Chapter 3

Characterisation of six microsatellite loci in the African Wild Silk Moth (*Gonometa postica*, Lasiocampidae)

"Have patience awhile; slanders are not long-lived. Truth is the child of time; erelong she shall appear to vindicate thee."

Immanuel Kant

Characterisation of six microsatellite loci in the African Wild Silk Moth (*Gonometa postica*, Lasiocampidae)

Wayne Delport[1], J. Willem H. Ferguson[2] & Paulette Bloomer[1]

[1]Molecular Ecology and Evolution Programme, Department of Genetics, University of Pretoria, Pretoria, 0002, South Africa

[2]Centre for Environmental Studies, Department of Zoology and Entomology, University of Pretoria, Pretoria, 0002, South Africa

Abstract

Six microsatellite markers were developed for the African Wild Silk Moth, *Gonometa postica* (Lasiocampidae) using an enrichment protocol. The total number of alleles ranged from 3 to 17 for a sample of 130 individuals across the distribution of the species. Observed levels of heterozygosity ranged from 0.17 to 0.78. Deviation from Hardy-Weinberg equilibrium detected in some loci is probably the result of large inter-annual population size fluctuations characteristic of this species. No evidence of linkage disequilibrium was detected among loci. These loci will be useful for the inference of demographic processes in a moth species that is of potential economic importance.

keywords: Lepidoptera, microsatellites, *Gonometa postica*, Lasiocampidae, FIASCO

Accepted for publication in *Molecular Ecology Notes*

The African Wild Silk Moth (*Gonometa postica*) is a species that could have considerable economic potential in rural southern Africa given that the larvae have been shown to yield a silk of high quality (Freddi *et al.* 1993). Indeed several small-scale cottage industries have been established using silk derived from these species. However, the potential for a rural-based silk industry utilising *G. postica* is largely dependent on annual population numbers. Given that *G. postica* shows large inter-annual fluctuations in population size the problem of insufficient supply of cocoons is evident. Therefore, population-based research is required such that an understanding of annual fluctuations in population size can be achieved. To this end we developed six microsatellite loci for the purpose of population genetic studies on *G. postica* in southern Africa.

Genomic DNA was extracted from wing muscle of one *G. postica* moth collected in the Northern Cape, South Africa. Thereafter a partial genomic library was constructed according to the FIASCO protocol (Zane *et al.* 2002). Approximately 50 ng of total genomic DNA was digested with *MseI* (New England Biolabs) and simultaneously ligated with T4 DNA Ligase (New England Biolabs) to *MseI* adaptors. The resulting fragments were PCR amplified with single-base 3' degenerate primers, and visualised with ethidium bromide staining under UV light on an agarose gel. Optimal PCR products were those that yielded a smear in the size range of 200-1000bp. The digested fragments were thereafter hybridised to 5' biotin labelled probes. Since it has proven difficult to isolate microsatellites from Lepidoptera (Meglecz *et al.* 2004) we used as many probes as feasible, yet combined non-complementary probes of the same motif length in single reactions. The following probes were used: $(gt)_{17}$, $(ca)_{15}$, $(ct)_8$, $(gc)_8$, $(tgc)_7$, $(cag)_5$, $(cca)_5$, $(cat)_5$, $(cac)_7$, $(ata)_8$, $(gtg)_5$, $(caa)_5$, $(aca)_5$, $(cga)_5$, $(cgca)_6$, $(tcca)_6$, $(tgtc)_6$, $(gata)_6$, $(tatc)_6$, $(gaaa)_6$, $(cagc)_6$. Following hybridization, enrichment for microsatellites was performed by the addition of streptavidin magnetic beads (Dynabeads) and separation of the probe-microsatellite-containing-fragments was achieved with three non-stringent and four stringent wash steps. Eluted DNA was thereafter PCR amplified and cloned into bacterial vectors using the TOPO10 chemically competent cloning kit for sequencing (Invitrogen). Recombinant plasmids were plated on ampicilin-supplemented agar plates and allowed to grow overnight at 37ºC. We picked positive clones and infected 1ml of LB medium, which was subsequently grown for 16 hours at 37ºC. Thereafter 1 µl of 1/10 dilution of infected medium was used as template in a 25 µl colony PCR amplification. Polymerase chain reactions were carried out in 25 µl reactions consisting of 1 x PCR buffer, 2.5 mM $MgCl_2$, 2 mM of each DNTP, 5 µM of each primer (T3: 5' ATTAACCCTCACTAAAGGGA 3', T7: 5' TAATACGACTCACTATAGGG 3') and 0.3 U of Supertherm Taq polymerase (Southern Cross Biotechnology). Reaction cycles were performed on a Hybaid multiblock (Thermofast) and consisted of denaturation at 94ºC for 2 minutes, followed by 35 cycles of denaturation at 94ºC for 30 seconds, primer annealing at 55°C for 30 seconds, and elongations at 72ºC for 30 seconds, and

