Molecular Ecology Notes (2003) doi: 10.1046/j.1471-8286 .2003.00366.x

PRIMER NOTE

Polymorphic microsatellite loci for the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) and some remarks on their isolation

Y.-J. JI,* D.-X. ZHANG,* G. M. HEWITT,† L. KANG* and D.-M. LI*
*State Key Laboratory of Integrated Management of Insect Pests & Rodents, Institute of Zoology, Chinese Academy of Sciences, 19 Zhongguancun Road, Beijing 100080, China; †School of Biological Sciences, University of East Anglia, Norwich, UK

Abstract

Five polymorphic tri- and tetranucleotide microsatellite loci suitable for population genetic analysis were identified in the cotton bollworm *Helicoverpa armigera* from two partial phagemid genomic libraries enriched for microsatellite inserts. The overall microsatellite-cloning efficiency in *H. armigera* is 2.5%, which is approximately eightfold lower than that for the gadoid fishes (20%) employing the same enrichment protocol, supporting the notion of a relative low frequency of microsatellite sequences in lepidopteran genomes. In addition, a large proportion of cloned microsatellite sequences turned out to be repetitive DNA, thus further increasing the difficulty of developing such markers in butterflies and moths.

Keywords: *Helicoverpa armigera*, insecticide resistance, microsatellite, migration, moth PCR primer, population dynamics

Received 15 August 2002; revision received 1 October 2002; accepted 7 November 2002

Microsatellite sequences remain the most powerful nuclear DNA markers available to date for population genetic analysis. However, their isolation in lepidopteran species has proven rather difficult (see Meglécz & Solignac 1998; Nève & Meglécz 2000). Limited success has been achieved in a dozen or so taxa investigated thus far (Harper *et al*. 2000; Nève & Meglécz 2000; Anthony *et al*. 2001; Williams *et al*. 2002 and references 3–7 listed therein), with only five or fewer loci being developed in each case, except for the silkworm (Reddy *et al*. 1999). Here we report the isolation and characterization of five di-allelic microsatellite loci from the cotton bollworm *Helicoverpa armigera* (Hübner) and discuss some observations which should have general relevance.

H. armigera is a highly polyphagous agricultural pest of worldwide significance. This pest occurs throughout Asia, Europe, Africa and Australia, and attacks a broad range of crops including cotton, maize, sorghum, tomatoes, tobacco, peppers, soybeans, chickpea and sunflower (Fitt 1989). For an integrated pest management strategy, two important aspects of the biology of this species require the investigation of their population genetic structure and dynamics. First, *H. armigera* has been thought to be a facultative migratory insect, with the potential to carry out long-distance migration (Fitt 1989). However, supporting genetic evidence is lacking. Second, *H. armigera* is able to rapidly develop resistance to many insecticides (Tang *et al*. 1988; Ahmad *et al*. 2001). Therefore, knowledge regarding population interconnections and gene flow should help in the management of both cotton planting and pest control.

As mitochondrial DNA (mtDNA) markers do not suffice for such studies (D-X Zhang, unpublished data), we decided to develop microsatellite markers for this organism. Two partial genomic libraries with short inserts (300–800 bp, on average) enriched for microsatellite DNA were constructed using a method modified from Karagyozov *et al*. (1993). Previously, the same procedure has been successfully employed for the isolation of microsatellite DNA markers from two gadoid fishes (Zhang *et al*. 2001). Briefly, two dinucleotide repeat, four trinucleotide repeat and two tetranucleotide repeat oligonucleotides were used as probes to select microsatellite-bearing genomic DNA fragments from *H. armigera* prior to cloning into pBluescript phagemid vector, viz: $(GT)_n$, $(GA)_n$, $(GTA)_n$, $(CCA)_n$, $(TTC)_n$, $(CAG)_n$, $(GTGA)_n$ and $(AGTT)_n$. For the sake of

Correspondence: De-Xing Zhang. Tel. +86 10 626 12962; Fax: +86 10 626 612962; E-mail: dxzhang@panda.ioz.ac.cn

