Bulletin of Entomological Research (2008) **98**, 509–518
doi:10.1017/S000748530800583X

Printed in the United Kingdom
First published online 16 June 2008

Exon-primed intron-crossing (EPIC) PCR markers of *Helicoverpa armigera* (Lepidoptera: Noctuidae)

W.T. Tay[1]*, G.T. Behere[1], D.G. Heckel[2], S.F. Lee[1] and P. Batterham[1]

[1]Centre for Environmental Stress and Adaptation Research, Department of Genetics, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria 3010, Australia: [2]Department of Entomology, Max-Planck Institute for Chemical Ecology, Beutenberg Campus, Hans-Knöll-Straße 8, Jena D-07745, Germany

Abstract

Applying microsatellite DNA markers in population genetic studies of the pest moth *Helicoverpa armigera* is subject to numerous technical problems, such as the high frequency of null alleles, occurrence of size homoplasy, presence of multiple copies of flanking sequence in the genome and the lack of PCR amplification robustness between populations. To overcome these difficulties, we developed exon-primed intron-crossing (EPIC) nuclear DNA markers for *H. armigera* based on ribosomal protein (Rp) and the Dopa Decarboxylase (DDC) genes and sequenced alleles showing length polymorphisms. Allele length polymorphisms were usually from random indels (insertions or deletions) within introns, although variation of short dinucleotide DNA repeat units was also detected. Mapping crosses demonstrated Mendelian inheritance patterns for these EPIC markers and the absence of both null alleles and allele 'dropouts'. Three examples of allele size homoplasies due to indels were detected in EPIC markers RpL3, RpS6 and DDC, while sequencing of multiple individuals across 11 randomly selected alleles did not detect indel size homoplasies. The robustness of the EPIC-PCR markers was demonstrated by PCR amplification in the related species, *H. zea*, *H. assulta* and *H. punctigera*.

Keywords: *Helicoverpa*, EPIC PCR markers, Lepidoptera, cross-species amplification, dopa decarboxylase, ribosomal protein genes

(Accepted 18 December 2007)

Introduction

The Lepidoptera are one of the world's most diverse insect groups with important roles within ecological, agricultural and horticultural settings. Developing effective and reliable DNA markers (e.g. microsatellites) is especially desirable in lepidopteran insect pests for population genetic structure studies, as well as for mapping economically and agriculturally important traits (e.g. insecticide and allelochemical resistance). Recent studies involving full genome sequence analysis of the silk moth *Bombyx mori* have identified numerous microsatellite DNA loci (Reddy *et al*., 1999; Prasad *et al*., 2005a,b), and about 550 of these were used for mapping purposes and establishing a genetic linkage map (Miao *et al*., 2005). The use of microsatellite DNA enrichment protocols in various lepidopteran species has also led to substantial microsatellite DNA markers being

*Author for correspondence
Fax: +61 2 62464000
E-mail: weetek.tay@csiro.au
Current address: CSIRO Entomology, Black Mountain Laboratories, Clunies Ross Street, Black Mountain, ACT 2601, Australia.

developed (e.g. Daly *et al*., 2004; Wardill *et al*., 2004; Zhou *et al*., 2005). The general perception that developing and designing microsatellite DNA markers is difficult for Lepidoptera nevertheless remains (Keyghobadi *et al*., 1999; Nève & Meglécz, 2000; Zhang, 2004) and appears, in part, to be due to redundancy in the microsatellite DNA flanking regions across different loci (Meglécz *et al*., 2004, 2007; Van't Hof *et al*., 2007), the lack of polymorphisms (e.g. Prasad *et al*., 2005b; Van't Hof *et al*., 2005) and possible evolutionary associations with mobile elements (Ji & Zhang, 2004; Meglécz *et al*., 2004, 2007; Zhang, 2004; Van't Hof *et al*. 2007).

The noctuid moth *Helicoverpa armigera* has a wide geographic distribution all over the Old World (Hardwick, 1965), and together with the closely related *H. assulta*, *H. punctigera*, and the New World *H. zea*, represents one of the world's most devastating agricultural lepidopteran pests (Hardwick, 1965; Mitter *et al*., 1993). Microsatellite markers for *H. armigera* have been developed by Tan *et al*. (2001), Ji *et al*. (2003, 2005) and Scott *et al*. (2004). A sub-set of markers developed by Scott *et al*. (2004) was applied to study Australian populations (Scott *et al*., 2003, 2005a,b, 2006). Endersby *et al*. (2007) utilised microsatellite markers reported by both Ji *et al*. (2003) and Scott *et al*. (2004) in further studies of Australian *H. armigera*. The authors reported amplification difficulties in samples collected both from Victoria and Queensland and observed varying degrees of allele dropouts (i.e. individuals scored as homozygotes once but subsequently as heterozygotes in replicate genotyping) in the markers tested, with null alleles (i.e. allele(s) that failed to amplify within an individual because of mutations at primer annealing sites) being the most likely underlying factor for these allele dropouts. Ji *et al*. (2005) also reported the likely presence of null alleles in microsatellite markers isolated (e.g. HarSSR5 and HarSSR7), as well as the presence of multiple-copy microsatellite DNA families (fig. 1 in Ji & Zhang, 2004). Technical problems in various *H. armigera* microsatellite DNA loci developed to-date are, therefore, evident although the underlying factors have not been examined in detail.

Due to the wide distribution and migration ability of *H. armigera*, reliable molecular genetic markers that enable population comparisons within and between countries are required. Furthermore, such markers would be even more useful if they could be applied to related pest species within the genus (i.e. *H. zea, H. assulta* and *H. punctigera*). Exon-Primed Intron-Crossing (EPIC) PCR markers (Lessa, 1992; Palumbi & Baker, 1994; Palumbi, 1996) have the potential to meet these requirements. Developing EPIC markers can be challenging, especially in organisms that lack sufficient coding DNA sequence data, although EPIC markers have been described for conserved nuclear genes in a limited taxonomic range of organisms (reviewed in Palumbi, 1996).

In additional to previously described 'universal' EPIC markers (Palumbi, 1996), ribosomal protein (Rp) genes are also suitable candidates. Rp genes function as house keeping genes and are highly conserved in genomes across a wide range of organisms. In Lepidoptera, the majority of Rp genes exist as single copy genes (Lee, 2006). Lee (2006) further showed that in *B. mori*, which represents a model organism in Lepidoptera genetics, at least 69 of the 80 Rp genes contain intron(s). The use of ribosomal protein genes as EPIC markers in population genetic studies has been reported (e.g. Gaffney, 2000). Polymorphic EPIC markers based on single copy nuclear genes are robust both for use as markers in evolutionary and population genetic studies (He & Haymer, 1997; Garrick & Sunnucks, 2006; Hubert *et al*., 2006) and for gene mapping purposes (Lee, 2006; Yasukochi *et al*., 2006; Pringle *et al*., 2007). The occurrence of allele dropout and null alleles in EPIC markers is expected to be minimal, since primer annealing sites are specifically designed from evolutionary conserved coding regions (Palumbi, 1996). We report here the development of *H. armigera* EPIC-PCR markers through comparative genomic analysis. We further provide molecular characterisation of EPIC length polymorphisms, establish Mendelian inheritance patterns and demonstrate robust PCR amplification in related pest *Helicoverpa* species.

