Bulletin of Entomological Research (2010) **100**, 445–450
doi:10.1017/S0007485309990460

First published online 11 December 2009

Genetic variation among *Helicoverpa armigera* populations as assessed by microsatellites: a cautionary tale about accurate allele scoring

A.R. Weeks[1]*, N.M. Endersby[1], C.L. Lange[2], A. Lowe[3,4], M.P. Zalucki[2] and A.A. Hoffmann[1,5]

[1]Centre for Environmental Stress and Adaptation Research (CESAR), Department of Genetics, The University of Melbourne, Parkville, Victoria, Australia 3010: [2]The University of Queensland, School of Integrative Biology, Queensland 4072, Australia: [3]The School of Earth and Environmental Science, University of Adelaide, North Terrace, SA 5000, Australia: [4]State Herbarium, Hackney Road, Adelaide, SA 5005, Australia: [5]CESAR, Department of Zoology, The University of Melbourne, Parkville, Victoria, Australia 3010

Abstract

The existence of genetic differences among Australian populations of the pest moth *Helicoverpa armigera* based on microsatellite markers is contentious. To resolve this issue, we analyzed microsatellite variation in moth samples from multiple locations simultaneously in two laboratories that have previously reported contrasting patterns. Alleles and allele numbers detected in the laboratories differed, as did the genetic differences found between the samples. The automated scoring system used in one of the laboratories combined with non-denaturing polyacrylamide gels led to inaccurate identification of alleles and high F_{ST} values between the populations. However, *H. armigera* in Australia is probably not structured geographically, with high gene flow between populations. This influences management of *H. armigera* and the development of area-wide control options, as populations need to be considered as one panmictic unit. The results also highlight potential problems of automated scoring systems when these are not checked carefully.

Keywords: *Helicoverpa armigera*, microsatellite, genetic variation, gene flow

(Accepted 2 September 2009)

Introduction

Helicoverpa armigera Hübner (Lepidoptera: Noctuidae) is a major old world agricultural pest of cotton and vegetable crops (Fitt, 1989). In Australia, Scott *et al*. (2004) developed microsatellites for *H. armigera* and used these to test for genetic structure among *H. armigera* populations from multiple locations in several states. Strong structure was found to exist among locations (Scott *et al*., 2003, 2005a,b) and crops (Scott *et al*., 2006) although genetic differences were not always found. This was thought to reflect movement patterns of the moth, with localised structure in some years and extensive movement in others.

These findings have implications for pest control and the evolution of resistance to pesticides and toxins in genetically

*Author for correspondence
Fax: 613 83442279
E-mail: aweeks@unimelb.edu.au

Table 1. Origin of samples of *Helicoverpa armigera* genotyped at eight microsatellite markers by the CESAR and UQ laboratories.

Population	Location	*n*	Date Collected	Collected from	Reference
Pukekohe	New Zealand	11	01-04-2004	Tomato (larvae)	Endersby *et al.* (2007)
Somerville	Victoria	35	23-03-2001	Trap (adult)	Endersby *et al.* (2007)
Werribee	Victoria	22	23-03-2001	Trap (adult)	Endersby *et al.* (2007)
Dalmore	Victoria	31	30-11-1999	Trap (adult)	Endersby *et al.* (2007)
Bowen	Queensland	1	24-05-2004	Sweetcorn (larvae)	Endersby *et al.* (2007)
Biloela	Queensland	21	15-11-2004	Cotton (larvae)	Unpublished
Bundaberg	Queensland	3	25-10-2002	Cotton (larvae)	Unpublished
Bundaberg	Queensland	31	23-11-2004	Cotton (larvae)	Unpublished
Dalby	Queensland	2	29-11-2004	Cotton (larvae)	Unpublished
Dalby	Queensland	2	18-1-2005	Cotton (larvae)	Unpublished
Goondiwindi	Queensland	2	31-1-2005	Cotton (larvae)	Unpublished
Kingaroy	Queensland	1	18-1-2005	Cotton (larvae)	Unpublished
Moree	Queensland	5	18-1-2005	Cotton (larvae)	Unpublished
Moree	Queensland	6	20-2-2004	Cotton (larvae)	Unpublished
Narrabri	NSW	1	30-11-2004	Cotton (larvae)	Unpublished
Narrabri	NSW	1	1-2-2005	Cotton (larvae)	Unpublished
Warren	NSW	4	15-12-2004	Cotton (larvae)	Unpublished
Warren	NSW	20	16-12-2004	Cotton (larvae)	Unpublished
Walgett	NSW	1	31-1-2005	Cotton (larvae)	Unpublished

