced using Geneious Pro 10.1.2 and PhyML 3.3.20180214 [23,24]. The analysis underwent bootstrap testing (100 replicates) with the optimal tree shown and drawn to scale. The evolutionary distances were computed using the maximum composite likelihood method [25] and diagrammed based on midpoint rooting.

3. Results

3.1. Characterization of the Mitochondrial sCOIB Segment

A pair of mitochondrial haplotypes were previously shown to detect a genetic structure in a survey of SBL in the U.S. [26]. These were identified from the 3' portion of the SBL mitochondrial *COI* gene, a fragment designated as sCOIB, from which nine SNPs were identified. Only two of the nine SNPs were polymorphic in more than one sample, sC1035 and sC1272, and these produced four pairwise combinations, with the C₁₀₃₅T₁₂₇₂ and T₁₀₃₅C₁₂₇₂ haplotype categories predominating (Figure 1A). Sequence analysis of 1,405 specimens revealed 25 variants that via maximum-likelihood phylogeny analysis could be separated into two groups based solely on the sC1035 polymorphism (Figure 1B). These are subsequently referred to as either the T₁₀₃₅_{COI} or C₁₀₃₅_{COI} haplotype depending on the presence of T₁₀₃₅ or C₁₀₃₅, respectively.

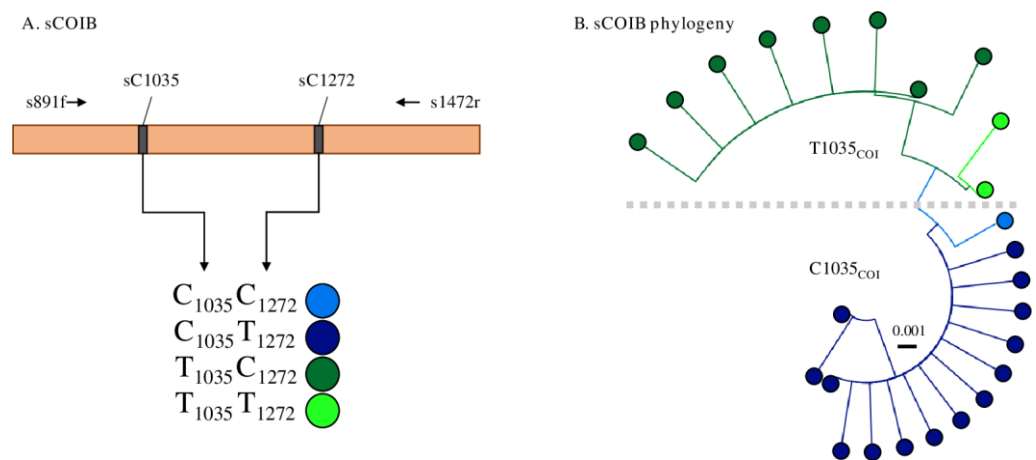


Figure 1. Description of the sCOIB SNPs and the T₁₀₃₅ and C₁₀₃₅ haplotypes. **(A)** Diagram of the sCOIB fragment with horizontal arrows indicating primers used for PCR amplification and DNA sequencing. The observed polymorphisms produced by the two most frequent polymorphic sites (sC1035 and sC1272) in the segments are listed. The pairings represent haplotypes and are color-coded for the rest of the figure. **(B)** Maximum-likelihood phylogeny of the sCOIB variants observed. Dashed line separates the T₁₀₃₅ and C₁₀₃₅ haplotypes.

3.2. Linkage between *i65del_{Tpi}* and T₁₀₃₅_{COI}

A highly polymorphic intron segment from the Z-linked *Tpi* gene had previously been used to study the FAW host strains [17,20], and the analogous fragment from SBL, sTpiI140, was recently isolated and characterized (Figure 2A). A 7-bp deletion, designated as *i65del_{Tpi}*, was found in the sTpiI40 segment that was present in 12% of the pooled collections (Figure 2A).

