Methodology

Microsatellite markers for genetic studies of the fall armyworm, Spodoptera frugiperda

V.A.C. Pavinato¹, S. Martinelli², P.F. de Lima³, M.I. Zucchi⁴ and C. Omoto¹

¹Laboratório de Resistência de Insetos a Táticas de Controle, Departamento de Entomologia e Acarologia, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, Brasil
²Monsanto do Brasil Ltda., São Paulo, SP, Brasil
³Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, Campinas, SP, Brasil
⁴Agência Paulista de Tecnologia dos Agronegócios, Piracicaba, SP, Brasil

Corresponding author: M.I. Zucchi
E-mail: mizucchi@apta.sp.gov.br

Received April 24, 2012
Accepted August 13, 2012
Published February 8, 2013
DOI http://dx.doi.org/10.4238/2013.February.8.1

ABSTRACT. We developed six microsatellite markers for the fall armyworm Spodoptera frugiperda (Lepidoptera: Noctuidae). The SSR loci were isolated with enriched genomic library protocol by using native individuals as a genome source for markers. These loci were characterized in 48 individuals and they were tested for the ability to identify candidate migrants exchanged among the samples. The number of alleles per locus ranged from 5 to 18 (10.8 on average). The observed polymorphism information content ranged from 0.172 to 0.891. Beside the lower efficiency to obtain SSR loci, the six microsatellites were polymorphic and sufficiently discriminant for the genetic studies of S. frugiperda; it allowed us to identify migrants with both NJ clustering and the Bayesian methods. These markers will be useful for molecular
ecology studies of this highly polyphagous species in order to understand the processes that determine genetic differentiation in the complex agro-ecosystems that it infests and improve local integrated pest management practices.

Key words: Fall armyworm; Genetic structure; Candidate migrant; SSRs; Molecular markers; Insect resistance management

INTRODUCTION

The fall armyworm, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera: Noctuidae), is an insect native to the Americas, occurring from the South and Central Americas to North America (Hardke et al., 2011). This species has a polyphagous feeding habit, and its larval diet includes a broad variety of host plants such as maize, cotton, and rice (Yu et al., 2003). In Brazil, *S. frugiperda* is a major insect pest, especially in maize (Sena Jr. et al., 2003) and cotton crops. This species has a great dispersion capability that enables its fast spread through host plants (Johnson, 1987) and determines its genetic complexities and migration patterns (Nagoshi et al., 2008). The occurrence of two host strains of the fall armyworm species has been reported - one associated with maize/cotton plants (host strain C) and the other with rice (host strain R) (Pashley et al., 1992). Differences in female pheromone composition (Groot et al., 2008) and differential timing of reproductive activity at night (Schöfl et al., 2009) might be the major forces related to the reproductive isolation detected between these host strains.

The fall armyworm has many characteristics that make it a model species in ecology and evolutionary biology. Challenges in these research areas are to understand host-plant adaptation, multi-scale geographic genetic differentiations, and population ecology of sympatric speciation. Basic geographic differentiations and host-range association studies have been performed using allozyme markers (Pashley et al., 1985; Pashley, 1986), random amplified polymorphic DNA and amplified fragment length polymorphism markers (Martinelli et al., 2006, 2007), and mitochondrial cytochrome oxidase subunit I genes (Nagoshi et al., 2006; Nagoshi, 2010). However, the use of highly variable and sensitive molecular markers as microsatellites has not yet been reported in this species in molecular ecology studies.

Microsatellite markers, also called simple sequence repeats (SSRs), are highly polymorphic and abundant genetic markers that exist throughout all eukaryotic genomes (Goldstein and Schlötterer, 1999). In addition to their early applications in genome mapping studies and paternity tests, the so-called microsatellites are still the most powerful molecular markers available in molecular ecology and evolutionary studies. Inherent characteristics of high locus variability, ease and reliability of scoring, co-dominant inheritance, short length, and association with powerful statistical analyses such as Bayesian and maximum likelihood methods make their use widespread among insect population geneticists and ecologists (Fuentes-Contreras et al., 2008; Meng et al., 2008; Lagisz et al., 2010). The major objective of this research was to develop a new microsatellite marker set by using genomic enriched library protocols for *S. frugiperda* collected in Brazil and test its applicability as a tool for further population genetic studies of this important pest species.
MATERIAL AND METHODS