modified crops. Genetic structure implies that the suppression of moth populations in one region or crop might occur independently of suppression in other genetically distinct regions, and spread of resistance alleles should be less likely across populations that are distinct genetically. Scott *et al.* (2005a) proposed that genetic markers can even be used to define the geographic origin of moths and assist in targeted control programs at a local and regional level.

However, other researchers have failed to find genetic structure among Australian *H. armigera* populations. Both allozyme (Daly & Gregg, 1985) and mitochondrial (Spackman & McKechnie, 1995) data suggest high levels of gene flow among populations. Moreover, Endersby *et al.* (2007) found no genetic differences among *H. armigera* populations from Australian locations when characterized with microsatellites. Microsatellite markers can be particularly problematic in lepidopteran species because of a high level of redundancy and a very low frequency of single copy loci (Zhang, 2004).

Endersby *et al.* (2007) suggested several reasons for the lack of consistency among studies. These included problems in scoring alleles, the presence of null alleles and moth sampling issues. Here, we test problems with allele scoring by characterizing microsatellite variation in two sets of moths with both the manual scoring system of Endersby *et al.* (2007) and the automated scoring system of Scott *et al.* (2003). We show that allele scoring rather than null alleles or other factors are likely to account for differences between the studies and find no evidence for geographic differences between samples when markers are scored manually.

Methods

Four microsatellite loci developed by Ji *et al.* (2003) (HarSSR1, HarSSR2, HarSSR3 and HarSSR4) and four microsatellite loci developed by Scott *et al.* (2004) (HaB60, HaC14, HaC87 and HaD47) were used to screen 100 specimens of *H. armigera* from the study by Endersby *et al.* (2007) and 100 specimens derived from samples collected at the same time as those in Scott *et al.* (2004) (table 1). Both sets of samples were screened by the Centre for Environmental Stress and Adaptation Research (CESAR) laboratory at the University of Melbourne, Victoria, Australia and at the School of Integrative Biology, the University of Queensland (UQ), Queensland, Australia.

The UQ genomic DNA samples were extracted with a modified salting out protocol developed by Miller *et al.* (1988), and DNA aliquots were diluted 1 : 20 for Polymerase Chain Reactions (PCR). UQ and CESAR DNA samples were first confirmed as *H. armigera* DNA using the internal transcribed spacer (ITS) PCR confirmation assay (Amornsak *et al.*, 1998). All positive samples were then screened for the eight microsatellite loci in a final volume of 10 µl containing: 1 × PCR Buffer (Fisher Biotech, Wembley, WA, Australia), 2 mM $MgCl_2$, 0.5 mM dNTPs (Fisher Biotech, Wembley, WA, Australia), 0.5 µM forward (Hex labelled) and reverse primer, 0.5 units Taq F1 DNA Polymerase (Fisher Biotech, Wembley, WA, Australia), and the final volume was adjusted with sterile water to a volume of 10 µl. For loci HarSSR1, HarSSR2, HarSSR3, HarSSR4 and HaB60, 5 µg of Bovine Serum Albumin (BSA, New England Biolabs, Ipswich, MA, USA) was also added to the PCR mix. PCR cycling conditions were: denaturation (4 min, 94°C), 35 cycles of 94°C (30 s), annealing (30 s) (59°C: HarSSR1, HarSSR3 and HarSSR4; 53°C: HarSSR2; 50°C: HaB60, HaC87, HaC14 and HaD47), and 72°C (45 s), with final extension at 72°C (5 min). PCRs were performed using a Palm Cycler (Corbett Research, Mortlake, NSW, Australia). Microsatellite PCR fragments and Genescan-500 (Tamra) standards (Applied Biosystems, California, USA) were separated electrophoretically on 5% non-denaturing polyacrylamide 0.1 mm thick gels (native gels) and digitally captured using the GS2000 Genetic Analyser (Corbett Research, Mortlake, NSW, Australia) at 1200V and 40°C. All microsatellite alleles were each examined and sized with ONE-Dscan V2.05 (Scanalytics, Fairfax, VA, USA).

