Chapter 2

Development of microsatellite markers for *Spirocerca lupi*

2.1. ABSTRACT

The parasitic nematode, *Spirocerca lupi*, is known to cause spirocercosis in canids. It occurs mostly in regions with a warm, tropical climate and has emerged as a potentially life-threating parasitosis in canine populations worldwide. Early detection is not possible due to diagnostic limitations and very few molecular studies have been conducted on *S. lupi* thus far. Microsatellites are ideal genetic markers for many areas of research since they are co-dominant, highly polymorphic, selectively neutral and show biparental inheritance. In this study, a set of microsatellite markers was developed for *S. lupi* to conduct fine-scale genetic analysis in order to gain an understanding of parasite dispersal. Microsatellite-enriched DNA sequences were isolated from adult worms using the FIASCO (fast isolation by AFLPs of sequences containing repeats) protocol and Roche 454 sequencing. This method has been shown to be time and cost effective compared to cloning and Sanger sequencing, with thousands of sequences that contain repeats being obtained. Using only half of a pyrosequencing lane, 36 482 reads were obtained, of which 21 390 sequences (58.63%) contained repeats. After applying certain restrictions to obtain optimal markers, 233 primer pairs were designed for *S. lupi*. Of these, 20 loci were tested and ten polymorphic loci were identified. A single multiplex PCR reaction was designed to amplify all ten loci, making genotyping very efficient and cost-effective. Nine loci were deemed suitable for population genetic analyses, with an average of 9.89 alleles per locus (range: 6-18). Since the microsatellite markers designed in this study amplified *S. lupi* in domestic dogs (*Canis domesticus*), they were tested on nematode samples obtained from jackal (*Canis mesomelas*) in South Africa. The results suggested that *S. lupi* infestation does occur in these jackals. The markers were also successful in cross-species tests of amplification in *Cylicospirura felineus*, *Philonema oncorhynchi* and *Gongylonema pulchrum*. These markers will be applied further to study the population structure and diversity of *S. lupi* populations.

Keywords: 454 sequencing, *Canis domesticus*, *Canis mesomelas*, marker development, microsatellites, parasitic nematode, *Spirocerca lupi*

2.2. INTRODUCTION

Major advances in molecular technologies have led to the development of new approaches that help to better understand parasitism (Gasser 2006). Genetic markers can be utilised to understand disease mechanisms through the elucidation of genetic structure so that ecological and evolutionary processes of parasitic populations can be determined (Criscione *et al.* 2005; Nadler 1995). Microsatellites are progressively becoming the marker of choice in many areas of research. These molecular markers are a class of repetitive DNA made up of tandem repeats that are 2-8 nucleotides in length (Pompanon *et al.* 2005). Microsatellite markers provide a multilocus approach for characterising individuals, whereas the use of a single locus may sometimes be misleading and unreliable (Redman *et al.* 2008). Due to high levels of polymorphism in both eukaryotes and prokaryotes, they can be used effectively to detect mating patterns, population structure and diversity.

Many methods are available for developing microsatellite-rich libraries and until recently, the development of microsatellite markers was complex, time-consuming and expensive (Wasimuddin *et al.* 2012). For organisms with little or no sequence data available, clone libraries are generated and sequenced in order to develop markers (Santana *et al.* 2009). If genome sequences are available for a species, bioinformatics tools are used to search for microsatellite sequences in these or expressed sequence tag (EST) databases and primers are designed to target these regions (La Rosa *et al.* 2012). However, very few species have whole genomes sequenced and to generate a whole genome sequence is very costly (Santana *et al.* 2009).

Many studies have developed microsatellite markers using classical enrichment techniques, together with cloning and sequencing (Belanger *et al.* 2011; Loiseau *et al.* 2010; Schwenkenbecher & Kaplan 2007), however this is time consuming and labour intensive. The high cost of Sanger sequencing results in only a few hundred sequences being obtained (Malausa *et al.* 2011) and only a subset of those contain repeats. Recently, next generation sequencing together with microsatellite enrichment protocols have advanced the development of microsatellite markers, by providing a more cost effective and rapid method of sequencing microsatellite-enriched sequence libraries (Castoe *et al.* 2010; Malausa *et al.* 2011; Reid *et al.* 2012; Santana *et al.* 2009). Pyrosequencing eliminates the need to spend a significant amount of time to clone and sequence potential microsatellite-containing sequences, since a significant amount of data is produced within a short period of time, from which hundreds of potential markers can be developed (Castoe *et al.* 2010). There are a few limitations to 454 sequencing. For example, sequences may have insufficient flanking region and there may be potential sequencing errors (Ekblom & Galindo 2011). However these shortcomings can be overcome with very little effort (Reid *et al.* 2012).