DNA extraction for the microsatellite-enriched library protocol

Fall armyworm individuals were collected in 2006 in three different maize-producing regions in Brazil: (1) São Paulo (G); (2) Minas Gerais/Goiás (U), and (3) Mato Grosso/Mato Grosso do Sul (M) (Figure 1). Genomic DNA was extracted from fresh adult thoraces using a Doyle and Doyle (1990) cetyltrimethylammonium bromide protocol with minor modifications. Before homogenization, the head, wings, and abdomen were removed from the collected adults. The separated thoraces were homogenized in preheated extraction buffer [1 M Tris-HCl, pH 8.0, 5 mM NaCl, 500 mM ethylenediaminetetraacetic acid, pH 8.0, and 2% (w/v) cetyltrimethylammonium bromide without β-mercaptoethanol]. After the homogenization, the DNA was isolated with chloroform/isoamyl alcohol (24:1) and precipitated with 100% isopropanol. The DNA pellets were washed with 70% ethanol and air dried, then resuspended in 50 µL TE buffer (10 mM Tris-HCl, pH 8.0 and 1 mM ethylenediaminetetraacetic acid, pH 8.0) and treated with 10 µg/µL RNase. Each sample was quantified on a 1% agarose gel with a λ DNA marker (22.2 ng/mL).

Figure 1. Map of Brazil. Red stars represent the approximate location of each collection. G, U and M represent the São Paulo population, Unaí, Minas Gerais/Goiás population and Mato Grosso/Mato Grosso do Sul population.
Microsatellite isolation and primer design

We obtained six markers following adapted microsatellite-enriched library protocols described by Billotte et al. (1999). Genomic DNA from one genotype of S. frugiperda was digested with RsaI (Invitrogen), and we performed the enrichment in microsatellite fragments using (CT)\(_n\) and (GT)\(_n\) motifs. The enriched fragments were cloned into pGEM-T vectors (Promega) and ligation products were used to transform Epicurian Coli XL1-Blue Escherichia coli competent cells (Stratagene). The positive clones were selected using the β-galactosidase gene and then grown overnight with ampicillin. We sequenced 109 clones in an ABI 377 automated sequencer (PE Applied Biosystems) with a BigDye terminator cycle sequencing kit (Applied Biosystems). The TROLL software (Castelo et al., 2002) was used to search SSR regions in the sequences obtained. Nine primer pairs were designed using Primer 3 (Rozen and Skaletsky, 2000), and a series of polymerase chain reaction (PCR) analyses were performed in 48 specimens to optimize the loci obtained.

Locus microsatellite characterization

Amplifications were performed in a final volume of 15 μL containing 20 ng genomic DNA, 1X PCR buffer (10 mM Tris-HCl, pH 8.5 and 50 mM KCl), 1.5 mM MgCl\(_2\), 0.2 μM of each designed primer pair (forward and reverse), 100 μM of each deoxyribonucleotide triphosphate, and 1 U Taq DNA polymerase (Invitrogen). The amplification cycle used in the thermocycler PTC100 (MJ Research, Inc.) was programmed to start at 94°C for 12 min, followed by 32 cycles of 1 min of denaturation at 94°C, 1 min at the specific annealing temperature for each primer (Table 1), and 1 min of extension at 72°C, finalizing with 20 min at 72°C. We separated the alleles for each locus by genotyping the individuals with 7% denaturing polyacrylamide gel electrophoresis visualized with silver staining. The 10-bp DNA Ladder (Invitrogen) was used as a size standard to score allele variability.

Data analysis

By genotyping 48 S. frugiperda individuals, we accessed descriptive statistics [expected (H\(_E\)) and observed heterozygosities (H\(_O\)) and polymorphic index content] using MSTools (http://animalgenomics.ucd.ie/sdepark/ms-toolkit/). The population parameters [Nei’s unbiased H\(_E\) estimate, H\(_O\), and Weir and Cockerham’s measure of the departure of genotype frequencies from Hardy-Weinberg (HW) expectations within population (aka inbreeding coefficient F\(_{IS}\))] and the Fisher exact test for departure from HW proportions for each locus into populations were performed using the Genetic Data Analysis software (Weir, 1996). For average F\(_{IS}\) estimative, we obtained the confidence interval across loci for each population over 10,000 bootstraps.

When departures from HW proportions were detected, the frequency of null alleles was calculated for each locus using maximum likelihood estimations via the expectation-maximization algorithm implemented in FREENA (Chapuis and Estoup, 2007).
Composite gametic disequilibrium was tested using Genetic Data Analysis (Weir, 1996). Bonferroni’s correction was used to adjust the nominal level for all multiple tests.

