

## Article

# Genetic Screening to Identify Candidate Resistance Alleles to Cry1F Corn in Fall Armyworm Using Targeted Sequencing

Katrina Schlum <sup>1</sup>, Kurt Lamour <sup>1,2</sup>, Peter Tandy <sup>2</sup>, Scott J. Emrich <sup>1,3</sup>, Caroline Placidi de Bortoli <sup>2</sup>, Tejas Rao <sup>2</sup>, Diego M. Viteri Dillon <sup>4</sup>, Angela M. Linares-Ramirez <sup>5</sup> and Juan Luis Jurat-Fuentes <sup>1,2,\*</sup> 

<sup>1</sup> Genome Science and Technology Graduate Program, University of Tennessee, Knoxville, TN 37996, USA; kschlum@vols.utk.edu (K.S.); klamour@utk.edu (K.L.); semrich@utk.edu (S.J.E.)

<sup>2</sup> Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, TN 37996, USA; ptandy@vols.utk.edu (P.T.); cplacidi@utk.edu (C.P.d.B.); trao2@utk.edu (T.R.)

<sup>3</sup> Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA

<sup>4</sup> Isabela Research Substation, Department of Agro-Environmental Sciences, University of Puerto Rico, Isabela, PR 00662, USA; diego.viteri@upr.edu

<sup>5</sup> Lajas Research Substation, Department of Agro-Environmental Sciences, University of Puerto Rico, Lajas, PR 00667, USA; angela.linares@upr.edu

\* Correspondence: jurat@utk.edu; Tel.: +1-(865)-974-5931



**Citation:** Schlum, K.; Lamour, K.; Tandy, P.; Emrich, S.J.; de Bortoli, C.P.; Rao, T.; Viteri Dillon, D.M.; Linares-Ramirez, A.M.; Jurat-Fuentes, J.L. Genetic Screening to Identify Candidate Resistance Alleles to Cry1F Corn in Fall Armyworm Using Targeted Sequencing. *Insects* **2021**, *12*, 618. <https://doi.org/10.3390/insects12070618>

Academic Editors: Ralf Nauen and Gaelle Le Goff

Received: 26 May 2021

Accepted: 30 June 2021

Published: 8 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Simple Summary:** Monitoring of resistance alleles is critical to the sustainability of transgenic crops producing insecticidal Cry proteins. Highly sensitive and cost-effective DNA-based methods are needed to improve current bioassay-based resistance screening. Our goal was to evaluate the use of targeted sequencing in detecting known and novel candidate resistance alleles to Cry proteins. As a model, we used field-collected fall armyworm (*Spodoptera frugiperda*) from Puerto Rico, the first location reporting continued practical field-evolved *S. frugiperda* resistance to corn producing the Cry1F insecticidal protein, and sequenced the *SfABCC2* gene previously identified as critical to Cry1F toxicity. Targeted sequencing of *SfABCC2* detected a previously reported Cry1F resistance allele and mutations originally identified in populations from Brazil. Importantly, targeted sequencing also identified nonsynonymous and frameshift mutations as novel candidate resistance alleles. These results advocate for the use of targeted sequencing in screening for resistance alleles to Cry proteins and support potential gene flow, including resistance alleles, between *S. frugiperda* from Brazil and the Caribbean.

**Abstract:** Evolution of practical resistance is the main threat to the sustainability of transgenic crops producing insecticidal proteins from *Bacillus thuringiensis* (Bt crops). Monitoring of resistance to Cry and Vip3A proteins produced by Bt crops is critical to mitigate the development of resistance. Currently, Cry/Vip3A resistance allele monitoring is based on bioassays with larvae from inbreeding field-collected moths. As an alternative, DNA-based monitoring tools should increase sensitivity and reduce overall costs compared to bioassay-based screening methods. Here, we evaluated targeted sequencing as a method allowing detection of known and novel candidate resistance alleles to Cry proteins. As a model, we sequenced a Cry1F receptor gene (*SfABCC2*) in fall armyworm (*Spodoptera frugiperda*) moths from Puerto Rico, a location reporting continued practical field resistance to Cry1F-producing corn. Targeted sequencing detected a previously reported Cry1F resistance allele (*SfABCC2mut*), in addition to a resistance allele originally described in *S. frugiperda* populations from Brazil. Moreover, targeted sequencing detected mutations in *SfABCC2* as novel candidate resistance alleles. These results support further development of targeted sequencing for monitoring resistance to Bt crops and provide unexpected evidence for common resistance alleles in *S. frugiperda* from Brazil and Puerto Rico.

**Keywords:** *Spodoptera frugiperda*; fall armyworm; resistance; Cry1F; genotyping; targeted sequencing; resistance screen; ABCC2

## 1. Introduction

The fall armyworm (*Spodoptera frugiperda*, J.E. Smith) (Lepidoptera: Noctuidae) is a global invasive pest affecting numerous food and fiber staple crops, although the highest damage is observed in corn (*Zea mays* L.) [1]. In North America, *S. frugiperda* moths display long-distance northward migration over several generations from overwintering sites in southern Texas and Florida, reaching Canada in the summer months [2]. Substantial gene flow also occurs between *S. frugiperda* populations from the Caribbean and Florida, but not between these populations and moths from South America [3].

Effective control of *S. frugiperda* in the Western hemisphere in addition to pesticides, has been provided by genetically modified corn and cotton producing insecticidal proteins from the bacterium *Bacillus thuringiensis* (*Bt*), namely Cry1F, Cry1Ab, Cry1A.105, Cry2Ab and Vip3Aa. In the past decade, resistance to Cry1F, Cry1Ab and Cry1A.105 was reported for *S. frugiperda* populations in Puerto Rico and the continental US (Florida and North Carolina) [4,5], Brazil [6] and Argentina [7]. Fall armyworm resistance to these Cry proteins is recessive and genetically linked to mutations in an ABC transporter superfamily C2 gene (*SfABCC2*) [8–10]. In Puerto Rico, two *SfABCC2* resistance alleles have been reported, a nonsense 2 bp insertion (*SfABCC2mut* allele) [8] and an insertion near the start of the fourth exon associated with aberrant splicing [9]. In contrast, resistance to Cry1F in Brazil is linked to missense mutations primarily localized to the fourth extracellular loop in *SfABCC2*, in support of this region containing a Cry1F binding domain [11] and being a hotspot for resistance mutations [10]. Genetic knockout of *SfABCC2* leads to resistance against Cry1F [12,13] further supporting that truncation of *SfABCC2* is predictive of a Cry1F-resistant phenotype.

Identification of resistance alleles in *S. frugiperda* is critical to developing highly sensitive DNA-based genotyping tools, which are less laborious and costly compared to F1 [14] or F2 [5] resistance screens. Successful *SfABCC2* genotyping efforts were described using TaqMan probes [8,9] and pyrosequencing [10] in populations from Puerto Rico, USA and Brazil. Results from these tests support high-resistance allele frequencies in areas where practical resistance had been reported, further associating identified *SfABCC2* alleles with field-relevant resistance. An unexpected observation, considering predicted migratory models [3], was not detecting the *SfABCC2mut* allele reported in Puerto Rico at migratory Caribbean destinations or in the Florida populations tested [8,9]. Relevant fitness costs are not associated with *SfABCC2mut* under laboratory conditions [15] and thus probably do not explain the limited distribution to Puerto Rico of *SfABCC2mut*.

Hi-Plex (HP) is a highly multiplexed PCR-based approach for targeted massive parallel sequencing with demonstrated efficacy in disease gene screening [16,17]. It involves amplifying small (e.g., 180 bp) overlapping portions of a larger target region (tiling) in one (or a few) multiplex PCR reaction(s) adding a unique sample-specific barcode sequence during the PCR reaction. The HP technology is attractive because it requires very little genomic DNA to amplify multiple target amplicons (e.g., 20–40 ng of template DNA to amplify a few hundred targets), the DNA can be highly complex (e.g., mixed with genomic DNA from many different organisms) and the overall cost in time and reagents is greatly reduced compared to F1 and F2 screens. The barcoded amplicons are pooled and sequenced on a next-generation sequencing (NGS) device, and the sample-specific amplicons mapped to a reference genome.

