Cross-species investigation of Helicoverpa armigera microsatellites as potential markers for other related species in the Helicoverpa - Heliothis complex

Authors: Grasela, James J., and McIntosh, Arthur H.

Source: Journal of Insect Science, 5(47) : 1-13

Published By: Entomological Society of America

URL: https://doi.org/10.1673/1536-2442(2005)5[1:CIOHAM]2.0.CO;2

Cross-species investigation of *Helicoverpa armigera* microsatellites as potential markers for other related species in the *Helicoverpa - Heliothis* complex

James J. Grasela and Arthur H. McIntosh

USDA, ARS, Biological Control of Insects Research Laboratory, 1503 S. Providence Rd., Research Park Columbia, Missouri USA 65203-3535

Abstract

Primers previously designed to amplify microsatellite DNA markers in the Old World bollworm, *Helicoverpa armigera*, larvae were tested in three closely related species: the corn earworm, *Helicoverpa* zea, tobacco budworm, *Heliothis virescens*, and *Heliothis subflexa*. Of the fourteen loci surveyed, only four loci (HaB60, HaC14, HaC87, HarSSR1) consistently demonstrated scorable single-copy microsatellite bands. Of these four, length polymorphism was identified only in the HaB60 marker (160 bp, 140 bp) of the *H. virescens* and *H. subflexa* sampled laboratory populations. Partial DNA sequences of all the identified single-copy microsatellites are presented as well as alignments to their respective *H. armigera* microsatellite.

Cite this paper as: Grasela JJ and McIntosh AH. 2005. Cross-species investigation of *Helicoverpa armigera* microsatellites as potential markers for other related species in the *Helicoverpa - Heliothis* complex. 13pp. *Journal of Insect Science* 5:47, available online: insectscience.org/5.47

Keywords:Lepidoptera, *Helicoverpa armigera*, *Helicoverpa zea*, *Heliothis virescens*, *Heliothis subflexa*, microsatellites
Correspondence: graselajj@missouri.edu
Received 16 July 2005 | Accepted 20 September 2005 | Published 31 December 2005

Introduction

The lengthy process and expertise required to isolate and identify potential microsatellite markers often precludes the use of this valuable technique in studies to determine genetic variation in natural populations. If microsatellite markers identified and developed from one biological source could be applied to other similar species, the usefulness of these genetic markers could be broadened. Fortunately, a number of microsatellites markers have been developed to study the population genetic variation in the Old World bollworm, *Helicoverpa armigera*, a serious insect pest of several agriculturally important grain and fiber crops (Tan et al., 2001; Ji et al., 2003; Scott et al., 2004). We therefore undertook a survey of some of the available lepidopteran species that are used routinely in our laboratory, namely *Helicoverpa zea*, *Heliothis virescens*, and *Heliothis subflexa*, to determine if previously designed microsatellite markers for *H. armigera* from several published sources could be applied to these closely related lepidopteran species. Selection of these species for study was also contingent on their importance as field crop pests. The host range of *H. zea*, the corn earworm, includes over 100 plants with the most significant crops being corn, cotton and tomato. Occasional hosts include bean, broccoli, cabbage, chrysanthemum, eggplant, head cabbage, green bean, lettuce, okra, pea, pepper, soybean, strawberry and watermelon. The tobacco budworm, *H. virescens*, is also principally a field crop pest, attacking such crops as alfalfa, clover, cotton, flax, soybean, corn, and tobacco. However, it sometimes feeds on such vegetables as cabbage, cantaloupe, lettuce, pea, pepper, pigeon pea, squash, and tomato, especially when cotton or other favored crops are not abundant. *H. virescens* is a common pest of geranium and other flower crops such as ageratum, bird of paradise, chrysanthemum, and gardenia, to name a few. In contrast, *H. subflexa* is of minor agricultural importance feeding on a few plant species such as *Solanum nigrum* and *Physalis* spp, but serves as a unique laboratory subject in studies to determine and compare host range infectivity and genetic resistance to baculoviruses.

Materials and Methods

Based on a previously published protocol (McIntosh et al., 1996), genomic DNA was extracted from 2nd or early 3rd instar *H. zea* and *H. virescens* larvae obtained from the North Carolina State University- Entomology Insectary, Raleigh, North Carolina, whereas 2nd or 3rd instar *H. subflexa* larvae were obtained in-house at the USDA, ARS, Biological Control of Insects Research Laboratory, Columbia, Missouri. Sample sizes are shown in Tables 1 and 2. Fourteen published primer sets designed to amplify the following microsatellite loci of *H. armigera* were employed in this study: (1) (HaB60) -- $(CTG)_2$ $(TTG)_3$ $(CTG)_5$ $(TTG)_2$, (2) (HaC14) -- $(ATTT)_5$, (3) (HaD47) -- $(CA)_5$ $(TCA)_4$, (4) (HaC87) - $(TC)_5$ (Scott et al., 2004); (5) Ham2 -- $(TTTTGA)_9$, (6) Ham3 -- $(TAAA)_2$ $(TAAAT)_4$, (7) Ham4 -- $(TCTG)_6$ TCTT $(TCTG)_6$, (8) Ham5 -- $(T)_n$ $(G)_n$, (9) Ham6 -- $(GAT)_2$ TT $(GAT)_2$ TT……$(AATA)_5$ (Tan et al., 2001); (10) HarSSR1 - $(TGC)_2$GAT $(TGY)_4$GAT$(TGY)_{35}(TGA)_2$ AGC$(TGY)_8$ (11) HarSSR2 - $(ATG)_7$, (12) HarSSR3 - $(TCA)_6$, (13) HarSSR4 - $(GYT)_{25}$, and (14) HarSSR5 - $[T(T)AA]_6$ (Ji et al., 2003). DNA microsatellite amplification was conducted under the following two polymerase chain reaction conditions using a Hybaid OmniGene thermal cycler (Midwest Scientific, www.midsci.com) in 25 l of puReTaq Ready-To-GoTM PCR bead reaction mixture (Amersham Biosciences, www.apbiotech.com), including 100-200 ng of genomic DNA template. First, after initial denaturing at 94° C for 5 min, the reaction mixture underwent 35 cycles at 94° C for 1 min, variable annealing temperature (see Tables 1 and 2) for 30 sec, 72° C for 40 sec, and a final extension at 72° C for 5 min (Tan et al., 2001). Second, after initial denaturing at 940 C for 1 min, the reaction mixture underwent 35 cycles at 94° C for 1 min, 50° C for 1 min, 73° C for 1 min, and a final extension at 72° C for 5 min (Scott et al., 2004). These two previously published PCR conditions with their respective primers were used to establish a comparative baseline for the three lepidopteran species examined in this study. However, if the expected fragment size(s) was not detected under the original PCR conditions for a particular microsatellite, empirical studies with various annealing temperatures were conducted in an attempt to resolve these problematic microsatellite markers (Table 1). A 10 l aliquot of each amplified sample was run on a 2.5% Metaphor™ agarose gel (10 mM Sodium hydroxide-Boric acid buffer, pH 8.5) for ca. 1 h at a constant 120 v using a Bio-Rad Wide Mini-Sub Cell-gel system.

