Chapter V

Isolation of microsatellite markers from *Scarabaeus (Pachysoma)* MacLeay (Scarabaeidae: Scarabaeinae)

Abstract

In this section of the study the isolation of polymorphic microsatellite loci for the genus *Scarabaeus* was attempted using the FIASCO protocol. The FIASCO protocol is an enrichment protocol based on the ability to recover microsatellite DNA by PCR amplification, after selective hybridisation. High quality genomic DNA is fragmented using a restriction enzyme (*MseI*) and then ligated to a known adaptor (*MseI* AFLP). Following the fragmentation-ligation step the DNA is then hybridised to specific selected 5' biotinylated probes and bound to streptavidin coated beads. After the hybridisation step and several washes to remove DNA that has bound non-specifically, the DNA is eluted and recovered by PCR. Enriched DNA fragments are then cloned into a plasmid vector using a restriction site on the known flanking regions. The resultant recombinant clones are in turn screened for microsatellite repeats by directly sequencing them using primers specific for the vector. Following the identification of clones containing repeats, primers are designed for marker optimisation. This protocol was optimised for *Scarabaeus* and four out of six potential loci were identified as being polymorphic with one being monomorphic and the other exhibiting unstable amplification reactions.

Introduction

"Microsatellite DNA" is the term used to describe tandem repeats of short sequence motifs between two (di-nucleotide) and six (hexa-nucleotides) bases in length. These repeats are arranged head to tail without interruption by any other base or motif and their functional significance is unknown. Microsatellites have been found in every organism studied so far and since they may be highly polymorphic are useful genetic markers. Adenine (A) and thymine (T) di-nucleotide repeats are the most common microsatellites in all genomes; however, they do show subtle differences in frequency distributions. Rates of mutation in microsatellites are high compared to rates of point mutations in non-repetitive DNA regions. High mutation rates in microsatellites are said to be due to either slipped strand mis-pairing during replication, which is thought to be the predominant mechanism, or recombination between DNA molecules (Hancock, 1999; Toth *et al.*, 2000).

Slippage/slipped strand mis-pairing occurs during replication when the nascent DNA strand dissociates from the template strand. When non-repetitive sequences are being replicated there is only one way in which the nascent strand can re-anneal precisely to the template strand before replication is recommenced. If the replicated sequence is repetitive the nascent strand may re-anneal out of phase with the template strand. When replication is continued after such a mis-annealing, the eventual nascent strand will be either shorter or longer than the template strand (Hancock, 1999).

Recombination could potentially alter the length of microsatellites in two ways, by unequal crossing-over or by gene conversion. Unequal crossing over results from crossover between misaligned chromosome strands. Unequal crossover gives rise to a deletion in one molecule and an insertion in the other. Gene conversion, which involves unidirectional transfer of information by recombination, probably as a response to DNA damage, can transfer sequence in an out of phase manner from one allele to another (Hancock, 1999).

The genetic architecture of a species can be interpreted as the result of historical biogeographic factors as well as contemporary ecologies and behaviours of the organism under investigation (Avise *et al.*, 1987; Avise, 2000). Phylogeography is the study of the principles and processes governing the geographical distributions of genealogical lineages, especially those within and amongst closely related species. Phylogeography therefore deals with historical and phylogenetic components of the spatial distributions of gene lineages. Time and space are joint axes of phylogeography (Avise, 2000). The majority of phylogeographic studies so far have employed mitochondrial DNA as the marker of choice, but recent developments in the field recommend the synchronous usage of nuclear-based

microsatellite markers (Avise, 1998) in order to check for congruence across several unlinked loci with different genealogical pathways.