Table 1 Characteristics of five polymorphic microsatellite loci in *Helicoverpa armigera*. Primer sequences, repeat unit structure, average expected (H_E) and observed (H_O) heterozygosities, number of alleles, PCR annealing temperatures (T_a) for Perkin-Elmer GeneAmp 9700 and the optimal $MgCl_2$ concentration are indicated for each locus. GenBank Accession nos: AJ504785–AJ504789

Locus	Primer name	Sequence (5′–3′)	Repeat type	T_a (°C)	MgCl2 (mM)	No. of alleles	Size range (bp)	H_E	H_O
HarSSR1	L3cF74	TAGGTGATTGTGGCTCAGTTTT	$(TGC)_2GAT(TGY)_4GAT$	58	1.25	20	228–288	0.8810	0.7372
	L3cB308	CAAACCCATCAGCAAATGCAAC	$(TGY)_{35}(TGA)_2AGC(TGY)_8$						
HarSSR2	B6hF72	AACACCCATTGAAGTCCCATGAA	$(ATG)_7$	52	1.5	6	156–171	0.2397	0.0877
	B6hB220	TTCCTATGTTCACTGCTAGTT							
HarSSR3	M9hF206	ATCTTTATGCTTTTAGCCGTTTA	$(TCA)_6$	59	1.75	6	129–147	0.3657	0.3322
	M9hB375	CAGTGGACTGCTATAGGCTGA							
HarSSR4	C53F542	TGTTACTTGGGTTTCCTGAATA	$(GYT)_{25}$	60	2.7	11	166–196	0.6779	0.4918
	C53B701	ACCACCGACACGTGCCGACTTC							
HarSSR5	E3fF106	GATAAGTTATTTCGGTTTAGTATT	$[T(T)AA]_6$	54	1.5	9	165–192	0.8836	0.1429
	E3fB272	AAGTACCTAATCCGTTTTTATTC							

fidelity, only one cycle of enrichment was performed, as with the gadoid fishes (Zhang *et al*. 2001). In total, 2000 recombinant clones were screened with P^{32}-labelled oligonucleotide probes, and ≈ 180 were positives producing strong or quite strong signals. Inserts of positive clones were isolated using polymerase chain reaction (PCR) amplification directly from bacterial colonies using M13 universal and reverse sequencing primers (−47 and −48, respectively, New England Biolabs), then sequenced with ABI BigDye™ Terminators Cycle Sequencing Kit (v2.0) in either an ABI PRISM3700 or ABI PRISM3100 automated sequencer. Sequencing analysis conclusively revealed that only 51 positive clones contain real microsatellite sequence. This gives an overall microsatellite-cloning efficiency of 2.5% in *H. armigera*, a value which is approximately eightfold lower than that for the gadoid fishes (20%, Zhang *et al*. 2001) using the same enrichment cloning strategy.

Among the 51 microsatellite clones, only 20 can be used for primer design, with the others either being repetitive sequences (i.e. the whole sequences, microsatellites *per se* and their flanking regions, are repeated in the genome) or having one or both flanking region too short. Oligonucleotide primers were designed using OLIGO®4 primer analysis software (National Biosciences Inc.) and synthesized by MWG Biotech-UK or Dingguo Biotech (Beijing).

Extensive optimization of the 20 pairs of primers was carried out on Perkin-Elmer GeneAmp 9700 systems using *Taq* DNA polymerase from Promega. It turned out that 11 pairs amplify repetitive sequences, and thus are not suitable for use as molecular markers in population genetic studies. With the remaining nine pairs, one primer of each pair was end-labelled with a fluorescent dye, either 6-FAM, NED or HEX. They were further tested by genotyping ≈ 60 cotton bollworm individuals on an ABI PRISM3100 machine with GeneScan-400HD (ROX) as the internal size standard. Three more loci were dropped because they amplify more than two alleles from some individuals. Five loci show typical characteristics of single-copy di-allelic nuclear locus, and thus are being used as genetic markers for population analysis. Table 1 shows the various characteristics of the five microsatellite loci of *H. armigera*, based on genotyping data from 50 to 280 individuals. The expected and observed heterozygosities were calculated using the program GENEPOP (Raymond & Rousset 1995). Heterozygote deficiency is obvious for these *H. armigera* samples. This has been seen in studies of several other species. It is worth noting that null alleles with varying frequency have been detected for all five loci, with the locus *HarSSR5* (E3f) having the highest frequency (23%). Because of the nature of the repeat type (Table 1), E3f does not appear to follow stepwise mutation.