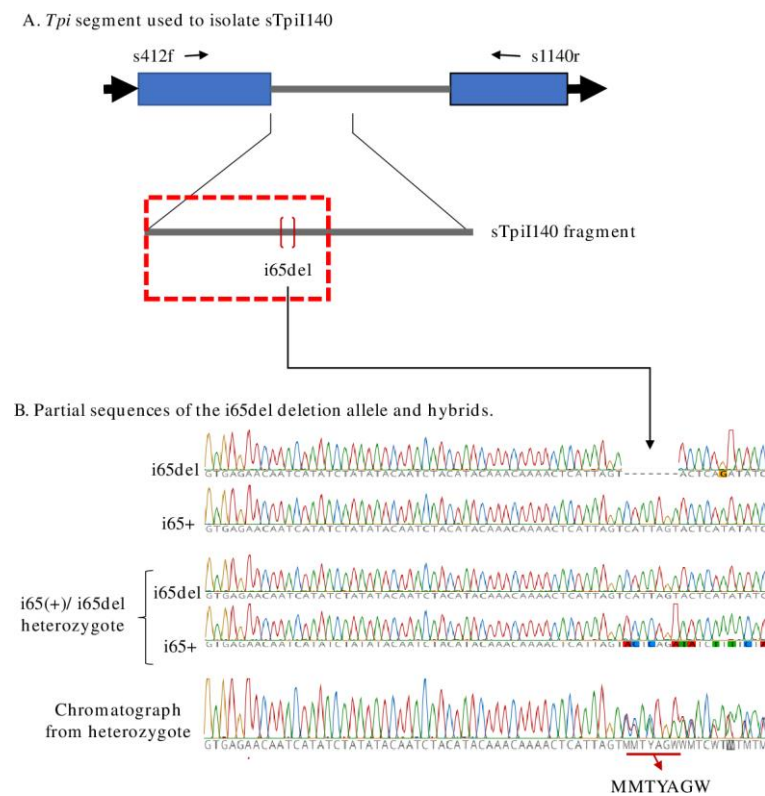


Figure 2. Description of the sTpiI140 segment from the SBL *Tpi* gene. (A) Map of portion of the SBL *Tpi* gene showing the two exons (blue boxes) separated by the intron containing the sTpiI140 fragment. One sTpiI140 variant contains a 7-bp deletion. Thin arrows depict the approximate location of primers used for the PCR amplification of sTpiI140. (B) Sequence of the sTpiI140 portion outlined by the red box that shows the consensus sequences associated with the i65del and i65+ alleles. In i65del/i65+ heterozygotes, DNA sequencing of the amplified *Tpi* segment using the s412f primer produces a frameshift in the chromatograph beginning at the start of the i65del deletion that starts with the diagnostic MMTYAGW sequence.

Comparisons of the distribution of the sCOIB mitochondrial haplotypes and sTpiI140 in our collections uncovered evidence of a possible linkage between i65del_{Tpi} and T1035_{COI}. This was unexpected given that these markers reside on different genetic elements, the Z-chromosome for the former and the mitochondrial genome for the latter. This can be visualized via a direct comparison of the frequencies of T1035_{COI} and i65del_{Tpi} for each of the SBL collections (Figure 3). The similarity of the overlapping curves was confirmed via the test for correlation (Pearson's $r = 0.630$; $p = 0.002$). The data indicate that the T1035_{COI} mitochondrial haplotype and i65del_{Tpi} allele segregate in combination at frequencies too high for what would be predicted from random mating.