Results

Initially, annealing temperatures previously published for the various microsatellites detected in *H. armigera* were employed in this study with resulting mixed success. Therefore, as indicated in Table 1, several annealing temperatures were tested for each locus in all three species in an attempt to determine the optimal running conditions for successful microsatellite amplification. Table 2 shows the microsatellite loci that failed to show distinct single-copy bands under the various PCR amplification conditions tested. Of the fourteen loci surveyed, only four loci (HaC14, HaB60, HaC87, and HarSSR1) consistently demonstrated scorable single-copy microsatellite bands that might lead to the potential detection of population polymorphism in subsequent studies (Fig.1). The phrase "potential detection" must be emphasized since the samples tested were limited to only laboratory reared insects. Of the four loci that consistently demonstrated scorable single-copy microsatellite bands, length polymorphism was identified only in the HaB60 marker (160 bp and 140 bp). The remaining microsatellites investigated showed multiple banding patterns, which have typically been observed in a number of lepidopteran species during the process of microsatellite clone development, and further indicate the repetitive nature of the flanking regions of microsatellites throughout the genome of Lepidoptera (Zhang, 2004). The HaC14 270 bp band (Fig. 2) detected among all three species, and first thought to be a microsatellite repeat variation, appears to be actually caused by a duplication of the downstream primer sequence used to amplify the microsatellite (sequence data not shown).

To obtain a more accurate picture of the nucleotide base composition of some of the detected single-copy microsatellites that showed either the expected allele size or a variant, direct DNA sequencing of PCR products was performed at the University of Missouri DNA Core Facility, Columbia, Missouri using an Applied Biosystems (www.appliedbiosystems.com) 3730 DNA Analyzer. Because of the known potential for amplification errors during the PCR reaction due the inherent nature of the *Taq* polymerase, 2-3 replicate samples of each locus were sequenced from individual insects and a single consensus sequence was generated employing VisCoSe (Spitzer et al., 2004). Partial sequence alignments of five alleles from four microsatellites are indicated in Figure 3. In addition to the generated sequence alignments, the T-coffee program also provides an index of Consistency of the Overall Residue Evaluation (CORE), an objective measure that identifies which regions of the compared sequences are correctly aligned by averaging the scores of each of the aligned pairs involving a base within a column (Notredame et al., 2000). A CORE value > = 3 would indicate a properly aligned base position and is considered the best compromise between a level of sensitivity and specificity required for proper base alignment. All of the aligned portions of the expected microsatellites showed reasonably high CORE scores for their individual alleles (70% for HaC14; 70%, 51% for HaB60 (160 bp, 140 bp, respectively); 91% for HaC87 (118bp); and 71% for

Table 1. Microsatellite markers previously published for *Helicoverpa armigera* found to successfully amplify similar microsatellite loci in three other related lepidopteran species.

Species	Locus	Annealing temp (Co) - DNA band fragment profile	Number of replicates	Sample size (n)	Reference for PCR running conditions
H. zea	HaB60	50° - multiple bands > 200 bp marker	1	7	
		55° - single 160 bp band	3	18	Tan et al., 2001(modified)*
		60° - multiple bands > 200 bp marker and a single 160 bp band	2	12	
H. zea	HaC87	50° - single-copy band 118 bp band	2	14	Scott et al., 2004;
		55°- multiple bands	1	3	Tan et al., 2001(modified)
		60° - single-copy 118 bp band	2	12	
H. zea	HaC14	50° - single-copy 160 bp band	1	12	Scott et al., 2004
		55° - single-copy 160 bp band	1	3	Tan et al., 2001(modified)
H. zea	HarSSR1	50° - single-copy 240 bp band	1	2	Scott et al., 2004
		58° - single-copy 240 bp band	2	5	
		50° - multiple bands	2	12	
H. virescens	HaC14	55° - single-copy 160 bp	2	15	Scott et al., 2004
		50° - multiple bands	2	4	Scott et al., 2004
H. virescens	HaB60	55° - 160 bp and 140 bp single-copy bands	3	11	Tan et al., 2001(modified)
		50° - 240 bp single-copy band	1	5	
H. virescens	HarSSR1	58° - multiple bands	2	4	Scott et al., 2004
		60° - multiple bands	1	2	
H. subflexa	HaB60	50° - multiple bands	2	4	Scott et al., 2004
		55° - 160 bp and 140 bpsingle-copy bands	3	14	
H. subflexa	HaC14	50° - 160 bp single-copy band	3	14	Scott et al., 2004
		55° - multiple bands	1	2	
H. subflexa	HaSSR1	58° - 240 bp single-copy band	2	4	Scott et al., 2004

Table 2. Microsatellite markers previously published for Helicoverpa armigera found to unsuccessfully amplify similar microsatellite loci in three other related lepidopteran species.

