Development and characterization of microsatellite loci for genetic studies of the sugarcane borer, *Diatraea saccharalis* (Lepidoptera: Crambidae)

V.A.C. Pavinato[1], K.L. Silva-Brandão[2], M. Monteiro[3], M.I. Zucchi[3], J.B. Pinheiro[4], F.L.F. Dias[3] and C. Omoto[2]

[1]Instituto de Biologia, Universidade Estadual de Campinas,
Cidade Universitária Zeferino Vaz, Campinas, SP, Brasil
[2]Departamento Entomologia e Acarologia,
Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo,
Piracicaba, SP, Brasil
[3]Agência Paulista de Tecnologia dos Agronegócios, Piracicaba, SP, Brasil
[4]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz",
Universidade de São Paulo, Piracicaba, SP, Brasil

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

ABSTRACT. We present polymorphic microsatellite markers isolated for genetic studies of the sugarcane borer, *Diatraea saccharalis* (Fabricius). We isolated 16 microsatellite loci through an enriched genomic library protocol. After characterization, 12 markers showed polymorphic information expressed in the observed number of alleles (ranging from 2 to 7; 5 on average) and in the polymorphism information content (ranging from 0.292 to 0.771; 0.535 on average).

These markers can be used in further studies to understand the basic ecological characteristics of the sugarcane borer, e.g., dispersion patterns and population genetic differentiation, associated with distinct geographic scales and host plants.

Key words: Microsatellite; Population genetics; Lepidoptera evolution; Genetic diversity; Gene flow; Management units

INTRODUCTION

The sugarcane borer, *Diatraea saccharalis* (Fabricius), is one of the major sugarcane (*Saccharum officinarum* L.) lepidopteran pests across the Western hemisphere (Pashley et al., 1990). This species has encountered suitable conditions for its development in grass-related crop fields and the outbreak of *D. saccharalis* populations is historically associated to sugarcane, maize (*Zea mays* L.), sorghum (*Sorghum bicolor* L.), and rice (*Oryza sativa* L.) agricultural expansion in the Southern United States of America and Central and South America (Botelho, 1992; Castro et al., 2004).

D. saccharalis presents important traits such as the use of different host plants (Long and Hensley, 1972; Moré et al., 2003) and pheromone composition (Cortés et al., 2010) that make it a suitable model in ecology and evolutionary biology studies. These facts challenge evolutionary studies of *D. saccharalis* conducted to understand multi-scale geographic genetic differentiations, host-plant adaptation, and population ecology of species hybridization.

Microsatellites, also called simple-sequence repeats, are highly polymorphic and abundant genetic markers. Their characteristics, e.g., high loci variability, easy determination, and good reliability of scoring and co-dominant inheritance, associated with powerful statistical analysis as Bayesian and maximum likelihood methods (Luikart and England, 1999), make them widely applied by insect molecular ecologists (Behura, 2006; Beadell et al., 2010; Aggarwal et al., 2011). Here, we present the efforts to develop primers to access microsatellite loci for future genetic studies of *D. saccharalis*.

MATERIAL AND METHODS

Total genomic DNA was extracted from fresh thoracic tissue of adults using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA). DNA was diluted in 50 µL TE buffer (10 mM Tris-HCl, pH 8.0, and 1 mM EDTA, pH 8.0) and stored at -20°C. Each sample was quantified by running on a 1% agarose gel with a 22.2-ng/mL λ DNA marker. A microsatellite-enriched library was obtained using adapted protocols from Billotte et al. (1999). Genomic DNA from 4 genotypes of *D. saccharalis* were digested with *Afa*I (Invitrogen, Carlsbad, CA, USA), enriched in microsatellite fragments using $(CT)_8$ and $(GT)_8$ motifs. The enriched fragments were cloned into *pGEM-T* (Promega) and ligation products were used to transform Epicurian Coli XL1-Blue *Escherichia coli* competent cells. The positive clones were selected using the *β-galactosidase* gene and then grown overnight in the presence of ampicillin. All clones were sequenced on an ABI

3730 automated sequencer (PE Applied Biosystems, Carlsbad, CA, USA) using a BigDye terminator cycle sequencing kit (Applied Biosystems). A total of 16 primer pairs were designed using Primer 3 v. 0.4.0 (Rozen and Skaletsky, 2000) and tested in 30 specimens of *D. saccharalis* collected on sugarcane plants from 1 sample site (Ribeirão Preto - 21°10'S, 47°49'W) in the State of São Paulo, Brazil.