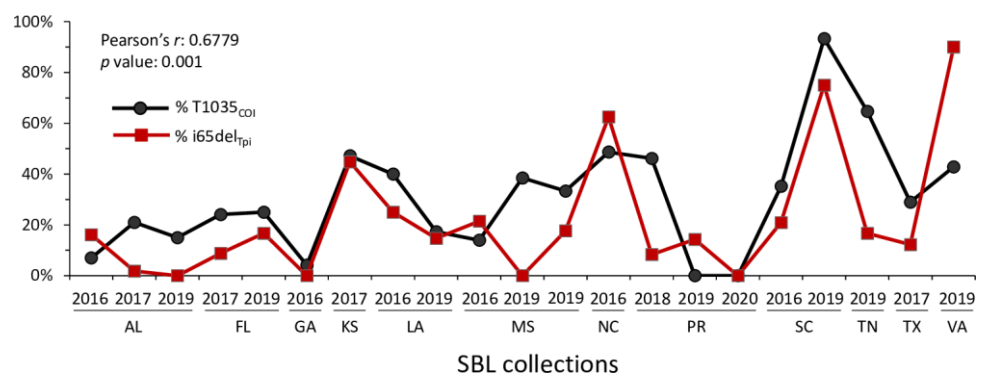


Figure 3. Frequencies of the T1035_{COI} mitochondrial haplotype and the i65del_{Tpi} allele in the SBL collections (Table 1). The test for correlation using Pearson's r is indicated.

Additional statistical support came from organizing the *i65del*_{Tpi} allele frequencies by state and comparing its frequencies in the C1035_{COI} and T1035_{COI} populations. Because the colony and field collections differed with respect to artificial rearing and source locations, the two data sets were analyzed separately as well as in combination. In each analysis, states with fewer than ten specimens of either the C1035_{COI} or T1035_{COI} haplotypes were pooled. Both the colony and field-collected specimens showed a significantly higher association of the *i65del*_{Tpi} allele with the T1035_{COI} mitochondrial haplotype (Table 2A,B).

Table 2. Comparisons of *i65del*_{Tpi} allele frequencies in the two sCOIB haplotype populations in the colony (A), field (B), and combined collections (C). Frequencies are in percentages, “n” refers to sample numbers, and an asterisk indicates that the observed difference between the mean frequencies is statistically significant.

Collection	n	<i>i65del</i> _{Tpi} Frequency (%)		<i>t</i> -Test Scores
		C1035 _{COI}	T1035 _{COI}	<i>p</i> (<i>t</i> , <i>df</i>)
A. Mean Colony ¹	410	18 ± 5	47 ± 11	<i>p</i> = 0.007 * (5.203, 4)
B. Mean Field ²	533	18 ± 5	27 ± 11	<i>p</i> = 0.037 * (3.075, 4)
C. Colony + Field				
AL	144	16	26	
FL	190	14	46	
KS	53	27	42	
LA	92	13	54	
MS	80	24	29	
NC	37	21	44	
SC	86	14	36	
TX	145	20	42	
GA-VA-TN-PR	116	13	48	
Mean ± sd		18 ± 5	41 ± 9	<i>p</i> = 0.004 * (5.808, 8)

¹LA2016, LA2019, MS2016, MS2019b, NC2016, SC2016, SC2019, and Pool (AL2016, AL2019, GA2016, PR2020, TN2019, and VA2019). ²AL2017, FL2017, FL2019, KS2017, TX2017, and Pool (MS2016, MS2019a, PR2018, and PR2019). *, significant difference.

Given the similarity in their results, the data from the colony and field collections were combined and the *i65del*_{Tpi} frequencies were then presented for each state, with the results for GA, VA, TN, and PR being pooled (Table 2C). Despite the considerable variability between states, the frequency of *i65del*_{Tpi} was consistently higher in the T1035_{COI} population, with the difference from the C1035_{COI} group ranging from a low one of 5% (MS) to as large as 32% (FL). As with the individual data set analyses, the observed difference between the means was statistically significant.

3.3. Frequency of *i65del*/₊ Heterozygotes

A test for potential reproductive barriers was possible because the methodology used to identify *i65del*_{Tpi} also allowed the detection of heterozygotes between *i65del*_{Tpi} and *i65+*_{Tpi} (designated as *i65del*/₊). The presence of both alleles in the same sequencing reaction will produce a chromatograph with a predictable superimposition of curves resulting from the frameshift beginning at the *i65del* breakpoint (Figure 2B). The overlapped region will begin with the diagnostic MMTYAGW sequence (M = T/A, Y = C/T, and W = A/T per the IUPAC convention).