Species	Locus	Annealing temperature (Co) - DNA band fragment profile	Number of replicates	Sample size	PCR running conditions
H. zea	HaD47	500 - single-copy 240 bp band	1	2	Scott et al., 2004; Tan et al. 2001 (modified)
		550 - multiple bands	3	18	
H. zea	Ham3	530 - multiple bands	1	6	Tan et al., 2001; Tan et al. 2001 (modified)
		600 - multiple bands > 200 bp marker	1	6	
		660 - multiple bands > 200 bp marker	2	12	
H. zea	HarSSR2	520 - fb	1	2	Scott et al., 2004; Tan et al. 2001 (modified)
		600 - nb	1	4	
H. zea	HarSSR3	590 - nb	1	2	Scott et al., 2004; Tan et al. 2001 (modified)
		600 - multiple bands	1	4	
		550 - multiple bands > 200 bp marker	1	3	
H. zea	HarSSR4	60 - single-copy 240	1	2	Scott et al., 2004; Tan et al. 2001 (modified)
		660 - multiple bands > 240 bp band	1	3	
H. zea	HarSSR5	660 - multiple bands	1	4	Tan et al., 2001
H. zea	Ham5	520 -multiple bands	1	2	Tan et al., 2001
H. virescens	HaC87	500 - multiple bands	1	2	Scott et al., 2004; Tan et al. 2001 (modified)
		550 - multiple bands	1	4	
		600 - nb	1	4	
H. virescens	HarSSR2	520 - fb	1	2	Tan et al. 2001(modified); Scott et al., 2004
		660 - multiple bands	2	7	
H. virescens	HarSSR3	590 - multiple bands	1	2	Scott et al., 2004
		600 - multiple bands	1	3	
		500 - multiple bands	2	7	
H. virescens	HarSSR4	600 - multiple bands	1	2	Scott et al., 2004
		530 - multiple bands	1	2	
H. virescens	Ham3	550 - multiple bands	1	4	Tan et al., 2001
H. virescens	Ham5	520 - multiple bands	1	2	Tan et al., 2001
		550 - nb	1	4	
H. virescens	Ham6	660 - nb	1	3	Tan et al., 2001
H. subflexa	HaC87	500 - multiple bands	1	2	Scott et al., 2004
		550 - multiple bands > 200 bp marker and strongly stained bands at 55, 70, 75 bp markers	2	18	
H. subflexa	HaD47	500 - multiple band	1	2	Scott et al., 2004
		550 – suspected single-copy band at 140 bp	2	10	
H. subflexa	HaD47	500 - multiple bands	1	2	Scott et al., 2004
H. subflexa	Ham3	530 - multiple bands >180 bp marker band	2	16	Tan et al., 2001; Tan et al. 2001 (modified)
		590 - multiple bands	1	4	
H. subflexa	Ham6	530 - multiple bands	1	4	Tan et al., 2001
H. subflexa	HaSSR2	520 - nb	1	2	Tan et al., 2001
H. subflexa	HaSSr3	590 - fb	1	2	Tan et al., 2001
H. subflexa	HaSSR4	600 - fb	1	2	Tan et al., 2001
H. subflexa	HaSSR5	540 - fb	1	2	Tan et al., 2001

nb= no bands detected; fb = faint bands

Repetitive sequences are defined as repeated genomic regions containing microsatellite motifs and their flanking regions.

HarSSR1 (240bp), indicating at least for the most part a good portion of the base positions were properly aligned. Based on the aligned regions generated by the T-coffee program (www.ch.embnet.org/software/TCoffee.html), the identity of the nucleotide sites of the partially sequenced microsatellites relative to *H. armigera* was found to be 78% for all three species at the HaC14 160 allele, 83% for all three species at the HaB60 160 allele, 41% for *H. virescens* and *H. zea* at the HaB60 140 allele, and 84% for *H. zea* and *H. virescens* at the HaC87 118 allele, and 76% for *H. zea* and *H. virescens* at the HarSSR1 240 allele. As indicated in Fig. 3 (A-D) only the downstream primer used in PCR amplification for each locus appeared in the sequence along with the microsatellite marker. However, the upstream primer that would typically be included as part of the 5'-end of the microsatellite marker was not sequenced during the automatic analysis.

Several reports have shown that comparing allele sizes can result in inaccurate allele size differences for microsatellites (Estoup et al., 1995; Haberl and Tautz, 1999). One can approach this potential problem of size homoplasy by either employing single-strand conformation polymorphism analysis (SSCP) or sequence analysis of the DNA fragments. However, Liepelt et al. (2001) has shown that even sequenced, aligned microsatellites can show differences in repeat numbers occurring among clones and samples from the same individual. Their solution was to split the analyzed complex locus into two new loci. Nevertheless, we chose sequence analysis to determine if our unknown fragments contained not only the microsatellite but also to obtain an overall view of the alignment patterns of the fragments relative to the *H. armigera* markers.