The goal of this project was to adapt and provide preliminary data for a new HP targeted screen for resistance alleles to *Bt* insecticidal proteins. We focused our analysis on *S. frugiperda* samples from Puerto Rico, as both *SfABCC2mut* and Cry1F resistance are highly frequent at this location [8,18]. Furthermore, mutations resulting in a truncated *SfABCC2* protein result in a Cry1F-resistant phenotype [8,9,12,13], potentially allowing for phenotypic predictions. Analysis of Taqman and targeted *SfABCC2* sequences detected the known *SfABCC2mut* allele in 34–37% of tested samples, depending on the technique used. In addition, the HP approach coupled to a bioinformatics pipeline detected one nonsense and nine frameshift *SfABCC2* mutations as novel candidate Cry1F resistance

alleles. Missense mutations (13) were also identified, but phenotype predictions are not possible in this case as the effect of these mutations on Cry1F susceptibility and resistance requires functional testing. We document, for the first time, that a *SfABCC2* resistance allele previously described in *S. frugiperda* from Brazil is present in Puerto Rico fall armyworm populations. These data provide proof of concept for the power of HP targeted sequencing for performing surveillance and field monitoring of well-established and novel candidate *Bt* resistance alleles.

## 2. Materials and Methods

### 2.1. *S. frugiperda* Sample Collection and Genomic DNA Extraction

Adult moths were captured at three locations in Puerto Rico (Lajas, Salinas and Santa Isabel) using sex pheromone-baited traps [19] at sites near field and sweet/field corn plantings in order to optimize the capture efficiency of C-strain males. Collection details of the 41 *S. frugiperda* individuals tested are in Table S1. Captured moths were confirmed as *S. frugiperda* by visual inspection [20] and kept refrigerated until shipped to the University of Tennessee. Samples were stored at  $-20^{\circ}\text{C}$  until processed for purification of genomic DNA. Samples were named and numbered according to trapping location as PRL (Lajas), PRS (Salinas) and PRSI (Santa Isabel). A sample from the previously characterized Cry1F-resistant 456LSD4 strain of *S. frugiperda*, which originated from collections in Puerto Rico [21], was also included as positive control for detection of *SfABCC2mut* [8].

Genomic DNA was isolated from moth legs using the Pure Link Genomic DNA mini kit (Invitrogen) following manufacturer's protocols, and then quantified using a Nanodrop spectrophotometer (Thermo Scientific).

### 2.2. TaqMan Genotyping for *SfABCC2mut*

Detection of the *SfABCC2mut* allele using Taqman probes was performed as previously described [8]. Briefly, 10  $\mu\text{L}$  (final volume) reactions included 10–20 ng of gDNA as template, a VIC-labeled probe specific to the *SfABCC2mut* allele (5' AAGCACATCGCCCACTT 3'), a FAM-labeled probe specific to the wild-type *SfABCC2* allele (5' CCAAGCACATCCCACTT 3'), and forward (5' TGGAGGCCGAAGAGAGACA 3') and reverse (5' AGGAGTTGACTGACTTCATGTACCT 3') primers. Plates (Micro Amp Fast optical 96 well reaction plate, Applied Biosystems) were loaded in the Quant studio 6 Real Time PCR instrument (Applied Biosystems) and amplified as follows: pre read stage at  $60^{\circ}\text{C}$  for 30 s, hold stage at  $95^{\circ}\text{C}$  for 10 min, PCR stage at  $95^{\circ}\text{C}$  for 15 s and  $60^{\circ}\text{C}$  for 1 min for 40 cycles, post read stage at  $60^{\circ}\text{C}$  for 30 s. The fluorescence in each well was measured in the post read stage of the PCR, and the software generated an allelic discrimination plot based on the post amplification intensity of the fluorescent probes.

The frequency of the *SfABCC2mut1* allele or any mutation was determined using the Hardy–Weinberg equation with the formula:  $F = (2 \times \text{ObsAa} + \text{Obsaa}) / [2 \times (\text{ObsAA} + \text{ObsAa} + \text{Obsaa})]$ , where "F" is the frequency of the "a" allele (*SfABCC2mut1*) and "Obs" the observed frequency of each of the three possible genotypes for the allele.

### 2.3. Targeted DNA Amplification and Sequencing

In early BLASTn searches of the *S. frugiperda* corn strain v3.1 genome from LepidoDB ([https://bipaa.genouest.org/sp/spodoptera\\_frugiperda\\_pub/](https://bipaa.genouest.org/sp/spodoptera_frugiperda_pub/), accessed on 11 July 2019), we found that the *SfABCC2* gene was split between Scaffolds 11087 and 7154. These scaffolds were manually combined and the resulting *SfABCC2* gene sequence was used as a template to design TILING primers, using PCRTiler [22] at default settings allowing 50 bp overlap on each end and producing 180–200 bp final amplicon sizes (Table S2). The two approximately 1 Kbp intronic regions at the beginning of that gene sequence (Scaffold\_11087) were not targeted. In a later version of the *S. frugiperda* corn strain genome (v6.0) from LepidoDB (accessed on 16 November 2020), the full-length *SfABCC2* gene appears in Scaffold 33. The initial *SfABCC2* gene assembly based on Scaffolds 11087 and 7154 used for TILING primer design did not include an inter-scaffold gap encompassing

nucleotides 7239–8976 of the *SfABCC2* gene assembly in the v6.0 genome. The TILING primers covered approximately 90% of the exonic regions of the assembled *SfABCC2* gene found in Scaffold 33 of the *S. frugiperda* corn strain v6.0 genome (see Figure S1 for a visual representation of typical coverage).

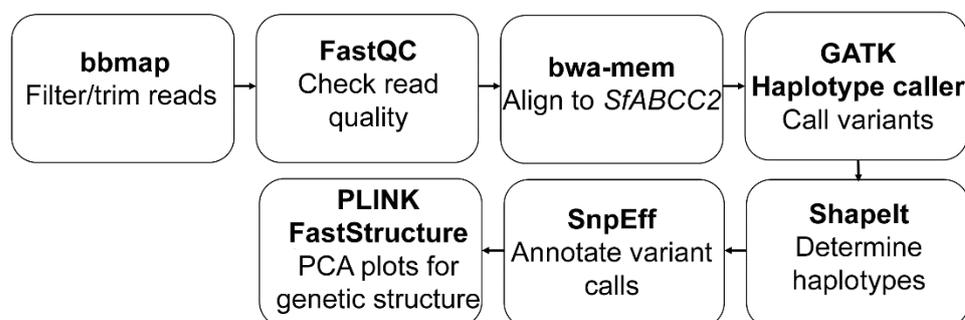
All TILING primers (238) were mixed into a single multiplex (119 overlapping targets) and PCR performed using the Hi-Plex targeted sequencing strategy, as previously described [17]. The resulting amplicons were pooled and sequenced on a HiSeqX device running a  $2 \times 150$  bp configuration and merged to produce single, high quality reads using PEAR [23]. Demultiplexing was accomplished using the sample-specific barcode sequences incorporated during the Hi-Plex PCR amplification.

#### 2.4. Genotype Assignment for *SfABCC2mut* Using a Targeted-Sequencing *k*-mer-Based Approach

To genotype the *SfABCC2mut* 2 bp insertion in the *SfABCC2* targeted sequencing data, a custom Python 3 script (dubbed Pgeno) was developed and is deposited in Github at <https://github.com/petertandy/pgeno> (accessed on 26 May 2021). This script scans each of the sequence reads for the known mutation by using 10 mer flanking sequences as a guide. To account for unknown variation in the flanking regions, each of the 10 mers was allowed up to two mismatches. Sequences with matching flanks were counted and samples assigned a *SfABCC2mut* genotype based on the proportion of the known 2 bp insertions. In order to reduce false-positive genotype assignment, we required at least 40X sequence coverage in the PCR targeted regions and >15% of the total number of matching reads to assign a genotype of hetero- or homozygous for the 2 bp insertion. The minimum sequence coverage refers to the number of high-quality reads covering a target region in the *SfABCC2* reference gene and depends on the efficiency of the primers in the multiplex. The primer efficiencies are highly reproducible and by over-sequencing the samples, most of the gene is covered at a high depth (up to 1000X for some targets). These parameters can be adjusted in the Pgeno script to increase or decrease the genotyping stringency.