Polymerase chain reactions were performed in a 20-µL volume containing: 1X reaction buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl, and 2.0 mM $MgCl_2$), 9 ng genomic DNA, 0.2 mM forward and reverse primers, 0.25 mM dNTPs, 50 ng bovine serum albumin, and 1 U recombinant *Taq* DNA polymerase (Invitrogen). All reactions were performed in the Applied Veriti 384 thermal cycler. The polymerase chain reaction program consisted of an initial denaturing step at 95°C for 5 min followed by 40 cycles of amplification (95°C for 30 s, 45 s at the specific annealing temperature of each primer pair (Table 1), and 72°C for 45 s), 8 additional cycles of amplification (94°C for 30 s, 53°C for 45 s, and 72°C for 45 s), and a final elongation step at 72°C for 10 min. Amplification products were resolved by electrophoresis on 7% denaturing polyacrylamide gels and visualized by silver staining. The allele scoring was done using the 10-bp DNA ladder (Invitrogen) as size standard.

Descriptive statistics (expected and observed heterozygosities and polymorphism information content) were calculated using the MSTools applicative (Stephen Park, http://animalgenomics.ucd.ie/sdepark/ms-toolkit/). The Fisher exact tests for deviation from Hardy-Weinberg proportions were performed for each locus using the R package "pegas", version 0.4.1 (Paradis, 2009). When deviations from Hardy-Weinberg proportions were detected, the frequency of null alleles was calculated for each locus using the maximum-likelihood estimation via the EM algorithm implemented in FREENA (Chapuis and Estoup, 2007). The composite gametic disequilibrium was tested using the Genetics Data Analysis II program (Weir, 1996). Bonferroni's correction was used to correct nominal level for all multiple tests.

RESULTS AND DISCUSSION

Twelve of 16 loci were polymorphic and informative for population genetic studies (Table 1). The Fisher exact test showed that only the locus Dsc3 deviated significantly from Hardy-Weinberg proportion after Bonferroni's correction (P < 0.004). The excess of homozygotes leading to deviations from Hardy-Weinberg proportions was caused by the frequency of the null allele (frequency = 0.248). No gametic disequilibrium was detected among all loci.

Molecular ecology studies with highly reliable and statistically powerful molecular markers as microsatellites can help us to identify patterns of dispersion/migration determining scales of genetic divergence (Torriani et al., 2010), population differentiation through geographic discontinuities (Abila et al., 2008), the degree of population differentiation related to host plants (Carletto et al., 2009), and the process of sympatric speciation (Santos et al., 2011). Microsatellite markers developed for *D. saccharalis* in this study can be applied in further population genetic studies to address ecological and evolutionary questions and also to improve our ability to manage populations of this species according to local integrated pest management practices.

ACKNOWLEDGMENTS

Table 1. Characteristics of the 12 microsatellite loci from *Diatraea saccharalis* (Fabr.).

Locus	GenBank accession	Primer nucleotide sequence (5'-3')	Repeat motif	Ta (°C)	Size range (bp)	N	N_A	H_O	H_E	PIC	Null allele frequency	P value H-W proportions
Dsc1	GFI11048	F: CGAGGCTATATTTGCGTGTG R: GATGATGGAGTTGGAAGGTGA	$(TG)_{10}$	56	180-192	25	4	0.640	0.688	0.610	0.004	0.228
Dsc2	GFI11061	F: GCGGTGCCTCTTTGTCATA R: TTGACCAACTACTGCAAGACG	$(CA)_{19}$	60	188-230	22	6	0.591	0.776	0.720	0.077	0.063
Dsc3	GFI11049	F: CCATCAAGCTCCTTCTAAGAGAC R: CCTTGCTCAGTTACCATTCG	$(AC)_{11}$	56	250-274	19	7	0.3684	0.818	0.771	0.248	0.000*
Dsc7	GFI11051	F: TGTCGAGCTACTCCATGCTT R: TGAGACTGAACACTGGCAAGA	$(ATG)_6$	60	214-250	28	5	0.786	0.631	0.562	0.000	0.160
Dsc9	GFI11052	F: AACCTTCGATGAGCTACTGC R: TGTGGTGATTTGTTTGCTTG	$(TG)_{16}$	56	160-182	22	4	0.455	0.562	0.511	0.096	0.047
Dsc10	GFI11060	F: GGTCCGCGTTTGTTATTGTT R: TCAAGTGCTCCTTAAAACACGA	$(GT)_7$	56	270-280	29	2	0.552	0.407	0.320	0.000	0.066
Dsc11	GFI11990	F: ATACGGCTTCATTCGCTTC R: GGTTTCGCACTCATCACG	$(GT)_{10}$	54	220-228	29	3	0.448	0.547	0.476	0.053	0.320
Dsc13	GFI11053	F: CGTGGACTAACCATAGAAGAT R: GGTTTAGCAGAACTTGGCATA	$(GT)_{18}$	54	220-270	26	7	0.538	0.719	0.666	0.079	0.061
Dsc14	GFI11991	F: CTATTCTCCGTTCCGCTGAT R: GAATGAGATTAT GTGTTATGTGTATGC	$(AC)_{16}$	60	84-104	16	6	0.688	0.794	0.739	0.057	0.111
Dsc16	GFI11055	F: TGTGGGTGAGTGCGTGTAA R: GCGTGGACTAACAGTTTTCG	$(TA)_5$	56	250-280	29	4	0.621	0.464	0.391	0.000	0.207
Dsc19	GFI11058	F: CACACACGAACACACACGA R: ATGGTTGGGTCTTTCCTTTT	$(CA)_{10}$	60	160-170	27	4	0.250	0.317	0.292	0.047	0.192
Dsc20	GFI11059	F: TTGGCAGAGTTGTGGGTAAC R: ACAGCAGCATCATCAGAAGG	$(AG)_8$	54	220-230	25	2	0.240	0.490	0.365	0.165	0.016
Mean							5	0.503	0.593	0.528		