The efficiency of the SSRs obtained as molecular markers for genetic studies of *S. frugiperda* was tested by estimating the genetic differentiation of the samples collected. First, Weir and Cockerham’s overall estimate of population differentiation ($F_{ST}$), its confidence interval with 1000-bootstrap resampling, and pairwise $F_{ST}$ were calculated using FSTAT (Goudet, 2001).

A model-based approach was used to identify populations as clusters of individuals. This method first determines whether the data fit a model with *k* populations. The algorithm searches the best *k* set of individuals that minimizes the amount of HW and gametic disequilibrium in the data. Once *k* is inferred, the software algorithm estimates the posterior probability (*Q*) of individual genotypes originating from each population (*k* clusters) (Allendorf and Luikart, 2007). This analysis was performed with the STRUCTURE software (Pritchard et al., 2000), assuming a model with admixture and correlated allele frequencies. Burn-in length was set at 50,000 iterations followed by a run phase of 500,000 iterations. Five independent runs were performed, with the number of potential populations (*k*) ranging from 1 to 7. The best *k* (number of clusters that best represents the data set) was determined using a method described by Evanno et al. (2005).

In addition, the genetic relationships of the individuals were explored through calculation of a neighbor-joining (NJ) dendrogram based on chord distance (Cavalli-Sforza and Edwards, 1967). The estimated genetic distances were accessed with a resampling of 10,000 bootstraps over loci using MICROSAT (http://genetics.stanford.edu/hpgl/projects/microsat/). A consensus tree with node support was obtained using the PHYLIP package of the CONSENSUS program (Felsenstein, 1989). We drew the dendrogram using the DARWIN software (http://darwin.cirad.fr/darwin) (Perrier and Jacquemoud-Collet, 2006). Both the model-based and the individual-based methods were used to access the number of migrants exchanged between the samples.

**RESULTS**

Of 109 sequenced inserts, 60 (55.05%) presented at least one microsatellite region. We designed, synthesized, and tested primer sequences through 9 sequenced inserts that showed SSR loci (8.26% of the total sequences), but only 6 showed satisfactory PCR products (5.50% of the total sequences). These 6 loci were then tested in 48 specimens of *S. frugiperda* to assess polymorphism information and characterize the population parameters.

The 6 isolated loci were polymorphic. The information of each locus is shown in Table 1. We observed departure from HW proportions in 4 loci (Spf01, Spf03, Spf06, and Spf09) when we analyzed all individual as one sample. Therefore, we separated the individuals in their original populations and performed the test again. This time, the Fisher exact test for HW equilibrium showed that the Spf01 locus departed significantly from HW proportions in the three populations studied, and the Spf06 and Spf08 loci departed from those proportion loci in Mato Grosso/Mato Grosse do Sul population (M). No linkage disequilibrium was detected among all loci in the three populations.

$H_e$, $H_o$ and $F_{IS}$ for each locus and each population are presented in Table 2. In all
### Table 1. Characteristic of the 6 microsatellite loci isolated for *Spodoptera frugiperda* (Smith).

<table>
<thead>
<tr>
<th>Locus</th>
<th>GenBank accession</th>
<th>Motif</th>
<th>Primer sequences</th>
<th>Ta (°C)</th>
<th>Allele size range (bp)</th>
<th>N_a</th>
<th>H_E</th>
<th>H_O</th>
<th>PIC</th>
<th>P value HWP</th>
</tr>
</thead>
</table>
| Spf01  | GF119035          | (TG)_{14} 14 bp (AAAGG) | F: TTTTGATTGGGTATTGATGCAT  
R: GGCTTTATGCAACCGGATA | 60     | 280-360         | 12  | 0.873 | 0.146 | 0.850 | 0.000* |
| Spf03  | GF100741          | (GT)_{14}  | F: AATCAAGCTATGCAACCGAG  
R: AGGCGCCAGGATGCTAGCT | 60     | 140-238         | 18  | 0.912 | 0.483 | 0.888 | 0.002* |
| Spf05  | GF100742          | (AC)_{6}   | F: AGCCACCCAACGACCTACAC  
R: GGAATCGCAGACGCTTGTC | 60     | 238-262         | 7   | 0.694 | 0.638 | 0.652 | 0.391 |
| Spf06  | GF100743          | (AC)_{12}  | F: CAGAGAAATGCCTTCTGACTCTT  
R: GGTGATTTATGCTAAAGGCTC | 58     | 220-292         | 12  | 0.917 | 0.154 | 0.891 | 0.000* |
| Spf08  | GF100740          | (GT)_{6}   | F: TCTTCTTACGCGTGGAGCTA  
R: GGAACCTGAGATCGTCTGCG | 59     | 178-190         | 5   | 0.178 | 0.125 | 0.172 | 0.040 |
| Spf09  | GF100744          | (AC)_{6}   | F: TCAAGCCCTACTCGTCACATC  
R: GCGTATTTATGCTTTGAG | 59     | 248-360         | 12  | 0.881 | 0.176 | 0.855 | 0.000* |

