Molecular Ecology Resources (2009) **9**, 268–270 doi: 10.1111/j.1755-0998.2008.02414.x

PERMANENT GENETIC RESOURCES

Microsatellite markers for the Indian golden silkmoth, *Antheraea assama* (Saturniidae: Lepidoptera)

K. P. ARUNKUMAR, L. KIFAYATHULLAH and J. NAGARAJU
Centre of Excellence for Genetics and Genomics of Silkmoths, Laboratory of Molecular Genetics, Centre for DNA Fingerprinting and Diagnostics, ECIL Road, Nacharam, Hyderabad 500 076, India

Abstract

***Antheraea assama*, an economically important and scientifically unexplored Indian wild silkmoth, is unique among saturniid moths. For this species, a total of 87 microsatellite markers was derived from 35 000 expressed sequence tags and a microsatellite-enriched sub-genomic library. Forty individuals collected from Tura and West Garo Hills region of Northeast India were screened for each of these loci. Ten loci from expressed sequence tags and one from genomic library were found to be polymorphic. These microsatellite markers will be useful resources for population genetic studies of *A. assama* and other closely related species of saturniids. This is the first report on development of microsatellite markers for any saturniid species.**

Keywords: *Antheraea assama*, EST-SSRs, genetic variation, genomic-SSRs

Received 17 June 2008; revision accepted 25 August 2008

Indian golden silkmoth (*Antheraea assama*) popularly known as muga silkmoth, is a semi-domesticated, polyvoltine and polyphagous lepidopteran insect, endemic to northeastern India. Unlike other silkmoths, *A. assama* has low chromosome number (n = 15) (Deodikar *et al*. 1962) and ZZ/ZO sex chromosome system (Gupta & Narang 1981). *Antheraea assama* is known for its production of quality silk with natural golden colour, glossy fine texture and durability. Large-scale production of the golden yellow silk from *A. assama* is practiced by farmers of the Brahmaputra valley of Assam and its bordering states of India. The muga silk has now secured geographical indication status, the recognition under the intellectual property rights that it has its origin in Assam region. Over the past two decades, there have been reports that the natural populations of *A. assama* have experienced declines that are attributed to continued deforestation activities, population fragmentation and inbreeding. The decline in muga silkworm population has raised concern among silkworm farmers and conservationists. The aim of the present study was to develop informative microsatellite markers that could be used in analysing genetic structure and phylogenetic status of this species. With the knowledge of information and importance of microsatellites that reside in the exonic parts of the genome of various organisms, we have developed expressed sequence tag–simple sequence repeats (EST–SSR) for *A. assama*. In addition, we have also developed microsatellite markers from repeat enriched genomic library constructed from *A. assama* DNA.

To develop EST–SSR markers, 8200 unique sequences derived from 35 000 *A. assama* ESTs (Arunkumar *et al*. 2008) were screened using WebTROLL (Tandem Repeat Occurrence Locator) software (Castelo *et al*. 2002) for SSRs harbouring di-, tri-, tetra- and pentanucleotide repeats. A total of 63 sequences was selected from unique ESTs and of which 35 were di-, 12 tri-, 4 tetra- and 2 pentanucleotide repeat containing sequences. Ten ESTs harboured compound repeat motifs.

For the construction of genomic library, DNA was isolated using the method described earlier (Prasad & Nagaraju 2003) from *A. assama* pupae. Genomic DNA was enriched using an oligonucleotide mix of $(ATT)_8$, $(GAGT)_2$, $(CA)_{10}$, $(GA)_{10}$, $(GATA)_{10}$, $(CAC)_7$ and $(AGC)_7$ following previously reported protocol (Glenn & Schable 2005). In short, DNA

Correspondence: J. Nagaraju, Fax: 91-40-27155610;
E-mail: jnagaraju@cdfd.org.in

was digested with *Rsa*I and *Xmn*I restriction enzymes (New England Biolabs Inc.), ligated to double-stranded superSNX linkers, hybridized with biotinylated microsatellite oligonucleotides and captured on streptavidin-coated magnetic beads. Unhybridized DNA was washed away and the captured DNA was recovered by polymerase chain reaction (PCR) using the single stranded superSNX-F as a primer. The PCR products were ligated into the pCR4-TOPO TA vector (Invitrogen) and transformed into XL1-Blue competent cells. Colony PCR was performed to select the amplicons of size 500–1000 bp and purified using the AMPure PCR purification systems (Agencourt Bioscience). In total, 134 of the 185 colonies screened contained inserts of more than 500 bp. These inserts were amplified using M13 primers and sequenced on an ABI PRISM 3100 Genetic Analyser (Applied Biosystems), with BigDye Terminator (version 3.1) chemistry. Only 24 of the 134 inserts were harbouring microsatellite repeats.