#### 2.5. Alignment, Variant Calling and Variant Analysis

The complete pipeline used for trimming, alignment, variant calling, and analysis using targeted *SfABCC2* sequences is summarized in Figure 1. Raw reads were quality trimmed at both ends and filtered for adapter sequences using BBDuk [24]. The quality of samples was checked before and after trimming/processing to confirm adapter and non-relevant (e.g., contaminants, primer artifacts) sequences were removed using FastQC [25]. The full-length *SfABCC2* gene interval was determined by aligning a *SfABCC2* cDNA from a Cry1F-susceptible *S. frugiperda* strain (GenBank accession number KY489760.1) to the *S. frugiperda* corn reference genome v6.0 [26] via the BLAST server hosted on LepidoDB. The full-length *SfABCC2* gene sequence was found included in Scaffold 33, and this reference gene sequence was used for further analyses.



**Figure 1.** Schematic pipeline of overall workflow used in this work from raw Fastq targeted sequencing reads to population genetics and phylogenetic analyses based on bi-allelic SNP variant calling.

Reads were aligned to the *SfABCC2* reference gene using a Burrows–Wheeler aligner algorithm (bwa-mem) with default options [27]. The mapping rate for all samples was determined via SAMtools flagstat [28]. Only samples with a mapping rate of 80% or higher were kept, which led to the removal of only three samples from further analysis. We determined the count of DNA sequence reads across the *SfABCC2* gene using bam-readcount (<https://github.com/genome/bam-readcount>, accessed on 20 June 2019). As noted in 2.3, two regions (nucleotides ~138–1518 and ~7195–7858) within the *SfABCC2* gene were not sequenced due to the TILING primers being designed based on an old assembly (Scaffolds 11,087 and 7154 from *S. frugiperda* corn genome v3.1) of the gene. The second region representing the inter-scaffold gap in the early gene assembly was filled with ‘Ns’, precluding alignment of any reads to this region.

We performed alignment cleanup by sorting the bam file via SAMtools [28], deduplicating via Genome Analysis Tool Kit (GATK) version 4.1.20 [29] with MarkDuplicates and then adding read groups to each bam file via GATK’s function addOrReplaceReadGroup. Variants were called via the Haplotypecaller of GATK with default parameters set based on the GATK Best Practice pipeline. The initial set included 2006 variants with a quality score greater than 40. After considering different filters based on depth and minor allele frequencies, the final filter applied included depth of coverage  $\geq 8$ , each individual allele depth  $\geq 2$ , alternate allele frequency  $\geq 5\%$ , and the allele being called in at least 25% of the samples ( $>10$ ). This filter produced 1861 variants (1428 SNPs and 433 indels), over 92% of the unfiltered set. We further filtered this variant set to include only bi-allelic variants using the GATK SelectVariants function, which resulted in 1333 bi-allelic variants (1143 SNPs and 190 indels) that were used in all variant-based analysis.

The intron/exon boundaries in *SfABCC2* were determined by alignment to a *SfABCC2* cDNA (GenBank accession number KY489760.1) using Exonerate with the est2genome model option [30]. To confirm the accuracy of the intron/exon boundaries, we performed the same alignment of the *SfABCC2* gene with three additional available cDNA sequences from Cry1F-susceptible *S. frugiperda* (GenBank accession numbers MG387043.1, MN399979.1 and KY646296.1). The same intron/exon boundaries were identified in each generic feature file (GFF) generated by Exonerate. The consensus protein sequence from aligning the *SfABCC2* sequences derived from translating the *SfABCC2* cDNA sequences above in CLC Genomics Workbench 21.0.4 (Figure S2) was used as a reference to locate amino acid changes from mutations. This consensus *SfABCC2* protein sequence differs in two amino acid positions in the notation for mutations located to extracellular loop 4 region presented by Boaventura et al. [10].

Functional annotation was performed with SnpEff [31], and SnpSift was used to filter annotations to a tabular format [32]. The final filtered set of 1333 variants contained 17 missense mutations and 1 nonsense mutation. Compared to the unfiltered variant set, the filtered set excluded one nonsense mutation because of both alternate allele frequency and read depth criteria, and two missense mutations because of the alternate allele frequency criterion.

We also computed PCA projections via PLINK [33], and visualized them in ggplot2 [34]. Given the highly polymorphic nature of the samples, we also determined haplotypes for PCA and generated neighbor joining trees using the ShapeIT software [35]. Since no recombination maps are available for *S. frugiperda*, the recombination rate in ShapeIT was set to 0.000023 cM/Mb based on previous research for *Spodoptera litura* [36].

Admixture model-based clustering was performed using FastStructure [37] to identify genetic structure within the *SfABCC2* gene. FastStructure is a generative model-based approach based on Hardy–Weinberg equilibrium assumptions between alleles and linkage disequilibrium between genotyped loci. The FastStructure script structure.py was used to determine  $k$ , the number of assumed populations or genetic groups that share a subset of allele frequencies from  $k = 2$  to 10. The choosek.py script in FastStructure chose  $k = 3$  for model complexity maximizing the marginal likelihood, and  $k = 4$  for model components

to explain structure in the data. The final output of FastStructure was visualized using *distruct.py* at  $k = 4$ .

### 3. Results

#### 3.1. Genotyping Results Using Taqman

Based on Taqman genotyping, of the 41 *S. frugiperda* samples tested, 18 carried the *SfABCC2mut* allele, with nine being homozygous. Overall, these detections suggest an allele frequency of 0.3292 for *SfABCC2mut* in the analyzed samples, similar to previous estimations for the frequency of this allele in Puerto Rico [8]. The relative proportion of individuals carrying the *SfABCC2mut* allele at each location was higher in Lajas (53% carriers, including 33% homozygotes and 20% heterozygotes) followed by the Salinas (41% carriers, 8% homozygotes and 33% heterozygotes) and Santa Isabel (21% homozygotes) locations.

#### 3.2. Genotyping Results Using Targeted Sequencing

Overall, the multiplexed TILING primers amplified 63% of the *SfABCC2* reference gene from the *S. frugiperda* v.6.0 corn genome, with 23 of the 25 predicted exons located in contig 33 being fully sequenced (Figure S1). The Hi-Plex strategy does not rely on equilibration of primer efficiencies (or any kind of primer optimization) and as such, some targets will amplify more efficiently than others (or fail completely) and thus receive more sequencing (or none at all). To accommodate these expected variations in template concentration, the samples were sequenced to high depth ( $>100\times$  per target amplicon).

Of the 41 samples genotyped via targeted sequencing, 14 contain the *SfABCC2mut* GC insertion, with 8 scored as heterozygous and 6 as homozygous, resulting in a 0.2683 allele frequency. There was a 73% (30/41) concordance rate between the Taqman and the k-mer-based *SfABCC2mut* genotyping (Table 1). Five samples identified as susceptible by the k-mer-based genotyping were identified as heterozygotes (3) or homozygote resistant (2) by the Taqman assay. Furthermore, three heterozygous samples, based on k-mer-based genotyping, appeared as homozygous resistant in Taqman tests. Finally, two samples identified as homozygous resistant by the k-mer method were classified as homozygous susceptible by the Taqman assay. Examination of the assembled targeted sequences for the discrepant samples detected a number of SNPs in the *SfABCC2mut* site targeted by Taqman primers, which likely affected Taqman amplifications and therefore explain the observed discrepancies. In fact, of the 58 sites covered by the Taqman primers and probe, 19% were polymorphic in at least one of the 41 samples (data not shown).

#### 3.3. Variant Analysis Pipeline and Genotyping

Functional annotation of variants detected using the bioinformatics pipeline shown in Figure 1 detected *SfABCC2mut* (2 bp GC insertion) as a frameshift mutation (D740A). There was a total of 18 samples in which the *SfABCC2mut* allele was detected using the GATK variant and filtering described earlier (see Methods). Within these 18 samples, 12 were heterozygous and 6 homozygous for the mutation, displaying 100% concordance with results from the k-mer-based genotyping approach. When a lower alternative allele threshold ( $\geq 3$ ) was utilized, some discrepancies occurred between GATK and the k-mer-based genotyping (data not shown).