Different protocols exist for microsatellite isolation. Traditionally microsatellites were isolated from partial genomic libraries of the species of interest which involved screening several thousand clones through colony hybridisation with repeat containing probes (Rassman *et al*., 1991). Although a simple method it can be tedious and very inefficient for species with low microsatellite frequencies. While many remain faithful to the traditional means of isolating microsatellites several other methods are being used with increased frequency in an attempt to decrease the time invested in the isolation process while increasing the yield of microsatellites. To avoid library construction and screening it was proposed to modify the randomly amplified polymorphic DNA (RAPD: Williams *et al*., 1990) approach by using either repeat-anchored random primers (Wu *et al*., 1994) or RAPD primers and subsequent Southern hybridisation of the PCR bands with microsatellite probes (Richardson *et al*., 1995). A second strategy, based on primer extension was also proposed for the production of libraries rich in microsatellite repeats. This method involves a high number of steps explaining the limited application and use by the scientific community (Ostrander *et al*., 1992). A third class of isolation was based on selective hybridisation, which appeared to be a very popular method for microsatellite isolation (Kijas *et al*., 1994). If the enrichment was successful sequencing recombinant clones alone could then identify microsatellites. Time is required to get the enrichment protocols running efficiently, but they are advantageous in that they are fast, efficient, require only basic skills in molecular biology and limited laboratory equipment as compared to that for traditional methods of microsatellite screening (for added details of these procedures see Zane *et al*., 2002 and references therein). With this in mind, and the fact that the selective hybridisation procedure had been used successfully in the Genetics Department at the University of Pretoria, it was decided to use this protocol, to isolate microsatellites, in preference to traditional or any other methods.

This part of the study was aimed at isolating microsatellites for the genus *Scarabaeus*. To examine the genetic variation between the populations of the different species both mitochondrial DNA and microsatellites need to be employed. The mitochondrial COI gene reveals relatively older genetic structuring whereas the microsatellites are thought to reveal more recent dynamics. The employment and use of DNA microsatellites is therefore expected to reveal patterns of variation that would be undetectable by other molecular markers (Kirchman *et al*., 2000). The primary goal of the molecular comparisons is to characterize genetic diversity within and between populations of the same species.

Methodology and Results

Taxonomic Samples

A single species, *Scarabaeus (Pachysoma) gariepinus*, was chosen from which to isolate microsatellites, as it is found on both sides of the Orange River and it represented populations that were more comprehensively sampled than the other species. Individuals were collected from five designated populations (see Chapter 4; Table 1), two populations in South Africa, Langhoogte/Kommagas and Holgat River, and three in Namibia, Hohenfels, Daberas/Obib and Klingharts Mountains. A minimum of 10 individuals per population, where possible, were collected otherwise as many individuals as possible were collected and stored in 99.8 % ethanol.

Isolation of Microsatellites

Isolation of microsatellites was performed following Zane *et al*., (2002), the Fast Isolation by AFLP of Sequences Containing Repeats (FIASCO protocol). This method is fast and simple and relies on the efficient digestion-ligation reaction of the amplified fragment length polymorphism (AFLP) procedure as described by Vos *et al*. (1995).

Total genomic DNA was extracted using the Roche High Pure PCR Template Preparation Kit (Roche Diagnostics). The extracted genomic DNA was digested with *MseI* and simultaneously ligated to *MseI* AFLP adaptor (5' TAC TCA GGA CTC AT 3'/5' GAC GAT GAG TCC TGA G 3'). Digestion-ligation was performed on a Hybaid Multiblock for 3 hours at 37°C in a total volume of 25µl containing 25 – 250 ng of DNA, 10 x NEB 2 buffer, 25 mM DTT, 10 x BSA, 5 mM ATP, 10 U/µl *MseI*, 2000 U/µl NEB T4 DNA ligase and 50 µM *MseI* adaptor.

The digestion-ligation mixture was subsequently diluted 10 fold- with SABAX water and amplified in a 20 µl reaction with adaptor specific primers (5' CAT GAG TCC TGA GTA AN 3') henceforth referred to as *MseI*-N where 'N' equals A, C, G and T. The 20 µl Polymerase Chain Reaction (PCR) contained 10 pmol *MseI*-N primer, 1.5 mM $MgCL_2$, 1 X *Taq* DNA polymerase buffer, 10 mM dNTP's (200 µM) in the presence of 0.4 units *Taq* DNA polymerase. Following optimisation, final thermal cycling parameters for *S. (P.) gariepinus* comprised an initial denaturation for 2 minutes at 94°C followed by 22 cycles of 94°C for 30 seconds, 53°C for 60 seconds and 72°C for 60 seconds with a final elongation step at 72°C for 7 minutes (Vos *et al*., 1995). A small amount of the PCR product (3 µl) was electrophoresed on a 1.5 % agarose gel (Fig. 1). Clear DNA smears (lanes 1 & 2) indicated that *MseI* cut the genomic DNA in fragments ranging from 250 to 1 200 base pairs (bp) in

length. Figure 1 also shows an over-representation of fragments as seen by the dark distinct band approximately 500 bp in size. This over-represented band is not ideal, as it is believed to represent multi-copy sequences in the original genome.