The frequency of heterozygotes was used to calculate Wright’s local inbreeding coefficient, *F*_{is}, which is a metric for quantifying observed heterozygosity (*H*_o) relative to what would be expected (*H*_e) from a population at the Hardy–Weinberg equilibrium, based on the equation $F_{is} = (H_e - H_o)/H_e$. A *F*_{is} equal to 0 indicates a state of equilibrium where *H*_e = *H*_o, with values becoming increasingly positive if heterozygosity is suppressed (high inbreeding) and more negative if the observed heterozygosity is greater than expected. An

added complication is that because *Tpi* is on the Z-chromosome, heterozygotes of two *Tpi* alleles can only occur in SBL males as females being ZW will only carry one copy of the gene. Our initial analysis was therefore limited to the pheromone trap collections that exclusively attracted adult males. The pooled data produced a F_{is} of 0.30 while the mean F_{is} of the six collections was 0.41 (Table 3A). Both highly positive values indicate the substantial suppression of hybrids. Only the Alabama collection (AL2017) gave a negative F_{is} (−0.11).

The subsequent analysis of the remaining collections (limited to those with sample numbers of > 20) gave similar results even though the gender of the specimens was not determined. The likely presence of female specimens reduces the opportunities for i65del/+ hybrids and therefore should lead to overestimates of F_{is} . The results are consistent with this reasoning with both pooled and mean F_{is} values being higher than that observed with the pheromone trap collections (Table 3B). In the GA2016 collection, the i65del_{Tpi} allele was not found either as a homozygote or heterozygote and so a calculation of F_{is} could not be performed. Otherwise, the larva results support the findings from the pheromone trap data set as a consistently positive F_{is} was observed that was indicative of suppressed heterozygosity.

Table 3. Comparisons of heterozygosity. Numbers of relevant genotypes and calculations of Wright’s inbreeding coefficient, F_{is} , for all collections with sample sizes greater than 20.

A. Male adults				
Collection	Genotype (number)			F_{is}
	i65del/i65del	+/+	i65del/+	
AL2017	1	56	25	−0.11
FL2017	11	125	26	0.36
FL2019	4	24	0	1.00
KS2017	13	29	11	0.54
MS2019a	3	17	4	0.49
TX2017	11	90	44	0.14
Total	43	341	109	0.30
			Mean	0.41

B. Larvae of unknown gender				
Collection	i65del/i65del	+/+	i65del/+	F_{is}
	or i65del/W	or +/W		
AL2016	5	31	7	0.49
GA2016	0	24	0	nd ¹
LA2016	7	28	5	0.65
LA2019	5	34	13	0.27
MS2016	6	28	9	0.43
NC2016	10	16	11	0.39
PR2019	3	21	3	0.60
SC2016	9	43	19	0.31
VA2019	9	10	2	0.81
Total	54	235	69	0.48
			Mean	0.49

¹ No i65del_{Tpi} alleles were detected so F_{is} could not be calculated.

3.4. The Phylogeny of the s*Tpi*1140 Segment

Neighbor-joining phylogenetic analysis of the SBL s*Tpi*1140 sequence from 550 specimens collected from multiple U.S. locations and times produced a tree that could be divided into two groups based on the distribution of i65del_{Tpi} allele (Figure 4A). The tree is unrooted because of the high variability of the intron segment that curtailed finding an appropriate outgroup. Therefore, the groups are designated as clans indicating a

phylogenetic split but with the direction of evolution unspecified. We identified one group as ClanB_{SBL}, which was made up of 66 individuals. The remaining sequences, designated as ClanA_{SBL}, emerge from a single branch out of ClanB_{SBL}, with 88% of the 484 sequences of a single haplotype. The i65del_{Tpi} allele was asymmetrically distributed between the two SBL groups, with 100% (66/66) carrying i65del_{Tpi} in ClanB_{SBL} compared to none out of 492 in ClanA_{SBL} (NJ, Table 4). Based on these observations, i65+_{Tpi} and i65del_{Tpi} are accurate molecular proxies for ClanA_{SBL} and ClanB_{SBL}, respectively. The sCOIB haplotypes also showed a disproportional distribution, though not to the same degree as that of the *Tpi* alleles. The T1035_{COI} haplotype made up 41% of ClanB_{SBL}, but only 17% of ClanA_{SBL}. Haplotype diversity was more than four-fold higher in ClanB_{SBL} than ClanA_{SBL} (0.935 vs. 0.229) and nucleotide diversity was eight-fold greater (0.024 vs. 0.003).