Materials and methods

EPIC marker design and optimisation

We used comparative genomic analysis to develop EPIC-PCR markers for *H. armigera*. Ribosomal protein gDNA sequences from *B. mori* were aligned with ESTs, cDNA and mRNA of Rp gene sequences from publicly available sequences from *B. mori*, *Mamestra brassicae*, *Lonomia oblique*, *Papilio dardanus*, *Plutella xylostella*, *Spodoptera frugiperda*, *H. zea* and *H. armigera*. *Helicoverpa* species dopa decarboxylase (DDC) cDNA sequences were as reported by Fang *et al*. (1997) and aligned with *B. mori* DDC gDNA. Intron/exon boundaries of gDNA versus cDNA were determined using the SPIDEY tool from NCBI <http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/>, selecting *Drosophila* option as the model organism for all input sequences. Introns identified in *B. mori* as having approximate lengths of between 80–350 bp were selected for EPIC primer design.

Primers were designed to adjacent exons of single copy nuclear genes. Primers were designed for minimal primer-dimer and heteroduplex formation, and for minimal false priming sites between primers and template DNA using the Oligo Primer Analysis Software v6.40 (Molecular Biology Insights, Inc., Cascade, Colorado, USA). Primers were tested by gradient PCR (at 45°C to 55°C; 12 increments, with 6th and 7th increments set at 50°C) in an EPPENDORF Master Cycler gradient PCR machine (5331). EPIC primers that gave distinct amplicons of expected sizes were subsequently tested for intron length polymorphisms in *H. armigera* samples from Australia (AD1–AD24) and China (CH1–CH24) on 6% polyacrylamide gels.

EPIC DNA marker allele characterisation

Randomly selected alleles from six EPIC markers exhibiting higher (DDC, RpL29 and RpS6) and lower (RpL11, RpL12 and RpL3) numbers of alleles were characterised by DNA sequencing from Dalmore (Australia) and China *H. armigera* samples using the appropriate PCR primer annealing temperature profiles (table 1). Amplicons were gel-purified, cloned and sequenced. Only indels (insertions/deletions) were scored, while single nucleotide polymorphisms (SNPs) identified were not used for allele characterisation.

Gene copy number, allele inheritance pattern and cross-species amplification

EPIC-PCR markers RpL12, RpL10, RpS15A, DDC, RpL29 and RpS6 were tested for Mendelian inheritance patterns in

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

Table 1. EPIC markers developed from Dopa Decarboxylase (DDC) and ribosomal protein (Rp) genes for *H. armigera*. GCG codes used (R = G/A, K = G/T, Y = C/T, M = C/A, W = A/T, S = G/C) in degenerate primers. Primer annealing temperature (Ta), observed heterozygosity (Ho), expected heterozygosity (He) for individual primer pairs are provided. GenBank Identifier (GI) for *Bombyx mori* gDNA and various lepidopteran ribosomal proteins mRNA used in comparative genomic method of primer design are also provided. The number of alleles detected for the characterised EPIC-markers in five *H. assulta* (Hs), 20 *H. zea* (Hz) and 20 *H. punctigera* (Hp) are reported.

Primer name and sequence (5′ – 3′)		Expected allele size (bp)	No. alleles	Ta (°C)	gDNA and mRNA GI used in EPIC primers design	Ho	He	*Hz*	*Hp*	*Hs*
EPIC-DDC-F1	CGTCGAAGGCGCAAGATGATGTT	206	11	50	2108063, 2108065, 2108077, 2108079	0.864	0.885	4	5†	2
EPIC-DDC-R1	GAGGAAGATATCCGCAATGGATTG				2108085, 2108091, 2108099, 112984465, 54102649					
EPIC-RpL11-F	CTTTTTGAAAAGTTGTAATCATGG	287	2	51	54097916, 15213757, 54609212	0.217	0.198	1	2	1
EPIC-RpL11-R	CAACGCRGGCGGTKGTACACG				148226459					
EPIC-RpL29-F	AAATGGCAAAGTCAAAGAATCA	301	10	50	46793771, 16566715, 54609250	0.609	0.535	2	8	1
EPIC-RpL29-R	CAAGGGTGGATTCGTGCCT									
EPIC-RpS6-F	AGCARGGCTTCCCSATGAARCAG	274	17	50	54098483, 1498732, 16566730,	0.750	0.908	1	16	1
EPIC-RpS6-R	CTTTGACATCARCARACGAACACG				112349997, 54609292					
EPIC-RpL3-F2	GCATTGGTGTYACMAAGGGYAAAG	154	4	50	54094334, 18253046, 112349932	0.455	0.511	2	6	3
EPIC-RpL3-R2	CARCTTCTTWGTGTGCCAACG									
EPIC-RpL12-F	TGCCSCCTAAGTTYGAYCCWAATG	146	2	55	54091590, 16566712, 112349873,	0.050	0.050	ND	ND	ND
EPIC-RpL12-R	GCACCGACTTCYCCRCCGACGCATC				112983545					
EPIC-RpS10-F	GATCCTTTTCTTKTGTYAARWTAG	239	2	50	54098115, 15213807, 56462193,	0.042	0.042	ND	ND	ND
EPIC-RpS10-R	TTYTGTTTAGGCATCAACATCTTG				54098115					
EPIC-RpS8-F2	GAGCARTTCCATACWGGRCGT	328	3	50	54104932, 49532847, 18253054,	0.083	0.082	ND	ND	ND
EPIC-RpS8-R2	CCRCACTGRCCKGGSCGRCT				112349996					
EPIC-RpS15A-F	GTTCTGTCAWTTCGGTTCGCT	295	2	50	46744252, 15213819, 56462195,	0.042	0.042	ND	ND	ND
EPIC-RpS15A-R	TCARTACRTTCATWCGCACCAT				54609314, 112349867					
EPIC-RpS2-F	TCCGMAGAGGTTACTGGGGTA	324	3	51	54098782, 50284379, 27260895,	0.100	0.185	ND	ND	ND
EPIC-RpS2-R	GCWGACACAATACCAGTACCAC				112349777					
EPIC-RpS4-F	ACCATCCGCTAYCCMGACC	304	2	50	46704897, 16566724, 112350104	0.125	0.120	ND	ND	ND
EPIC-RpS4-R	CTCRCGGGACACRATGGTGC									
EPIC-RpL30-F	GCTATTGCTCCTCGTCCTGC	337	3	51	54107725, 54609252, 15213779	0.083	0.082	ND	ND	ND
EPIC-RpL30-R	TCAACTGCAACACCAATGGTTC									

'ND', not determined, '†', possible gene duplication due to detection of >2 alleles within individual *H. punctigera* samples.

the *H. armigera* mapping families Ha304 or Ha929 (for details of mapping crosses, see Lee, 2006). The single copy gene status and Mendelian inheritance patterns for all other Rp genes listed in table 1 have been ascertained by Lee (2006) using southern hybridization, single strand conformation polymorphism (SSCP), PCR-RFLP and PCR heteroduplex analysis. The EPIC-PCR markers DDC, RpS6, RpL29, RpL3 and RpL11 were further tested for cross-species PCR amplification in *H. assulta* ($n = 5$, Ha1-5), *H. punctigera* ($n = 20$, HP1-20) and *H. zea* ($n = 20$, HZBR1-20) sampled from India, Australia and Brazil respectively.