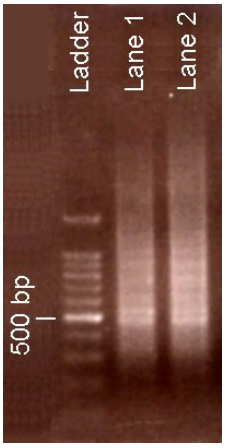

Figure 1. Electrophoresis of the PCR products amplified with the *MseI*-N primers following the digestion-ligation step. Lanes 1 & 2 contain 3 µl of PCR product. Product sizes range from 250 – 1200 bp in length.

In the AFLP protocol the *MseI*-N primer has a selective nucleotide at the 3' end, which matches the first nucleotide beyond the original restriction site, allowing for pre-selective amplification. For this protocol all four primers are mixed (*N* = A, C, G, T) allowing for the amplification of all fragments flanked by *MseI* sites, providing they have an appropriate size for PCR. The advantage of this is that if undesired bands appear within the PCR, one can go back one step and try different combinations of the selective primers, i.e. remove one of the bases 'A, C, G, T from the primer mix, and re-amplify, and in doing so preclude amplification of the unwanted bands. Over representation of bands which probably correspond to multi-copy sequences in the original genome are not favourable as they tend to be carried over during enrichment and when cross-hybridised with the biotinylated probes will represent a significant fraction of the obtained recombinant clones (Zane *et al*., 2002). A single base (either A, C, G or T) from the *MseI*-N primer was systematically removed and four different amplification reactions were preformed, the results of which can be seen in Figure 2. The reaction lacking an over represented band can be seen in reaction 4 - lanes 3 &

4 - (Fig. 2), which lacked the base 'C'. All subsequent PCR's using the *MseI*-N primer, were therefore performed without the base 'C'.

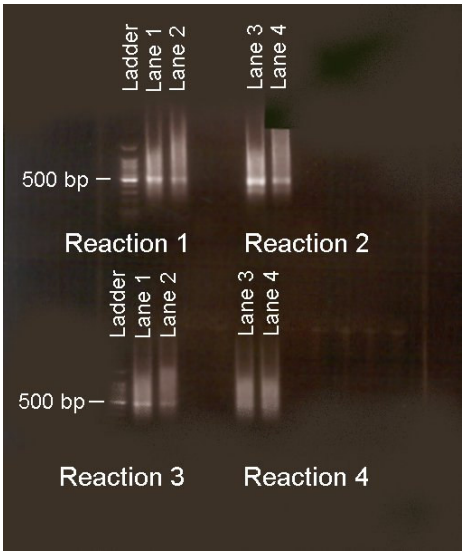

Figure 2. The four experimental amplification reactions (labelled from the top left) after the removal of a single base A (reaction 1), T (reaction 2), G (reaction 3) and C (reaction 4).

The DNA from the PCR was then used as template for hybridisation to selected biotinylated probes using the Travis Glen Protocol (http://129.252.89.20/Msats/Microsatellites.html). The DNA-probe hybrid molecules were prepared in the following way: 250 – 500 ng of the amplified DNA, 50 – 80 pmol of the selected biotinylated probe made up to a total volume of 100 µl with 4.2 X SSC, 0.7 % SDS. Probes were denatured at 95°C for 3 minutes followed by annealing at room temperature for 15 minutes. Three hundred µl of TEN_{100} were added to the prepared DNA-probe hybrid molecules.

Table 1 shows the probes used in different combinations to isolate microsatellites. Different probes can be combined in a single hybridisation reaction only if the designated probes have the same length i.e. only di-nucleotide or only tri-nucleotide probes can be combined in a single reaction.

Table 1. Biotinylated probes used to isolate microsatellites.