In addition to identifying the known *SfABCC2mut* allele, we also detected 374 coding synonymous mutations (data not shown) and 23 nonsynonymous and frameshift mutations in *SfABCC2* in the filtered dataset. After mutations already observed in at least one *SfABCC2* cDNA from Cry1F-susceptible *S. frugiperda* were removed, 14 nonsynonymous and 9 frameshift mutations (apart from *SfABCC2mut* being synonymous with D740A) were identified as new potential Cry1F resistance alleles (Table 2 and Table S1).

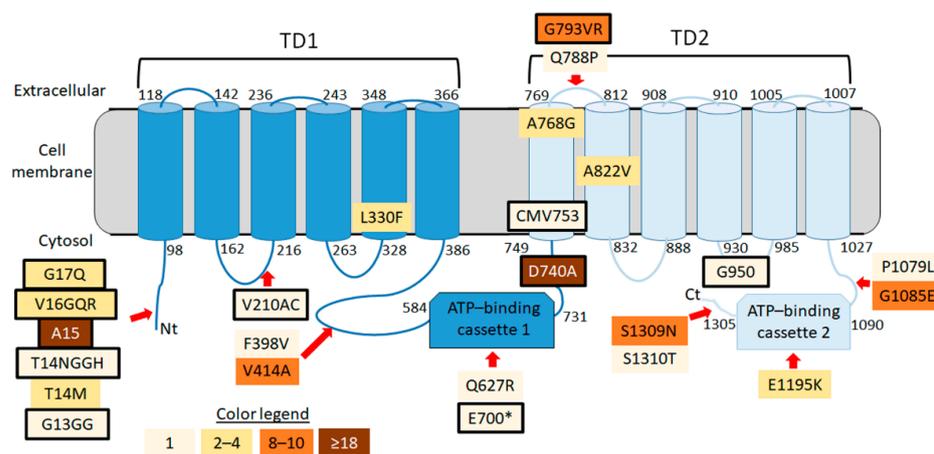
**Table 1.** Concordance between results using Taqman and k-mer-based genotyping for the *SfABCC2mut* (GC insertion) allele, represented as the “r” allele. Samples with discrepant genotypes are highlighted in bold font.

Sample Name	Taqman	K-mer
PRL_164	SS	SS
<b>PRL_165</b>	<b>rr</b>	<b>Sr</b>
PRL_166	SS	SS
PRL_167	SS	SS
PRL_168	Sr	Sr
<b>PRL_169</b>	<b>rr</b>	<b>Sr</b>
PRL_170	Sr	Sr
PRL_171	Sr	Sr
PRL_172	rr	rr
PRL_173	rr	rr
PRL_174	SS	SS
PRL_175	SS	SS
PRL_176	SS	SS
PRL_177	rr	rr
<b>PRL_178</b>	<b>SS</b>	<b>SS</b>
<b>PRS_6</b>	<b>SS</b>	<b>SS</b>
PRS_7	Sr	Sr
PRS_8	SS	SS
<b>PRS_9</b>	<b>Sr</b>	<b>SS</b>
PRS_10	rr	rr
PRS_11	SS	SS
<b>PRS_12</b>	<b>Sr</b>	<b>SS</b>
PRS_13	SS	SS
PRS_14	SS	SS
<b>PRS_17</b>	<b>Sr</b>	<b>SS</b>
<b>PRS_18</b>	<b>SS</b>	<b>rr</b>
PRS_20	SS	SS
PRSI_1	SS	SS
PRSI_2	SS	SS
PRSI_4	SS	SS
<b>PRSI_5</b>	<b>SS</b>	<b>Sr</b>
<b>PRSI_6</b>	<b>rr</b>	<b>Sr</b>
<b>PRSI_7</b>	<b>rr</b>	<b>SS</b>
PRSI_8	SS	SS
<b>PRSI_9</b>	<b>SS</b>	<b>rr</b>
<b>PRSI_10</b>	<b>rr</b>	<b>SS</b>
PRSI_11	SS	SS
PRSI_12	SS	SS
PRSI_13	SS	SS
PRSI_14	SS	SS
PRSI_15	SS	SS

All missense (13) and nonsense (1) nonsynonymous mutations, except for an homozygous sample for T14M, were detected in heterozygous individuals. The most abundant missense mutations (10 individuals each, 24% of samples) were G1085E and S1209N, which occur in the C terminal intracellular region of SfABCC2 (Figure 2). The other common missense mutation was V414A (carried by 9 individuals, 22% of samples), which occurs in the loop connecting the first SfABCC2 transmembrane domain and ATP-binding cassette. Mutation P1079L was localized to a predicted splice site region. Only one missense mutation (Q788P, found in 1 sample) was localized to the fourth extracellular loop, a region hypothesized to be vital to Cry1F toxicity [10,11]. Missense mutations had a predicted moderate effect on the SfABCC2 protein, while the only nonsense mutation (E700 \*) was predicted to result in a truncated SfABCC2 protein encompassing up to the first ATP-binding cassette (Figure 2).

**Table 2.** Nonsynonymous and frameshift mutations identified by SnpEFF across 41 *S. frugiperda* *SfABCC2* bi-allelic variants. The nucleotide position in the *SfABCC2* gene reference (Position), the functional class of mutation (MISS = missense, NON = nonsense, and SHIFT = frameshift), the mutation with capitalized reference and alternative alleles, the corresponding amino acid change (Aa change), and the number of samples positive for that mutation (#Samples) are shown. Asterisks denote that the mutation leads to a premature stop codon. The *SfABCC2mut* allele is synonymous with D740A.

Position	Class	Mutation	Aa Change	#Samples
38	SHIFT	ggc/ggTGGTc	G13GG	1
40	SHIFT	acg/aATGGTGGTCATcg	T14NGGH *	1
41	MISS	aCg/aTg	T14M	4
42	SHIFT	gct/	A15 *	20
45	SHIFT	gtg/gGTCAACG	V16GQR *	4
48	SHIFT	ggc/CAggc	G17Q *	4
3421	SHIFT	gtc/gCATGtc	V210AC *	1
4077	MISS	Ctc/Ttc	L330F	2
4357	MISS	Ttt/Gtt	F398V	1
4540	MISS	gTg/gCg	V414A	9
6020	MISS	cAa/cGa	Q627R	1
6915	NON	Gaa/Taa	E700 *	1
7034	SHIFT	gat/gCGat	D740A *	18
7070	SHIFT	tgcatggtg/	CMV753	1
7120	MISS	gCt/gGt	A768G	3
9098	MISS	cAa/cCa	Q788P	1
9122	SHIFT	gga/GTAAgga	G793VR	8
9335	MISS	gCc/gTc	A822V	3
10491	SHIFT	gga/	G950	1
11708	MISS	cCg/cTg	P1079L	1
11726	MISS	gGa/gAa	G1085E	10
12507	MISS	Gaa/Aaa	E1195K	2
12748	MISS	aGt/aAt	S1209N	10
13285	MISS	Tcc/Acc	S1310T	1