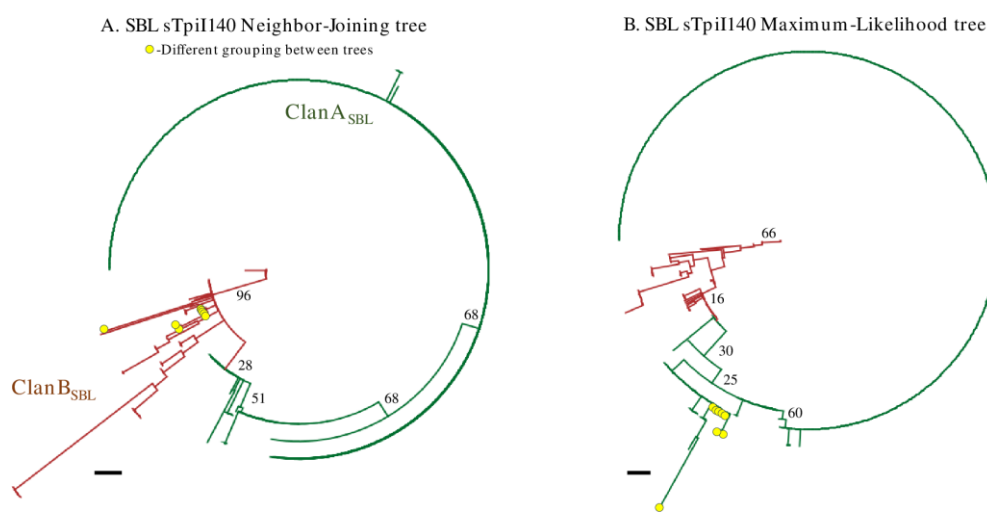


Figure 4. Phylogenetic analysis of the SBL sTpiI140 segment identifying taxonomic groups such as those found using the analogous fTpi100 segment from FAW. (A) Neighbor-joining phylogenetic tree produced from 550 SBL sTpiI140 sequences. The ClanA_{SBL} and ClanB_{SBL} groups are identified by color. Numbers indicate bootstrap frequency. (B) Maximum-likelihood treatment of same data set with colors indicating analogous groups. Ovals identify sequences found in the ClanB group in the neighbor-joining tree. Scale bar denotes a phylogenetic distance of 0.01 nucleotide substitutions per site.

Table 4. Metrics from neighbor-joining (N-J) and maximum-likelihood (M-L) phylogenetic analyses of the sTpiI140 intron segment (sd = standard deviation).

Metric	Tree Type	ClanA	ClanB
<i>Tpi</i> marker: i65del _{Tpi}	NJ	0%	100%
	ML	2%	98%
<i>COI</i> marker: T1035 _{COI}	NJ	17%	41%
	ML	18%	41%
Haplotype diversity (±sd)	NJ	0.229 ± 0.025	0.935 ± 0.014
	ML	0.194 ± 0.024	0.943 ± 0.011
Nucleotide diversity (±sd)	NJ	0.003 ± 0.000	0.024 ± 0.144
	ML	0.002 ± 0.001	0.026 ± 0.001

Because the neighbor-joining tree was poorly supported by bootstrapping, the same dataset was analyzed via maximum-likelihood analysis (Figure 4B). The results were similar with respect to the identification of groupings with asymmetric distributions of the i65del_{Tpi} and T1035_{COI} markers (ML, Table 4). Eight specimens were grouped differently with the two methods resulting in only minor variations in marker frequencies and genetic diversity metrics. Bootstrap support was again weak, indicating that the details of the phylogenetic pattern are uncertain. However, both tree building methods identified

two phylogenetic groups with similar differences in marker distributions, consistent with the identification of two distinct populations.