Probes	Length
$(ca)_{15}$	30mer
$(ct)_{8}$	16mer
$(gc)_{8}$	16mer
$(tg)_{15}$	30mer
$(ata)_{8}$	24mer
(gtg)5	15mer
$(caa)_{5}$	15mer
$(aca)_{5}$	15mer
$(cga)_{5}$	15mer
$(cca)_{5}$	15mer
$(cat)_{5}$	15mer
$(cac)_{5}$	15mer
$(ccagt)_{10}$	40mer
$(gaaa)_{6}$	24mer
$(tgtc)_{6}$	24mer
$(tatc)_{6}$	24mer
$(gata)_{6}$	24mer
$(cagc)_{6}$	24mer
$(tcca)_{6}$	24mer

The DNA molecules attached to the biotinylated probes were captured using Streptavidin coated beads (Streptavidin Magnetic Particles, Boehringer Mannheim) (Kandpal *et al*., 1994;Kijas *et al*., 1994; Mcrae *et al*., 2005). Firstly, the beads were prepared by washing 1 mg of beads three times in 500 µl TEN_{100} (10mM Tris-HCL, 1 mM EDTA, 100 mM NaCL at pH 7.5) after which they were re-suspended in 40 µl of $TEN_{100.}$ Ten µl of an unrelated PCR product were added to the prepared beads so as to prevent non-specific binding. The DNA-probe hybrid molecules were then added to the prepared beads and allowed to incubate for 30 minutes at room temperature with constant gentle agitation.

The non-specific DNA was removed by 7 non-stringency and 7 stringency washes. Non-stringency washes were performed by adding 400 µl of TEN_{1000} (10 mM Tris-HCL, 1 mM EDTA, 1M NaCL, at pH 7.5). The stringency wash was performed by adding 400 µl 0.2 SSC, 0.1 % SDS to the DNA. All washes were carried out at room temperature for 5 minutes, recovering the DNA with magnetic field separation, except for the last stringency wash which

was left at 40°C for 5 minutes. The last non-stringency and stringency washes were kept for further amplification.

The DNA was then separated from the beads-probe complex by means of two denaturation steps. The first step involved adding 50 µl of TE (10 mM Tris-HCL, 1 mM EDTA at pH 8) prior to incubation at 95°C for 5 minutes, after which the supernatant was removed and stored. The second step involved treating the beads with 12 µl of 0.15 M NaOH, removing the supernatant and neutralizing it with 12µl of 0.1667 M acetic acid. TE was then added to a final volume of 50 µl. The last non-stringency, stringency and two elutions from the denaturation step should contain an increasing amount of DNA fragments containing the repeat and should carry the *MseI*-N primer target at each end.

The DNA was then precipitated from the washing and denaturation steps by adding 1 volume of ethanol and sodium acetate (0.15 M final concentration) and then re-suspending in 50µl of water. The precipitated DNA was amplified in a 50 µl reaction containing 10 pmol *MseI*-N primer, 2 mM $MgCL_2$, 1 X *Taq* DNA polymerase buffer, 10mM dNTP's (200 µM) in the presence of 0.4 units *Taq* DNA polymerase. Thermal cycling parameters comprised an initial denaturation for 2 minutes at 94°C followed by 35 cycles of 94°C for 30 seconds, 53°C for 60 seconds and 72°C for 60 seconds with a final elongation step at 72°C for 7 minutes.

The PCR products from the washing and denaturation steps yielded a product with smears above 200 bp with the last stringency wash yielding an order of magnitude less product showing the distinct removal of non-specifically bound DNA. Figure 3 clearly indicates that the more non-stringency and stringency washes done, the more non-specifically bound DNA is removed. Figure 3 shows the effect of eight washes with lane 8 having the least amount of DNA.

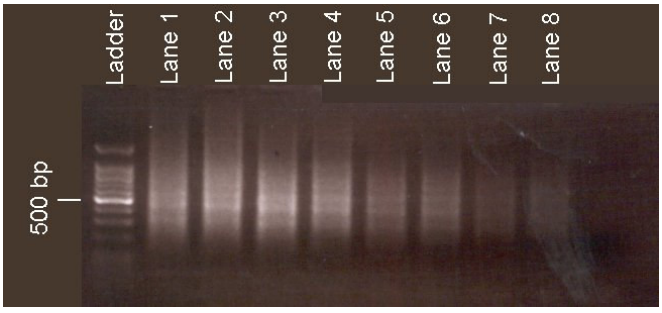

Figure 3. Amplification products of eight stringent wash steps, loaded in lanes 1-8. Lane 8 represents the last stringency wash clearly indicating that the more washes done the more non-specifically bound DNA was removed.