DNA extraction, PCR amplification, cloning and sequencing

DNA samples from Australia (AD1–AD24), China (CH1–CH24) and India (ICY1–ICY6) as reported in Behere *et al.* (2007) were used to determine PCR robustness of EPIC markers. PCR amplification was carried out in a 25 µl reaction volume using 50–75 ng gDNA; 2.5 mM $MgCl_2$, 2.5 µl of 10× PCR reaction buffer; 0.625 units *Taq* DNA polymerase (Promega, cat. #M186A); 0.5 µM of each forward and reverse primers (table 1) and 0.2 mM dNTPs. The PCR amplification profile consisted of a 5 min DNA denaturing step at 95°C (1 cycle); 35 cycles of 95°C, 50°C and 72°C (1 minute each) and a final 10 min extension at 72°C (1 cycle). All PCR products were resolved on 1.5% or 1.8% agarose gels stained with Ethidium Bromide and visualized over a UV light illuminator. PCR amplicons were gel-purified using Wizard® SV Gel and PCR Clean-Up System (Promega, cat. #A9282) according to the suppliers' manuals. Purified DNA was ligated into pGEM-T easy vector (Promega, cat. #A1380) prior to transformation in JM109 competent cells (Promega, cat. #L1001). Transformed cells were grown on LB agar plates supplemented with 0.005% Ampicillin, 0.008% X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and 0.5 mM ITPG (Isopropyl β-D-1-thiogalactopyranoside) at 37°C for 16–18 h. Up to seven positively transformed colonies from each transformation experiment were lysed in 75 µl DNAse-free water (Sigma®) in individual sterile 1.7 ml Eppendorf tubes. Lysed colonies were kept frozen at −20°C until needed in PCR amplification as DNA template using universal SP6 and T7 primers in the standard 25 µl PCR amplification reaction. Sequencing of purified PCR amplicons used the universal sequencing primers SP6 and T7 and the Big Dye Terminator sequencing reaction V3.1 kit (Applied Biosystems). Post sequencing reaction clean up of samples followed the instructions as supplied by the Australian Genome Research Facility (AGRF) in Melbourne, Australia, where all samples were sequenced.

Polyacrylamide gel electrophoresis (PAGE)

Size polymorphisms were visualized on 6% PAGE gels using γ-^{33}P dATP end-labelled forward primers following the protocol of Mohra *et al.* (2000) and Tay & Crozier (2001). Gels were vacuum-dried at 70°C for 30 min prior to overnight exposure onto storage phosphor screens (Molecular Dynamics, Amersham BioSciences, cat. #63-0034-79) at room temperature and scanned using a Typhoon 8600 High Performance Laser Scanning System (Amersham Pharmacia Biotech). The numbers of alleles detected from individual loci on the 6% polyacrylamide gels were determined.

Post sequencing and PAGE EPIC DNA marker analyses

DNA sequences were analysed using the Pregap4 and Gap4 programs within the Staden Molecular DNA analysis software (Staden *et al.*, 2000). Gap4 assembled consensus sequences were imported into CLC Free Workbench 3.2.1 (CLC bio, Aarhus, Denmark) for sequence alignment. Sequence identity was determined by BLASTN search (Altschul *et al.*, 1997) against the nr DNA database deposited in GenBank. Expected levels of heterozygosity (He), observed heterozygosity (Ho) and deviation from Hardy-Weinberg equilibrium in EPIC markers DDC, RpS6, RpL3, RpL11 and RpL29 were calculated using a web version of the population genetic software Genepop v3.4 (Raymond & Rousset, 1995; <http://genepop.curtin.edu.au>). We used EnzymeX v3.1 <http://mekentosj.com/enzymex/> to further characterise homoplasious alleles detected within the Rp and DDC EPIC-PCR markers.

Results

Ribosomal protein EPIC-PCR markers

A total of 45 ribosomal protein genes from *H. armigera* were tested in EPIC-PCR marker design. We identified a total of 11 Rp genes with suitable, polymorphic intron lengths. A subset of these EPIC-PCR markers that exhibited comparable levels of observed heterozygosity as microsatellite DNA loci were further characterised by DNA sequence analyses and cross-species PCR amplification tests (table 1). Of the 45 Rp genes tested, six were monomorphic in the samples tested, while the remaining 28 either had intron sizes that were too large (i.e. >400 bp PCR amplicons) or failed to amplify reliably. A list of all failed or monomorphic Rp genes tested and their primer sequences are available on request.

Allele characterisation in Rp EPIC DNA markers

Examples of allelic size homoplasy (i.e. alleles sharing identical amplicon size but differing in indel and SNP composition) were detected in RpL3, RpS6 and DDC by DNA sequence characterisation (table 2). In RpL3, three 145 bp alleles from *H. armigera* AD4, AD6 and AD21 samples were characterised that showed the presence of four indels between AD4 (EU190362), AD6 (EU190361) and AD21 (EU190360). In RpS6, allele size homoplasy was identified for the 269 bp allele from AD24 (EU190411) and AD20 (EU190410) due to two indels at nucleotide positions 63 and 165. In DDC, size homoplasy was detected for the 207 bp allele between samples AD3 (EU190407), AD5 (EU190406), AD7 (EU190405), AD8 (EU190408) and AD1 (EU190409) due to two indels at nucleotide positions 80 and 118. Two examples of allele size homoplasy detected in RpL3 and RpS6 are provided (fig. 1). Homoplasious alleles in RpL3, RpS6 and DDC EPIC markers are all differentiable by RFLPs (table 3). No allele size homoplasy, due to indels, was detected for the remaining 11 alleles in the six EPIC markers where multiple *H. armigera* samples were sequenced. All alleles characterised by DNA sequencing have been deposited in GenBank (Accession Numbers EU190360–EU190424).

Cross-species amplification of EPIC markers

EPIC-PCR markers DDC, RpL11, RpL29, RpL3 and RpS6 were tested in *H. zea* ($n = 20$), *H. punctigera* ($n = 20$) and

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

Table 2. *Helicoverpa armigera* EPIC DNA markers and alleles characterised by sequencing. The numbers of clones sequenced from individual *H. armigera* samples are indicated. Allele size homoplasy detected in RpL3 (allele 145 bp: AD6, AD4 and AD21), RpS6 (allele 269 bp: AD20 and AD24) and DDC (allele 207: AD3, AD5, AD7, AD8 and AD1). Gaps indicated by [-], Dalmore (Australia) and China samples have the sample ID prefix 'AD' and 'CH', respectively.