At CESAR, amplification of microsatellites by PCR took place in a volume of 10 µl with 2 µl of genomic DNA extracted using a Chelex® 100 Resin (BioRad, Hercules, CA) method (Endersby *et al.*, 2005). Primer concentrations of 0.03 µM (forward primer end-labelled with [γ^{33}P]-ATP), 0.1 µM (unlabelled forward primer) and 0.4 µM (reverse

https://doi.org/10.1017/S0007485309990460 Published online by Cambridge University Press

Table 2. Characteristics of *Helicoverpa armigera* microsatellites in samples from Victoria and Queensland scored using two different methods (manual and automated).

Samples	Marker	Number of alleles		Allele size range (bp)		Most common allele (bp)	
		Manual	Automated	Manual	Automated	Manual	Automated
Victoria	HarSSR1	15	5	233–278	162–174	257	168
	HarSSR2	3	6	159–172	132–174	169	138
	HarSSR3	8	12	126–141	117–147	135	138
	HarSSR4	13	11	166–191	133–196	185	145
	HaC14	20	10	125–175	113–161	154	133
	HaC87	4	9	116–120	106–157	118	108
	HaB60	4	4	152–171	126–138	165	135
	HaD47	20	70	117–160	50–239	138	142
Queensland	HarSSR1	15	9	233–281	162–270	257	168
	HarSSR2	4	7	160–172	129–177	169	171
	HarSSR3	6	8	129–141	120–147	135	138
	HarSSR4	11	10	173–191	136–205	182	145
	HaC14	22	17	125–174	109–185	154	133
	HaC87	5	7	111–118	104–122	118	108
	HaB60	3	3	159–168	132–138	165	135
	HaD47	17	42	125–150	64–168	138	125

primer) were used. The PCR reagent mix contained 2.0 mM $MgCl_2$, 0.1 mM dNTPs, 0.5 mg ml^{-1} purified bovine serum albumin (New England Biolabs, Ipswich, MA), 2.5 µl 10 × PCR amplification buffer and 0.4 units of *Taq* polymerase (New England Biolabs, Ipswich, MA). PCR cycling conditions were: denaturation (4 min, 94°C), 35 cycles of 94°C (30 s), annealing (30 s) (59°C: HarSSR1, HarSSR3 and HarSSR4; 55°C: HaB60; 53°C: HarSSR2 and HaC87; 51°C: HaC14 and HaD47) and 72°C (45 s), with final extension at 72°C (5 min). Fragments derived from PCR were separated through 5% denaturing polyacrylamide gels at 65 W for 2.5–3.5 h and exposed for 15 h to autoradiograph film. Allele sizes were derived from manual comparisons with λgt11 ladders (fmol® DNA Cycle Sequencing System, Promega, Madison, WI, USA).

Bands taken from four of the loci (HarSSR2, HarSSR3, HaB60 and HaC87) were excised from the polyacrylamide gel, rehydrated and subjected to a second PCR with the same microsatellite primers. The resulting products were sequenced in the forward and reverse directions by Macrogen Inc. (Seoul, Korea) to obtain accurate sizes (bp). Sequences were manipulated with Sequencher v4.8 (Gene Codes Corporation, Michigan, USA).

Data analysis

A comparison of the results obtained from two microsatellite allele scoring methods (manual at CESAR vs. automatic at UQ) was made with respect to number of alleles in the sample, allele size range and size of most common allele. The mean allele size of each individual scored manually was plotted against the mean allele size of each individual scored using an automated system. *F* statistics (F_{ST}) were calculated with the program FSTAT (Goudet, 2001).

Results and discussion

Contrasting patterns of allele numbers, size distributions and allele frequencies were found between the laboratories for the samples of *H. armigera* derived from the Endersby *et al*. (2007) and Scott *et al*. (unpublished) studies (table 2). To compare the results, we undertook two types of comparisons. The first involved plotting the mean allele size of the individual between the studies. These were expected to match across studies given that at least rank order should be preserved across laboratories. If the allele sizes are a little different in the labs, as can happen when different techniques and equipment are used, we would not expect these analyses to be affected. The graphs (fig. 1) indicate that the consistency in the results is poor. In no case was agreement observed. Rank correlation coefficients varied from 0 to 0.9 and numerous correlations were lower than 0.5.