Cloning and optimisation

Cloning of the PCR amplicons was carried out using the TOPO TA Cloning Kit for Sequencing (Invitrogen). The PCR products from the two-denaturation steps were used as they were considered the best candidates for producing a highly enriched library.

Pre-mixed agar media containing Ampicillin (Invitrogen) was used to prepare the agar plates on which to grow up colonies. Pouring of the plates was done in a laminar flow cabinet after which the plates were allowed to cool for 30 minutes. Colonies were plated out using 50 µl and 100 µl from each transformation and incubated upside down at 37°C for 12 – 16 hours. Colonies were then cultured for 16 – 24 hours in 1000 µl of LB Broth (Invitrogen Corporation), containing 100 µg/ml of Ampicillin at 37°C, with constant agitation. Once the colonies had been cultured 10 µl of the colony was added to 10 µl of SABAX water and denatured at 96°C for 7 minutes. The balance (850 µl) of the cultured colony was added to 150 µl of glycerol, creating a 15 % glycerol solution for long-term storage at -80°C. One µl of the denatured colony was amplified in a 50 µl reaction containing the following 10 pmol of each of the T3 and T7 universal primers, 2 mM $MgCL_2$, 1 X *Taq* DNA polymerase buffer, 10 mM dNTP's (200 µM) in the presence of 0.4 units *Taq* DNA polymerase. Thermal cycling parameters were the same as mentioned above. The colony PCR products when electrophoresed on a 1.5 % agarose gel showed PCR products of different sizes (Fig. 4).

Different sized PCR products confirmed that vectors had different sized cloned products incorporated into them. A total of 260 colonies were picked all of which were sequenced. Sequencing reactions were performed at an annealing temperature of 53°C with version 3.1 of the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer). Each amplicon was sequenced with the T3 primer (Invitrogen) to screen for microsatellite (Msat) repeats. Primer pairs flanking the microsatellite repeats were then designed using Primer Designer 4 (Scientific and Educational Software). Of the 260 sequences a total of 15 possible repeat sequences were identified. These were narrowed down to six possibilities due to either poor repeats or the inability to design primers suitable for amplifying the repeat region (see Table 2 for microsatellite repeats).

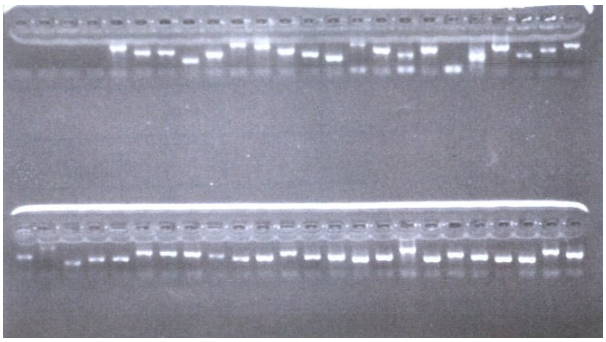

Figure 4. Electrophoresis of the size-discrete PCR products obtained from the colonies grown in LB broth containing potential Msat repeats.

To determine the optimal PCR conditions at which the primer pairs amplify the repeat DNA region, a set of 'cold tests' were performed under varying annealing temperatures, Mg-concentration and DNA-extract concentrations on a Hybaid Multiblock. Amplifications were performed in 7 µl final volume containing approximately 30-50 ng of DNA, 1-2.5 mM $MgCL_2$, 0.2 mM of each dNTP, 0.35 units of Expand High Fidelity *Taq* (Roche), 1 X buffer and 4 pmol of each primer. For each primer pair an optimal annealing temperature was obtained and cycling conditions were optimised as comprising an initial denaturation for 3 minutes at 94°C followed by 35 cycles of 94°C for 30 seconds, variable primer annealing temperatures (ranging from 44.8ºC to 53.4ºC – see Table 2) for 60 seconds and 72°C for 30 seconds with a final elongation step at 72°C for 45 minutes. The full 7 µl of PCR product was electrophoresed on an 8 % polyacrylamide gel, at room temperature overnight, against a 100 bp ladder (Promega), to test for polymorphism.