Enrichment before pyrosequencing is essential in order to gain maximum value from the sequences, otherwise very few sequences with repeats are obtained, as in (Scoble *et al.* 2011) where only 0.88% of

sequences contained microsatellites. A study that compared the efficiency of enrichment found that on average, 4.6% of SCRs were obtained without enrichment and 21% of SCRs were obtained with enrichment (Reid *et al.* 2012). Besides increasing the overall number of microsatellite loci isolated, enrichment also reduces the number of unwanted motifs that may cause difficulties in amplification when genotyping (Malausa *et al.* 2011). It should be noted that in this, as well as other studies, the initial number of microsatellites that are detected is most likely inflated since multiple reads that cover the same sequence may be present (Perry & Rowe 2011).

The development of markers is essential in order to conduct studies on the genetic diversity and structure of parasitic populations. *Spirocerca lupi* is a parasitic nematode which is known to cause spirocercosis (Chandrasekharon *et al.* 1958), a disease which affects canids. It is a severe and debilitating disease (Kok *et al.* 2011) for canine populations worldwide and shows a higher prevalence in regions with a warm, tropical climate (Bailey 1972; Mazaki-Tovi *et al.* 2002; Van der Merwe *et al.* 2008). The aim of this study was to develop microsatellite-containing markers for *S. lupi* using the FIASCO (fast isolation by AFLPs of sequences containing repeats) protocol (Zane *et al.* 2002) together with Roche 454 sequencing, to obtain a microsatellite-enriched library. The markers were assessed for their utility and tested for cross-species transferability. They were also used to verify the finding of *S. lupi* in jackal from South Africa.

2.3. MATERIALS AND METHODS

2.3.1. Sample collection

Spirocerca lupi worms were obtained from post-mortems conducted by veterinarians and pathologists in the Tshwane Metropole (Gauteng), Grahamstown (Eastern Cape) and Durban (KwaZulu-Natal) areas in South Africa. Adult worms (130) were collected opportunistically between 2005 and 2012 from dogs and stored in 70% ethanol. It should be noted that two worm samples obtained from veterinarians were stored in formalin and later transferred to 70% ethanol upon receipt of the samples.

Seven worms were collected in 2012 from three jackals in Kimberley (Northern Cape, South Africa) and were obtained from veterinary pathologists at the Onderstepoort Veterinary Institute.

Cylicospirura subaequalis, *Cylicospirura felineus*, *Philonema oncorhynchi*, *Gongylonema pulchrum* and *Toxocara canis* worm samples were obtained from parasitologists in the USA and Iran to test for cross-species amplification.

2.3.2. DNA extraction

The DNeasy Blood and Tissue kit (Qiagen, Hilden) was used to extract total genomic DNA from a small segment of each nematode. The Spin-column Protocol for Purification of Total DNA was followed. Ten worms, each from a different dog, were washed three times each with 70% ethanol and transferred to separate tubes. Worms were crushed with liquid nitrogen and incubated overnight at 56 °C in buffer containing protein kinase. RNase A (100 mg/ml) was added to each sample. Two elutions were performed, each in 200 µl of elution buffer supplied with the extraction kit. Extracted DNA was stored at 4 °C. The concentration and quality of the DNA was determined by absorbance values read on the Nanodrop ND-1000 Spectrophotometer v 3.2 (Nanodrop Technologies, Wilmington, USA).

DNA was also extracted from one nematode of each of the samples obtained for cross-species amplification. The same extraction method was followed for the worms obtained from jackal.

2.3.3. Microsatellite marker development

Seven samples that had the highest DNA quality and concentration were pooled together in order to increase the total amount of DNA and to avoid the development of markers from a single individual which would in turn avoid an ascertainment bias (Glenn *et al.* 2005; Leach *et al.* 2012).