F = forward primer sequences; R = reverse primer sequences; Ta = annealing temperature; N = individual successfully genotyped; N_A = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content. P value from the exact test for Hardy-Weinberg proportions. *Departure significantly from H-W proportions after Bonferroni's correction (P < 0.004).

REFERENCES

Abila PP, Slotman MA, Parmakelis A, Dion KB, et al. (2008). High levels of genetic differentiation between Ugandan *Glossina fuscipes fuscipes* populations separated by Lake Kyoga. *PLoS Negl. Trop. Dis.* 2: e242.

Aggarwal RK, Allainguillaume J, Bajay MM, Barthwal S, et al. (2011). Permanent genetic resources added to Molecular Ecology Resources Database 1 August 2010-30 September 2010. *Mol. Ecol. Resour.* 11: 219-222.

Beadell JS, Hyseni C, Abila PP, Azabo R, et al. (2010). Phylogeography and population structure of *Glossina fuscipes fuscipes* in Uganda: implications for control of tsetse. *PLoS Negl. Trop. Dis.* 4: e636.

Behura SK (2006). Molecular marker systems in insects: current trends and future avenues. *Mol. Ecol.* 15: 3087-3113.

Billotte N, Lagoda PJR, Risterucci AM and Baurens FC (1999). Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits* 54: 277-287.

Botelho PSM (1992). Quinze anos de controle biológico da *Diatraea saccharalis* utilizando parasitóides. *Pesq. Agropec. Bras.* 27: 254-262.

Carletto J, Lombaert E, Chavigny P, Brevault T, et al. (2009). Ecological specialization of the aphid *Aphis gossypii* Glover on cultivated host plants. *Mol. Ecol.* 18: 2198-2212.

Castro BA, Leonard BR and Riley TJ (2004). Management of feeding damage and survival of southwestern corn borer and sugarcane borer (Lepidoptera: Crambidae) with *Bacillus thuringiensis* transgenic field corn. *J. Econ. Entomol.* 97: 2106-2116.

Chapuis MP and Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* 24: 621-631.

Cortés AM, Zarbin PH, Takiya DM, Bento JM, et al. (2010). Geographic variation of sex pheromone and mitochondrial DNA in *Diatraea saccharalis* (Fab., 1794) (Lepidoptera: Crambidae). *J. Insect Physiol.* 56: 1624-1630.

Long WH and Hensley SD (1972). Insect pests of sugarcane. *Annu. Rev. Entomol.* 17: 149-176.

Luikart G and England PR (1999). Statistical analysis of microsatellite DNA data. *Trends Ecol. Evol.* 14: 253-256.

Moré M, Trumper EV and Prola MJ (2003). Influence of corn, *Zea mays*, phenological stages in *Diatraea saccharalis* F. (Lep. Crambidae) oviposition. *J. Appl. Entomol.* 127: 512-515.

Paradis E (2009). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419-420.

Pashley DP, Hardy TN, Hammond AM and Mihm JA (1990). Genetic evidence for sibling species within the sugarcane borer (Lepidoptera, Pyralidae). *Ann. Entomol. Soc. Am.* 83: 1048-1053.

Rozen S and Skaletsky H (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132: 365-386.

Santos H, Burban C, Rousselet J, Rossi JP, et al. (2011). Incipient allochronic speciation in the pine processionary moth (*Thaumetopoea pityocampa*, Lepidoptera, Notodontidae). *J. Evol. Biol.* 24: 146-158.

Torriani MV, Mazzi D, Hein S and Dorn S (2010). Structured populations of the oriental fruit moth in an agricultural ecosystem. *Mol. Ecol.* 19: 2651-2660.

Weir BS (1996). Genetics Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer, Suderland.