Molecular Ecology Notes (2006) **6**, 683–685 doi: 10.1111/j.1471-8286.2006.01309.x

PRIMER NOTE

Variable nuclear markers for *Melanoplus oregonensis* identified from the screening of a genomic library

B. C. CARSTENS and L. L. KNOWLES
Department of Ecology and Evolutionary Biology, 1109 Geddes Rd., Museum of Zoology, Rm. 1089, University of Michigan, Ann Arbor, Michigan 48109-1079, USA

Abstract

We report the isolation of seven single-copy nuclear loci from the montane grasshopper *Melanoplus oregonensis*. With an average length of approximately 930 bp and between five and 67 variable sites, these loci are useful for population genetic analyses within *M. oregonensis*. Amplification of nuclear loci in additional species of western *Melanoplus* grasshoppers suggests that they will also be useful for a variety of population genetic, phylogeographic and phylogenetic studies.

Keywords: genomic library, *Melanoplus oregonensis*, single-copy nuclear polymorphic sequences

Received 27 September 2005; revision received 21 November 2005; accepted 8 January 2006

Since the statistical power in a variety of parameter estimates relies on the number of loci (e.g. Brumfield *et al*. 2003), the lack of variable nuclear loci limits the resolution of many empirical investigations conducted on nonmodel organisms. Genomic libraries enable biologists working in these systems to collect single-copy nuclear polymorphic sequences (SCNPs), allowing parameters such as species phylogeny, divergence dates, and genetic diversity to be estimated with more accuracy. We describe the creation of one such library using genomic DNA from the grasshopper *Melanoplus oregonensis*, a species where phylogeographic inference has been complicated by the incomplete sorting of ancestral polymorphism (Knowles 2001).

Total genomic DNA was extracted from a single individual of *M. oregonensis* using QIAGEN DNeasy kits and concentrated via ethanol precipitation to an initial concentration of ~4.2 μg/μL. A restriction enzyme digest (using *Hin*dIII) was used to shear the genomic DNA. *Hin*dIII was selected in part because this enzyme did not produce any fragments smaller than 1.6 kb in the mitochondrial genome of the related orthopteran *Locusta migratoria* (X80245; Flook *et al*. 1995), thereby reducing the likelihood of cloning mitochondrial DNA fragments. Sheared DNA was visualized on an agarose gel stained with EtBr, and fragments of between 1.0 and 1.5 kb were excised by gel extraction using the QIAquick Gel Extraction kit. Ethanol precipitation was used to concentrate the 1.0–1.5 kb fragments to a concentration of 8.0 ng/μL, and sticky ends were repaired with the Novagen Single dA Tailing Kit (EMD Biosciences). Fragments were cloned using the QIAGEN PCR*plus* Cloning kit.

Two-hundred colonies were picked, grown overnight in SOC media, and sequenced using an ABI 3730 Automated Sequencer at the University of Michigan DNA Sequencing Core. Forty-three sequences were discarded, either because they aligned with other cloned sequences ($n = 23$), had undesirable genomic properties such as Poly A runs or extreme AT richness ($n = 7$), or did not sequence well ($n = 13$). Sequences from the remaining clones were then translated into amino acid sequence. We selected 45 sequences that were apparently from noncoding regions of the *Melanoplus* genome, and used these sequences to design *Melanoplus*-specific polymerase chain reaction (PCR) primers. Primer design was conducted using PRIMER 3 (Rozen & Skaletsky 2000) and OLIGO 4.0 (Molecular Biology Insights, Inc.). Approximately 10 ng of genomic DNA was used per PCR. One unit of Invitrogen *Taq* DNA polymerase in a 1× concentration of buffer (supplied with the *Taq*) was used per PCR with the following protocol: initial denature 94°C for 3:00, 34 cycles at 94°C for 0:30, X°C for 0:45 (see Table 1 for locus-specific annealing temperatures and reagent concentrations), 72°C for 1:30 and a final extension at 72°C for 10:00. PCR products were subcloned in one individual per locus and five clones were sequenced to verify that these loci were single copy elements. PCR subcloning was conducted with the QIAGEN PCR*plus* Cloning kit and template