The FIASCO protocol (Zane *et al.* 2002) was followed to isolate microsatellite-rich sequences. Slight modifications were made in order to optimise the protocol for *S. lupi*. Ten µM of each *MseI* AFLP adaptor (Inqaba Biotec) was combined and placed in a thermal cycler under the following conditions: 96 °C for 2 minutes and 94 °C for 1 minute and then allowed to cool to room temperature. Genomic DNA was digested with restriction enzyme *MseI* (New England Biolabs) and simultaneously ligated to the prepared *MseI* AFLP adaptor. The reaction mix contained approximately 1 µg of genomic DNA, 10x enzyme buffer NEB2 (New England Biolabs), 100x BSA, 20 units *MseI* enzyme, 1 mM ATP, 20 units T4 DNA ligase and 10 µM *MseI* AFLP adaptor in a total volume of 100 µl. This was incubated overnight at 37 °C. The reaction mixture was then heated at 65 °C for 20 minutes in order to inactivate the *MseI* restriction enzyme and the T4 DNA ligase. When analysing the result of the digestion-ligation reaction on a 1% agarose gel, a visible product in the form of a smear, was considered optimal.

The digestion-ligation mixture was diluted 10-fold with SABAX water. Digested-ligated products were amplified by PCR with AFLP adaptor-specific primers (5' GATGAGTCCTGAGTAAN-3'), referred to as *Mse1*-N, where N represents A, C, G and T. The 25 µl PCR reaction consisted of 10x *Ex Taq* Buffer, 1.5 mM $MgCl_2$, 10 mM dNTPs, 10 pmol *Mse1*-N primers and 2 U TaKaRa Ex Taq™ DNA polymerase (Takara Bio Inc.). The optimal primer combination determined for *S. lupi* was a mixture of *Mse1*-A, *Mse1*-G and

Mse1-T. After optimisation, thermal cycling conditions for *S. lupi* were an initial denaturation for 2 min at 94 °C followed by 25 cycles of 94 °C for 30 seconds, 53 °C for 1 minute and 72 °C for 1 minute with a final elongation step at 72 °C for 7 minutes (Vos *et al.* 1995). In order to increase DNA yield, PCR amplification under optimal conditions was replicated several times. The PCR product (5 µl) was visualised on a 1% agarose gel by electrophoresis. Clear smears indicated that all fragments in the genome were equally represented. A negative control reaction, where *S. lupi* DNA was substituted with UHQ water, was always included to control for the possible contamination of reaction components. All reactions were carried out on a PxE 0.2 thermocycler (Thermo Electron Corporation).

PCR-amplified DNA was purified following the protocol in the Roche High Pure PCR Product Purification kit and eluted in 50µl elution buffer. Purified DNA was hybridised to selected biotinylated oligonucleotide probes: $(AC)_{15}$, $(AG)_{15}$, $(AT)_{15}$, $(TAT)_{10}$, $(TGT)_{10}$, $(AAG)_{10}$, $(ATC)_{10}$, $(AAAT)_8$ and $(TTTG)_{8,}$ which were chosen due to the fact that they were the most commonly found repeats in other nematode species (Belanger *et al.* 2011; Criscione *et al.* 2007; Grillo *et al.* 2006; Hoekstra *et al.* 1997; Johnson *et al.* 2006; Otsen *et al.* 2000; Schwenkenbecher & Kaplan 2007; Tóth *et al.* 2000). Di-, tri- and tetra- nucleotide probes were combined in separate reactions. The three reactions each consisted of 250 ng of DNA mixed with 10 µM of biotinylated oligonucleotide made up to a total volume of 100 µl with hybridisation solution (SSC 4.2X, SDS 0.7%). Denaturation of DNA was performed at 95 °C for 5 minutes followed by annealing of probes at room temperature for 30 minutes. Then 300 µl of TEN_{100} was added to the DNA-probe hybridisation reaction. One milligram of Streptavidin-coated magnetic beads (Roche) was washed three times with 100 µl TEN_{100} (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl, pH 7.5) and resuspended in 40 µl TEN_{100}. Ten micrograms tRNA was added to minimise non-specific binding. The DNA-probe hybridisation reaction was added to the beads and incubated at room temperature for 30 minutes. The beads-probe-DNA complex was separated from the hybridisation buffer with a magnet. Seven non-stringent and seven stringent washes were performed. All washes were stored and their products amplified by PCR, which was visualised on an agarose gel to determine whether non-specific DNA was in fact removed.