Marker	Alleles (bp)	# of clones sequenced	*H. armigera* samples	INDEL position (nt)				GenBank Accession
RpL12	146	1, 1	CH13, AD10	no Indel				EU190387, EU190388
	148	2, 1, 2, 4, 1, 3, 2, 2	CH17, CH13, CH7, CH6, AD10, AD17 AD12, AD11	no Indel				EU190389, EU190390, EU190391, EU190392, EU190393, EU190394, EU190395, EU190396
RpL3	145	1	AD21	34 [-]	77 [-]	98 [A]	99 [A]	EU190360
	145	2, 1	AD6, AD4	34 [G]	77 [A]	98 [-]	99 [-]	EU190361, EU190362
	146	3, 5	AD2, AD21	no Indel				EU190364, EU190363
	147	6, 4, 6, 3	AD13, AD6, AD4, AD2	no Indel				EU190365, EU190366, EU190367, EU190368
RpL29	278	1	AD7					EU190378
	283	2	AD6					EU190379
	287	1	AD8					EU190380
	291	3, 3, 2, 4, 2	AD8, AD7, AD6, AD2, AD11	no Indel				EU190381, EU190382, EU190383, EU190384, EU190385
	301	1	AD11					EU190386
RpS6	219	4	AD7					EU190419
	225	2	AD8					EU190417
	244	2	AD17					EU190415
	260	1	AD20					EU190412
	263	2, 1	AD15, AD17	no Indel				EU190414, EU190413
	265	2	AD16					EU190416
	267	2	AD8					EU190418
	268	2, 3	AD15, AD7	no Indel				EU190421, EU190420
	269	3	AD20	106 [T]	209 [-]			EU190410
	269	3	AD24	106 [-]	209 [G]			EU190411
	270	1	AD6					EU190422
	274	4, 1	AD2, AD16	no Indel				EU190424, EU190423
RpL11	292	1, 1, 2	AD22, AD20, AD17	no Indel				EU190369, EU190370, EU190371
	294	2, 1, 2, 1, 3, 3	AD22, AD20, AD17, AD12, AD11, AD10	no Indel				EU190372, EU190373, EU190374, EU190375, EU190376, EU190377
DDC	212	1	AD6					EU190397
	211	1	AD6					EU190398
	206	1	AD2					EU190402
	217	1	AD5					EU190399
	210	1	AD4					EU190400
	209	1	AD3					EU190401
	216	1, 1	AD2, AD4	no indel				EU190403, EU190404
	207	1, 1, 1, 1	AD3, AD5, AD7, AD8	80 [T]	118 [-]			EU190407, EU190406, EU190405, EU190408
	207	1	AD1	80 [-]	118 [T]			EU190409

H. assulta ($n = 5$) using conditions specified for *H. armigera* (table 1). No further optimisation, of PCR and 6% PAGE conditions, was necessary. The DDC-EPIC f1/r1 marker is also one of the most polymorphic markers in *H. zea* with four easily identifiable alleles. In two other highly polymorphic EPIC-PCR markers (RpS6 and RpL29), relatively high numbers of alleles were detected in *H. punctigera* (16 and 8 alleles, respectively); however, in *H. zea* of similar sample size, numbers of alleles from these two EPIC-PCR markers were surprisingly low (one and two alleles, respectively) (fig. 2, note that not all alleles detected were shown).

Mendelian inheritance and gene copy number

With the exception of three EPIC-PCR markers (RpL12, RpS15A and RpL10) that did not show observable length polymorphisms in the *H. armigera* mapping crosses Ha304 and Ha929, all remaining EPIC-PCR markers (table 1) demonstrated Mendelian inheritance patterns. Single copy gene status for individual loci was inferred from the parental alleles being present in single copy gene ratio among the F2 individuals (Lee, 2006). No allele dropout or null alleles were detected in these EPIC-PCR markers in the mapping families.

Discussion

We have shown that alternative nuclear DNA molecular genetic markers, such as EPIC-PCR markers, can be successfully designed in the noctuid moth *H. armigera*, where developing microsatellite DNA markers has been difficult. Screening of intron regions detected length polymorphisms

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

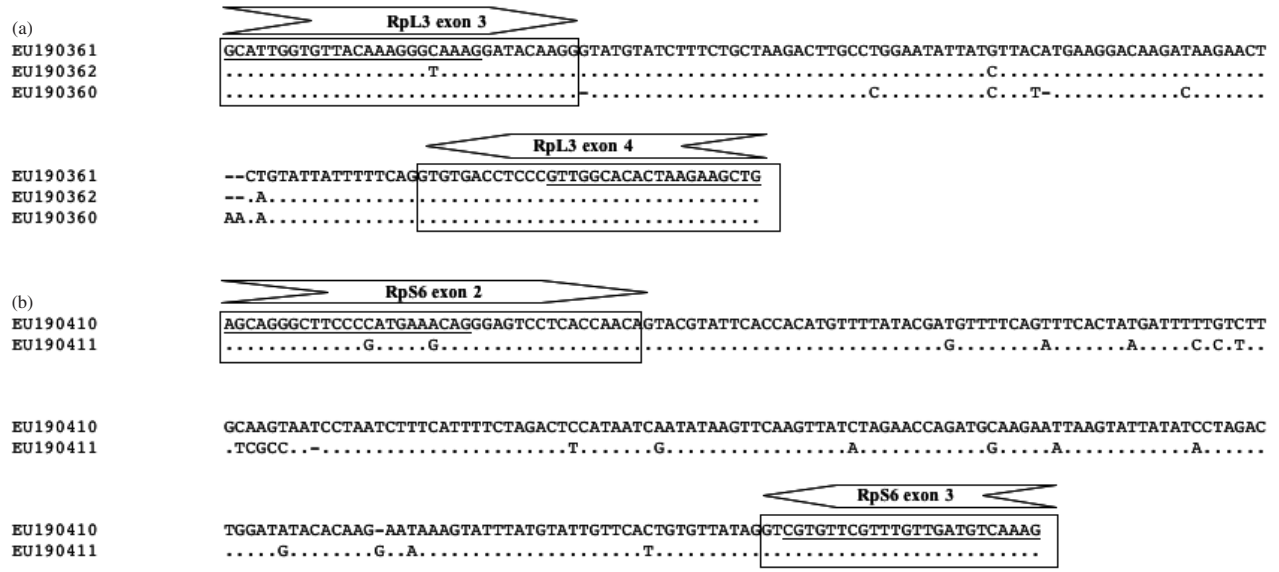

Fig. 1. Examples of size homoplasies in the *Helicoverpa armigera* EPIC markers. (a) RpL3 EPIC-DNA marker detected size homoplasy in the 145 bp amplicons from *H. armigera* samples AD6 (EU190361), AD4 (EU190362) and AD21 (EU190360). (b) RpS6 EPIC-DNA marker detected size homoplasy in the 269 bp alleles in *H. armigera* AD20 (EU190410) and AD24 (EU190411) samples. Identical bases are represented by (.), gaps inserted for alignment of sequences due to presence of indels are represented by '-'. SNPs detected between the homoplasious alleles are shown. Boxed regions are partial RpL/RpS exons, primer sequences are underlined.