Overall, the alignments of the four microsatellite loci detected in the three species, but with the

Figure 1. An assortment of PCR amplifications depicting several potential microsatellite primer pairs. (A) PCR amplification of three single-copy microsatellites from 10 individual *Heliothis subflexa* larvae; (B) PCR amplification of two single-copy microsatellites from eight *Heliothis virescens* larvae; (C) three single-copy microsatellites detected in *H. zea*, the more closely related of the three species to *Helicoverpa armiger* a. Base pair markers are indicated on the left of each gel. The size of specific bands that were sequenced is indicated for each of the microsatellite loci. nc = negative control.

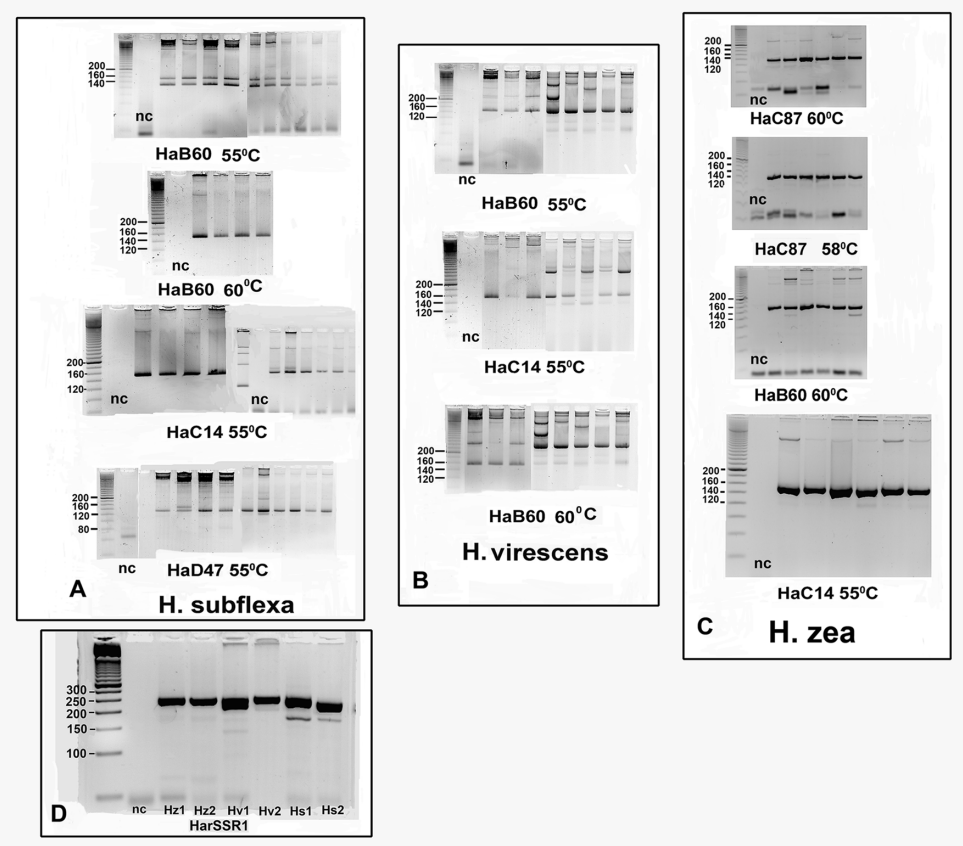


Figure 2. Successful identification of PCR amplified single-copy microsatellites from sampled individuals of the three species. (A) HaC14; (B) HaB60; (C) HaC87; and (D) HarSSR1. Hz = *Helicoverpa zea*; Hv = *Heliothis virescens*; Hs = *Heliothis subflexa*. Base pair markers are indicated on the left of each gel. nc = negative control. The size of specific bands that were sequenced is also indicated for each of the microsatellite loci.

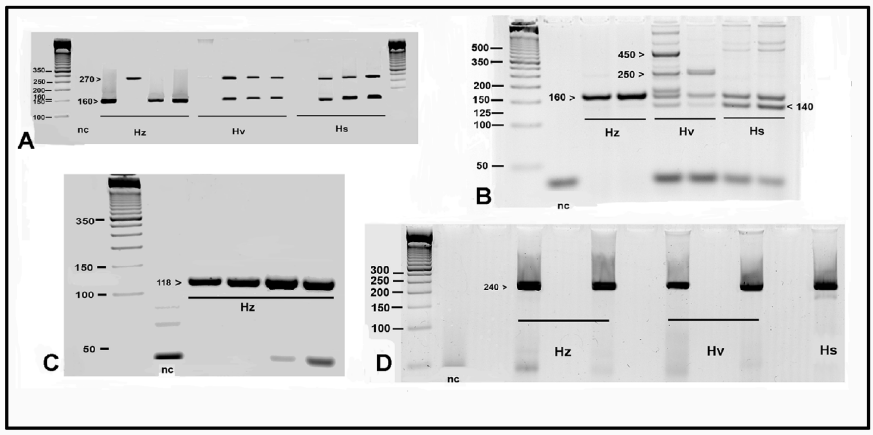


Figure 3. Partial sequences of the four simple sequence loci. All sequences were aligned employing the T-Coffee multiple sequence alignment package. Microsatellite alleles are shown for (A) HaC14, (B,C) HaB60, (D) HaC87 and (E) HarSSR1. Bold letters indicate the location of the simple sequence repeat and the box-shaded regions indicate identities. A CORE index for each base position is indicated in the outlined box below each alignment. The primer sequences flanking the loci are shown in lowercase letters.