Heat denaturation with 1x TE buffer (10 mM Tris HCl, 1 mM EDTA, pH 8) was performed at 95 °C for 5 minutes. Chemical denaturation was performed at room temperature for 5 minutes with 0.15 M NaOH. The supernatant was removed and stored with an equal volume of 0.1667 M acetic acid. A final volume of 50 µl was made up with 1x TE buffer. DNA was precipitated and stored at -20 °C.

PCR amplification of the enriched product was performed with the same conditions as above, however with 30 cycles. Enriched PCR products were purified with the Roche High Pure PCR Product Purification Kit and eluted in 50 µl elution buffer. Equal concentrations of purified PCR products from each di-, tri- and tetra-nucleotide enrichment were pooled together and sent to Inqaba Biotec (Pretoria, South Africa) for

sequencing using 454 LifeSciences/ Roche GS-FLX pyrosequencing. *Spirocerca lupi* DNA was tagged with a unique multiplex identifier and placed on 1/16th of a PicoTiter Plate with another, unrelated library.

2.3.4. Primer design

BioEdit version 5.0.6. (Hall 1999) was used to detect and discard all sequences shorter than 60 bp in length and to remove adaptors. The program MSATCOMMANDER (Faircloth 2008) was used to search for sequences containing repeats (SCRs). The minimum number of repeats for each repeat type was as follows: di- (8), tri- (8), tetra- (6), penta- (5), hexa- (5) and a maximum of 30 repeats was set for the search. Data generated from the search was analysed in Microsoft Office Excel 2007 by using the data filter to sort according to fragment size, repeat types and repeat lengths. Sequences were aligned and reverse complemented on Clustal X version 1.81 (Thompson *et al.* 1997) to identify and remove duplicated repeat fragments, which may represent non-unique loci (Scoble *et al.* 2011). Microsatellite sequences with the same flanking regions were considered as the same locus. Primers were designed using default parameters (as automated within MSATCOMMANDER) using the Primer3 version 1.1.1 program (Rosen & Skaletsky 2000). Primers that were not immediately adjacent to the repeat unit and that had a length of expected amplicon between 100-400 bp were selected.

2.3.5. Marker testing

A set of twenty primer pairs that amplified eight or more pure repeats, had similar melting temperatures (Tm) and a GC content of 60%, were chosen as candidates for marker development. PCR amplification reactions consisted of 10x *Ex Taq* Buffer, 1.5 mM $MgCl_2$, 10 mM dNTPs, 10 pmol each of the forward and reverse primers (Inqaba Biotec), 0.25 U TaKaRa Ex Taq™ DNA polymerase and approximately 10 ng of genomic DNA in a 10 µl volume. The PCR conditions were 95 °C for 2 min followed by 30 cycles of 95 °C for 1 min, 50-63 °C for 30 s and 72 °C for 30 s and a final elongation step at 72 °C for 5 min. Primer pairs that produced consistent amplification of the expected fragment lengths were tested on six different *S. lupi* individuals to determine whether loci were polymorphic, by visualising PCR products on 3% agarose gels with ethidium bromide staining. PCR reactions were repeated with 0.02 pmol ChromaTide Alexa Fluor® 488-5-dUTP (Invitrogen), which are fluorescently labelled dNTPs that incorporate during PCR allowing the fragments to be visualised on an automated sequencer (Reid *et al.* 2012). Sequencing was performed at the DNA sequencing facility (University of Pretoria, South Africa) on an ABI 3500xl automated sequencer (Applied Biosystems) with GeneScan Liz™ 500 Size Standard (Applied Biosystems).

Dye set G5 consisting of 6-FAM, VIC, NED and PET (Applied Biosystems) was used to fluorescently label the forward primers of loci that were found to be polymorphic. The Quantitect Multiplex PCR kit (Qiagen) was used to genotype 130 *S. lupi* individuals following the protocol as described by the manufacturer.

Multiplex reactions of 14 µl each were set up each containing approximately 50 ng of genomic DNA, 7 µl Quantitect Multiplex PCR Master Mix and 0.2 pmol each of forward and reverse labelled primers for each primer set. Cycling conditions were 95 °C for 15 min followed by 40 cycles of 94 °C for 60 s and 60 °C for 90 s. All ten primer sets could be combined in a single multiplex reaction according to the dye colour and length of the product to avoid overlapping of loci with the same label. Amplification was repeated for all microsatellite loci in 22 samples to determine genotyping error. Allele sizes were separated and measured on an ABI3500xl Genetic Analyser (Applied Biosystems) using GeneScan LIZ500 Internal Size Standard (Applied Biosystems). The genotype of each individual was determined by analysing chromatograms in GeneMarker™ version 2.4.0. (SoftGenetics LLC).