Table 3. Examples of expected restriction fragment length polymorphisms (RFLPs) of homoplasious alleles detected in EPIC-PCR markers of *Helicoverpa armigera*. For each homoplasious allele within each of the three markers, RFLP patterns generated by two different restriction enzymes are provided. RFLPs are not determined for non-homoplasious alleles. Note that visualisation of the small fragment sizes generated by DdeI restriction digest of EU190362 should be carried out using 6% polyacrylamide gel electrophoresis.

Marker	*H. armigera* samples	Accession numbers	Allele size (bp)	Restriction enzyme 1	Fragment size generated (bp)	Restriction enzyme 2	Fragment size generated (bp)
RpL3	AD21	EU190360	145 bp	DdeI	87, 48, 10	NciI	86bp, 59bp
	AD6	EU190361	145 bp		87, 48, 10		uncut
	AD4	EU190362	145 bp		49, 47, 39, 10		uncut
RpS6	AD20	EU190410	269 bp	DdeI	uncut	MseI	177 bp, 92 bp
	AD24	EU190411	269 bp		184, 85		uncut
DDC	AD3, AD5, AD7, AD8	EU190408	207 bp	DraI	uncut	MseI	107 bp, 100 bp
	AD1	EU190409	207 bp		125, 82		uncut

in 11 Rp genes and the nuclear gene DDC. We further carried out sequence characterisation of various alleles in three of the most polymorphic loci (DDC, RpL29 and RpS6) and in three loci that showed lower levels of allele length polymorphisms (RpL3, RpL11 and RpL12). Mutation characterisation of randomly selected alleles from these six loci indicated that sequence length polymorphisms were generally due to random insertions and deletions, although size homoplasy was detected in three instances.

DNA sequencing allows the identification and fine-scale characterisation of homoplasious alleles with respect to the presence of SNPs and/or indels. Homoplasious alleles, thus identified, could be further differentiated through RFLPs patterns (table 3) or through DNA fragment conformation analytical methods, such as SSCP coupled with denaturing gradient gel electrophoresis (Palumbi, 1996; Estoup *et al.*, 2002). We have provided examples of differentiating homoplasious alleles by RFLPs; however, the treatment of SNPs within introns in individual EPIC DNA loci as distinctive alleles should proceed with caution. Differentiating between real mutations and DNA polymerase introduced errors (e.g. O'Mahony *et al.*, 2007) would require intensive multiple sequencing and was beyond the scope of our study. Allele length homoplasies, due to SNPs and/or indels, are also common in microsatellite DNA markers, both at the flanking sequence regions and within the microsatellite DNA repeat units (reviewed in Jarne & Lagoda, 1996). Size homoplasies in microsatellite alleles have also been identified through DNA sequencing in Australian *H. armigera* individuals from the Dalmore population (G.T. Behere, data not shown).

Microsatellite allele size homoplasy may not significantly affect various population genetic parameter estimations such as Wright's (1951) *F*-statistics (but see O'Reilly *et al.* 2004), relatedness estimates, the model of population isolation by distance and phylogenetic reconstruction of closely related populations (reviewed in Estoup *et al.*, 2002). Size homoplasy can, however, lower the level of observed heterozygosity, thereby affecting the power of exact tests on genotypic

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

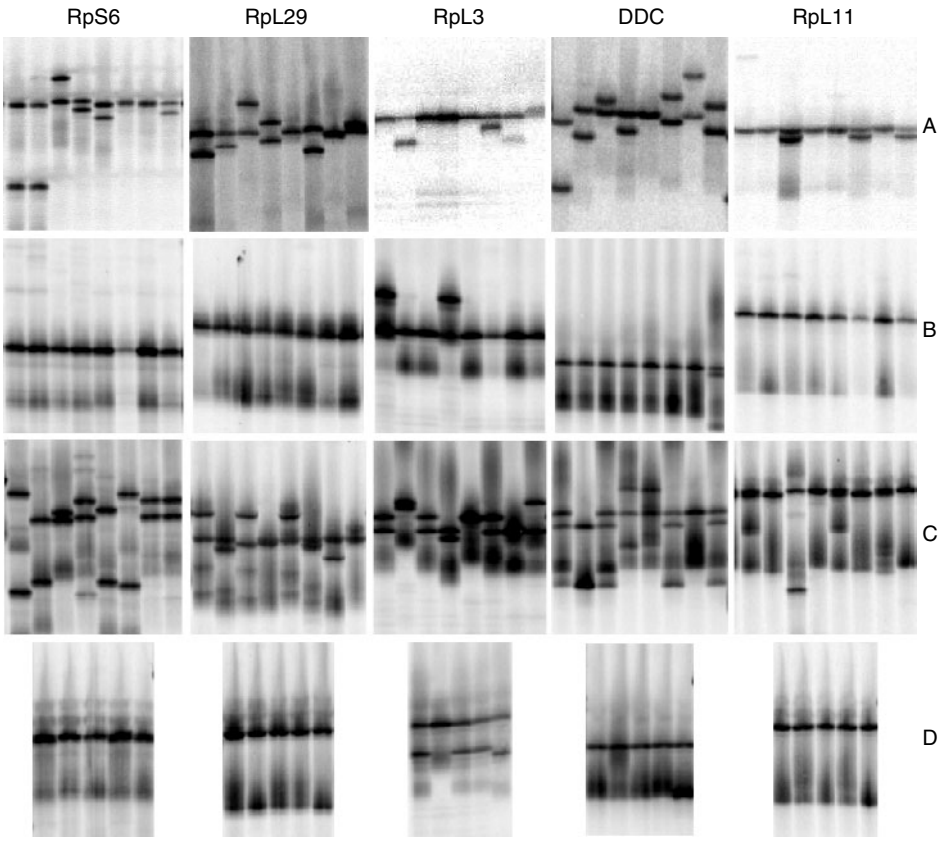

Fig. 2. Six percent PAGE analysis of five EPIC markers in pest *Helicoverpa* species. Rows A, B, C and D are *H. armigera*, *H. zea*, *H. punctigera* and *H. assulta*, respectively. EPIC markers are least polymorphic in *H. zea* (row B) and *H. assulta* (row D). Note that the DDC gene in *H. punctigera* is potentially duplicated due to the detection of multiple (>2) bands per individual by the DDC EPIC DNA marker.

linkage disequilibrium and Hardy-Weinberg equilibrium (see Estoup *et al*., 2002). Furthermore, strong constraints on allele size can increase size homoplasy leading to inaccurate gene flow estimates between sub-populations (Paetkau *et al*., 1997; Gaggiotti *et al*., 1999) and can cause downward bias in F_{ST} and/or R_{ST} estimates of population differentiation (O'Reilly *et al*., 2004). We observed a random pattern of indels within the intron regions of Rp and DDC EPIC markers, low frequencies of size homoplasy, the absence of null alleles in mapping crosses and Hardy-Weinberg equilibrium for all markers. These observations indicate that these EPIC markers will be useful in population and evolutionary genetic studies in *H. armigera* and related species. The accuracy of this assumption should be further examined based on complementary data from different types of molecular markers (Queney *et al*., 2001; Zhang & Hewitt, 2003).