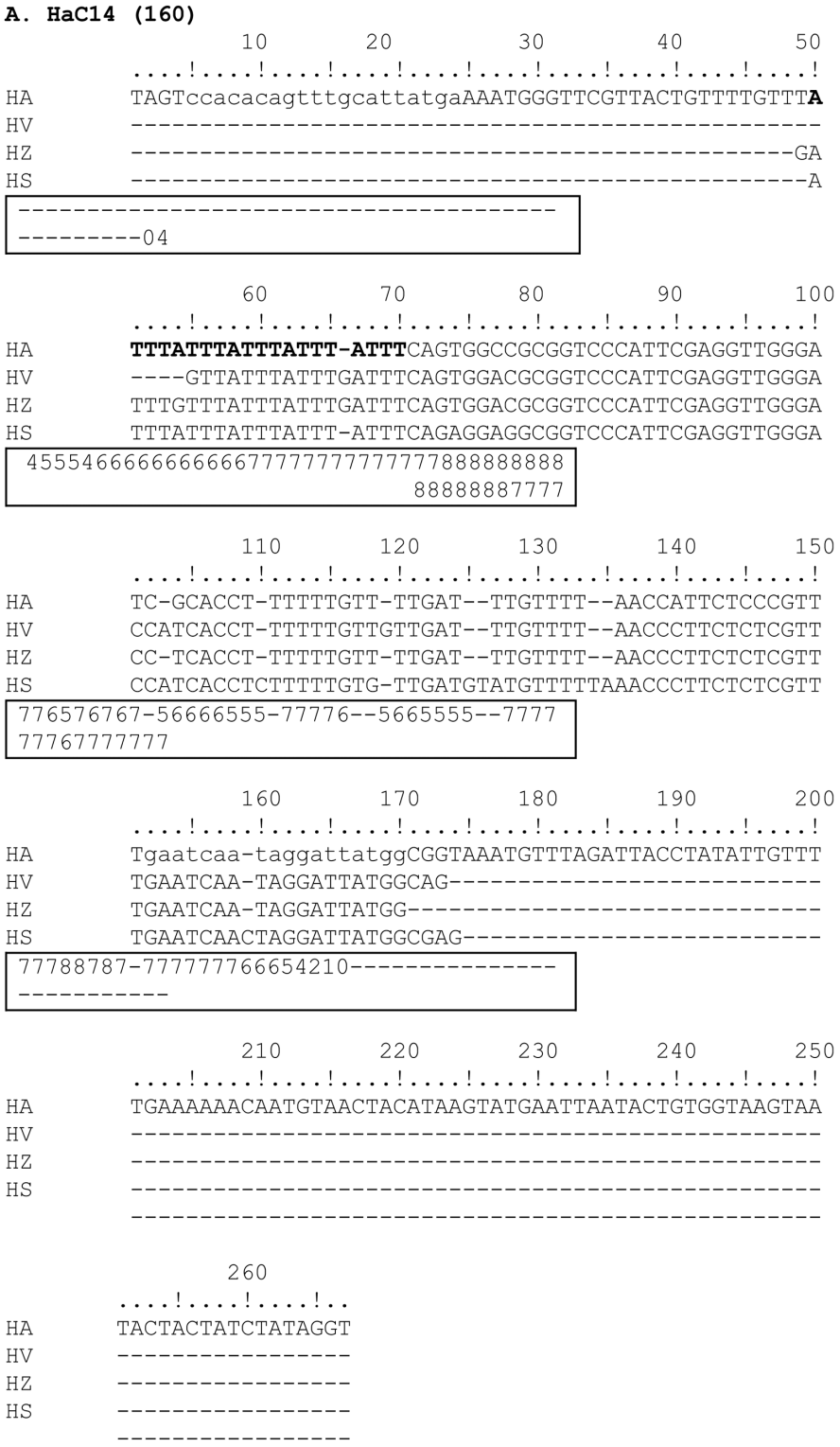


Figure 3 (B).

B. HaB60 (160)

```
                10        20        30        40        50
         ....!....!....!....!....!....!....!....!....!....!
HA       ACGCcaccacctgacataacgcTCACAGGTTGCTGCAACTGTTGTTGTTG
HV       -------------------------------------GCTG-TGTTG
HZ       ----------------------------------------TGTTGTTG
HS       -------------------------------------CTG-TGTTG
         ------------------------------------
         --233334666

                60       70        80        90        100
         ....!....!....!....!....!....!....!....!....!....!
HA       CTGCTGTTGCAATTGCGCAACTTGCTGCTGTTGTTGTTGCTGCTGCTGCT
HV       CTGCTGTTGCAGTTGCGCCACTTGCTGCTGTTGTTGTTGCTGCTGCTGCT
HZ       CTGCTGTTGCAATTGCGCAACTTGCTGCTGTTGTTGCTGCTGCTGCTGCT
HS       CTGCTGTTGCAGCTGCGCAACTTGCTGCTGCTGTTGTTGCTGCTGCTGCT
         7777777777777666666676777777777777777
         7777777

                110       120       130       140      150
         ....!....!....!....!....!....!....!....!....!....!
HA       GCTGTTGTTGCTGTTGCTGCTGCTGTT-GCGCCGCCTGTTGCTGTT-GAg
HV       GTTGCTGTTGCTGTTGCTGCTGCTGTT-GCGCCGCTTGTTGCTGTTAGAG
HZ       GTTGTTGTTGCTGTTGCTGCTGCTGTT-GCGCTGCCTGTTGCTGTTAGAG
HS       GTTGCTGTTG---TTGCTGCTGCTGTTAGAGCCGCTTGTTGCTGTTAGAG
         766676666666676677777777777-67777666677
         77777776777

                160       170       180       190       200
         ....!....!....!....!....!....!....!....!....!....!
HA       cttgcaattgctgctccttCCGTTTTTGTCTCTTTTCCTCGAGTTCTTTC
HV       CTTGCAATTGCTGCTCCTTAAA--------------------------
HZ       CTTGCAATTGCTGCTCCTTAAA--------------------------
HS       CTTGCAATTGCTGCTCCTTA----------------------------
         7878876787777775554222----------------
         -----------

                210       220      230
         ....!....!....!....!....!...
HA       TGTATTTTGTATATCTTTTCGGCAAGTAGATGA
HV       --------------------------------
HZ       --------------------------------
HS       --------------------------------
```

Figure 3 (C).

C. HaB60 (140)

```
                 10        20        30        40        50
          ....!....!....!....!....!....!....!....!....!....!
HA        ACGCcaccacctg-acataacgcTCACAGGTTGCTGCAA-CTGTTGTTGT
HV        A---CATC-TCGGTACTTGG----------GAGCATCTAGCCCACGGGGA
HS        -------A-TCGGTACTTGG-----------AGCATCTAGCCAACGGGGA
1---1112-23337444443----------144444444
73333333334

                60        70        80        90        100
         ....!....!....!....!....!....!....!....!....!....!
HA       TGCTGCTGTTGCAATTGCGCAACTTGCTGCTGTTGTTGTTGCTGCTGCTG
HV       GAC-GCTATTACCATGGC-CTTCAT-CT-CTGA-----------------
HS       GAC-GCTATTACCATGGC-CTTCAT-CT-CTGA-----------------
         444-45555544445444-344444-44-4443-----------------
444-45555544445444-344444-44-4443------
----------

                110      120      130      140      150
         ....!....!....!....!....!....!....!....!....!....!
HA       CTGCTGTTGTTGCTGTTGCTGCTGTTGCGCCGCCTGTTGCTGTTGAg
HV       ---CAACAGGAACTA-------------------GATT------GTTGC-
HS       ---CAA-AGGAACTA-------------------GATT------GTTGCG
         ---222234444443-------------------2333------545442
---222234444443-------------------2333-
-----545442

                160       170       180       190       200
         ....!....!....!....!....!....!....!....!....!....!
HA       cttgcaattgctgctCCTTCCGTTTTTGTCTCTTTTCCTCGAGTTCTTTC
HV       -TTGCAATTGCTGCTCCTTAA-----------------------------
HS       -TTGCAATTGCTGCTCCTTAA-----------------------------
         -66666667777777654322-----------------------------
-66666667777777654322------------------
----------

                210       220       230
         ....!....!....!....!....!...
HA       TGTATTTTGTATATCTTTTCGGCAAGTAGATGA
HV       --------------------------------
HS       --------------------------------
```

Figure 3 (D).

D. HaC87 (118)

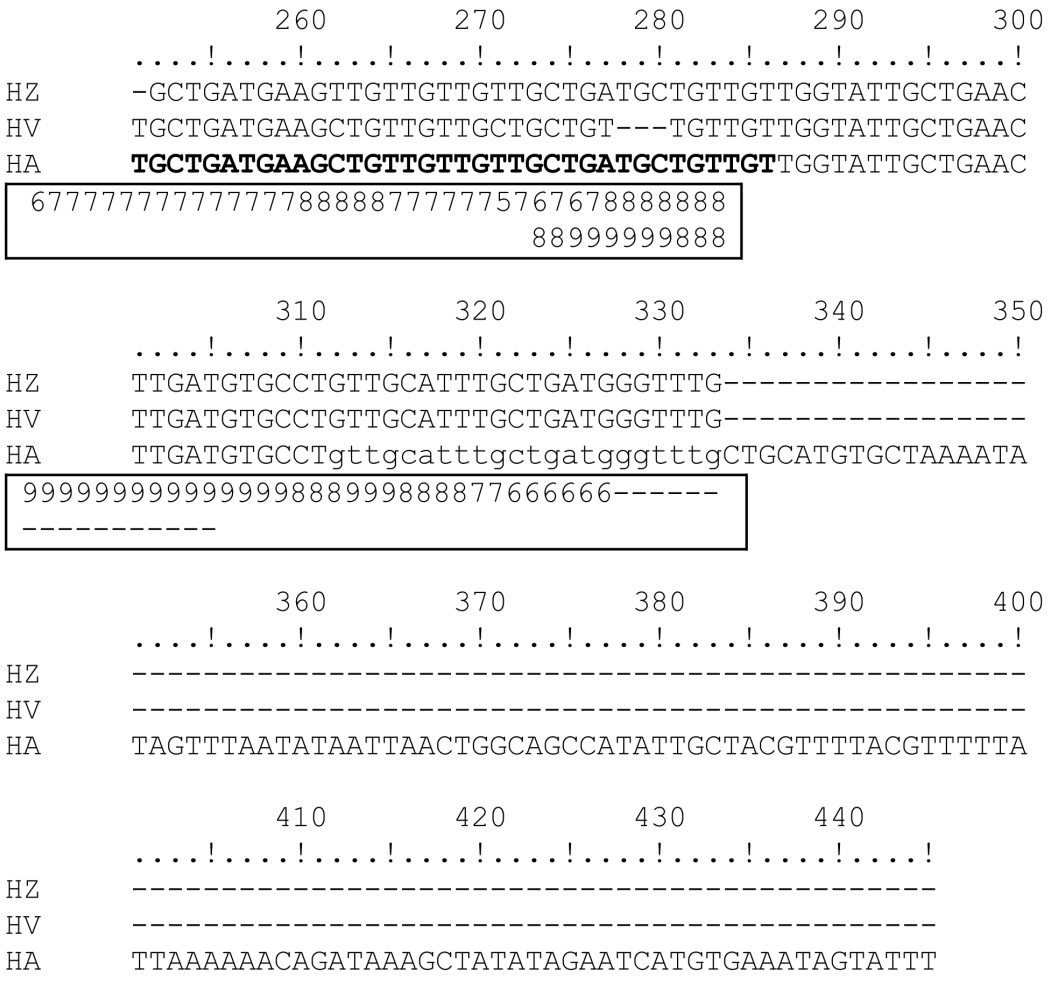
```
                10        20        30        40        50
           ....!....!....!....!....!....!....!....!....!....!
HA         ACCTTCCAGCTCTACGAGCACAGCACCAGGAACTCCAacgcgagcaccaa
HZ         --------------------------------------------------
           --------------------------------------------------

                60        70        80        90       100
           ....!....!....!....!....!....!....!....!....!....!
HA         ctgtaAATATTACTCTCATTTTATGCCGCTCTTCGAACTTTCTTCACTTT
HZ         ------------------------CAGCTCTTCGAACTTTCTTCACTTT
           ------------------------66999999999999
           99999999999

                110       120       130       140       150
           ....!....!....!....!....!....!....!....!....!....!
HA         --CTTTTTCTCTCTCTCAACTCCTTG-TTATTTTTgagaactactg-cga
HZ         TTCTTTTTTTCTCTCTCAACTCCTTGGTCATTTTTGAGAACTACTGGCGT
           --99999999999999999999999-996696666966
           6666666-666

                160       170       180       200       210
           ....!....!....!....!....!....!....!....!....!....!
HA         ttggtctcACCCACAAGGTCTTCAATCATTATTACGGCGTTCGATCTTTT
HZ         ATTGGGTTACTCACA---------------------------------
           66666666666666------------------------
           ----------

                220       230       240       250
           ....!....!....!....!....!....!....!....!...
HA         TAAATTCACTTCAAAGACTCTCATTACACACCAGTATTCTATT
HZ         -------------------------------------------
          -----------------------------------------
```