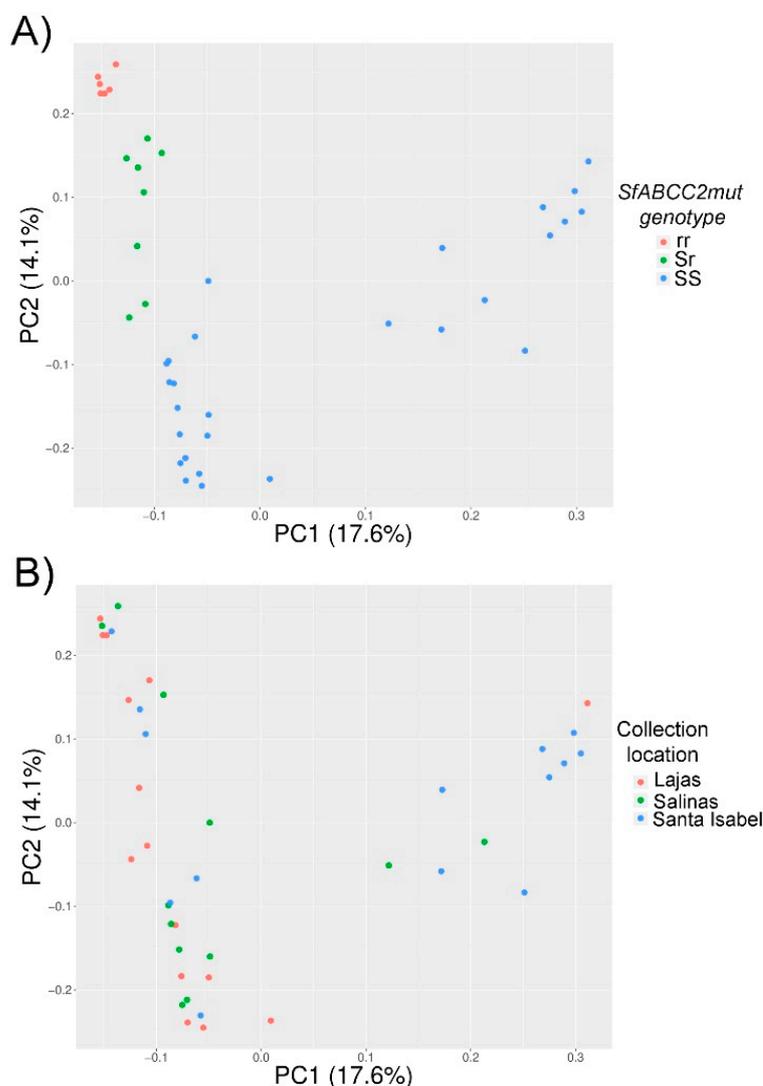


**Figure 2.** Schematic representation (not drawn to scale) of *SfABCC2* topology on the cell membrane and location of detected nonsynonymous, nonsense and frameshift mutations in sequenced *S. frugiperda* samples from Puerto Rico. Amino acid numbers are indicated at the start and end of transmembrane domains and ATP-binding cassettes as predicted based on detection of conserved domains in BLAST searches. Distinct colors in boxes indicate the number of individuals carrying each mutation, as detailed in the color legend. The only nonsense mutation (E700 \*) and all frameshift mutations are shown in a boxes with bold borders, including the *SfABCC2mut* allele (D740A). TD = transmembrane domain.

Frameshifts were comparatively more frequent than nonsynonymous mutations, with homozygous individuals detected for all the frameshift mutations for which more than one positive sample was detected. These frameshift mutations generated aberrant proteins and predicted premature stop codons, leading to truncated SfABCC2 proteins, which in homozygous individuals would correspond to a Cry1F-resistant phenotype [8,12]. The A15 mutation was the most common (49% of individuals), with a slightly lower allele frequency (0.3170) than *SfABCC2mut* (synonymous with D740A mutation). Only three samples (PRL\_168, PRL\_170, PRL\_171) from the same location (Lajas) were heterozygous for both A15 and D740A. These samples could be resistant to Cry1F based on complementation of these two alleles.

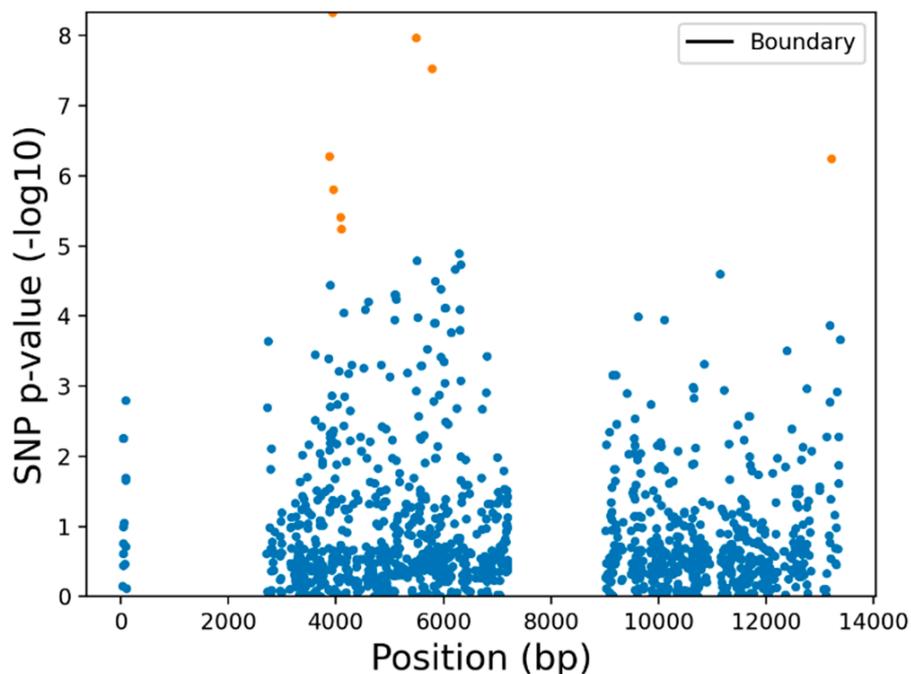
### 3.4. *SfABCC2* Sequence Variation

We also considered population-level diversity in *SfABCC2* sequences from Puerto Rico. Interestingly, a Principal Component Analysis (PCA) of all samples showed two distinct groups, one of them without *SfABCC2mut* in any individual (Figure 3A, right). Individuals homozygous for *SfABCC2mut* had the most compact cluster, suggesting more *SfABCC2* sequence variation within the SS compared to rr samples.



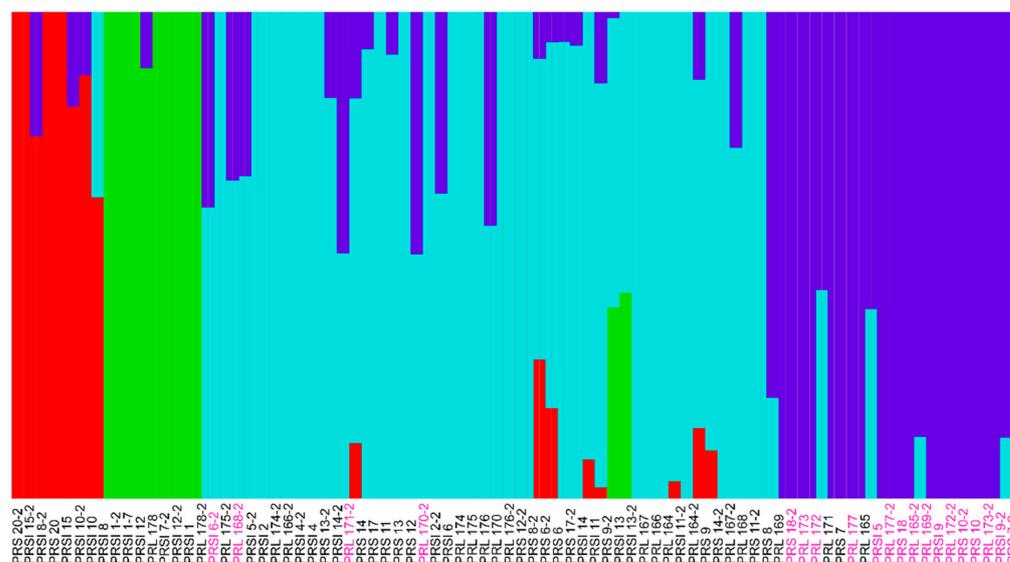
**Figure 3.** PCA projection plot of 1333 bi-allelic variants found within *SfABCC2* of 41 *S. frugiperda* samples and colored based on genotype for *SfABCC2mut* (A) and location (B). The *SfABCC2mut* genotype was determined as described in the Methods.

To determine whether location was driving differentiation of the PC clusters observed given different frequencies of *SfABCC2mut* in the three Puerto Rico locations, we also plotted these samples colored by location (Figure 3B). No clusters correlated with sample location. Further, a logistic regression association test on PC1 vs. PC2 (Figure 4) generated eight SNPs with significant  $-\log_{10} p$ -value  $> 5$  at positions 3874, 3936, 3944, 4085, 4097, 5485, 5788, and 13206. None of these variants, however, were predicted to have any functional effect change by SnpEff or corresponded to previously reported Cry1F-resistant alleles [8,10]. These data suggest that there may be barriers to Bt resistance gene flow, for example between host strains, even within the island of Puerto Rico.