Table 2. Potential microsatellites and primers designed to determine whether the loci were polymorphic or not.

Probes	MSAT repeats	Ta	Clone	Primers					Polymorphic
				Name	Sequence	Tm	Pos	L	
1. (ca)15(ct)8	(tg)24	53.4°C	27	tg36_f	atg agt ggg tgt gtg tcg tg	61°C	36	20	Unreliable
				tg257_r	tct tcc ttg gtc ttt att ttg g	57°C	257	22	Unreliable
2. (ata)8(gtg)5									
	(cta)12	51.3°C	31	ap42_f	ggt cac gct tta gga cta ga	60°C	83	20	√
3. (caa)5(aca)5(cga)5				ap42_r	ggt tga taa ggt aga tgc cc	59°C	340	20	√
	(tg)8	44.8°C	23	ap35_f1	gcc tct tcg agt att gtg	56°C	277	18	√
				ap35_r1	cgt taa caa gga gct gca	60°C	399	18	√
	(ca)5 cg (ca)7	62°C	13	ap25_f	cgt gaa tcg acg acg tga aa	62°C	116	20	No
				ap25_r	gtg tat gta tgt gcg ggt gt	62°C	181	20	No
	(ca)3 ta (ca)6	46.8°C	15	ap27_f	cgt tat cac gcg ctc gca ca	68°C	254	20	√
				ap27_r	ccg tat ggt gcc gct tcc tt	67°C	334	20	√
	(gggtt)3 gtgatgtgtt (gggtt)2	50°C	6	ap19_f	cgt cag aga ggg tat gta ac	58°C	29	20	√
				ap19_r	cct atc ttg tag aca ggt gc	59°C	337	20	√

[£] Ticks in the polymorphic column indicate that the locus was polymorphic

Discussion

Microsatellites are found everywhere in both prokaryote and eukaryote genomes and are present within coding and non-coding regions. Their distribution, however, is not homogenous within a single genome due to different constraints on coding vs. non-coding DNA (Toth *et al*., 2000), historical processes (Wilder & Hollocher, 2001) and possible different functional roles of the repeats (Valle, 1993). The frequency of microsatellites also varies across taxa, in terms of both absolute numbers of the loci and repeats (Hancock, 1999). As microsatellites have a high mutability they are thought to play an important role in genome evolution by creating and maintaining quantitative genetic variation (Toth *et al*., 2000). The aim of this part of the study was to create a microsatellite library for *Scarabaeus.* However, due to time constraints and technical problems encountered this goal was not satisfactorily achieved.

Technical problems resulting in low yield of polymorphic loci were two-fold. Firstly, problems were experienced with cloning the DNA fragments into the vector cells. This can be seen by the small number of colonies that were available to select for growth (namely 260 across 10 different attempts). This contrasts markedly with the more than 400 colonies obtained by P. Bloomer after three attempts with Avian DNA under the same laboratory conditions and using the same reagents (pers. comm.). Probes based on the microsatellites isolated from other families of Coleoptera were taken into account and used first when attempting to increase the cloning success rate but this did not seem to solve the problem. Different agar was tried in case the vector cells were sensitive to certain agar media. The TOPO cloning manual suggested leaving the cloning reaction for different lengths of time to allow for maximising the PCR products. Different combinations of these times were tried but this did not improve the cloning reaction. The competent cells from the same kit were tested using bird and mammal DNA to see whether the competent cells were of a poor quality. However, both these sets of DNA produced a magnitude more clones on the agar as opposed to beetle DNA. After trying different combinations of times, probes, agar and testing the competent cells for competency it is concluded that *Scarabaeus* may be comparatively poor in microsatellite repeats.