finally 72ºC for 7 minutes. PCR products were subsequently sequenced using only the T7 (forward) primer with ABI PRISM® BigDye™ chemistry according to the manufacturer's instructions. The nucleotide sequences of probed fragments were subsequently resolved on an ABI 3100 Capillary Sequencer (Applied Biosystems). Sequences were scanned for microsatellite repeats using the Staden software package (Bonfield *et al.* 2002), and primers were designed for clones containing microsatellite repeats using the Primer Designer software package (Sci Ed Software, http://www.scied.com/ses_pd5.htm). Potential microsatellite loci were scored for polymorphism, across ten individuals sampled widely from the distribution of the species, on polyacrylamide gels stained with GelStar (Cambrex Bio Science Rockland, Inc.), and labelled primers were ordered for polymorphic loci. Thereafter fragment size analysis of 130 individuals genotyped for the polymorphic loci was performed on an ABI 3100 Capillary Sequencer, using Genescan 3.1 and the LIZ-500 size standard (Applied Biosystems). Representative alleles from each locus were subsequently sequenced from homozygotes and have been deposited with Genbank (http://www.ncbi.nih.nlm.gov) under the accession numbers DQ020593-DQ020611. Calculation of levels of heterozygosity, and tests for deviation from Hardy-Weinberg equilibrium and the non-random association of alleles among loci (linkage disequilibrium) were performed in Arlequin v2.001 (Schneider *et al.* 2000).

A total of 323 positive colonies were picked and scanned for microsatellite repeats. Of the 323 positive colonies, 22 contained microsatellite repeats of sufficient length and for which primers could be designed. Polymorphism tests revealed that seven of these 22 were polymorphic. Seven pairs of labelled primers were subsequently ordered, yet one locus (Table 1, Gon 26.3) proved difficult to amplify consistently and was removed from further analyses. Of the six remaining loci two showed high levels of polymorphism (Gon120.3 & Gon65, Table 1), whereas the remaining four had intermediate levels of polymorphism. Sequencing of alleles indicated substitutions to be stepwise (Figure 1), yet Gon65 may exhibit a complicated mutation model evident in the homoplasy observed for allele size 224. Clearly such a complex repeat may have high degrees of homoplasy and might not be useful for inferences of population history. Levels of heterozygosity are low in four of the six loci as might be expected for a population that experiences large inter-annual population size fluctuations. Furthermore, there was no evidence for linkage disequilibrium (Exact test, 10000 permutations, mean P across loci-pairs = 0.456), yet significant deviations from Hardy-Weinberg equilibrium are evident in two of the six loci (Table 1). Clearly, demographic processes would have a profound effect on the observed levels of genotypic diversity and heterozygosity. Meglecz *et al.* (2004) have attributed the low levels of heterozygosity observed in many Lepidopteran microsatellite loci to the presence of null alleles. Since microsatellites in Lepidoptera have been shown to be inconsistent in amplification across individuals, the flanking

Table 1: Microsatellite loci isolated and optimized for *Gonometa postica*. Amplification conditions for each locus is provided, as are the levels of allelic diversity, observed and expected heterozygosity (H_O and H_E respectively), and statistical significance (P) of deviation from Hardy-Weinberg equilibrium. *conc.* = concentration. The Bonferroni corrected alpha rejection level is 0.008.