Although allele homoplasy may not necessarily affect the applicability of microsatellite loci in estimating various population genetic parameters, other problems, such as presence of null alleles, allele dropout or the occurrence of multiple alleles for a specific locus within an individual (due to high copy number of microsatellite flanking sequences) can adversely affect population genetic data analysis (Endersby *et al*., 2007), problems that are especially pronounced in certain organisms (e.g. *Aedes egypti*, Chambers *et al*., 2007; *Anopheles gambiae*, Meglécz *et al*., 2007) including various lepidopteran species. In *H. armigera*, microsatellite loci that exhibited multiple alleles within single individuals have been detected (e.g. HaB60, HaD47, HarSSR9, HarSSR10; G.T. Behere, data not shown) and reported (Ji & Zhang, 2004; Grasela & McIntosh, 2005; Ji *et al*., 2005; Endersby *et al*., 2007). The presence of multiple (i.e. >2) alleles from specific microsatellite loci in a diploid individual suggested that the loci may be present in multiple copies and, therefore, unsuitable for use in population genetic analysis. Multi-copy microsatellite loci, due to similarity at the flanking sequence regions in many organisms, have been linked to or suspected to have evolutionary associations with mobile elements or are due to unequal recombination events (Arcot *et al*., 1995; Nadir *et al*., 1996; Ramsay *et al*., 1999; Akagi *et al*., 2001; Wilder & Hollocher, 2001; Meglécz *et al*., 2004, 2007; Van't Hof *et al*., 2007).

Null alleles in *H. armigera* microsatellite loci have represented a major challenge to investigating population genetic structures of this pest species (Ji *et al*., 2005; Endersby *et al*., 2007). Anonymous nuclear DNA markers, such as microsatellite DNA markers, are especially prone to null alleles (Tay *et al*., 2003; Ji *et al*., 2005) because the non-coding primer annealing sites are unlikely to be evolutionary conserved, resulting in higher frequencies of random mutations such as SNPs and indels. EPIC-PCR markers can overcome the problems of multi-copy loci and null alleles associated with microsatellite loci. However, designing

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

EPIC-PCR markers requires several steps, including isolation of genes, determination of gene copy numbers and identification of introns of appropriate size suitable for population genetic studies. These obstacles have hindered the development of widespread lepidopteran EPIC markers in the past. With the increasing availability of ESTs for comparative genome analysis, developing EPIC markers is expected to become more achievable in many non-model organisms.

The use of intron sequences in phylogenetic inferences has been demonstrated (Hedin & Maddison, 2001; Allen & Omland, 2003; Kawakita *et al*., 2003; Fujita *et al*., 2004), and applying EPIC markers to population genetic structure investigations has been reported in many organisms including aquatic organisms and insects (Palumbi & Baker, 1994; He & Haymer, 1997; Villablanca *et al*., 1998; Berrebi *et al*., 2005, 2006; Hubert *et al*., 2006; see also Roderick, 1996; Zhang & Hewitt, 2003); however, its use in population genetics of Lepidoptera is yet to be demonstrated. Similar substitution rates between microsatellite DNA flanking regions and introns have been reported (Brohede & Ellegren, 1999). Mutation rates of *H. armigera* EPIC-PCR intron regions should be compared with the mutation rates of *H. armigera* microsatellite DNA loci not affected by issues such as multicopy alleles, the presence of null alleles or allele dropout. Similar mutation rates between EPIC DNA intron regions and non-coding simple sequence repeat regions (i.e. neutral microsatellite loci regions) will further support the suitability of EPIC markers as alternative markers for investigating *H. armigera* population genetic structures. Through *H. armigera* family crosses, Mendelian inheritance patterns in DDC, RpL29 and RpS6 EPIC markers were confirmed in this study. Furthermore, Lee (2006) has shown by southern hybridization, heteroduplex conformation, and EPIC-PCR analyses that the majority of Rp genes analysed in this study (table 1) also existed as single copy genes. The single copy gene status of RpL12, RpS10 and RpS15A has not been determined, as these EPIC markers were monomorphic in the *H. armigera* family crosses. These Rp genes are likely to exist as single copy genes, as shown by chromosome synteny analysis in *B. mori*, *Ostrinia nubilalis* and *Heliothis virescens* (Lee, 2006). Nevertheless, utilising these three EPIC markers in population and evolutionary genetic studies should proceed with caution, as these genes may be affected by the presence of pseudogenes and/or gene duplication, both of which are known to affect some Rp genes (Lee, 2006) while duplication of the DDC gene and surrounding region has been reported in spiders (Hedin & Maddison, 2001) and *Drosophila* (Eveleth & Marsh, 1986) and suspected, based on 6% PAGE, in *H. punctigera* (table 1, fig. 2).

We have shown that five of the EPIC-PCR markers developed for *H. armigera* also amplified in *H. assulta*, *H. punctigera* and *H. zea*, and the numbers of alleles detected were generally higher in *H. punctigera*. That all *H. assulta* samples were found to be monomorphic in three of the five EPIC markers characterised was unexpected, although this is likely to reflect the small sample size (table 1). Low numbers of alleles were also detected in *H. zea*, as compared with *H. armigera* and *H. punctigera*, all of which consisted of similar numbers of individuals. This may reflect the finding of Behere *et al*. (2007) that *H. zea* populations in North and South America were likely the results of *H. armigera* founder events at approximately 1.5 million years ago. Microsatellite DNA markers remained the marker system of choice and will continue to be powerful genetic markers in evolutionary, molecular ecological and population genetic studies for the majority of organisms. However, in organisms where the development and utilisation of microsatellite markers have consistently been problematic, EPIC DNA markers offer an alternative option. EPIC markers developed in this study will be valuable to the investigation of *H. armigera* population genetic structure, mating behaviour and population evolutionary history, thereby providing a much needed basis to better understand the population dynamics of one of the most significant lepidopteran pests in the world.