4. Discussion

The experimental strategy applied in this study on SBL was based on methods used to investigate the FAW host strains, with the results suggesting similarities in the population structure of the two species. One example came from comparisons of the two groups identified via the phylogenetic analysis of sTpiI140 (Figure 4) with that derived from a portion of the analogous *Tpi* intron in FAW that differentiates the C-strain and R-strain [20]. Like the SBL ClanA and ClanB, the phylogenetic groups associated with the FAW host strains differ in the frequencies of *Tpi* and *COI* markers as well as with haplotype diversity and nucleotide diversity [20]. The FAW strains are believed to diverge via a process that involves differences in host usage, and it is tempting to hypothesize that something similar may be happening with SBL, though we have no evidence yet of any population structure correlated with plant hosts or habitats. Testing for these will require systematic sampling from different plant hosts over multiple locations and times.

Additional evidence for two distinct SBL populations was derived from the fact that the nuclear Z-chromosome and the mitochondrial genome are separate genetic entities that segregate independently of the other. This means that under conditions of random mating, the i65del_{Tpi} allele should display no linkage with a mitochondrial haplotype and therefore should be found in equal proportions in the T1035_{COI} and C1035_{COI} populations. Our observations of linkage disequilibrium between these markers (Table 2) indicate non-random mating behavior that we believe is most simply explained by two populations with limited capacity for successful cross-hybridization, one disproportionately predominated by T1035_{COI} i65del_{Tpi} (ClanB_{SBL}) and the other predominated by C1035_{COI} i65+_{Tpi} (ClanA_{SBL}).

This suggestion is supported by our ability to identify i65del/+ heterozygotes, which allowed the quantification of the hypothesized reproductive restrictions between the i65del_{Tpi} and i65+_{Tpi} populations. The F_{is} calculations showed consistent evidence of high levels of inbreeding diagnostic of the Wahlund effect that describes a reduction in observed heterozygosity as populations diverge into multiple groups [27] (Table 3). The reduced frequency of cross-hybridization between i65del_{Tpi} and i65+_{Tpi} populations as indicated by the F_{is} data provides compelling explanation for the asymmetric distribution of the T1035_{COI} and C1035_{COI} mitochondrial haplotypes across these groups.

Given the similarities between FAW and SBL with respect to their overlapping host range, geographical distribution, and climate suitability, it is plausible that SBL is facing similar selective pressure to diverge into two populations. This can be difficult to demonstrate, however, as populations at intermediate stages of speciation are, by definition, still capable of some cross-hybridization with potentially high levels of gene flow. In addition, the genes driving divergence are likely to be low in number early in the speciation process and many of the polymorphisms that become asymmetrically distributed by chance will likely vary by region. Given these issues, our phylogenetic identification of ClanA_{SBL} and ClanB_{SBL} that seem by analogy to FAW to be good candidates for diverging populations represents an important advance in understanding SBL population structure.

In summary, evidence was presented based on the distribution of *COI* and *Tpi* polymorphisms for two sympatric SBL populations with only a limited capability for cross-hybridization. These findings complement results from a companion paper demonstrating geographical differences in haplotype distribution that may reflect how SBL migration is impacted by the timing of soybean agriculture. Together, they describe a more complicated population biology in the United States than previously thought with potential relevance to the dispersion of resistance traits and projections of SBL migration.

Author Contributions: Conceptualization, R.N.N. and G.P.H.; methodology, R.N.N.; formal analysis, R.N.N.; investigation, R.N.N., J.A.D., R.L.M., F.R.M., H.P., and H.T.; resources, R.N.N., J.A.D.,

R.L.M., F.R.M., H.P., and H.T.; data curation, R.N.N.; writing—original draft preparation, R.N.N.; writing—review and editing, R.N.N., G.P.H., J.A.D., R.L.M., F.R.M., H.P., and H.T.; visualization, R.N.N.; supervision, R.N.N.; project administration, R.N.N.; funding acquisition, R.N.N., and G.P.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a Trust Fund Cooperative Research Agreement with the Insecticide Resistance Action Committee (IRAC) of CropLife (58-6036-9-004).