The second approach involved looking at the number of heterozygotes in the studies. Again, one would expect that these would match, ideally without any errors. When patterns were examined, we found that mismatches constituted a significant proportion of cases. Averaged across the two collections, percentage mismatches were 43, 55, 22 and 22% for the HarSS1 to HarSS4 markers, respectively, and 55, 9, 32 and 27% for HaC14, HaC87, HaB60 and HaD47 markers, respectively.

Both these results indicate that the alleles do not line up, and scoring is therefore presumably most likely to reflect the different results in studies by the two groups (Scott *et al*., 2003, 2004, 2005a,b; Endersby *et al*., 2007). We strongly suspect that manual scoring is accurate, whereas the automated system used by Scott and colleagues is not, and this is likely due to the different gel chemistry systems used (denaturing vs. non-denaturing) combined with the scoring method. Allele sizes from manual scoring match up with those published in the independent and unrelated paper by Ji *et al*. (2003), who developed the HarSSR primers (on an automated system, but with denaturing), whereas the automated allele sizes do not match. The automated allele sizes do, however, appear similar to Scott *et al*. (2004) for HaC14, HaC87, HaB60 and HaD47.

Similarly, when we look at the distribution of alleles in the automated scoring system, these do not match expectations for microsatellites. The alleles have bimodal distributions in some cases (e.g. SSR3, HaC14), whereas unimodal distributions are expected. Moreover, there are occasional alleles that are outliers with respect to allele sizes (e.g. SSR4).

https://doi.org/10.1017/S0007485309990460 Published online by Cambridge University Press