**Figure 4.** Manhattan plot of SNPs associated with PC2 (differentiating clusters of Sr and rr genotyped samples) on the filtered bi-allelic dataset (1333 variants).

To confirm this observation, we re-ran the PCA analysis without the *SfABCC2mut* homozygous individuals and observed the same clusters (data not shown). We next haplified our *SfABCC2* data using ShapeIt2 (1333 variants, see Methods) and ran FastStructure (Figure 5). Based on the targeted sequence data, we observed four populations, with the *SfABCC2mut* allele always associating with one population (purple color in Figure 5). Furthermore, even though there is little observed sharing of *SfABCC2* between the two primary PCA-based groups, admixture did occur between the *SfABCC2mut*-associated population and the three other populations. The frequency of the *SfABCC2mut* allele was highest in Lajas, indicating that this allele likely originated from this location and spread to other fall armyworm populations in Puerto Rico, as seen in the Structure plot. The observed moderate levels of admixture between the fall armyworms studied across all three Puerto Rican populations suggest that—contrary to the PCA results—the *SfABCC2mut* has the potential to spread to other subpopulations.



**Figure 5.** Genetic structure inferred by FastStructure on 82 haplotypes representing 41 samples at  $k = 4$ . Samples in pink text contain the *SfABCC2mut* allele.

#### 4. Discussion

Bioassay-based testing of F1 or F2 generations has traditionally been used to screen for resistance alleles in *S. frugiperda* to insecticidal proteins from *B. thuringiensis* produced in transgenic crops [5,14,38,39]. As an alternative, DNA-based screening methods provide higher sensitivity and are amenable to high-throughput applications and can considerably reduce overall labor and costs. Initial DNA-based resistance monitoring efforts for *S. frugiperda* focused on Taqman PCR targeting known mutations in the *SfABCC2* gene [8,9] and pooled population sequencing [10]. More recently, whole-genome resequencing was used to detect the presence of five reported resistance alleles in *S. frugiperda* populations from 12 geographic locations [40]. As a more cost-effective alternative to whole-genome sequencing, we provide proof of concept for the use of targeted highly multiplexed PCR for monitoring a known resistance allele (*SfABCC2mut*) and using bioinformatics to identify novel candidate *SfABCC2* resistance alleles. An advantage of the Cry1F-*SfABCC2* system is that available knockout [12,13], mutational [11] and resistance [8,9] data in *S. frugiperda* support that a truncated *SfABCC2* protein cannot function as Cry1F receptor. Thus, homozygote individuals for nonsense or frameshift mutations leading to truncated *SfABCC2* proteins display a Cry1F-resistant phenotype, as observed for the *SfABCC2mut* resistance allele [8,9]. In addition, the amplification strategy of Hi-Plex allows the use of highly fragmented DNA, suggesting that poorly preserved samples, as is often the case for adult moths collected with pheromone traps, are amenable to analysis.

Both a k-mer-based and a bioinformatics pipeline provided the same genotyping results for a known mutation (*SfABCC2mut*). The genotypes of some samples differed between the Hi-Plex and Taqman methods. We observed that 19% of tested samples had polymorphisms around the priming/probe site for Taqman, which suggests an advantage of Hi-Plex over Taqman when dealing with highly polymorphic resistance genes. Most likely, the overlapping targets (and thus different priming sites) used in the TILING approach helps reduce the potential for erroneous genotypes and conclusions.

We found that among the individuals carrying the *SfABCC2mut* allele, 14.6% were homozygotes, which are expected to produce a truncated *SfABCC2* and thus be resistant to Cry1F [8,9,12,13]. The allele frequency of *SfABCC2mut* (0.3292) was in line with previous screening efforts in fall armyworm from Puerto Rico [8,9]. Based on the Faststructure and PCA-based analysis, no clear selective sweep was observed. The higher relative frequency of *SfABCC2mut* suggests the allele may have emerged in *S. frugiperda* in or around Lajas. Our population-level analysis suggests that the genomic region around *SfABCC2mut* is

subject to both recombination and introgression, and that this occurs frequently, especially between populations at Santa Isabel and Lajas. In contrast, we still observed two main populations in Puerto Rico with statistically significant markers that do not associate with known or candidate resistance alleles. Further work is needed to determine whether there is at least one barrier to large-scale gene flow on the island, e.g., the well-established Corn vs. Rice strains in *S. frugiperda*. It is important to consider that Lajas has had rice winter nurseries (Univ. of Puerto Rico and Ricetech) for more than 30 years, while corn is the most prevalent crop at Santa Isabel. Furthermore, organic corn is planted in Lajas, which requires the frequent use of *B. thuringiensis* insecticides (e.g., Dipel) to control Lepidopteran pests [41].

In addition to *SfABCC2mut*, a second Cry1F-resistant allele was previously described in *S. frugiperda* from Puerto Rico [9]. In this case, insertion of a highly repetitive sequence at the beginning of the *SfABCC2* gene resulted in aberrant splicing and resistance. This type of allele would not be detected in our current Hi-Seq strategy as our TILING primers were not designed to amplify the highly repetitive sequence. Addition of primers amplifying this region and re-sequencing may allow detection and estimation of frequency for this allele. However, the reported highly repetitive nature of the insert [9] may hinder successful detection.

Targeted sequencing combined with a custom bioinformatics pipeline allowed detecting, in addition to 374 coding synonymous mutations, 23 nonsynonymous and frameshift *SfABCC2* mutations as novel candidate Cry1F resistance alleles. A Cry1F-resistant phenotype similar to homozygous individuals for *SfABCC2mut* would be predicted for homozygous frameshift and nonsense mutations, as they would result in lack of a functional *SfABCC2* protein [9,12,13]. Interestingly, some of the frameshift mutations were noticeably frequent. Importantly, a number of individuals positive for these frameshift mutations were homozygous samples predicted to produce truncated *SfABCC2* proteins and thus having a Cry1F-resistant phenotype similar to *SfABCC2mut* [8]. In addition, heterozygous individuals could also be phenotypically resistant to Cry1F in cases of complementation between alleles involving nonsense and frameshift mutations.

Significant gene flow between *S. frugiperda* populations from Brazil and Puerto Rico is not expected based on haplotype marker and meteorological observations [3]. In agreement with this observation, previous genotyping efforts did not detect *SfABCC2mut* among Brazilian *S. frugiperda* populations [9]. However, whole-genome comparisons suggest panmixia among *S. frugiperda* populations [42]. We did find two heterozygote individuals from Lajas (PRL\_168 and PRL\_164) carrying a GY deletion (DGY784D in our model, allele frequency 0.0571) previously described in Cry1F-resistant and field-collected *S. frugiperda* from Brazil [10]. However, detection of this GY deletion in Cry1F-susceptible individuals from Brazil [40] questions the relevance of this mutation to resistance.

The extracellular loops 1 and 4 in *SfABCC2* have been proposed as relevant to Cry1F toxicity [11], with loop 4 being most critical and also a hot spot for resistance mutations [10]. Only the Q788P mutation localized to that loop in our dataset. Interestingly, samples containing this mutation were detected in pooled sequencing of Brazilian *S. frugiperda* populations in one sample each from four locations (Balsas, Dourados, Santa Cruz das Palmeiras, and Uberlandia) [10]. These particular collection sites in Brazil are separated by >1200 km, suggesting that the Q788P mutation may be common in Brazilian *S. frugiperda* populations. Among our samples, the mutation was rare and a single heterozygote individual (PRS\_14) was detected. Functional tests would be needed to test the effect of Q788P on Cry1F toxicity.

Detection of common mutations may be suggestive of gene flow, including resistance alleles, between *S. frugiperda* populations from Brazil and the Caribbean. Although this flow is not expected based on predicted migratory movements [3], commercial exchange of contaminated plant materials could result in unintended dissemination of resistance alleles. This hypothesis should be further tested with increased sample sizes to better understand the potential paths for the spread of *S. frugiperda* resistance to *Bt* crops.