Mean: 10.8 0.742 0.287

F = forward; R = reverse; Ta = annealing temperature; N_a = number of alleles per locus; H_E = expected and H_O = observed heterozygosities; PIC = polymorphic informative content; P value for departures from Hardy-Weinberg proportions (HWP). *Depart significantly from HWP after Bonferroni’s correction (P < 0.008).
population analyzed, the $H_O$ were smaller than the $H_E$, which indicated a high $F_{IS}$ for the loci obtained. In addition, Weir and Cockerham’s overall estimate of $F_{IS}$ (see Table 2) showed high degree of inbreeding in the populations. One aspect of $F_{IS}$ can be explained by the small sample size. Additional explanations are the presence of null alleles and non-random mating. To assess the effect of null alleles in the determination of deficiencies in heterozygosity, we estimated the null allele frequency for each loci and population using FREENA. These approaches allowed disentangle the effects of non-random mating from null allele presence leading to homozygous excess. Genome-wide effects as inbreeding occurred on all genome regions equally, although the presence of null alleles is a locus-specific event affecting one or more populations.

To assess the Wahlund effect on HW departures, we assessed the genetic structure through three approaches. The estimate of overall $F_{ST}$ was 0.102 (95% confidence interval = 0.056-0.136), showing that 10% of the genetic variability was structured between the populations. To determine which samples contributed to the total genetic differentiation, we calculated the pairwise $F_{ST}$, which showed that populations M and U were the most similar ($F_{ST} = 0.0696$). The $F_{ST}$ between populations G and U was 0.0844, and between G and M populations, its estimate was 0.1635, demonstrating a degree of genetic differentiation among geographically distant samples. The Bayesian method showed the same pattern of genetic structure. Figure 2 shows that the best $k$ to explain the data was 3. Most of the individuals were grouped at cluster $k = 1$ (red bars). Populations G and U were composed mostly of individuals from cluster 1. The individuals collected from the M population were divided onto two clusters (blue and green bars). Individuals from the two other clusters also composed the U population. This sample site could be a convergent region, a hypothesis supported by the presence of candidate migrants correlated with individuals from the M sample (U20, U26, U27, U30, and U32; see Figure 1) and by the lower estimate of pairwise $F_{ST}$ between the G/U and U/M population pairs. The NJ cluster showed the same genetic structure obtained using the Bayesian approach and revealed the kinship relationships between the candidate migrants found (Figure 3).

### Table 2. Populations’ parameters of 6 microsatellite loci obtained for *Spodoptera frugiperda* in three different sample collections.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$H_E$</th>
<th>$H_O$</th>
<th>$F_{IS}$</th>
<th>Frequency of null alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G</td>
<td>U</td>
<td>M</td>
<td>G</td>
</tr>
<tr>
<td>Spf01</td>
<td>0.703</td>
<td>0.903</td>
<td>0.727</td>
<td>0.125</td>
</tr>
<tr>
<td>Spf03</td>
<td>0.616</td>
<td>0.927</td>
<td>0.823</td>
<td>0.375</td>
</tr>
<tr>
<td>Spf05</td>
<td>0.717</td>
<td>0.723</td>
<td>0.610</td>
<td>0.625</td>
</tr>
<tr>
<td>Spf06</td>
<td>0.848</td>
<td>0.835</td>
<td>0.675</td>
<td>0.272</td>
</tr>
<tr>
<td>Spf08</td>
<td>0.121</td>
<td>0.340</td>
<td>0.065</td>
<td>0.125</td>
</tr>
<tr>
<td>Spf09</td>
<td>0.732</td>
<td>0.877</td>
<td>0.818</td>
<td>0.223</td>
</tr>
<tr>
<td>Total</td>
<td>0.607</td>
<td>0.740</td>
<td>0.598</td>
<td>0.324</td>
</tr>
<tr>
<td>Lower</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$G =$ São Paulo; $U =$ Minas Gerais/Goiás; $M =$ Mato Grosso/Mato Grosso do Sul; $H_E =$ expected heterozygosity; $H_O =$ observed heterozygosity; $F_{IS} =$ the inbreeding coefficient. *Departure significantly from HW proportions after Bonferroni’s correction ($P < 0.008$); ‘Weir and Cockerham’s estimative of $F_{IS}$ for each population.'
DISCUSSION