Figure 3 (E).

E. HarSSR1 (240)

```
               10        20        30        40        50
       ....!....!....!....!....!....!....!....!....!....!
HZ     --------------------------------------------------
HV     --------------------------------------------------
HA     AAACAAGGACATAGGTTAACAAAGTTATTTACATCAGTAGTTTGTTGTGG

               60        70        80        90        100
       ....!....!....!....!....!....!....!....!....!....!
HZ     --------------------------------------------------
HV     --------------------------------------------------
HA     GACTCCTGAGTTCCCATTACTGTTaggtgattgtggctcagTTTTTGGAA

               110       120       130       140       150
       ....!....!....!....!....!....!....!....!....!....!
HZ     --------------TGTGTGAGTTGCTGCGAATTGCTGTTGCTGTGATTG
HV     -------CTG----TGTGATGGTTGCTGTGAATTGCTGTTGTTGTGAA--
HA     TTTGATTCTGCTGTTGAGATGGTTGCTGCGA-TTGCTGTTGCTGTGATTG
       -------222----34444335666665555456666666777667664
       88

               160       170       180       190       200
       ....!....!....!....!....!....!....!....!....!....!
HZ     TTGTTGCTGCTGCTGCTGCTGTTGTTGTTGTTGCTGCTGCTGTTGTTGTT
HV     TTGCTGCTGCTGCTGTTGTTGTTGTTGCTGCTGCTGCTGCTGCTGTTGTT
HA     TTGTTGCTGCT---GCTGCTGTTGTTGTTGTTGCTGCTGCTGTTGTTGCT
       577777777776667666666666666666666677777
                                     76666666667

               210       220       230       240       250
       ....!....!....!....!....!....!....!....!....!....!
HZ     GCTGCTGCTGC---TGTTGTTGCTGCTGCTGTTGCT--------------
HV     GTTGTTGTTGC---TGTTGTTGCTGCTGCTGTTGCTGTTGTTGTTGCTGC
HA     GCTGCTGCTGCTGTTGTTGTTGCTGCTGCTGCTGCTGTTGCTGCTGCTGC
       77777777777---777777666666666666666555555555556666
```

Figure 3 (E, con't).



occurrence of some inversions at HaB60, a substitution in HaC87, and deletions in HaC14 and HarSSR1, showed a high number of identical nucleotide sites with the *H. armigera* repetitive motifs (Table 3). The length polymorphism detected in *H. virescens* and *H. subflexa* at the HaB60 locus revealed a large deletion of the repetitive array in the 140 bp allele of both species. However, with the complete sequence of one primer and a partial of the other 5'-end primer contained in the sequence read, it was still deemed to be a factual allele (Fig.3C).

The occurrence of null alleles in microsatellites is known to be an impediment to their successful application as markers in population genetic studies (Pemberton et al., 1995; Schlötterer and Pemberton, 1998; Liewlaksaneeyanawin et al., 2002), and have been implicated as a possible cause for the low levels of heterozygosity found in Lepidoptera (Meglecz et al., 2004). Since only samples collected from laboratory populations were employed in this study, we probably restricted ourselves from determining some level of polymorphism, if any, in the loci studied from the three species, though the number of polymorphic microsatellites to date has been found to be typically low in Lepidoptera (Ji and Zhang, 2004). Given the inherent variability of the microsatellite flanking regions in Lepidoptera, further work, in particular controlled mating studies, will be needed to elucidate the frequency of null alleles in these species.

The specific repetitive nature of the microsatellite flanking regions found in Lepidoptera demonstrates the difficulty of isolating similar microsatellites from closely related species.

Table 3. Comparison between four *H. armigera* microsatellites and the repetitive sequences identified in three related lepidopteran species.