Locus	*Motif*	*Primer sequence (5'-3')*	*Size Range*	*Dye*	*Annealing Temp (°C)*	*$MgCl_2$ conc. (mM)*	*No. of alleles*	*H_O*	*H_E*	*P*
Gon6	$(caaa)_n caat(caaa)$	F : AGCCCATGTTACTCGTGAAG R: GGGTGGAAAGCCAGTTATCT	161-173	TM PET	59	2	4	0.78	0.69	0.02974
Gon60	$(ay)_6$	F: CTGAAGAATAGCCAGCTAGG R: TGTGAATCGTGCCAGCAATG	213-217	TM VIC	59	1.5	3	0.17	0.41	**0.00001**
Gon65	$(ac)_6(gc)_4(ac)_n(ag)_2(ac)_n at(ac)_n$	F: ACGTTCGTATAGGTTGACAT R: GAGCTAATTGGTGCATTCAT	192-242	TM PET	59	1.75	17	0.49	0.89	**0.00001**
Gon107	$(gtct)_n gtrt$	F: GAGCAACAGAAGCCAAAG R: CACCTTTCTTCTATGGCC	157-177	TM 6-FAM	59	1.5	5	0.50	0.51	0.59564
Gon26.3	$(gaaa)_{32}$	F: TTCGCTGTGGAGAACGGAAG R: CTCGTCTGTTGTGATGAG	274	TM 6-FAM						
Gon 120.3	$(tc)_3 tg(tc)_n$	F: CCCCTAACCTAACTGATG R: CTCGTCTGTTGTGATGAG	87-115	TM VIC	45	2	11	0.66	0.69	0.00218
Gon55.3	$(gtct)_n$	F: GCCTTCACACATGCAGTA R: GCCTTCACACATGCAGTA	100-116	TM PET	47	2	4	0.27	0.27	1.00

Gon65 $(ac)_6(gc)_4(ac)_n(ag)_2(ac)_nat(ac)_n$

224 (41.4) GCACACGCACACAAGT**ACACACACACACGCGCGCGCACACACACACACACACAC--AGAGACACACACACACACAC--ATACACACACACACAC**----------GTGCGCGCACGCGTGTGTGCACC : 37
224 (66.2) GCACACGCACACAAGT**ACACACACACACGCGCGCGCACACACACACACACACACACAGAGACACACACACACAC----ATACACACACACACAC**----------GTGCGCGCACGTGTGTGTGCACC : 37
226 (10.11) GCACACGCACACAAGT**ACACACACACACGCGCGCGCACACACACACACACACACACAGAGACACACACACACAC--ATACACACACACACAC**----------GTGCGCGCACGCGTGTGTGCACC : 38
238 (23.2) GCACACGCACACAAGT**ACACACACACACGCGCGCGCACACACACACACACACAC--AGAGACACACACACACACACACATACACACACACACACACACACAC**GTGCGCACACGCGTGTGTGCACC : 44

Gon60 $(ay)_6$

215 (10.1) CAGATGTACACTAACTCGTATATTGATAGCGTACTATAGCTCCGTACAATGGGTAGCTACGCTGGTTCTACCTCA**ACACACATAC--ATAT**CTTAGTTAGATAACTCAAATCACAGAAACCACATTGCTGGCACGATTCACA : 7
217 (40.3) CAGATGTACACTAACTCGTATATTGATAGCGTACTATAGCTCCGTACAATGGGTAGCTACGCTGGTTCTACCTCA**ACACACACACACACACAT**CTTAGTTAGATAACTCAAATCACAGAAACCACATTGCTGGCACGATTCACA : 8
217 (41.4) CAGATGTACACTAACTCGTATATTGATAGCGTACTATAGCTCCGTACAATGGGTAGCTACGCTGGTTCTACCTCA**ACACACACACACACACAT**CTTAGTTAGATAACTCAAATCACAGAAACCACATTGCTGGCACGATTCACA : 8

Gon107 $(gtct)_ngtrt$

161 (17.4) AATTTTTTGAATTTTT**GTCTGTCTGTCTGTCT----GTGT**GTTTGTTACATCATCACGTGAAAATTACTAGATGGATTGAATGCAATTTTCACTGATGTATTGCTAAAGGT : 5
161 (22.1) AATTTTTTGAATTTTT**GTCTGTCTGTCTGTCT----GTAT**GTTTGTTACATCATCACGTGAAAATTACTAGATGTATTGAATGCAATTTTCACTGATGTATTGCTAAAGGT : 5
165 (18.4) AATTTTTTGAATTTTT**GTCTGTCTGTCTGTCTGTCTGTAT**GTTTGTTACATCATCACGTGGAAATTACTAGATGTATTGAATGCAATTTTCACTGTTGTATTGCTAAAGGT : 6