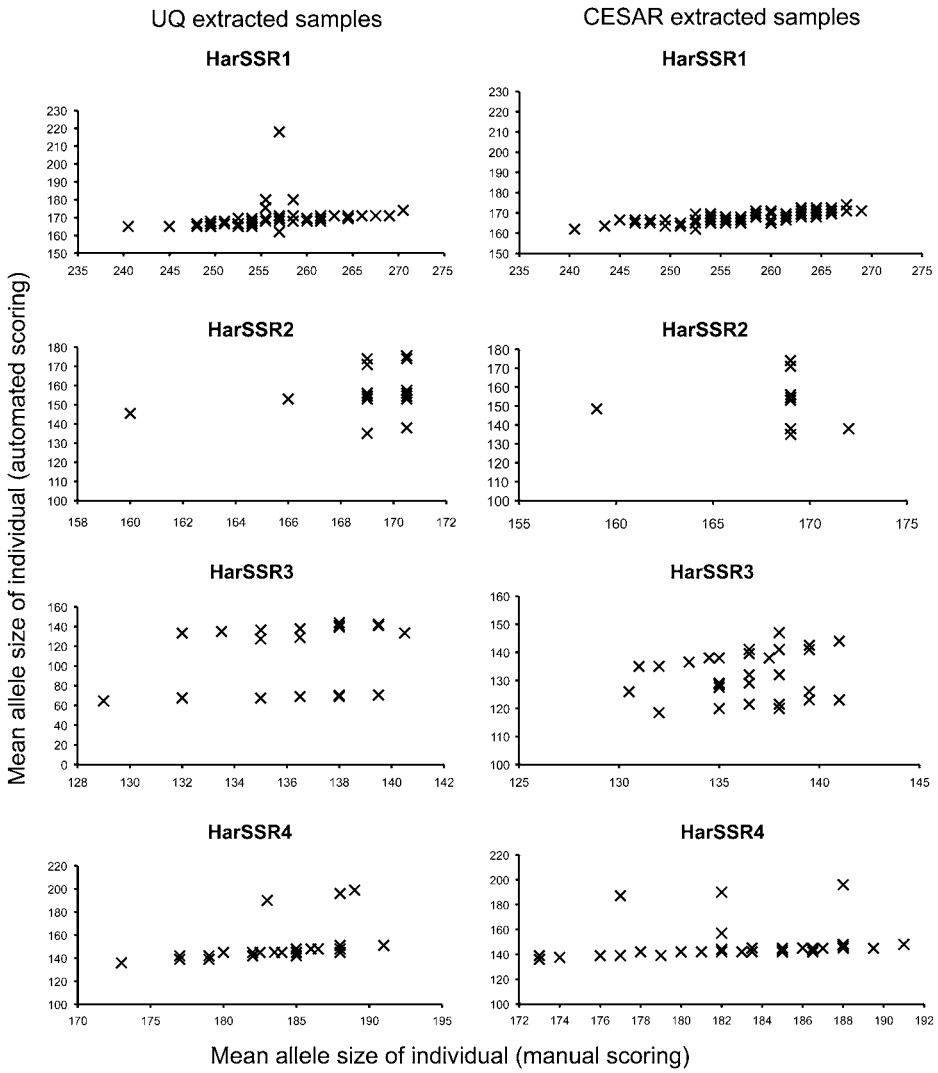

Fig. 1. (Continued)

Also, when we only compare the samples from Queensland and Victoria, the manual allele scoring does not produce population genetic differences, whereas the automated scoring system produces such differences between these populations (e.g. HaD47, SSR2). F_{ST} values between Queensland and Victoria were an order of magnitude higher when the automated scoring system was used (0.0716) compared with the manual method (0.0035). The values reflect the differences found between the Endersby *et al.* (2007) study and the Scott *et al.* (2003, 2005a,b) studies.

To directly determine the accuracy of the scoring methods, we sequenced alleles from four loci (HarSSR3, HaB60, HaC87 and HarSSR2) for several individuals that differed in allele sizes between the laboratories (table 3). Stark differences were apparent between the methods. For instance, for locus HarSSR3, the manual scoring method for one individual was a homozygote with an allele length of 142 bp, whereas the automated method scored that same individual as a homozygote with an allele length of 123 bp; sequencing revealed this allele to be 142 bp in length. Similarly, for another individual at the same locus, the manual method scored a homozygote with an allele length of 135 bp, whereas the automated method scored this as a heterozygote with alleles 120 bp and 138 bp in length; sequencing showed that the correct allele size was 136 bp (homozygote). This pattern was reflected across each of the four loci (table 3). Clearly, the manual scoring method consistently sizes alleles at the same or approximately the same size as direct sequencing, whereas the automated scoring method, combined with the use of non-denaturing polyacrylamide gels, is consistently inaccurate and can be out by over 30 bp. This is not surprising given that non-denaturing polyacrylamide gels can vary in fragment mobility by more than 10% compared with denaturing polyacrylamide gels (Sambrook & Russell, 2001).

There are undoubtedly problems in developing suitable microsatellite loci for evaluating population structure in Lepidoptera, including *H. armigera* (Zhang, 2004). In particular, primers may amplify families of loci rather than single loci. However, there are several reasons why we think that

https://doi.org/10.1017/S0007485309990460 Published online by Cambridge University Press