Overall, our analyses detected 374 synonymous and 24 (including *SfABCC2mut*) non-synonymous or frameshift *SfABCC2* mutations in 41 tested *S. frugiperda* individuals from Puerto Rico. These results indicate the *SfABCC2mut* and novel mutations predicted to result in truncated or aberrant SfABCC2 proteins and thus a Cry1F-resistant phenotype are not rare in *S. frugiperda* from Puerto Rico. Out of the 41 samples tested, the majority (38, 93%) carried at least one copy of a nonsense or frameshift *SfABCC2* mutation. There were 13 (32% of the total) homozygotes for these mutations, including 6 for *SfABCC2mut* and 6 for the A15 allele. Based on similar mutations in knockout and resistant strains of *S. frugiperda* [8,9,11,13], these homozygotes would display a Cry1F-resistant phenotype. The relatively high frequency of mutations predicted to result in resistance to Cry1F through SfABCC2 truncation is in line with the high levels of resistance to Cry1F maintained in Puerto Rico [18,43] and the lack of relevant fitness costs in disruptive *SfABCC2* mutants [15,44]. Our targeted approach does not allow us to identify additional Cry1F-resistant individuals carrying alternative resistance genes, which could potentially increase the frequency of Cry1F-resistant individuals. In addition, while missense *SfABCC2* mutations were also detected, a Cry1F-resistant phenotype cannot be predicted as their role in intoxication is unknown. For instance, the Q788P and A768G mutations could affect the predicted Cry1F-binding ability of the SfABCC2 loop 4 region [10]. Functional assays would be needed to test the effect of these mutations on susceptibility to Cry1F.

## 5. Conclusions

We provide proof of concept for the use of targeted sequencing in the monitoring of known mutations and the discovery of candidate resistance alleles to Cry1F in *S. frugiperda*. Furthermore, based on solid experimental evidence for the critical role of SfABCC2 in susceptibility to Cry1F, we predict the resistance phenotype of homozygous individuals for nonsense and frameshift mutations. This strategy provides higher sensitivity and should reduce labor and costs compared to traditional F1 and F2 screens to determine the frequency of diverse resistance alleles and, in specific mutations, susceptibility to Cry1F. Based on our observations, targeted sequencing overcomes the specificity limitations of Taqman genotyping, especially when targeting highly polymorphic genes. In addition, Taqman assays are limited to known alleles and cannot provide detection of novel candidate resistance alleles. Targeted sequencing provided preliminary evidence of a common mutation associated with Cry1F resistance in *S. frugiperda* from Brazil and Puerto Rico, which may suggest genetic flow between these populations as supported by whole genome-level comparisons [42]. Targeted sequencing such as Hi-Plex is adaptable to other pest-toxin resistance gene models, as long as known or suspected resistance loci are available, and the method can be expanded to sequence multiple genes within a single sample. While targeted sequencing provides an interesting list of candidate resistance alleles, further functional testing is needed to determine the role of missense mutations and their linkage with resistance to Cry1F.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/insects12070618/s1>, Figure S1: Representative partial visualization of targeted sequencing reads aligned to the *SfABCC2* reference gene; Figure S2: Consensus SfABCC2 protein sequence used in this study; Table S1: List of sequenced *S. frugiperda* samples and their genotype for detected mutations; Table S2: List of TILING primers used in PCR amplification of *SfABCC2*.

**Author Contributions:** Conceptualization, K.S., K.L., S.J.E. and J.L.J.-F.; methodology, K.S., K.L., P.T., S.J.E., C.P.d.B., T.R. and J.L.J.-F.; software, K.S., K.L., P.T. and S.J.E.; formal analysis, K.S., K.L., S.J.E. and J.L.J.-F.; resources, D.M.V.D., A.M.L.-R. and J.L.J.-F.; data curation, K.S., K.L., S.J.E., C.P.d.B. and J.L.J.-F.; writing—original draft preparation, K.S., K.L., P.T., S.J.E. and J.L.J.-F.; writing—review and editing, all authors.; visualization, K.S., S.J.E. and J.L.J.-F.; funding acquisition, S.J.E. and J.L.J.-F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Agriculture and Food Research Initiative Foundational Program competitive grant No. 2018-67013-27820 from the USDA National Institute of Food and Agriculture.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** All targeted sequencing data used for this study are available through NCBI BioProject ID PRJNA729034.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Montezano, D.G.; Specht, A.; Sosa-Gómez, 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. [[CrossRef](#)]
- Westbrook, J.; Fleischer, S.; Jairam, S.; Meagher, R.; Nagoshi, R. Multigenerational migration of fall armyworm, a pest insect. *Ecosphere* **2019**, *10*, e02919. [[CrossRef](#)]
- Nagoshi, R.N.; Fleischer, S.; Meagher, R.L.; Hay-Roe, M.; Khan, A.; Murua, M.G.; Silvie, P.; Vergara, C.; Westbrook, J. Fall armyworm migration across the Lesser Antilles and the potential for genetic exchanges between North and South American populations. *PLoS ONE* **2017**, *12*, e0171743. [[CrossRef](#)]
- Storer, N.P.; Babcock, J.M.; Schlenz, M.; Meade, T.; Thompson, G.D.; Bing, J.W.; Huckaba, R.M. Discovery and characterization of field resistance to Bt maize: *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Puerto Rico. *J. Econ. Entomol.* **2010**, *103*, 1031–1038. [[CrossRef](#)] [[PubMed](#)]
- Huang, F.; Qureshi, J.A.; Meagher, R.L., Jr.; Reisig, D.D.; Head, G.P.; Andow, D.A.; Ni, X.; Kerns, D.; Buntin, G.D.; Niu, Y.; et al. Cry1F resistance in fall armyworm *Spodoptera frugiperda*: Single gene versus pyramided Bt maize. *PLoS ONE* **2014**, *9*, e112958. [[CrossRef](#)]
- Farias, J.R.; Andow, D.A.; Horikoshi, R.J.; Sorgatto, R.J.; Fresia, P.; dos Santos, A.C.; Omoto, C. Field-evolved resistance to Cry1F maize by *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Brazil. *Crop Protect.* **2014**, *64*, 150–158. [[CrossRef](#)]
- Chandrasena, D.I.; Signorini, A.M.; Abratti, G.; Storer, N.P.; Olaciregui, M.L.; Alves, A.P.; Pilcher, C.D. Characterization of field-evolved resistance to *Bacillus thuringiensis*-derived Cry1F delta-endotoxin in *Spodoptera frugiperda* populations from Argentina. *Pest Manag. Sci.* **2018**, *74*, 746–754. [[CrossRef](#)] [[PubMed](#)]
- Banerjee, R.; Hasler, J.; Meagher, R.; Nagoshi, R.; Hietala, L.; Huang, F.; Narva, K.; Jurat-Fuentes, J.L. Mechanism and DNA-based detection of field-evolved resistance to transgenic Bt corn in fall armyworm (*Spodoptera frugiperda*). *Sci. Rep.* **2017**, *7*, 10877. [[CrossRef](#)]
- Flagel, L.; Lee, Y.W.; Wanjugi, H.; Swarup, S.; Brown, A.; Wang, J.; Kraft, E.; Greenplate, J.; Simmons, J.; Adams, N.; et al. Mutational disruption of the ABCC2 gene in fall armyworm, *Spodoptera frugiperda*, confers resistance to the Cry1Fa and Cry1A.105 insecticidal proteins. *Sci. Rep.* **2018**, *8*, 7255. [[CrossRef](#)] [[PubMed](#)]
- Boaventura, D.; Ulrich, J.; Lueke, B.; Bolzan, A.; Okuma, D.; Gutbrod, O.; Geibel, S.; Zeng, Q.; Dourado, P.M.; Martinelli, S.; et al. Molecular characterization of Cry1F resistance in fall armyworm, *Spodoptera frugiperda* from Brazil. *Insect Biochem. Mol. Biol.* **2020**, *116*, 103280. [[CrossRef](#)]
- Liu, Y.; Jin, M.; Wang, L.; Wang, H.; Xia, Z.; Yang, Y.; Bravo, A.; Soberón, M.; Xiao, Y.; Liu, K. SfABCC2 transporter extracellular loops 2 and 4 are responsible for the Cry1Fa insecticidal specificity against *Spodoptera frugiperda*. *Insect Biochem. Mol. Biol.* **2021**, *135*, 103608. [[CrossRef](#)]
- Jin, M.-H.; Tao, J.-H.; Li, Q.; Cheng, Y.; Sun, X.-X.; Wu, K.-M.; Xiao, Y.-T. Genome editing of the SfABCC2 gene confers resistance to Cry1F toxin from *Bacillus thuringiensis* in *Spodoptera frugiperda*. *J. Integr. Agric.* **2020**, in press. [[CrossRef](#)]
- Abdelgaffar, H.; Perera, O.P.; Jurat-Fuentes, J.L. ABC transporter mutations in Cry1F-resistant fall armyworm (*Spodoptera frugiperda*) do not result in altered susceptibility to selected small molecule pesticides. *Pest Manag. Sci.* **2021**, *77*, 949–955. [[CrossRef](#)] [[PubMed](#)]
- Vélez, A.M.; Spencer, T.A.; Alves, A.P.; Moellenbeck, D.; Meagher, R.L.; Chirakkal, H.; Siegfried, B.D. Inheritance of Cry1F resistance, cross-resistance and frequency of resistant alleles in *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Bull. Entomol. Res.* **2013**, *103*, 700–713. [[CrossRef](#)] [[PubMed](#)]
- Jakka, S.R.; Knight, V.R.; Jurat-Fuentes, J.L. Fitness costs associated with field-evolved resistance to Bt maize in *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *J. Econ. Entomol.* **2014**, *107*, 342–351. [[CrossRef](#)] [[PubMed](#)]
- Nguyen-Dumont, T.; Teo, Z.L.; Pope, B.J.; Hammet, F.; Mahmoodi, M.; Tsimiklis, H.; Sabbaghian, N.; Tischkowitz, M.; Foulkes, W.D.; Giles, G.G.; et al. Hi-Plex for high-throughput mutation screening: Application to the breast cancer susceptibility gene PALB2. *BMC Med. Genom.* **2013**, *6*, 48. [[CrossRef](#)] [[PubMed](#)]
- Nguyen-Dumont, T.; Pope, B.J.; Hammet, F.; Southey, M.C.; Park, D.J. A high-plex PCR approach for massively parallel sequencing. *BioTechniques* **2013**, *55*, 69–74. [[CrossRef](#)]
- Storer, N.P.; Kubiszak, M.E.; Ed King, J.; Thompson, G.D.; Santos, A.C. Status of resistance to Bt maize in *Spodoptera frugiperda*: Lessons from Puerto Rico. *J. Invertebr. Pathol.* **2012**, *110*, 294–300. [[CrossRef](#)]