2.3.6. Sequencing homozygotes

Automated sequencing was applied to confirm repeat-motif variation and allele scoring in 33 individuals. Selected homozygous alleles were amplified using unlabelled primers (using the protocol described above). Ethanol precipitation was used to purify the PCR products which were visualised on a 1% agarose gel. The precipitated DNA product was sequenced in the forward direction only. The ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Applied Biosystems) was used to perform the cycle sequencing reaction on a Thermo Electron Corporation PxE 0.2 thermocycler. The reactions contained approximately 100 ng of purified template DNA, 3.2 pmol of the forward primer, 2 µl of BigDye reaction mix and 5x BigDye sequencing buffer. Cycle sequencing conditions were according to the manufacturer's recommendations and precipitated as described above. Sequences were analysed on an ABI 3130 automated sequencer (Applied Biosystems) at the DNA sequencing facility, University of Pretoria, South Africa. The chromatograms were visualised with CLC Main Workbench 6 (CLC Bio).

2.3.7. Cross-species amplification

Microsatellite markers that were identified as polymorphic in *S. lupi* were used to evaluate their success in amplifying DNA from other species within the same class. One worm from each of the following species (hosts in brackets) were tested: *Cylicospirura subaequalis* (cougars), *Cylicospirura felineus* (bobcats), *Philonema oncorhynchi* (sockeye salmon), *Gongylonema pulchrum* (sheep) and *Toxocara canis* (canids).

2.3.8. Microsatellite data analysis

Genemarker version 2.4.0 (SoftGenetics, State College, Pennsylvania, USA) was used to score genotypes. Observed heterozygosity (H_O), expected heterozygosity (H_E) and Weir and Cockerheim's estimate of F_{IS} were computed using Genetix version 4.05.2. (Belkhir *et al.* 1996-2004). Deviations from Hardy-Weinberg equilibrium for each locus and globally, were calculated with Arlequin version 3.5.1 with 1 000 000 Markov

chain steps (Excoffier *et al.* 2005). A test for linkage disequilibrium was conducted in Genepop v 1.2. (Raymond & Rousset 1995; Rousset 2008). Microchecker version 2.2.3 (Van Oosterhout *et al.* 2004) was used to check for null alleles and scoring errors due to stuttering and large-allele dropout (i.e. short-allele dominance).

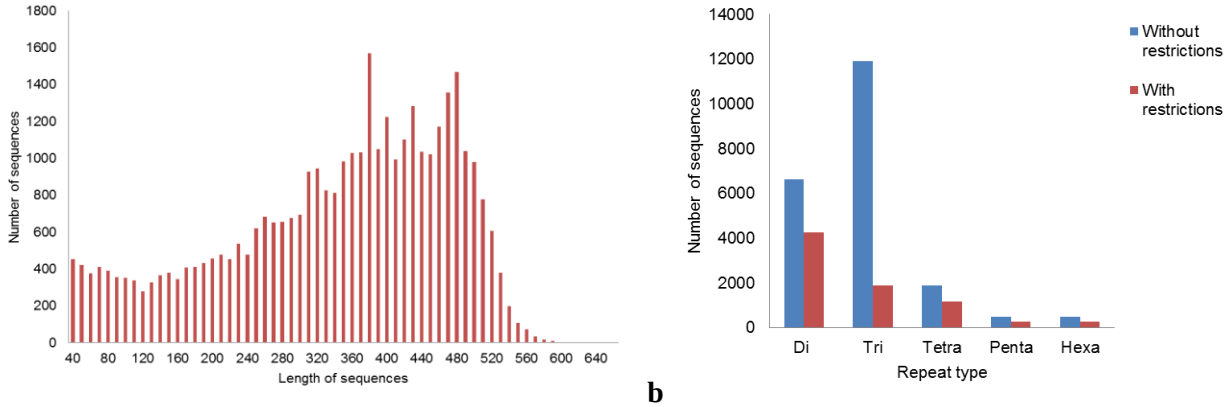
2.4. RESULTS

2.4.1. DNA extraction and PCR amplification

Sufficient amounts of DNA were extracted from each of the worm samples stored in 70% ethanol. DNA could be also be extracted from the two worm samples that were previously stored in formalin. Slight degradation of the DNA sample was observed when the DNA concentration was measured, however the DNA yield was sufficient since the required loci could be amplified by PCR.