Correspondence: Bryan C. Carstens, Fax: (734) 763-4080;
E-mail: bcarsten@umich.edu

Table 1 Primer sequences for seven variable loci within *Melanoplus oregonensis*. Shown (from left) are the identification numbers of the loci, the forward and reverse primer sequences, the annealing temperature (T_a), the final $MgCl_2$, primer and dNTP concentrations and the GenBank Accession no. of the cloned sequence. For some loci, additional sequencing primers are also shown

Locus	Primer sequences (5′–3′)	T_a (°C)	$MgCl_2$ (mM)	Primer (μM)	dNTPs (μM)	Accession no.
2	F: TGAGAAGAAGGAAGAACGAAGTTAG	52°	1.125	3.25	125	DU635373
	R: TTTCATCTCCTCACGTGTAAYA					
	R_{SEQ}: TCCTCACGTGTAAYATTTAG					
6	F: TGCTACGCCGAGGAGATACT	50°	1.1875	3.25	112.5	DU635374
	R: GGTTACATTCCCTACTTTCTCACG					
	F_{SEQ}: ATTATTCGKTATGGGTCSAAC					
	R_{SEQ}: AATACTTTTGTTGCTTCTTGT					
73	F: CTACCATAGCAGTGCGAATGTC	52°	1.125	3.25	125	DU635376
	R: TTGGTTTTGAACCGGTATAGTG					
85	F: GTTCTCAATCCTCCACATCCTC	54°	1.0625	3.25	112.5	DU635377
	R: CATCTTGGTGTGCGTATTGTTT					
	R_{SEQ}: TGGTGTGCGTATTGTTTTAG					
102	F: TCGCTTATCCTTCAGTTGACAC	54°	1.25	3.25	125	DU635378
	R: AGACCTCACAACTCCACTGCAT					
	F_{SEQ}: TATCCTTCAGTTGACACTTG					
	R_{SEQ}: GACAACTCCACTGCATTCGG					
140	F: CCTCCTCCCACCTACGCATG	54°	1.0	3.0	125	DU635379
	R: TTCCTCGGCCTCATATTTG					
211	F: GTAACATTCTTCAACCAAA	52°	1.25	3.29	125	DU635380
	R: TTCTCGGGCTTCTTGACCT					

Table 2 Descriptive statistics for *Melanoplus oregonensis* SCNPs. Shown for each locus are the length (in bp), the number of segregating sites (S), the alleles (with frequencies shown in subscripts), the GenBank Accession nos for the alleles, Tajima's D (D; Tajima 1989). Also shown are the per-site nucleotide diversity (π), Watterson's theta (θ_w; Watterson 1975), population recombination (γ) of Hey & Wakeley (1997), and the mean linkage disequilibrium (LD_D; Hedrick 2005), calculated by dividing the total sequence into 10 equally sized regions. All parameters were calculated with the program SITES (Hey & Wakeley 1997)