19. Meagher, R.L. Trapping fall armyworm (Lepidoptera: Noctuidae) adults in traps baited with pheromone and a synthetic floral volatile compound. *Fla. Entomol.* **2001**, *84*, 288–292. [CrossRef]
20. Capinera, J.L.; University of Florida. Fall Armyworm. Publication Number EENY-98. 2017. Available online: [http://entnemdept.ufl.edu/creatures/field/fall\\_armyworm.htm#host](http://entnemdept.ufl.edu/creatures/field/fall_armyworm.htm#host) (accessed on 20 May 2021).
21. Blanco, C.A.; Portilla, M.; Jurat-Fuentes, J.L.; Sanchez, J.F.; Viteri, D.; Vega-Aquino, P.; Teran-Vargas, A.P.; Azuara-Dominguez, A.; Lopez, J.D.J.; Arias, R.; et al. Susceptibility of isofamilies of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) to Cry1Ac and Cry1Fa proteins of *Bacillus thuringiensis*. *Southwest. Entomol.* **2010**, *35*, 409–415. [CrossRef]
22. Gervais, A.L.; Marques, M.; Gaudreau, L. PCRTiler: Automated design of tiled and specific PCR primer pairs. *Nucleic Acids Res.* **2010**, *38*, W308–W312. [CrossRef] [PubMed]
23. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2013**, *30*, 614–620. [CrossRef] [PubMed]
24. Bushnell, B. BMap (BMap Short Read Aligner, and Other Bioinformatic Tools). Available online: <https://jgi.doe.gov/data-and-tools/bbtools/> (accessed on 20 June 2019).
25. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequencing Data [Online]. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 21 December 2017).
26. Gimenez, S.; Abdelgaffar, H.; Goff, G.L.; Hilliou, F.; Blanco, C.A.; Hänniger, S.; Bretaudeau, A.; Legeai, F.; Nègre, N.; Jurat-Fuentes, J.L.; et al. Adaptation by copy number variation increases insecticide resistance in the fall armyworm. *Commun. Biol.* **2020**, *3*, 664. [CrossRef]
27. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
28. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [CrossRef]
29. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]
30. Slater, G.S.; Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* **2005**, *6*, 31. [CrossRef]
31. Jombart, T.; Ahmed, I. ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **2011**, *27*, 370–3071. [CrossRef]
32. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **2012**, *6*, 80–92. [CrossRef]
33. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [CrossRef] [PubMed]
34. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
35. Delaneau, O.; Marchini, J.; Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **2011**, *9*, 179–181. [CrossRef]
36. Cheng, T.; Wu, J.; Wu, Y.; Chilukuri, R.V.; Huang, L.; Yamamoto, K.; Feng, L.; Li, W.; Chen, Z.; Guo, H.; et al. Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. *Nat. Ecol. Evol.* **2017**, *1*, 1747–1756. [CrossRef]
37. Raj, A.; Stephens, M.; Pritchard, J.K. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* **2014**, *197*, 573–589. [CrossRef] [PubMed]
38. Niu, Y.; Qureshi, J.A.; Ni, X.; Head, G.P.; Price, P.A.; Meagher, R.L.; Kerns, D.; Levy, R.; Yang, X.; Huang, F. F2 screen for resistance to *Bacillus thuringiensis* Cry2Ab2-maize in field populations of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) from the southern United States. *J. Invertebr. Pathol.* **2016**, *138*, 66–72. [CrossRef]
39. Farias, J.R.; Andow, D.A.; Horikoshi, R.J.; Bernardi, D.; Ribeiro, R.D.S.; do Nascimento, A.R.; Dos Santos, A.C.; Omoto, C. Frequency of Cry1F resistance alleles in *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Brazil. *Pest Manag. Sci.* **2016**, *72*, 2295–2302. [CrossRef] [PubMed]
40. Yainna, S.; Nègre, N.; Silvie, P.J.; Brévault, T.; Tay, W.T.; Gordon, K.; dAlençon, E.; Walsh, T.; Nam, K. Geographic monitoring of insecticide resistance mutations in native and invasive populations of the fall armyworm. *Insects* **2021**, *12*, 468. [CrossRef] [PubMed]
41. Viteri, D.M.; Linares, A.M.; Cabrera, I.; Sarmiento, L. Presence of corn earworm and fall armyworm (Lepidoptera: Noctuidae) populations in sweet corn and their susceptibility to insecticides in Puerto Rico. *Fla. Entomol.* **2019**, *102*, 451–454. [CrossRef]
42. Schlum, K.A.; Lamour, K.; de Bortoli, C.P.; Banerjee, R.; Meagher, R.; Pereira, E.; Murua, M.G.; Sword, G.A.; Tessnow, A.E.; Viteri Dillon, D.; et al. Whole genome comparisons reveal panmixia among fall armyworm (*Spodoptera frugiperda*) from diverse locations. *BMC Genom.* **2021**, *22*, 179. [CrossRef] [PubMed]
43. Gutierrez-Moreno, R.; Mota-Sanchez, D.; Blanco, C.A.; Chandrasena, D.; Difonzo, C.; Conner, J.; Head, G.; Berman, K.; Wise, J. Susceptibility of fall armyworms (*Spodoptera frugiperda* j.e.) from Mexico and Puerto Rico to Bt proteins. *Insects* **2020**, *11*, 831. [CrossRef] [PubMed]
44. Vélez, A.M.; Spencer, T.A.; Alves, A.P.; Crespo, A.L.B.; Siegfried, B.D. Fitness costs of Cry1F resistance in fall armyworm, *Spodoptera frugiperda*. *J. Appl. Entomol.* **2013**, *138*, 315–325. [CrossRef]