According to Zane *et al*. (2002) arthropod DNA does not appear to be microsatellite rich and the general trend appears to be that the success rate for isolating positive repeats is approximately 2 %. This is exceedingly low, indicating that the success rate achieved in this study was not unrealistic particularly as the six good repeats obtained out of a total of 260 clones sequenced, works out at a 2.3 % success rate. Furthermore, approximately 50 % of

positive clones (i.e. those containing repeat motifs) are eventually discarded due to a lack of suitable sequence for primer design, or the absence of the repeat in the amplification product, or due to unreliable amplification (Zane *et al*., 2002). It has been confirmed that four of the six loci were polymorphic, a single one ($(tg)_{24}$ repeat) exhibited unreliable amplification while the last one was monomorphic. An additional consideration is that the expected frequency of tri- and tetra- nucleotide repeats is below 1 % of any clone analysed across all taxa (Zane *et al*., 2002).

The second problem we encountered was with the amplification step of a single locus, the $(tg)_{24}$ repeat, once the primers had been designed. In the 'cold tests' primer-specific annealing temperatures, DNA and $MgCl_2$ concentrations were identified, which gave the best amplification of genomic DNA. In some instances, however, the PCR would give double bands in one PCR and not in another using the same DNA and reagents, and under the same cycling parameters. If one knows the length of the fragment amplified then should any double bands occur that differ significantly in size, it could still in theory be possible to identify and select the correct band containing the repeat, based on size. However, in most cases the bands were too close together to permit adequate separation, hence it was not possible to ascertain whether the repeat was polymorphic or not. The reason for temperamental PCR's was unknown. Different types of *Taq* DNA polymerase (i.e. High Fidelity, Supertherm, Supertherm Gold and Biotools) were evaluated, between-thermal cycler variation was avoided by using the same PCR machine for optimisation as for amplification, new reagents were tried for PCR reactions and different individuals of the species were amplified but the problem still persisted. As a last resort new primers were designed for this locus but this did not seem to solve the problem.

In many cases obtaining 'well behaved' microsatellites requires considerable time and effort and even then some microsatellites may still have null alleles or single primer pairs that amplify more than a single locus (Meglécz *et al*., 2004). Difficulties arise during isolation and characterisation of microsatellites leading to few well-resolved loci (Nève & Meglécz, 2000). Problems appear during the design of primers and setting up PCR conditions. Reasons for the presence of null alleles and varying amplification intensities between individuals has been largely speculated upon but suggestions are that the flanking regions of microsatellites may be variable (Meglécz *et al*., 2004). A frequently observed problem is the amplification of more than two bands with a single primer pair (Meglécz *et al*., 2004). Two possible reasons primarily given are (i) the duplication or multiplication of microsatellite containing regions or (ii) that microsatellites lie within a minisatellite repeat unit and have microsatellite length

variations between the minisatellite repeat units. These problems appear to be common within the order Lepidoptera (Meglécz *et al*., 2004), but we have experienced them within *Scarabaeus* i.e. order Coleoptera, indicating that they may be more common across unrelated taxa. However, as failed attempts at microsatellite isolation are generally not published, the extent of the problem can only be speculated upon.

One of the major drawbacks of microsatellites is that they need to be isolated *de novo* from most species being examined for the first time. Most microsatellites are found in non-coding regions were the substitution rate is higher than in coding regions (Hancock, 1995). To design 'universal primers' matching conserved regions is therefore often problematic (Zane *et al*., 2002). Different taxa exhibit different preferences for microsatellite repeat types (Lagercrantz *et al*., 1993) hence attempting amplification across the generic, familial or order level in many instances proves fruitless.

The task of microsatellite isolation involves a large amount of effort in the time and money invested in isolation, compared to the results obtained. One has to screen genomic libraries with many appropriate probes, optimise amplification reactions at numerous steps, design primers and eventually gene scan individuals from the respective species. The number of positive clones (those containing microsatellite repeats) ranges from 12 % to as low as 0.04 % (Zane *et al*., 2002). Such isolation strategies will therefore only be successful in a limited time in taxa with high numbers of repeats e.g. fish or if a low number of microsatellite loci are needed.

Microsatellites are inherently unstable and undergo constant mutation. The abundance of certain repeat types varies with the genomic region and their distribution is often dependent on the taxonomic group examined (Hancock, 1996; Toth *et al*., 2000). Mean density for microsatellites within a species varies widely for reasons unknown, and therefore resulting in no *a priori* rule that can be forged for their predictability (Jarne & Lagoda, 1996). Moreover, overall microsatellite content within a genome is often correlated to genome size of the organism (Hancock, 1996). After taking all the above points into account the 18 months of laboratory work required to obtain four polymorphic loci exemplifies the amount of the labour involved as well as the poor success rate in the isolation process. The number of loci scored, degree of polymorphism of each locus and sample size are of paramount importance for the statistical power of microsatellites to be effective (Zane *et al*., 2002). With this in mind the process of optimising the loci will continue until such time as a minimum of five or more good polymorphic loci are obtained with which to work.