Concerning the isolation of microsatellite markers in this lepidopteran moth, a few remarks seem appropriate. First, in line with the observation of Meglécz & Solignac (1998) and Nève & Meglécz (2000), our experience in cloning microsatellite sequences from various organisms (cod, poor cod, locust, cotton bollworm, pine moth) suggested that the frequency of simple sequence repeats in lepidopteran genomes is considerably lower. Second, false positives produced during library screening are pronounced in lepidopteran species, a common phenomenon observed from several independent studies (e.g. Meglécz & Solignac 1998, Keyghobadi *et al*. 1999). Third, over half of the microsatellite loci finalized in cotton bollworm have a small repeat number (n = 6–7; Table 1), they are nevertheless polymorphic. Finally, a significant proportion (> 70%) of microsatellite sequences cloned in the cotton bollworm are 'repetitive' DNA, and thus not suitable for use as genetic markers. Note that the aforementioned value was obtained after preventive measures against repetitive DNA were taken throughout the cloning procedure, and also that this appears to be a general phenomenon in lepidopteran

species. This problem deserves close attention — it may introduce errors during microsatellite marker development, leading to, for example, unintentional inclusion of repetitive microsatellite loci in genetic analysis.

Acknowledgements

This work was supported by MOST and NSFC grants to DXZ (G2000016207, 30025008) and the CAS 'Bai Ren Ji Hua' professorship. Part of the manuscript was written during an academic visit of DXZ to the University of East Anglia in summer 2002, under the Royal Society UK–China Joint Project Grant Q776.

References

Ahmad M, Arif MI, Ahmad Z (2001) Resistance to carbamate insecticides in *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Pakistan. *Crop Protection*, **20**, 427–432.

Anthony N, Gelembiuk G, Raterman D, Nice C, Ffrench-constant R (2001) Isolation and characterization of microsatellite markers from the endangered Karner blue butterfly *Lycaeides melissa samuelis* (Lepidoptera). *Hereditas*, **134**, 271–273.

Fitt GP (1989) The ecology of *Heliothis* species in relation to agroecosystems. *Annual Review of Entomology*, **34**, 17–52.

Harper GL, Piyapattanakorn S, Goulson D, Maclean N (2000) Isolation of microsatellite markers from the Adonis blue butterfly (*Lysandra bellargus*). *Molecular Ecology*, **9**, 1919–1952.

Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acid Research*, **21**, 3911–3912.

Keyghobadi N, Roland J, Strobeck C (1999) Influence of landscape on the population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology*, **8**, 1481–1495.

Meglècz E, Solignac M (1998) Microsatellite loci for *Parnassius mnemosyne* (Lepidoptera). *Hereditas*, **128**, 179–180.

Nève G, Meglécz E (2000) Microsatellite frequencies in different taxa. *Trends in Ecology and Evolution*, **15**, 376–377.

Raymond M, Rousset F (1995) GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reddy KD, Abraham EG, Nagaraju J (1999) Microsatellite in the silkworm, *Bombyx mori*: abundance, polymorphism, and strain characterization. *Genome*, **42**, 1057–1065.

Tang ZH, Gong KY, You ZP (1988) Present status and counter measures of insecticide resistance in agricultural pests in China. *Pesticide Science*, **23**, 189–198.

Williams BL, Brown JD, Paige KN (2002) Highly polymorphic microsatellite loci for *Speyeria idalia* (Lepidoptera: Nymphalidae). *Molecular Ecology Notes*, **2**, 87–88.

Zhang DX, Smith JA, Rico C, Hewitt GM (2001) Isolation and characterization of 10 microsatellite loci in poor cod *Trisopterus minutus* (L). *Molecular Ecology Notes*, **1**, 50–52.