Primer 3 program (Rozen & Skaletsky 2000) was used to design primers flanking SSRs in the EST clusters and microsatellite-enriched sequences. PCR was carried out on a Mastercycler Gradient (Eppendorf) in a 10-μL reaction containing 1× PCR buffer (10 mM Tris-HCl, pH 8.8, 50 mM KCl) (MBI Fermentas), 100 μM dNTPs, 1.5 mM $MgCl_2$, 5 pmol of each primers, 10 ng of genomic DNA as template and 0.5 U *Taq* polymerase (MBI Fermentas). The PCR conditions were initial denaturation of 94 °C for 3 min, 35 cycles of 94 °C denaturation for 20 s, appropriate annealing temperature (established empirically) for 10 s and 72 °C extension for 45 s, and 72 °C for 10 min as final extension. Of the 63 EST–SSRs and 24 genomic SSRs, 51 and 19 gave amplification, respectively, as assessed 1.5% agarose gel electrophoresis. These 70 primer pairs were then tested on 40 individual *A. assama* genomic DNA samples extracted from pupae collected from Tura (20 nos.) and West Garo Hills (20 nos.) region of Northeast India. The sampling locations are 60 km apart. The amplified products were then analysed on 3.0% agarose gel (Hi Resolution, Sigma-Aldrich) along with 50 bp size standard (MBI Fermentas) at 90 V for 3.5 h and visualized by ethidium bromide staining. The alleles were scored using the Bio-Rad Quantity One Software (Bio-Rad Laboratories). A sample of individuals was also screened on denaturing PAGE using the same primer sets to check for the allele sizes.

Only 10 of the 51 EST–SSRs were polymorphic upon screening 40 individuals. Twenty-eight loci were monomorphic with single band pattern. Thirteen primer pairs gave bands with size range more than expected and were monomorphic. The possible reasons for unexpected size range of EST–SSR loci are, one or both primers of the EST–SSR extend across a splice site and presence of introns and insertions/deletions (indels) in the corresponding genomic sequence (Varshney *et al*. 2005). Of the 19 genomic SSRs, only one locus was polymorphic with three alleles across 40 individuals. The software GENALEX version 6 (Peakall & Smouse 2006) was used to calculate allele frequency, observed mean and expected heterozygosity. The observed and expected heterozygosities ranged from 0 to 0.737 and from 0.095 to 0.825, respectively, with the allele numbers of two to eight. Pairwise tests for linkage disequilibrium were performed using Web-based program GenePop version 3.4 (Raymond & Rousset 1995). In Tura population, we observed a significant linkage disequilibrium for three of the 35 pairwise comparisons between loci after sequential Bonferroni correction, particularly involving loci AaSat002, AaSat020, AaSat044 and AaSat065. However, only one pair (AaSat001 and AaSat053) out of 27 combinations showed significant deviation from linkage disequilibrium in West Garo Hills population. The linkage disequilibrium observed at certain loci may be due to substructure of population or bottleneck. Further studies are needed to clarify this discrepancy. Only three of the nine loci polymorphic in Tura population showed no significant deviation from Hardy–Weinberg equilibrium (HWE). The remaining six deviated significantly from HWE and three of them (AaSat002, AaSat020 and AaSat053) showed large heterozygote deficiencies. Eight microsatellite markers were found to be polymorphic in West Garo Hills population, of which only two were in HWE (Table 1). The presence of null alleles at each locus was tested using Micro-Checker version 2.2.3 (van Oosterhout *et al*. 2004). The three loci AaSat002, AaSat020 and AaSat053 that deviated from HWE also exhibited overall significant excess of homozygotes with null allele frequency of 0.435, 0.366 and 0.198, respectively, possibly indicating the presence of null alleles in Tura population. In the West Garo Hills population, four loci AaSat008, AaSat14 AaSat044 and AaSat040 that also deviated from HWE showed homozygote excess with null allele frequency of 0.2401, 0.3714, 0.4527 and 0.4438, respectively. Furthermore, in the West Garo Hills population, only less than half of the samples was successfully genotyped with markers AaSat040 and AaSat044 suggesting a high frequency of null alleles in this population. Taken together, these results suggest that the deviation from HWE of many loci may be due to Wahlund's effect caused by subpopulation structure (Wahlund 1928), which has to be studied further. The loci which showed significant linkage disequilibrium and deviation from HWE in Tura region population did not exhibit same features in West Garo Hills population, indicating that the 11 microsatellite markers developed in the present study are independent and useful for population genetic studies.

The polymorphic EST and genomic SSRs developed in the present study will be particularly useful to study inbreeding effects, population structure and founder effects. As there are no microsatellite markers developed in other saturniid moths, these markers can also be tested for their cross-species amplification.