Loc.	bp	S	Alleles	Accession	D	π	θ_w	γ	LD_D
2	956	48	$A_{(0.2)}$, $B_{(0.3)}$, $C_{(0.1)}$, $D_{(0.1)}$, $E_{(0.1)}$, $F_{(0.2)}$	DQ269689–DQ269694	0.8585	0.0181	0.0149	0.0032	0.0311
6	1009	15	A(0.3), B(0.2), C(0.05), D(0.05), E(0.2), F(0.2)	DQ269695–DQ269600	1.3353	0.0057	0.0042	0.0015	0.0022
73	854	5	$A_{(0.1)}$, $B_{(0.1)}$, $C_{(0.1)}$, $D_{(0.15)}$, $E_{(0.05)}$, $F_{(0.1)}$, $G_{(0.5)}$	DQ269710–DQ269716	0.4939	0.0019	0.0017	0.000	0.0135
85	707	28	$A_{(0.5)}$, $B_{(0.05)}$, $C_{(0.05)}$, $D_{(0.1)}$, $E_{(0.05)}$, $F_{(0.05)}$, $G_{(0.15)}$, $H_{(0.05)}$	DQ269717–DQ269724	−0.6363	0.0093	0.0111	0.0716	0.0284
102	755	7	$A_{(0.35)}$, $B_{(0.1)}$, $C_{(0.1)}$, $D_{(0.1)}$, $E_{(0.1)}$, $F_{(0.05)}$, $G_{(0.1)}$, $H_{(0.1)}$	DQ269725–DQ269732	1.4393	0.0027	0.0061	0.0070	0.0892
140	1002	22	$A_{(0.4)}$, $B_{(0.1)}$, $C_{(0.1)}$, $D_{(0.1)}$, $E_{(0.05)}$, $F_{(0.1)}$, $G_{(0.05)}$, $H_{(0.05)}$, $I_{(0.05)}$	DQ269733–DQ269741	−0.8429	0.0048	0.0061	0.0007	0.0445
211	1229	67	$A_{(0.05)}$, $B_{(0.05)}$, $C_{(0.05)}$, $D_{(0.05)}$, $E_{(0.1)}$, $F_{(0.2)}$, $G_{(0.1)}$, $H_{(0.05)}$, $I_{(0.05)}$, $J_{(0.1)}$, $K_{(0.1)}$, $L_{(0.1)}$	DQ269742–DQ269753	0.6332	0.0195	0.0003	0.0169	0.0821

concentrations of 4–6 ng/μL. We used an initial screening set that consisted of one individual from each of three closely related montane melanoplines: *M. oregonensis*, *Melanoplus montanus* and *Melanoplus marshalli*. We identified 12 loci with at least one variable site in this screening set. Percent sequence divergence averaged 1.1% between *M. montanus* and *M. marshalli*, 2.3% between *M. oregonensis* and *M. montanus*, and 1.4% between *M. marshalli* and *M. oregonensis*. We then sequenced nine additional *M. oregonensis* in seven loci (Table 2).

Construction of genomic libraries is an efficient way to generate a large amount of sequence data that are informative for both intraspecific and interspecific studies. However, there are some potential complications. Ascertainment bias can occur when a screening set is used to screen for polymorphisms (Rogers & Jorde 1996; Wakeley *et al*. 2001; Wang *et al*. 2003). When an intraspecific sample is used as a screening set, loci with intermediate levels of polymorphism are more easily identified as polymorphic, potentially biasing parameter estimates. An interspecific screening set consisting of three species was used to avoid problems with ascertainment bias. PCR subcloning of individuals with multiple heterozygous sites was also used to establish phase and to provide evidence that these loci were single copy.

Acknowledgements

We thank M. Keat and C. Davis for advice and assistance in the laboratory. This project was funded in part by the National Science Foundation grant (DEB-04-47224) to LLK and the Elizabeth Caroline Crosby Fund, NSF ADVANCE Project, University of Michigan, to LLK.

References

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphism in inferences of population history. *Trends in Ecology & Evolution*, **18**, 249–256.

Flook PK, Rowell CH, Gellissen G (2005) The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome. *Journal of Molecular Evolution*, **41**, 928–941.

Hedrick PW (2005) *Genetics of Populations*. Jones and Bartlett, Sudbury, MA.

Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics*, **145**, 833–846.

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, **10**, 691–701.

Rogers AR, Jorde LB (1996) Ascertainment bias in estimates of average heterozygosity. *American Journal of Human Genetics*, **58**, 1033–1041.

Rozen S, Skaletsky HJ (2000) PRIMER 3 on the www for general users and biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, NJ. http://www.genome.wi.mit.edu/genome_software/other/primer3.html.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Wakeley J, Nielsen R, Liu-Cordero SN, Ardlie K (2001) The discovery of single-nucleotide polymorphisms and inferences about human demographic history. *American Journal of Human Genetics*, **69**, 1332–1347.

Wang Z, Baker AJ, Hill GE, Edwards SV (2003) Reconciling actual and inferred population histories in the house finch (*Carpodacus mexicanus*) by AFLP analysis. *Evolution*, **57**, 2852–2864.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–275.