Data Availability Statement: Relevant DNA sequence data are openly available in Genbank. All other data are present in this study.

Acknowledgments: J.M.G. Thomas is recognized for technical assistance in preparing the specimens and for comments about the manuscript. Thanks go to Roy Parker, Brian McCornack, and Beverley Catchot for the specimen collections. The use of trade, firm, or corporation names in this publication is 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.

Conflicts of Interest: The authors declare no conflicts of interest. The authors affiliated with the private companies Corteva, FMC Agricultural Solutions, and Bayer provided SBL specimens.

References

- Vandermeer, J.; Pascual, M. Competitive coexistence through intermediate polyphagy. *Ecol. Complex.* **2006**, *3*, 37–43. <https://doi.org/10.1016/j.ecocom.2005.05.005>.
- Stireman, J.O. The evolution of generalization? Parasitoid flies and the perils of inferring host range evolution from phylogenies. *J. Evol. Biol.* **2005**, *18*, 325–336. <https://doi.org/10.1111/j.1420-9101.2004.00850.x>.
- Specht, A.; de Paula-Moraes, S.V.; Sosa-Gomez, D.R. Host plants of *Chrysodeixis includens* (Walker) (Lepidoptera, Noctuidae, Plusiinae). *Rev. Bras. Entomol.* **2015**, *59*, 343–345. <https://doi.org/10.1016/j.rbe.2015.09.002>.
- Luginbill, P. *The Fall Armyworm*; U.S. Department of Agriculture Technical Bulletins: Philadelphia, PA, USA, 1928; Volume 34, pp. 1–91.
- Montezano, D.G.; Specht, A.; Sosa-Gomez, D.R.; Roque-Specht, V.F.; Sousa-Silva, J.C.; Paula-Moraes, S.V.; Peterson, J.A.; Hunt, T.E. Host plants of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in the Americas. *Afr. Entomol.* **2018**, *26*, 286–300. <https://doi.org/10.4001/003.026.0286>.
- Groot, A.T.; Unbehend, M.; Hänniger, S.; Juárez, M.L.; Kost, S.; Heckel, D.G. Evolution of reproductive isolation of *Spodoptera frugiperda*. In *Pheromone Communication in Moths: Evolution, Behavior, and Application*; Jeremy, D.A., Ring, T.C., Eds.; University of California Press: Berkeley, CA, USA, 2016.
- Nagoshi, R.N.; Meagher, R.L. The *Spodoptera frugiperda* host strains: What they are and why they matter for understanding and controlling this global agricultural pest. *J. Econ. Entomol.* **2022**, *115*, 1729–1743. <https://doi.org/10.1093/jee/toac050>.
- Martin, S.H.; Most, M.; Palmer, W.J.; Salazar, C.; McMillan, W.O.; Jiggins, F.M.; Jiggins, C.D. Natural selection and genetic diversity in the butterfly *Heliconius melpomene*. *Genetics* **2016**, *203*, 525–541. <https://doi.org/10.1534/genetics.115.183285>.
- Palma, J.; Maebe, K.; Guedes, J.V.C.; Smagghe, G. Molecular variability and genetic structure of *Chrysodeixis includens* (Lepidoptera: Noctuidae), an important soybean defoliator in Brazil. *PLoS ONE* **2015**, *10*, e0121260. <https://doi.org/10.1371/journal.pone.0121260>.
- Silva, C.S.; Cordeiro, E.M.G.; de Paiva, J.B.; Dourado, P.M.; Carvalho, R.A.; Head, G.; Martinelli, S.; Correa, A.S. Population expansion and genomic adaptation to agricultural environments of the soybean looper, *Chrysodeixis includens*. *Evol. Appl.* **2020**, *13*, 2071–2085. <https://doi.org/10.1111/eva.12966>.
- Coyne, J.A.; Orr, H.A. Patterns of Speciation in *Drosophila*. *Evolution* **1989**, *43*, 362–381. <https://doi.org/10.2307/2409213>.
- Haldane, J.B.S. Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* **1922**, *12*, 101–109. <https://doi.org/10.1007/Bf02983075>.
- Oppenheim, S.; Hopper, K. Genetics of host range in Lepidoptera. In *Molecular Biology and Genetics of the Lepidoptera*; Goldsmith, M.R., Marec, F., Eds.; CRC Press: Boca Raton, FL, USA, 2010; pp. 195–217.
- Sperling, F.A.H. Sex-linked genes and species-differences in Lepidoptera. *Can. Entomol.* **1994**, *126*, 807–818.
- Prowell, D.P. Sex linkage and speciation in Lepidoptera. In *Endless Forms: Species and Speciation*; Howard, D., Berlocher, S., Eds.; Oxford Press: New York, NY, USA, 1998; pp. 309–319.
- Nagoshi, R.N. The fall armyworm Triosephosphate isomerase (Tpi) gene as a marker of strain identity and interstrain mating. *Ann. Entomol. Soc. Am.* **2010**, *103*, 283–292. <https://doi.org/10.1603/An09046>.
- Nagoshi, R.N.; Meagher, R.L. Using intron sequence comparisons in the Triose-phosphate isomerase gene to study the divergence of the fall armyworm host strains. *Insect Mol. Biol.* **2016**, *25*, 324–337. <https://doi.org/10.1111/imb.12223>.
- Dopman, E.B.; Bogdanowicz, S.M.; Harrison, R.G. Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn borer (*Ostrinia nubilalis*). *Genetics* **2004**, *167*, 301–309. <https://doi.org/10.1534/genetics.167.1.301>.