The overall efficiency in isolating microsatellite loci from the *S. frugiperda* genome...
was low (5.38% sequence inserts obtained with a genome-enriched library protocol). This low efficiency was unrelated to problems in the protocol or the execution stage of library enrichment. Instead, it was related to difficulties inherent in Lepidoptera. Microsatellite isolation from Lepidoptera has been difficult, and the major characteristics of its genome, which cause difficulties in isolating microsatellite markers, are the low frequency of SSR loci and high redundancy of the regions flanking them, which lead to nonspecific amplification of alleles (Nève and Meglécz, 2000). Despite the difficulties, many SSRs have recently been reported in moths (Scott et al., 2004; Dalecky et al., 2006; Kim et al., 2008; Abdoulaye et al., 2011).

In the present study, 6 polymorphic and informative SSR loci were successfully isolated and tested in a population genetic framework. Microsatellite markers for *S. frugiperda* were first developed and characterized by Arias et al. (2011) using individuals from sample sites in the United States and Puerto Rico. Some SSR markers have been shown ineffective in populations that are genetically distant; therefore, efforts to develop microsatellite markers that work properly in populations collected in Brazil are justified.

The HW deviation observed at Spf01, Spf03, Spf06, and Spf09 was consistent owing to the presence of null alleles at the population level and the presence of genetic structure in the data. The presence of null alleles in microsatellite loci is the major cause of deviations in HW proportions because it confounds genotyping and leads to an accounting of more homozygotes (Chakraborty et al., 1992). In addition, HW deviations are also determined by inbreeding; the Wahlund effect arises from population subdivisions (Weir, 1996).

Clusters analyses are strongly affected by the presence of null alleles in estimates of genetic distances, although Cavalli-Sforza chord distances are less affected than Nei’s standard genetic distance (Chapuis and Estoup, 2007). To increase the power of the estimation of genetic differentiation, we conducted a Bayesian estimate of $F_{ST}$. This approach permits more accurate determinations of individual differentiation compared with that using traditional NJ clustering, which makes individual assignments when previous population structure information is lacking and is a powerful means of estimating $F_{ST}$ even with a small number of microsatellite loci (Pritchard et al., 2000).

We isolated 6 high polymorphic and informative SSR markers for *S. frugiperda*. These markers were tested in small populations. Even under these conditions, we found some degrees of genetic structure among populations and identified putative migrants exchanged among sample sites. The 6 microsatellites developed for *S. frugiperda* showed high polymorphism and degree of diversity in the populations analyzed. These characteristics are expected for informative microsatellite loci that can be used in further population genetic studies of this species.

These genetic studies will be important for obtaining information about basic ecology. Time-scale and host-range micro-evolutionary patterns of *S. frugiperda* in the context of agro-ecosystems aid in the development of best pest management practices. The use of discriminatory markers capable of identifying migrants permits quantification of the number of genes exchanged between populations. In the context of insect resistance, gene flow is the most potent evolutionary force for spreading resistant alleles throughout over the pest populations (Caprio et al., 1992). In addition, understanding the patterns of genetic structure and gene flow can provide guidance for designing superior management programs for *S. frugiperda* resistant to insecticides and genetically modified crops.
ACKNOWLEDGMENTS

Research supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (Process #06/03373-9) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (Process #578509/2008-3).

REFERENCES


Hardke JT, Leonard BR, Huang F and Jackson RE (2011). Damage and survivorship of fall armyworm (Lepidoptera: Noctuidae) on transgenic field corn expressing Bacillus thuringiensis Cry proteins. Crop Protect. 30: 168-172.


Meng XF, Shi M and Chen XX (2008). Population genetic structure of Chilo suppressalis (Walker) (Lepidoptera:


