Hereditas 136: 247–250 (2002)

Brief report
Isolation of novel microsatellite loci in the Rocky Mountain apollo butterfly, *Parnassius smintheus*

NUSHA KEYGHOBADI, JENS ROLAND and CURTIS STROBECK

Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9 (present address of NK: Academy of Natural Sciences, 1900 Benjamin Franklin Pkwy., Philadelphia, PA 19103, USA. E-mail: keyghobadi@acnatsci.org

(Received May 30, 2002. Accepted July 28, 2002)

Microsatellite DNA loci have emerged as arguably the most popular molecular markers for ecological and evolutionary studies. They are prized for their high variability, co-dominant expression, and broad distribution throughout the genome. They have been widely used to study population genetic structure, kinship, and paternity. Butterflies have served as model systems in ecological and evolutionary studies, in part because they are often intimately associated with specific habitats, they are easily monitored and handled, and knowledge of their natural history is extensive (BOGGS et al. in press). Given the utility of butterflies as model systems in ecology and evolution, and the value of microsatellite markers as analytical tools in these fields, the application of microsatellite markers to studies of butterfly biology is very promising. Unfortunately, progress in this area has been hampered by the difficulty of isolating useful microsatellite markers from butterfly genomes (MEGLÉCZ and SOLIGNAC 1998; KEYGHOBADI et al. 1999; NÈVE and MEGLÉCZ 2000). Very few microsatellite loci have been reported for butterflies or in fact for any Lepidoptera (PALO et al. 1995; BOGDANOWICZ et al. 1997; MEGLÉCZ and SOLIGNAC 1998; KEYGHOBADI et al. 1999; REDDY et al. 1999; HARPER et al. 2000; ANTHONY et al. 2001) and isolation and characterization of microsatellite markers is clearly more difficult in these insects than in most other organisms (MEGLÉCZ and SOLIGNAC 1998; NÈVE and MEGLÉCZ 2000). Here, we describe four novel microsatellite loci useful for population studies in the butterfly, *Parnassius smintheus* (Doubleday). This brings the total number of usable loci for this species to eight, the largest number currently reported for any species of butterfly.

MATERIALS AND METHODS

Microsatellite isolation

Novel microsatellites were isolated using the PCR-based protocol described by GRIST et al. (1993). Details of the methodology are the same as in KEYGHOBADI et al. (1999). *Parnassius smintheus* genomic DNA was digested with *Sau*3A1 and fragments 200–600 bp long were ligated into *Bam* H1-digested and dephosphorylated M13mp18 vector. Ten-fold serial dilutions of the products of ligation were used as template in PCR amplifications with a microsatellite-specific primer $(GT)_{11}$ and a modified M13 universal forward sequencing primer (forward: 5′CGACGTTGTAAAACGACGGCCAG3′). Amplifications were in 25 μl with 1 × PCR buffer (10 mM Tris, pH 8.8, 0.1% Triton X100, 50 mM KCl and 0.16 mg/ml BSA), 2 mM $MgCl_2$, 0.16 μM of each primer, 120 μM dNTP, 0.5 units of *Taq* DNA polymerase (purified as described by ENGELKE et al. (1990)), and 1 μl of diluted ligation reaction. The cycling conditions were 1 min at 94 °C, followed by 35 cycles of 30 sec at 94 °C, 20 sec at 56 °C, and 7 sec at 72 °C, followed by 30 sec at 72 °C, on a 9600 thermal cycler (Perkin-Elmer). Products were run on agarose gels to determine the level of dilution at which discrete bands would be produced. At dilutions lower than this critical level smears would be observed and at higher dilutions no product would be observed. A large number (i.e. hundreds) of individual ''working stocks'' (GRIST et al. 1993) were made by using 1, 2, 5, or 10 μl of the critical ligation dilution as template in a similar PCR, but with only 20 cycles and with the M13 universal forward and reverse sequencing primers (reverse: 5′CAGGAAACAGCTATGACC3′). Working stocks were then diluted 100-fold and 1 μl of this dilution was used as template in a PCR with the $(GT)_{11}$ microsatellite-specific primer and the M13 forward sequencing primer under the same amplification conditions as in the first round of PCRs. The products of these reactions were run on agarose gels and individual microsatellite loci were identified as sharp bands. DNA from each such band was purified and sequenced from the forward sequencing primer with a dRhodamine Terminator Cycle Sequencing Kit (Applied Biosystems). Exten-

sion products were resolved on a 377 automated sequencer (Applied Biosystems). This provided flanking sequence for one side of the microsatellite. A primer was designed from this sequence and used with the reverse M13 sequencing primer to amplify across the microsatellite. These products were sequenced from the reverse sequencing primer to obtain the flanking sequence on the other side of the microsatellite. Optimal primer pairs for amplification of the microsatellite region were designed with OLIGO software (National Biosciences Inc., Version 5.0).

Sample collection and DNA extraction

Tissue samples of 557 adult butterflies were collected in 1995 and 1996 from seventeen different sites along three ridges in the Kananaskis region of Alberta, Canada (KEYGHOBADI et al. 1999). Sampling was at a very fine scale, with only 12 km separating the most distant sampling sites. Sample size per site ranged from 12 to 43 individuals as reported in KEYGHOBADI et al. (1999). Genomic DNA was extracted with the QIAamp™ tissue extraction kit (QIAGEN). The samples consisted of either small wing clippings (approximately 0.15 cm^2) or the entire body. In the former case, the entire sample was used for DNA extraction. In the latter case, approximately 25 mg of abdominal or thoracic tissue was used for males and females, respectively.

Microsatellite amplification and analysis

All novel loci were amplified separately in 15 µl reactions containing 1 × PCR buffer (10 mM Tris, pH 8.8, 0.1% Triton X100, 50 mM KCl and 0.16 mg/ml BSA), 2 mM $MgCl_2$, 0.16 µM of each primer, 120 µM dNTP, 0.3 units of *Taq* DNA polymerase, and approximately 75 ng of genomic DNA. Cycling conditions were 1 min at 94 °C, followed by three cycles of 30 sec at 94 °C, 20 sec at 54 °C, and 5 sec at 72 °C, followed by 33 cycles of 15 sec at 94 °C, 20 sec at 54 °C, and 1 sec at 72 °C, followed by a 30 min final extension at 72 °C, on a 9600 thermal cycler (Perkin-Elmer). One primer of each primer pair was end-labeled with either 6-FAM, HEX, or TET fluorescent dye (Applied Biosystems). Products were electrophoresed and detected on a 373A automated sequencer (Applied Biosystems) and electrophoretic data were collected and analyzed with GENESCAN and GENOTYPER software (Applied Biosystems).

RESULTS AND DISCUSSION

Three-hundred and twenty "working stocks" (GRIST et al. 1993) were screened for $(GT)_n$ repeats and 23 putative microsatellite loci were identified and sequenced. Of these 23 microsatellites, there was sufficient flanking sequence on both sides to design locus-specific primers for nine of them. One of these nine was identical to a microsatellite clone that had previously been characterized (Ps162 from KEYGHOBADI et al. (1999)). Thus, we designed PCR primers for eight novel loci (GenBank Accession numbers AY048089–AY048096). Four of these amplified cleanly and were polymorphic, and were therefore considered useful for genotyping individuals and for population studies. Primer sequences for these loci are presented in Table 1. Three of the four loci

Table 1. *Novel microsatellite loci useful for genotyping in Parnassius smintheus. FAM, HEX, and TET are fluorescent dye labels (Applied Biosystems). Data are for 557 individuals from 17 sampling sites. Size range of polymerase chain reaction (PCR) product is given in base pairs (bp). Expected heterozygosity (H_e) and observed heterozygosity (H_o) were averaged over sampling sites. Unbiased H_e was calculated as in NEI and ROYCHOUDHURY (1974). * GTTT was added to the 5′ end of the unlabelled primer of Ps163 to promote non-templated 3′ adenine addition in PCR (BROWNSTEIN et al. 1996). † For locus Ps165 all alleles greater than 250 bp in length were pooled into the 250-bp size class, as they could not be sized unambiguously due to excessive stutter.*

Locus name	Primer sequences (5′ to 3′)	Cloned repeat	Size Range	No. of non-null alleles	Mean H_e	Mean H_o	Accession #
Ps76	**FAM** GGCAAATACCCTCCCTA GTAACGCTCAGTAAATCTGC	$(TTG)_3(TG)_8$	260–364	29	0.860	0.399	AY048089
Ps163	**TET** CATTACCGAAACACGCACTT GTTTGCCAGGTCACGTTTAGGA *	$(GT)_4$	283–316	9	0.221	0.234	AY048091
Ps165	**HEX** CATGCGTAAATGTTGTAA CTAAACTAGGCGACGAAC	$(GT)_{12}$	180– > 250 †	36	0.878	0.472	AY048092
Ps262	**TET** TTTGGTGTGTGCAAATGAAA TGCGACTGGATGGGATT	$(TG)_{13}$	71–175	33	0.911	0.830	AY048096

(Ps76, Ps262, and Ps165) were particularly variable having 29, 33, and over 36 alleles, and mean expected heterozygosities of 0.86, 0.91, and 0.88, respectively (Table 1). Redesigning primers or altering PCR stringency did not improve amplification of the other four loci.

Exact tests for Hardy-Weinberg equilibrium were conducted for each new locus in each of the 17 sample sites with the program GENEPOP, version 3.1d (RAYMOND and ROUSSET 1995). Significant homozygote excess was observed at 16 sites for Ps76, at all sites for locus Ps165, and at seven sites for Ps262. With a conservative Bonferroni adjustment of the α-value to correct for total experimentwise error ($\alpha = 0.05/68 = 0.0007$), a total of 31 tests at loci Ps76, Ps165, and Ps262 were still significant. No significant deviations from Hardy-Weinberg equilibrium were observed for locus Ps163. However, this may reflect a lack of power to detect deviations from expected proportions at this locus because at all sites a single allele appeared at very high frequency (mean frequency = 0.88) and most individuals were homozygous for that allele.