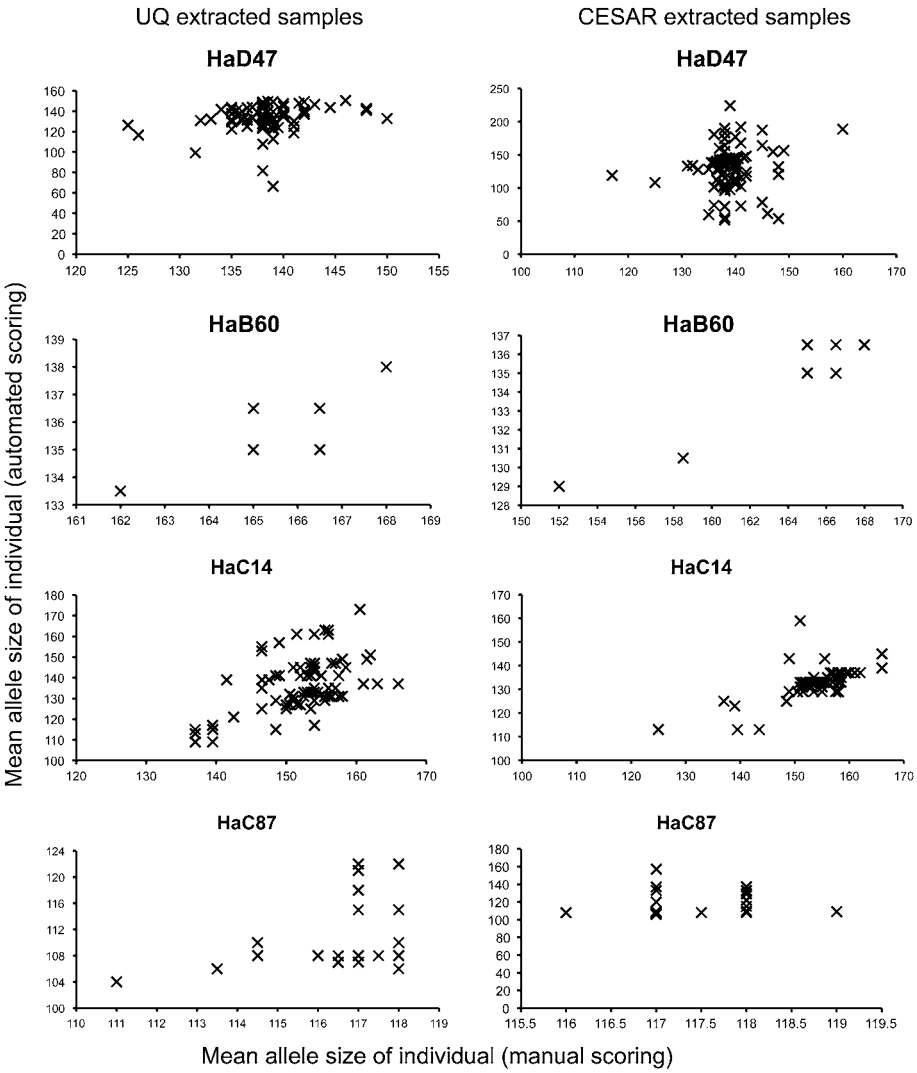

Fig. 1. Mean allele size of each *Helicoverpa armigera* individual scored manually plotted against the mean allele size of each individual scored using an automated system at eight microsatellite loci for samples extracted in the CESAR and UQ laboratories.

Table 3. Allele sizes at four loci determined by direct sequencing, and manual and automated genotyping methods.

Locus	Individual	Sequenced allele length (bp)	Genotype automated method (bp)	Genotype manual method (bp)
HarSSR3	1	142	123/123	142/142
	2	136	120/138	135/135
HaB60	1	165	135/135	165/165
	2	166	138/138	168/168
HaC87	1	117/115	137/108	118/116
HarSSR2	1	169	138/138	169/169
	2	169	171/171	169/169
	3	169	135/135	169/169

this is not a large issue with the loci used here. Firstly, if the presence of similar flanking regions had meant that multiple microsatellite loci were being amplified and scored, we would have expected 'alleles' to be present at a locus with very different sizes and odd (multimodal) distributions. While these patterns were occasionally seen with the automated system, the manual approach always provided a distribution of allele sizes that matched expectations typical for microsatellite loci. We also directly confirmed allele sizes by sequencing the actual alleles identified. Secondly, in all of the individuals genotyped in Endersby *et al*. (2007) and the present study (over 700 individuals), we have never encountered an individual that has more than two bands at any locus. If loci were being amplified from a family of loci, then we would expect that some individuals would have more than two bands amplified for a locus (reflecting alleles from multiple 'loci' being amplified). Finally, the patterns of population differentiation found using the manual method for scoring microsatellites matches patterns observed previously based on allozymes (Daly & Gregg, 1985).

These results indicate that genetic structure does not exist among Australian *H. armigera* populations. Instead

https://doi.org/10.1017/S0007485309990460 Published online by Cambridge University Press

H. armigera appears to comprise one panmictic population. This result matches what is found in *Plutella xylostella*, another lepidopteran pest species that inhabits a similar area in Australia (Endersby *et al.*, 2005). This means that population structure should not be considered when developing control strategies, as has previously been advocated (Scott *et al.*, 2003, 2005a,b).

In summary, investigators need to be very cautious when scoring genetic markers, particularly on automated systems or non-denaturing polyacrylamide systems. The results here are specific to the studies performed by Scott and colleagues; however, they also highlight potential problems that may be encountered in other studies. Ideally, genotypes need to be validated and often this is best performed on a manual system as described above. Automated systems have many benefits in terms of speed; however, this may come at a cost. Also, it is important for investigators to provide information on allele sizes when describing patterns of genetic differentiation. In an ideal world, inheritance patterns of markers should be tested through family studies (Endersby *et al.*, 2005).