19. Nagoshi, R.N.; Meagher, R.L.; Flanders, K.; Gore, J.; Jackson, R.; Lopez, J.; Armstrong, J.S.; Buntin, G.D.; Sansone, C.; Leonard, B.R. Using haplotypes to monitor the migration of fall armyworm (Lepidoptera: Noctuidae) corn-strain populations from Texas and Florida. *J. Econ. Entomol.* **2008**, *101*, 742–749.
20. Nagoshi, R.N. Observations of genetic differentiation between the fall armyworm host strains. *PLoS ONE* **2022**, *17*, e0277510. <https://doi.org/10.1371/journal.pone.0277510>.
21. Nagoshi, R.N.; Vizuete, J.L.A.; Murua, M.G.; Garces-Carrera, S. Comparisons of fall armyworm haplotypes between the Galapagos Islands and mainland Ecuador indicate limited migration to and between islands. *Sci. Rep.* **2021**, *11*, 3457. <https://doi.org/10.1038/s41598-021-83111-5>.
22. Drummond, A.J.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Duran, C.; Field, M.; Heled, J.; Kearse, M.; Markowitz, S.; et al. Geneious v5.5; Geneious: Auckland, New Zealand, 2010.
23. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. <https://doi.org/10.1093/sysbio/syq010>.
24. Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **2003**, *52*, 696–704. <https://doi.org/10.1080/10635150390235520>.
25. Tamura, K.; Nei, M.; Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11030–11035. <https://doi.org/10.1073/pnas.0404206101>.
26. Nagoshi, R.N.; Davis, J.A.; Meagher, R.L.; Musser, F.R.; Head, G.P.; Portillo, H.; Teran, H. Evidence for two soybean looper strains in the United States with limited capacity for cross-hybridization. *Genes* **2023**, *submitted*.
27. Garnier-Géré, P.; Chikhi, L. Population Subdivision, Hardy–Weinberg Equilibrium and the Wahlund Effect. In *The Encyclopedia of Life*; John Wiley & Sons: Hoboken, NJ, USA, 2013. <https://doi.org/10.1002/9780470015902.a0005446.pub3>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.