Acknowledgements

This project was supported by The Australian Research Council (ARC) through its funding of the Special Research Centre CESAR (Centre for Environmental Stress and Adaptation Research) and funding from The State Government of Victoria, Australia to WTT. GTB was supported by the Melbourne International Research Scholarship (MIRS) and Melbourne International Fee Remission Scholarship (MIFRS). SFL and WTT were supported by the Max-Planck-Gesellschaft. Ary Hoffman, Steve McKechnie, Adam Williams, Tamar Stzal and Nancy Endersby provided helpful discussions during the course of this study. *Helicoverpa* samples were kindly provided by Nancy Endersby, Stephen Cameron, Keshav Kranthi, Yidong Wu and Derek Russell.

References

Akagi, H., Yokozeki, Y., Inagaki, A., Mori, K. & Fujimura, T. (2001) Micron, a microsatellite-targeting transposable element in the rice genome. *Molecular Genetics and Genomics* **266**, 471–480.

Allen, E.S. & Omland, K.E. (2003) Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). *The Auk* **120**, 961–969.

Altschul, S.F., Madden, T.L., Shäfer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.

Arcot, S.S., Wang, Z., Weber, J.L., Deininger, P.L. & Batzer, M.A. (1995) *Alu* repeats: a source for the genesis of primate microsatellites. *Genomics* **29**, 136–144.

Behere, G.T., Tay, W.T., Russell, D.A., Heckel, D.G., Appelton, B.R., Kranthi, K.R. & Batterham, P. (2007) Mitochondrial DNA analysis of field populations of *Helicoverpa armigera* (Lepidoptera: Noctuidae) and of its relationship to *H. zea*. *BMC Evolutionary Biology* **7**, 117.

Berrebi, P., Boissin, E., Fang, F. & Cattaneo-Berrebi, G. (2005) Intron polymorphism (EPIC-PCR) reveals phylogeographic structure of Zacco platypus in China: a possible target for aquaculture development. *Heredity* **94**, 589–598.

Berrebi, P., Retif, X., Fang, F. & Zhang, C.-C. (2006) Population structure and systematics of *Opsariichthys bidens* (Osteichthyes: Cyprinidae) in south-east China using a new nuclear marker: the introns (EPIC-PCR). *Biological Journal of the Linnean Society* **87**, 155–166.

Brohede, J. & Ellegren, H. (1999) Microsatellite evolution: Polarity of substitutions within repeats and neutrality of flanking sequences. *Proceedings of the Royal Society of London Series B, Biological Sciences* **266**, 823–833.

Chambers, E.W., Meece, J.K., McGowan, J.A., Lovin, D.D., Hemme, R.R., Chadee, D.D., McAbee, K., Brown, S.E., Knudson, D.L. & Severson, D.W. (2007) Microsatellite isolation and linkage group identification in the yellow fever mosquito *Aedes aegypti*. *Journal of Heredity* **98**, 202–210.

Daly, D., Waltham, K., Mulley, J., Watts, P.C., Rosin, A., Kemp, S.J. & Saccheri, I.J. (2004) Trinucleotide microsatellite loci for the peppered moth (*Biston betularia*). *Molecular Ecological Notes* **4**, 197–181.

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.

Estoup, A., Jarne, P. & Corneut, J.-M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology* **11**, 1591–1604.

Eveleth, D.D. & Marsh, J.L. (1986) Evidence for evolutionary duplication of genes in the Dopa Decarboxylase region of *Drosophila*. *Genetics* **114**, 469–483.

Fang, Q.Q., Cho, S., Regier, J.C., Mitter, C., Matthews, M., Poole, R.W., Friedlander, T.P. & Zhao, S. (1997) A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within Heliothinae (Lepidoptera: Noctuidae). *Systematic Biology* **46**, 269–283.

Fujita, M.K., Engstrom, T.N., Starkey, D.E. & Shaffer, H.B. (2004) Turtle phylogeny: insights from a novel nuclear intron. *Molecular Phylogenetics and Evolution* **31**, 1031–1040.

Gaffney, P.M. (2000) Molecular tools for understanding population structure in Antarctic species. *Antarctic Science* **12**, 288–296.

Gaggiotti, O.E., Lange, O., Rassmann, K. & Gliddon, C. (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* **8**, 1513–1520.

Garrick, R.C. & Sunnucks, P. (2006) Development and application of three-tiered nuclear genetic markers for basal Hexapods using single-stranded conformation polymorphism coupled with targeted DNA sequencing. *BMC Genetics* **7**, 11.

Grasela, J.J. & McIntosh, A.H. (2005) Cross-species investigation of *Helicoverpa armigera* microsatellites as potential markers for other related species in the *Helicoverpa* – *Heliothis* complex. *The Journal of Insect Science* **5**, 47.

Hardwick, D.F. (1965) The corn earworm complex. *Memoirs of the Entomological Society of Canada* **40**, 1–248.

He, M. & Haymer, D.S. (1997) Polymorphic intron sequences detected within and between populations of the Oriental fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America* **90**, 825–831.

Hedin, M.C. & Maddison, W.P. (2001) Phylogenetic utility and evidence for multiple copies of Elongation Factor-1a in the spider genus *Habronattus* (Araneae: Salticidae). *Molecular Biology and Evolution* **18**, 1512–1521.

Hubert, N., Duponchelle, F., Nunez, H., Rivera, R. & Renno, J.-F. (2006) Evidence of reproductive isolation among closely related sympatric species of *Serrasalmus* (Ostariophysii, Characidae) from the Upper Madeira River, Amazon, Bolivia. *Journal of Fish Biology* **68**, 1–21.

Jarne, P. & Lagoda, P.J.L. (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**, 424–429.

Ji, Y.-J. & Zhang, D.-X. (2004) Characteristics of microsatellite DNA in lepidopteran genomes and implications for their isolation. *Acta Zoologica Sinica* **50**, 608–614.

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.

Ji, Y.-J., Wu, Y.-C. & Zhang, D.-X. (2005) Novel polymorphic microsatellite markers developed in the cotton bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae). *Insect Science* **12**, 331–334.

Kawakita, A., Sota, T., Ascher, J.S., Ito, M., Tanaka, H. & Kato, M. (2003) Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in bumble bees (*Bombus*). *Molecular Biology and Evolution* **20**, 87–92.

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.

Lee, S.F. (2006) Genome mapping in Lepidoptera. PhD thesis, University of Melbourne. Melbourne, Australia.

Lessa, E.P. (1992) Rapid survey of DNA sequence variation in natural populations. *Molecular Biology and Evolution* **9**, 323–330.

Meglécz, E., Petenian, F., Danchin, E., D'Acier, A.C., 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.

Meglécz, E., Anderson, S.J., Bourguet, D., Butcher, R., Caldas, A., Cassel-Lundhagen, A., d'Acier, A.C., Dawson, D.A., Faure, N., Fauvelot, C., Franck, P., Harper, G., Keyghobadi, N., Kluetsch, C., Muthulakshmi, M., Nagaraju, J., Patt, A., Péténian, F., Silvain, J.-F. & Wilcock, H.R. (2007) Microsatellite flanking region similarities among different loci within insect species. *Insect Molecular Biology* **16**, 175–185.