Gon55 $(gtct)_n$

108 (10.1) GCCTTCACACATGCAGTAAAAGCTAAAAACGTGGTTTTAGAAACTTTT**GTCTGTCTGTCTGTCT**GTTGTACGCGATTCGCACTAAAACTATTGGGCGTGTTACTCAG : 4
104 (59.2) GCCTTCACACATGCAGTAAAAGCTAAAAACGTGGTTTTAGAAACTTTT**GTCTGTCTGTCT----**GTTGTACGCGATTCGCACTAAAACTATTGGGCGTGTTACTCAG : 3

Gon6 $(caaa)_ncaat(caaa)$

161 (22.1) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAA----------CAATCAAA**TTTTTCCTCTTTATAATATTAGTATAGACATAG : 4
161 (72.8) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAA----------CAATCAAA**TTTTTCCTCTTTATAATATTAGTATAGACATAG : 4
169 (10.4) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAACAAACAAA----CAATCAAA**TTTTTCCTCTTTATAATATTAGTATAGACATAG : 6
169 (11.1) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAACAAACAAA----CAATCAAA**TTTTTCCTCTTTATAATATTAGTATAGACATAG : 6
173 (18.2) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAACAAACAAACAAACAATCAAA**ATTTTCCTCTTTATAATATTAGTATAGACATAG : 7
173 (58.1) GTTTTTGAGCCAAATCATTATATACCCA**CAAACAAACAAACAAACAAACAATCAAA**ATTTTCCTCTTTATAATATTAGTATAGACATAG : 7

Gon120 $(tc)_3tg(tc)_n$

97 (117.1) CAGAAGT**TCTCTCTGTCTCTCTCTCTCTCTCTCTC**--TATTTTTCCTCCTCACAACAGACGAG : 14
97 (41.3) CAGAAGT**TCTCTCTGTCTCTCTCTCTCTCTCTC**--TATTTTTCCTCCTCACAACAGACGAG : 14
99 (40.2) CAGAAGT**TCTCTCTGTCTCTCTCTCTCTCTCTCTC**TATTTTTCCTCCTCACAACAGACGAG : 15

Figure 1: Sequenced alleles for each of the microsatellite loci isolated. For each sequence the following are provided: size on GeneScan, (sampled individual) DNA sequence and : number of repeats.

regions may be highly variable. We, however, believe the observed levels of heterozygosity in *G. postica* microsatellites are the result of extreme fluctuations in population size, since PCR amplification of the loci optimized in this study was consistent across individuals, and flanking sequences in the alleles we sequenced did not show unusual levels of variation as found by Meglecz *et al.* (2004). We are currently investigating the effects of a complex demographic history on microsatellite variability, and the ability to infer population processes, in southern African Wild Silk Moth populations.

Acknowledgements

This work was funded by the Mellon Foundation Grant to J. W. H. Ferguson, P. Bloomer and W. Delport, and by a National Research Foundation grant to P. Bloomer (Gun: 2053653). The opinions and views presented in this article are however, not necessarily those of the National Research Foundation. We would also like to thank Dr. J. Nagaraju, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India for hosting WD and affording him the opportunity to work with *Bombyx mori* microsatellite loci.

References

Bonfield J, Beal K, Jordan M, Cheng Y & Staden R (2002) The Staden Package. Medical Research Council. Laboratory of Molecular Biology. http://www.mrc-lmb.cam.ac.uk

Freddi G, Bianchi Svilokos A, Ishikawa H, Tsukada M (1993) Chemical composition and physical properties of *Gonometa rufobrunnea* silk. *Journal of Applied Polymer Science,* **48**, 99-106.

Meglecz E, Petenian F, Danchin E, Coeur D'Acier A, 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.

Schneider S, Roessli D & Excoffier L (2000) ARLEQUIN version 2: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1-16.