Acknowledgements

We thank Gary Fitt for help initiating this study to help resolve inconsistencies in the literature and the Grains Research and Development Corporation for providing funding for this research through the National Invertebrate Pest Initiative. This research was supported by the Australian Research Council via their Special Research Centre Scheme.

References

Amornsak, W., Gordh, G. & Graham, G.C. (1998). Detecting parasitised eggs with polymerase chain reaction and DNA sequence of *Trichogramma australicum* Girault (Hymenoptera: Trichogrammatidae). *Australian Journal of Entomology* **37**, 174–179.

Daly, J.C. & Gregg, P. (1985) Genetic variation in *Heliothis* in Australia: species identification and gene flow in the two pest species *H. armigera* (Hubner) and *H. punctigera* Wallengren (Lepidoptera: Noctuidae). *Bulletin of Entomological Research* **75**, 169–184.

Endersby, N.M., McKechnie, S.W., Vogel, H., Gahan, L.J., Baxter, S.W., Ridland, P.M. & Weeks, A.R. (2005) Microsatellites isolated from diamondback moth, *Plutella xylostella* (L.), for studies of dispersal in Australian populations. *Molecular Ecology Notes* **5**, 51–53.

Endersby, N.M., Hoffmann, A.A., McKechnie, S.W. & Weeks, A.R. (2007) Is there genetic structure in populations of *Helicoverpa armigera* from Australia? *Entomologia Experimentalis et Applicata* **122**, 253–263.

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

Goudet, J. (2001) FSTAT: a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www2.unil.ch/popgen/softwares/fstat.htm (accessed 08 April 2009).

Ji, Y.-J., Zhang, D.-X., Hewitt, G.M., 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.

Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215.

Sambrook, J. & Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual, vol. 2, protocol 9, pp. 5.40–5.46. Cold Spring Harbor, New York, USA, Cold Spring Harbor Laboratory Press.

Scott, K.D., Wilkinson, K.S., Merritt, M.A., Scott, L.J., Lange, C.L., Schutze, M.K., Kent, J.K., Merritt, D.J., Grundy, P.R. & Graham, G.C. (2003) Genetic shifts in *Helicoverpa armigera* Hubner (Lepidoptera: Noctuidae) over a year in the Dawson/Callide Valleys. *Australian Journal of Agricultural Research* **54**, 739–744.

Scott, K.D., Lange, C.L., Scott, L.J. & Graham, G.C. (2004) Isolation and characterization of microsatellite loci from *Helicoverpa armigera* Hübner (Lepidoptera: Noctuidae). *Molecular Ecology Notes* **4**, 204–205.

Scott, K.D., Lawrence, N., Lange, C.L., Scott, L.J., Wilkinson, K.S., Merritt, M.A., Miles, M., Murray, D. & Graham, G.C. (2005a) Assessing moth migration and population structuring in *Helicoverpa armigera* (Lepidoptera: Noctuidae) at the regional scale: example from the Darling Downs, Australia. *Journal of Economic Entomology* **98**, 2210–2219.

Scott, K.D., Wilkinson, K.S., Lawrence, N., Lange, C.L., Scott, L.J., Merritt, M.A., Lowe, A.J. & Graham, G.C. (2005b) Gene-flow between populations of cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) is highly variable between years. *Bulletin of Entomological Research* **95**, 381–392.

Scott, L.J., Lawrence, N., Lange, C.L., Graham, G.C., Hardwick, S., Rossiter, L., Dillon, M.L. & Scott, K.D. (2006) Population dynamics and gene flow of *Helicoverpa armigera* (Lepidoptera: Noctuidae) on cotton and grain crops in the Murrumbidgee Valley, Australia. *Journal of Economic Entomology* **99**, 155–163.

Spackman, M.E. & McKechnie, S.W. (1995) Assessing the value of mitochondrial DNA variation for detecting population subdivision in the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae), in Australia. pp. 811–813 in *Proceedings Beltwide Cotton Conferences*. National Cotton Council of America, San Antonio, TX, USA.

Zhang, D.-X. (2004) Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology & Evolution* **19**, 507–509.