Microsatellite markers are excellent for population structure studies as they are highly variable, more likely to be neutral than other genetic markers and the results are reproducible (Jarne & Lagoda, 1996). These advantages tend to outweigh the long and expensive isolation process and establishment of appropriate amplification conditions. Even though the problems experienced, such as low microsatellite frequency and frequent PCR failure, do not appear to be unique to Coleoptera i.e. they are seen in at least one other order of insects, Lepidoptera, (Meglécz & Solignac, 1998), we are positive that with perseverance, repetition and patience, successful isolation of microsatellite loci will be obtained that will provide the desired statistical power to answer the original questions posed.

Acknowledgements

CS would like to thank Wayne Delport and Carel Oosthuizen from the MEEP laboratory (Genetics Department, University of Pretoria) for their patience and advice over the past 18 months. Paulette Bloomer is thanked for making the MEEP lab available to CS for laboratory work. The NRF and University of Pretoria are thanked for partial funding of this project.

References

Avise JC, Arnold J, Martin Ball R, Birmingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific Phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecological Systematics*, **18**, 489-522.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371-379.

Avise JC (2000) *Phylogeography: The history and formation of species.* Cambridge, Massachusetts, London, England : Harvard University Press.

Hancock JM (1995) The contribution of slippage-like processes to genome evolution. *Journal of Molecular Evolution*, **41**, 1038-1047.

Hancock JM (1996). Simple sequences and the expanding genome. *Bioassays,* **18**, 12467-12472.

Hancock JM (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites: Evolution and Application* (eds. Goldstein DB, Schlötterer C), Chp 1. Oxford University Press.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424-429.

Kandpal RP, Kandpal G, Weissman, SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Science if the United States*, **91**, 88-92.

Kijas JM, Fowler JC, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques*, **16**, 656-662.

Kirchman JJ, Whittingham LA, Sheldon FH (2000) Relationships among cave swallow populations (*Petrochelidon fulva*) determined by comparisons of microsatellite and cytochrome *b* data. *Molecular Phylogenetics and Evolution*, **14**, 107-121.

Lagercrantz U, Ellergren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acid Research*, **21**, 1111-1115.

Mcrae SB, Emlen ST, Rubenstein DR, Bogdanowicz SM (2005) Polymorphic microsatellite loci in a plural breeder, the grey-capped social weaver (Pseudonigrita arnaudi), isolated with an improved enrichment protocol using fragment size selection. *Molecular Ecology Notes*, online.

Meglécz 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.

Meglécz E, Solignac M (1998) Microsatellite loci for *Parnassius mnemosyne* (Lepidoptera). *Hereditas*, **128**, 170-180.

Nève G, Meglécz E (2000) Microsatellite frequencies in different taxa. *Trends in Ecology and Evolution*, **15**, 376-378.

Ostrander EA, Jong PM, Rine J, Duyk G (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proceedings of the National Academy of Sciences of the USA*, **89**, 3419-3423.

Rassmann K, Schlötterer C, Tautz D (1991) Isolation of simple-sequence loci for use on polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, **12**, 113-118.

Richardson T, Cato S, Ramser J, Kahl G, Weising K (1995) Hybridization of microsatellites to RAPD: a new source of polymorphic markers. *Nucleic Acid Research*, **23**, 549-557.

Tóth G, Zoltán G, Jurka J (2000) Microsatellites in different Eukaryotic genomes: survey and analysis. *Genome Research*, **10**, 967-981.

Valle G (1993) TA-repeat microsatellites are closely associated with ARS consensus sequences in yeast chromosome III. *Yeast*, **9**, 753-759.

Vos P, Hogers R, Bleeker M, *et al*. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

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

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18**, 6531-6535.

Wu K, Jones R, Danneberger L, Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research*, **22**, 3257-3258

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