2.4.2. Marker development

Half of a pyrosequencing lane yielded 36 482 reads from the microsatellite-enriched library. Read lengths ranged from 40-600 bp, with most of the fragments between 310-500 bp (Figure 2.1a), which is ideal for amplification by PCR (Malausa *et al.* 2011). Before applying any restrictions to the data, 21 390 sequences (58.63%) contained repeats. Most of the probes that were used for enrichment hybridised to regions of the *S. lupi* genome, however it is assumed that no hybridisation to $(TGT)_{10}$ and $(TAT)_{10}$ probes occurred since these repeats were not detected after sequencing (Table 2.1). After removing adaptors, removing sequences less than 60 bp and setting the stipulated restrictions for the number of repeats, 7835 (21.48%) sequences were identified as having a di-, tri-, tetra-, penta- or hexa- nucleotide repeat (Figure 2.1b), of which the total number of repeats that were found was 9505 (Table 2.2).



a **b**

Figure 2.1: a) Distribution of fragment lengths obtained from Roche 454 sequencing for *S. lupi* after enrichment for microsatellites. **b)** Comparison of number of sequences containing repeats before and afer applying restrictions to the data.

Table 2.1: Probes that were used for enrichment and the respective repeat types that were obtained without any restrictions set on the data. Percentages indicate the number of sequences obtained for a specific repeat from the total number of 36 482 reads.

Dinucleotide

Probe	Repeat obtained	No. of sequences
$(AC)_{15}$	AC/GT	5920 (16.23%)
$(AG)_{15}$	AG/CT	2541 (6.97%)
$(AT)_{15}$	AT	85 (0.23%)
	CG	11 (0.03%)

Trinucleotide

Probe	Repeat obtained	No. of sequences
$(ATC)_{10}$	ATC/GAT	6690 (18.33%)
$(AAG)_{10}$	AAG/CTT	720 (1.97%)
$(TGT)_{10}$	-	0 (0%)
$(TAT)_{10}$	-	0 (0%)

Tetranucleotide

Probe	Repeat obtained	No. of sequences
$(AAAT)_8$	AAAT/ATTT	292 (0.8%)
$(TTTG)_8$	TTTG/CAAA	109 (0.3%)

Table 2.2: Record of the numbers of repeat units obtained for each repeat type after restrictions were applied to the data.

Dinucleotide		**Trinucleotide**		**Tetranucleotide**		**Pentanucleotide**		**Hexanucleotide**	
Repeat unit	No. found	Repeat unit	No. found	Repeat unit	No. found	Repeat unit	No. found	Repeat unit	No. found
AC/GT	3553	AAC/GTT	1103	CTTT	789	AAAAG/CTTTT	185	AACAGC/GCTGTT	127
AG/CT	1475	ATC/GAT	591	ACCT/AGGT	668	AAAAC/GTTTT	92	AAAAAC/GTTTTT	36
AT	20	AAG/CTT	314	AAAT/ATTT	186	AAACC/GGTTT	5	ACTGCT/AGCAGT	25
CG	3	ACT/AGT	72	AGAT/ATCT	30	AACAC/GTGTT	4	ACCATC/GATGGT	23
		AGC/GCT	33	ATCC/GGAT	29	AAGTG/CACTT	4	AAGGAG/CTCCTT	12
		ACG/CGT	7	AATG/CATT	16	ACACT/AGTGT	4	ACACAT/ATGTGT	12
		AAT/ATT	5	AAAC/GTTT	12	ATTTT	4	AAAAAG/CTTTTT	10
				ACAT/ATGT	8	CGATT	3	CCTTTT	5
				AAGG/CCTT	4	CTATT	3	AAAAGT	2
				AATC/GATT	3	CTCTT	3	ATTCTT	2
				GGTT	3	AAGAT	2	CTTGTT	2
				AACT	1	ACACC/GGTGT	2	AAAACT	1
				ACCG	1	ACACG	1	ACGATG	1
				ACGT	1	ACCGT	1	AGGGTT	1
				CTGT	1	ATCCT	1	CTCTGT	1
						CCATT	1		
						CCTTT	1		
						GGGTT	1		
	5051		**2125**		**1752**		**317**		**260**

Total number of repeats found: 9505

Primers could be designed for 381 repeat-containing sequences with suitable flanking regions. After applying certain selection criteria, the final list was narrowed down to 233 sequences with suitable primers (11 sequences had compound repeat motifs). Repeat types for which primers were designed consisted of 183 dinucleotides, 39 trinucleotides, 5 tetranucleotides, 1 pentanucleotide and 5 hexanucleotides (Figure 2.2).