Table 1 Characteristics of 10 expressed sequence tag-derived and one genomic simple sequence repeat markers of *Antheraea assama*

Locus (GenBank)	Source	Forward/reverse primers (5′–3′)	Repeat motif	Size range (bp)	T_a	Tura population			West Garo Hills population		
						N/N_A	H_E	H_O	N/N_A	H_E	H_O
AaSat001 (EU597692)	EST	GTGTTCATTTCACGGAACATT CATTCGCTGTTCGTCTGAGAT	$(TA)_7$	182–200	53	20/2	0.455	0.700*	16/3	0.521	0.438*
AaSat002 (EU597693)	EST	TCTGGACAAATTGTAAAAGCTGTAG ACAAAACGAAAATCGCGTGT	$(GTCT)_5$	130–149	51	19/3	0.609	0.000*	19/6	0.770	0.632
AaSat008 (EU597695)	EST	CACGAAATGCCTCTGTCGTA GGTGTCTGTGGATGATGTGC	$(TA)_6$ Nn $(CA)_9$ Nn $(CA)_5$	199–249	51	Mono	—	—	19/4	0.432	0.211*
AaSat014 (EU597696)	EST	ATCTCTACCTACGCCGACGA AATTCGGCACGAGGAGTTC	$(GAT)_5$ $(TAA)_4$	260–264	48	Mono	—	—	19/2	0.388	0.000*
AaSat020 (EU597697)	EST	TTTCTTCGGTTCGTTTGGTT GACACGCGTTGCTTTGAGTA	$(TCGTG)_5$	164–226	53	20/2	0.375	0.000*	19/8	0.825	0.737
AaSat040 (EU597698)	EST	CGGACGTAACATTTGTCTGG CCACATGACTCTCATCAGCA	$(AT)_{17}$	133–221	60	16/2	0.430	0.625*	7/4	0.694	0.000*
AaSat044 (EU597699)	EST	CACCAGCTTCCAAAGAATTG CTAAAGCCCACGGGTTCATA	$(AT)_{20}$	194–226	51	17/3	0.657	0.588	9/3	0.642	0.000*
AaSat053 (EU597700)	EST	GAGTTCGGGTCGGACGTAAT TCTCTACCTACGCCGACGAC	$(ATT)_4$ N_7 $(ATC)_6$	204–225	50	20/2	0.095	0.000*	18/5	0.648	0.333*
AaSat059 (EU597701)	EST	CGAATAGCCGATTCCCTTTG TGCAAGCACGCACGTATC	$(TGCG)_{16}$	103–187	55	17/4	0.715	0.529*	Mono	—	—
AaSat065 (EU597702)	EST	GTCGAGCTGTCATAATTCCT AGTCTGACGTCGCTATAACC	$(AT)_{11}$ Nn $(AT)_{11}$ $(TA)_5$	100–168	54	18/2	0.375	0.500	Mono	—	—
AaGSat019 (EU597703)	Genomic	GATGGACTGGACCTCAATCG CCTGAGGAGAGAGGCGATG	$(TGA)_2$ $(AGA)_3$	151–172	55	18/3	0.248	0.278	Mono	—	—

T_a, annealing temperature; N, number of individuals scored; N_A, number of alleles; H_o, observed and H_E, expected heterozygosities; *significant deviation from Hardy–Weinberg equilibrium ($P < 0.001$). Data obtained upon screening 40 individuals (20 individuals each from Tura and West Garo Hills populations).

Acknowledgements

J.N. acknowledges Centre of Excellence grant from the Department of Biotechnology, Government of India. K.P.A. is recipient of the fellowship from Council of Scientific and Industrial Research (CSIR), India.

References

Arunkumar KP, Tomar A, Daimon T, Shimada T, Nagaraju J (2008) WildSilkbase: an EST database of wild silkmoths. *BMC Genomics*, **9**, 338.

Castelo AT, Martins W, Gao GR (2002) TROLL-tandem repeat occurrence locator. *Bioinformatics*, **18**, 634–636.

Deodikar GB, Chowdhury SN, Bhuyan BN, Kshirsagar KK (1962) Cytogenetic studies in Indian silkworms. *Current Science*, **31**, 247–248.

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**, 202–222.

Gupta ML, Narang RC (1981) Karyotype and meiotic mechanism in muga silkmoths, *Antheraea compta* Roth. and *A. assamensis* (Helf.) (Lepidoptera: Saturniidae). *Genetica*, **57**, 21–27.

van Oosterhout C, Hutchison WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Prasad MD, Nagaraju J (2003) A comparative phylogenetic analysis of full-length mariner elements isolated from the Indian tasar silkmoth, *Antheraea mylitta* (Lepidoptera: saturniidae). *Journal of Biosciences*, **28**, 443–453.

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

Rozen S, Skaletsky H (2000) Primer 3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, **23**, 48–55.

Wahlund S (1928) Zusammensetzung von Population und Korrelationserscheinung vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65–106.