Miao, X.-X., Xub, S.-J., Li, M.-H., Li, M.-W., Huang, J.-H., Dai, F.-Y., Marino, S.W., Mills, D.R., Zeng, P., Mita, K., Jia, S.-H., Zhang, Y., Liu, W.-B., Xiang, H., Guo, Q.-H., Xu, A.-Y., Kong, X.-Y., Lin, H.-X., Shi, Y.-Z., Lu, G., Zhang, X., Huang, W., Yasukochi, Y., Sugasaki, T., Shimada, T., Nagaraju, J., Xiang, Z.-H., Wang, S.-Y., Goldsmith, M.R., Lu, C., Zhao, G.-P. & Huang, Y.-P. (2005) Simple sequence repeat-based consensus linkage map of *Bombyx mori*. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 16303–16308.

Mitter, C., Poole, R.W. & Matthews, M. (1993) Biosystematics of the Heliothinae (Lepidoptera: Noctuidae). *Annual Review of Entomology* **38**, 207–225.

Mohra, C., Fellendorf, M., Segelbacher, G. & Paxton, R.J. (2000) Dinucleotide microsatellite loci for *Andrena vaga* and other andrenid bees from non-enriched and CT-enriched libraries. *Molecular Ecology* **9**, 2189.

Nadir, E., Margalit, H., Gallily, T. & Ben-Sasson, S.A. (1996) Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6470–6475.

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

https://doi.org/10.1017/S000748530800583X Published online by Cambridge University Press

O'Mahony, E.M., Tay, W.T. & Paxton, R.J. (2007) Multiple rRNA variants in a single spore of the microsporidium *Nosema bombi*. *Journal of Eukaryotic Microbiology* **54**, 103–109.

O'Reilly, P.T., Canino, M.F., Bailey, K. & Bentzen, P. (2004) Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye Pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology* **13**, 1799–1814.

Paetkau, D., Waits, L.P., Clarkson, P.L., Craighead, L. & Strobeck, C. (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147**, 1943–1957.

Palumbi, S.R. (1996) Nucleic acids II: the polymerase chain reaction. pp. 205–247 *in* Hillis, D.M., Moritz, C. & Mable, B.K. (*Eds*) *Molecular Systematics*. Sunderland, MA, USA, Sinauer Associates.

Palumbi, S.R. & Baker, C.S. (1994) Contrasting population structure from nuclear intron sequences and mtDNA of Humpback whales. *Molecular Biology and Evolution* **11**, 426–435.

Prasad, M.D., Muthulakshmi, M., Arunkumar, K.P., Madhu, M., Sreenu, V.B., Pavithra, V., Bose, B., Nagarajaram, H.A., Mita, K., Shimada, T. & Nagaraju, J. (2005a) SilkSatDb: a microsatellite database of the silkworm, *Bombyx mori*. *Nucleic Acids Research* **33**, D403–D406.

Prasad, M.D., Muthulakshmi, M., Madhu, M., Archak, S., Mita, K. & Nagaraju, J. (2005b) Survey and Analysis of Microsatellites in the silkworm, *Bombyx mori*: frequency, distribution, mutations, marker potential and their conservation in heterologous species. *Genetics* **169**, 197–214.

Pringle, E.G., Baxter, S.W., Webster, C.L., Papanicolaou, A., Lee, S.F. & Jiggins, C.D. (2007) Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in *Heliconius melpomene*. *Genetics* **177**, 417–426.

Queney, G., Ferrand, N., Weiss, S., Mougel, F. & Monnerot, M. (2001) Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Molecular Biology and Evolution* **18**, 2169–2178.

Ramsay, L., Macaulay, M., Cradle, L., Morgante, M., degli Ivanissevich, S., Maestri, E., Powell, W. & Waugh, R. (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *The Plant Journal* **17**, 415–425.

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

Reddy, K.D., Abraham, E.G. & Nagaraju, J. (1999) Microsatellites in the silkworm, *Bombyx mori*: abundance, polymorphism, and strain characterization. *Genome* **42**, 1057–1065.

Roderick, G.K. (1996) Geographic structure of insect populations: Gene flow, phylogeography, and their uses. *Annual Review of Entomology* **41**, 325–352.

Scott, K.D., Wilkinson, K.S., Merritt, M.A., Scott, J.L., Lange, C.L., Schutze, M.K., Kent, J.K., Merritt, D.J., Grundy, P.R. & Graham, G.C. (2003) Genetic shifts in *Helicoverpa armigera* Hübner (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. & Gahan, L.J. (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.

Staden, R., Beal, K.F. & Bonfield, J.K. (2000) The Staden package, 1998. *Methods in Molecular Biology* **132**, 115–130.

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

Tay, W.T. & Crozier, R.H. (2001) Mating behaviour of *Rhytidoponera* sp. 12 ants inferred from microsatellite analysis. *Molecular Ecology* **10**, 167–173.

Tay, W.T., Miettinen, M. & Kaitala, A. (2003) Do male golden egg bugs carry eggs they have fertilized? A microsatellite analysis. *Behavioral Ecology* **14**, 481–485.

Van't Hof, A.E., Zwaan, B.J., Saccheri, I.J., Daly, D., Bot, A.N.M. & Brakefield, P.M. (2005) Characterization of 28 microsatellite loci for the butterfly *Bicyclus anynana*. *Molecular Ecology Notes* **5**, 169–172.

Van't Hof, A.E., Brakefield, P.M., Saccheri, I.J. & Zwaan, B.J. (2007) Evolutionary dynamics of multi-locus microsatellite arrangements in the genome of the butterfuly *Bicyclus anynana*, with implications for other Lepidoptera. *Heredity* **95**, 320–328.

Villablanca, F.X., Roderick, G.K. & Palumbi, S.R. (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology* **7**, 547–560.

Wardill, T.J., Scott, K.D., Graham, G.C. & Zalucki, M.P. (2004) Isolation and characterization of microsatellite loci from *Chiasmia assimilis* (Warren, 1899) (Lepidoptera: Geometridae). *Molecular Ecology Notes* **4**, 358–360.

Wilder, J. & Hollocher, H. (2001) Mobile elements and the genesis of microsatellites in Dipterans. *Molecular Biology and Evolution* **18**, 384–392.

Wright, S. (1951) The genetical structure of populations. *Ann Eugenics* **15**, 323–354.

Yasukochi, Y., Ashakumary, L.A., Baba, K., Yoshido, A. & Sahara, K. (2006) A second generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. *Genetics*, **173**, 1319–1328.

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

Zhang, D.-X. & Hewitt, G.M. (2003) Nuclear DNA analyses in genetic studies of populations practice, problems and prospects. *Molecular Ecology* **12**, 563–584.

Zhou, Y., Gu, H. & Dorn, S. (2005) Isolation of microsatellite loci in the colding moth, *Cydia pomonella* (Lepidoptera: Tortricidae). *Molecular Ecology Notes* **5**, 226–227.