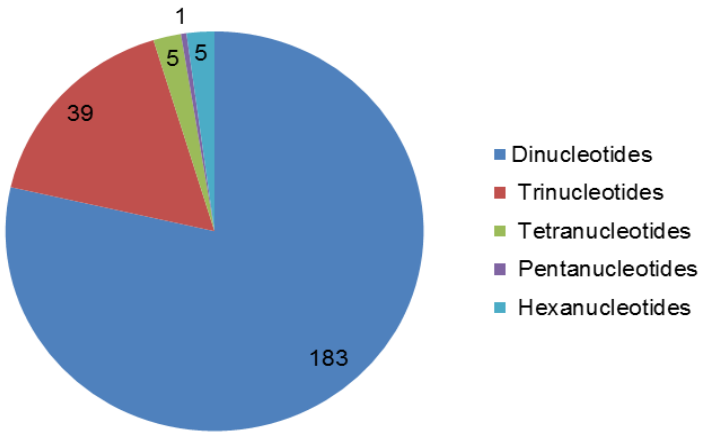


Figure 2.2: Distribution of repeat types obtained for 233 primer sets that could be designed for *S. lupi*

Of the twenty loci tested, ten loci were found to be highly polymorphic and could be combined into a single multiplex reaction, with markers at least 30 bp apart.

2.4.3. Marker suitability

Genotyping of all the loci demonstrated that out of the ten loci, nine scored reliably. 5.23% of the data could not be scored when marker SL20 was included whereas 2.91% of the data could not be scored when this marker was excluded. Since marker SL20 scored ambiguously (Figure 2.3) and resulted in a high amount of missing data, it was excluded from further analyses. This marker contained a dinucleotide repeat, which is known to sometimes be problematic when calling allele sizes, since these repeat types tend to stutter more than larger repeat classes (Gardner *et al.* 2011). The rest of the markers produced clear, unambiguous single or double peaks, as can be expected for a diploid organism. However, the 2.91% of the missing data was due to ambiguous amplification for the following numbers of individuals at the respective loci: SL02 (6), SL04 (3), SL06 (7), SL10 (2), SL13 (6), SL14 (4), SL15 (5), SL17 (5) and SL18 (1). Out of 29 individuals that displayed ambiguous amplification for some of the loci, 24 (87%) of these individuals failed only to amplify at one locus but amplified clearly across all other loci. The nine markers that were developed will be deposited in Genbank/ EMBL-Bank (See Appendix 2A for marker sequences).