Excess homozygosity may be caused by allelic drop-out, by the erroneous combination of distinct populations in a single sample (Wahlund effect), by inbreeding via assortative mating of related individuals, or by null alleles. Allelic drop-out is not a probable explanation for the observed homozygote excess at these loci. Any samples that were amplified several times yielded consistent results and decreasing PCR stringency did not reveal new alleles. For loci where a very high proportion of homozygous individuals were initially observed (e.g. Ps76), only new primers and not adjustment of assay conditions revealed new alleles. The Wahlund effect is also not a likely explanation given the very fine spatial scale of our sampling. Finally, a mating structure leading to inbreeding is unlikely given various aspects of the life history and behaviour of this univoltine butterfly. Males patrol in search of mates and most females are mated very soon upon eclosing. The offspring of a given female should be highly dispersed because females lay eggs individually and can travel tens of meters between successive ovipositions (FOWNES 1999), and overwintering, which occurs in the egg stage, must be associated with high mortality given that a female can lay hundreds of eggs (FOWNES 1999). Thus, there is no reason to suspect a disproportionate number of matings between related individuals. On the contrary, mating is likely to be close to random. The most probable cause of the excess homozygosity observed in these populations is therefore null alleles. Furthermore, we observed 10 apparent null homozygote individuals at locus Ps76, six at Ps163, and 66 at Ps165. For these individuals, no PCR product was obtained at that locus, even with repeated attempts, though genotypes were readily obtained at other loci. The occurrence of such individuals strongly suggests that these loci suffer from null alleles.

We have previously hypothesized that regions flanking microsatellite repeats in this species are highly variable (KEYGHOBADI et al. 1999). Observations of excess homozygosity and null homozygote individuals at these novel loci further support this hypothesis. Interestingly, significant homozygote excess has also been noted in all other reports of butterfly microsatellites (PALO et al. 1995; MEGLÉCZ and SOLIGNAC 1998; HARPER et al. 2000; ANTHONY et al. 2001). Thus, high variability in the regions flanking microsatellite repeats may be widespread in butterflies and may contribute to the difficulty of microsatellite marker development in this group.

ACKNOWLEDGEMENTS

We are grateful to Corey Davis, Isabelle Delisle and Renee Polziehn for technical advice. Sample collection was assisted by Norine Ambrose, Mark Caldwell, Sue Cotterill, Marian Forrester, Sherri Fownes, Fiona Johnson, Chris Schmidt, Lynnette Scott, Kirsty Ward, and Arliss Winship. Financial support was provided by scholarships from NSERC and the Killam foundation (NK), a Challenge Grant in Biodiversity from Alberta Dept. of Environmental Protection (NK & JR), and grants from Alberta Sports Recreation Parks and Wildlife Foundation (JR), NSERC (JR & CS), and Parks Canada (CS).

REFERENCES

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

Bogdanowicz SM, Mastro VC, Prasher DC and Harrison RG, (1997). Microsatellite DNA variation among Asian and North American gypsy moths (Lepidoptera: Lymantriidae). Ann. Ent. Soc. Am. 90: 768–775.

Boggs C, Watt W, and Ehrlich P, (in press). Ecology and evolution taking flight: butterflies as model study systems. University of Chicago Press, Chicago.

Brownstein MJ, Carpten JD and Smith JR, (1996). Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. BioTechniques 20: 1004–1010.

Engelke DR, Krikos A, Bruck ME and Ginsburg D, (1990). Purification of Thermus aquaticus DNA polymerase expressed in Escherichia coli. Analyt. Biochem. 191: 396–400.

Fownes SL, (1999). Meadow suitability and quality for the alpine butterfly, Parnassius smintheus, in the east slopes

of the Canadian Rockies (M.Sc Thesis). University of Alberta, Edmonton, Alberta.

Grist SA, Firgaira FA and Morley AA, (1993). Dinucleotide repeat polymorphisms isolated by the polymerase chain reaction. BioTechniques 15: 304–309.

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

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

Meglécz E and Solignac M, (1998). Microsatellite loci for Parnassius mnemosyne (Lepidoptera). Hereditas 128: 179–180.

Nei M and Roychoudhury AK, (1974). Sampling variances of heterozygosity and genetic distance. Genetics 76: 379–390.

Nève G and Meglécz E, (2000). Microsatellite frequencies in different taxa. Trends Ecol. Evol. 15: 376–377.

Palo J, Varvio S, Hanski I and Vainola R, (1995). Developing microsatellite markers for insect population structure: complex variation in a checkerspot butterfly. Hereditas 123: 295–300.

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

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