Locus	Species	Microsatellite sequence
HaC14 (160 bp)	*H. armigera*	ATTT ATTT ATTT ATTT ATTT
	H. zea	ATTT GTTT TATT ATTT
	H. virescens	GTT ATTT ATTT
	H. subflexa	ATTT ATTT ATTT ATTT ATTT
HaB60 (160 bp)	*H. armigera*	CTG CTG TTG TTG TTG $(CTG)_5$ $(TTG)_2$
	H. zea	CTG CTG TTG TTG TTG $(CTG)_4$ TTG CTG TTG
	H. virescens	CTG CTG TTG TTG TTG $(CTG)_4$ TTG CTG TTG
	H. subflexa	CTG CTG CTG TTG TTG $(CTG)_4$ TTG CTG TTG
HaB60 (140 bp)	*H. armigera*	CTG CTG TTG TTG TTG $(CTG)_5$ $(TTG)_2$
	H. zea	-
	H. virescens	(CT-) (CTG) (A--)$(---)_6$(CAA)(CAG)(GAA)
	H. subflexa	(CT-) (CTG) (A--)$(---)_6$(CAA) (-AG)(GAA)
HAC87 (118 bp)	*H. armigera*	($TC)_5$
	H. zea	TT $(TC)_4$
	H. virescens	-
	H. subflexa	-
HarSSR1 (240 bp)	*H. armigera*	$(TGC)_2$ GAAT $\mathbf{(TGY)_4}$ GAT $\mathbf{(TGY)_{30}}$ AGT $\mathbf{(TGY)_8}$
	H. zea	$(TGC)_2$ GAAT $(TGY)_4$ GAAT $(TGY)_{31}$ AGC $(TGY)_7$
	H. virescens	TGC TGT GAAT $(TGY)_4$ GAAT $(TGY)_{31}$ AGC $(TGY)_7$
	H. subflexa	-

$\mathbf{(TGY)_4}$ =	TGC TGT TGC TGT
$\mathbf{(TGY)_{30}}$ =	TGT TGT TGC TGC TGC TGC TGC TGT TGT TGT TGT TGC TGC TGC TGT TGT TGC TGC TGC TGC TGT TGT TGC TGC TGC TGT TGC TGC TGA TGA
$\mathbf{(TGY)_8}$ =	TGT TGT TGT TGC TGA TGC TGT TGT
$(TGY)_4$ =	TGC TGT TGT TGT
$(TGY)_{31}$ =	TGT TGT TGT TGT TGC TGC TGC TGC TGC TGC TGT TGT TGT TGT TGT TGC TGT TGT TGC TGC TGC TGT TGC TGT TGT TGT TGC TGC TGC TGA TGA
$(TGY)_7$ =	TGT TGT TGC TGC TGT TGT TGT
(-) =	not detected

However, some of the data presented here extends the utility of previously developed microsatellites of one species to closely related members, and has the potential to be used as population genetic markers in other related lepidopteran species.

Disclaimer

Mention of trade names or commercial product in the publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

Estoup A, Tailliez C, Cornuet Jean-Marie, Solignac M. 1995. Size homolplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution* 12(6): 1074-1084.

Haberl M, Tautz D. 1999. Comparative allele sizing can produce inaccurate allele size differences for microsatellites. *Molecular Ecology* 8: 1347-1350.

Ji Y-J, Zhang D-X, Hewitt GM, Kang L, Li D-M. 2003. Polymorphic microsatellite loci for the cotton bollworm *Helicoverpa armigera* (Lepidoptera:Noctuidae) and some remarks on their isolation. *Molecular Ecology Notes* 3: 102-104.

Ji Y-J, Zhang D-X. 2004. Characteristics of microsatellite DNA in lepidopteran genome and implications for their isolation. *Acta Zoology Sinica*. 50: 608-614.

Liepelt S, Kuhlenkamp V, Anzidel M, Vendramin G, Ziegenhagen B. 2001. Pitfalls in determining size homoplasy of microsatellite loci. *Molecular Ecology Notes*. 1: 332-335.

Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA. 2002. Inheritance of null alleles for microsatellites in the white pine weevil (*Pissodes strobe* [Peck] [Coleopterqa:Curculionidae]). *Journal of Heredity* 93(1): 67-70.

McIntosh AH, Grasela JJ, Matteri RL. 1996. Identification of insect cell lines by DNA amplification fingerprinting (DAF). *Insect Molecular Biology* 5(3): 187-195.

Meglecz E, Petenian F, Danchin E, Coeur D'Acier A, Rasplus J-Y, Faure E. 2004. High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. *Molecular Ecology* 13:1693-1700.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302: 205-217.

Pemberton JM, Slate J, Bancroft DR, Barrett JA. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4: 249-252.

Schlötterer C, Pemberton J. 1998. The use of microsatellites for genetic analysis of natural populations – a critical review. In: *Molecular approaches to ecology and evolution*. (DeSalle, R. and Schierwater, B. eds.). Berlin: Birkhäuser; 71-86.

Scott KD, Lange L, Scott LJ, Graham GC. 2004. Isolation and characterization of microsatellite loci from *Helicoverpa armigera* Hubner (Lepidoptera: Noctuidae). *Molecular Ecology Notes* 4: 204-205.

Spitzer M, Fuellen G, Cullen P, Lorkowski S. 2004. VisCoSe: visualization and comparison of consensus sequences. *Bioinformatics* 20: 433-435.

Tan S, Chen X, Zhang A, Li D. 2001. Isolation and characterization of DNA microsatellites from cotton bollworm (*Helicoverpa armigera*, Hubner). *Molecular Ecology Notes* 1: 243-244.

Zhang De-Xing. 2004. Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology and Evolution*. 19: 507 -509.