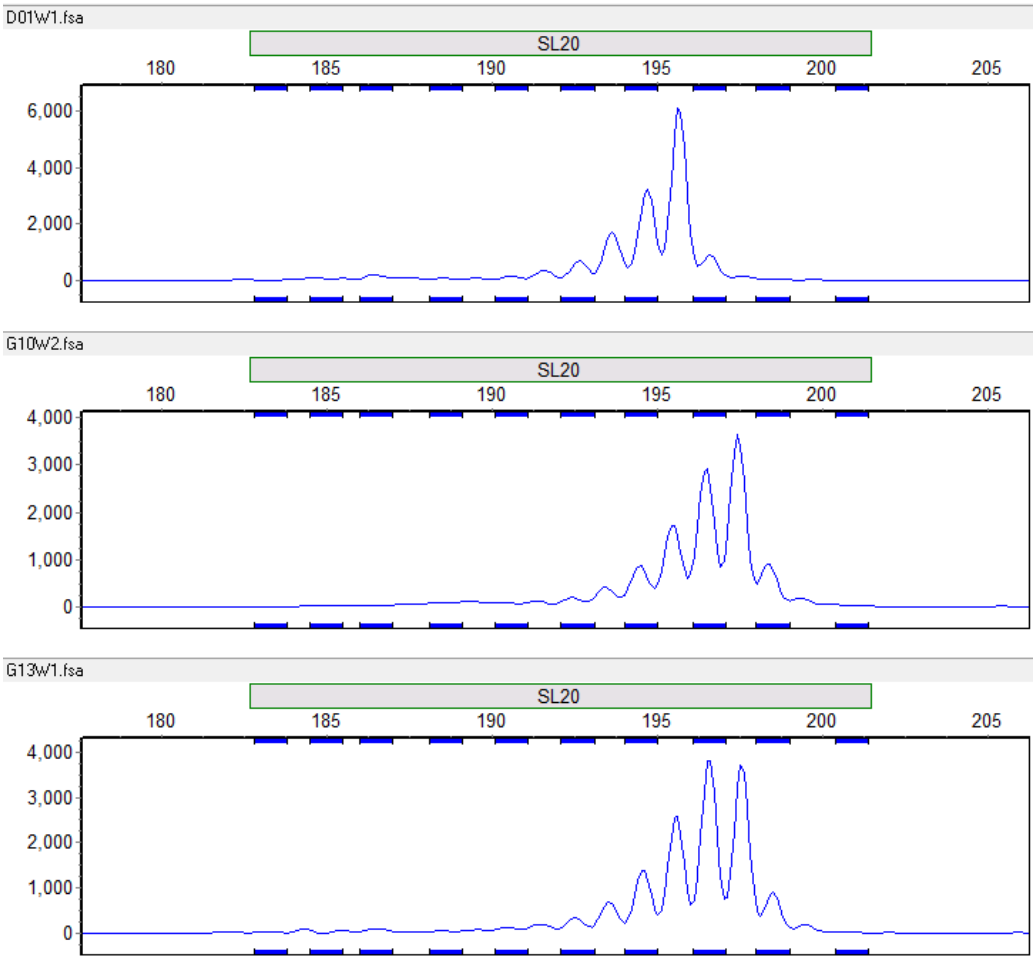


Figure 2.3: Electropherogram of the scoring pattern of marker SL20 indicating that alleles produced ambiguous amplification, making it difficult to distinguish between homozygotes and heterozygotes.

High levels of polymorphism were observed with the average number of alleles being 9.89 and ranging from 6-18, with allele sizes between 147-380 bp. Average observed (H_O) and expected (H_E) heterozygosities were 0.66 (range: 0.46 – 0.88) and 0.70 (range: 0.47 – 0.83), respectively (Table 2.3). Marker SL14 seemed to be the most genetically polymorphic since it had 18 alleles and an H_E of 0.83, whereas marker SL04 seemed to be the least genetically polymorphic with 6 alleles and an H_E of 0.47. Null alleles were detected at the following loci: SL06, SL10, SL15 and SL17, where a general excess of homozygotes were detected for most allele size classes. No loci showed significant deviation from Hardy-Weinberg equilibrium and F_{IS} values were low (Table 2.3). However, some of the locus pairs showed significant linkage disequilibrium. Genotyping was consistent since 17 samples that were tested twice showed the same genotypes. BLAST searches did not detect matches to any of the locus sequences, thus supporting the possibility that these microsatellites are not located near coding regions (Perry & Rowe 2011).

Table 2.3: Characteristics of ten microsatellite markers developed for *Spirocerca lupi*, with measures of genetic diversity. (N: number of sampled individuals, N_A: number of alleles, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: inbreeding coefficient)

Locus	Fluorescent label	Primer sequence (5'-3')	Tm (°C)	Repeat motif	Allele range (bp)	N	N_A	H_O	H_E	F_{IS}	P-values
SL02	PET	F: CCATCCTCTTGCGTTGCAC	60	$(ATC)_{10}$	344-380	124	9	0.65	0.68	0.05	0.209
		R: TCCTGTGCAGGCCATTACC									
SL04	PET	F: GAACGGTTTCCGCGAACTC	59	$(AGC)_8$	147-162	127	6	0.46	0.47	0.02	0.431
		R: CCGAAAGTCTGAACGTTGTC									
SL06	VIC	F: GAGATTGGCCGGAAAGGTG	59	$(AAC)_9$	352-373	128	8	0.65	0.79	0.19	0
		R: TACAGCATTGCCCGAAAGC									
SL10	6-FAM	F: CTCCCTGGAATCTATTTGCCC	58	$(AG)_{10}$	220-242	128	10	0.60	0.73	0.18	0
		R: AGCACTGTTAGGGATCAGC									
SL13	VIC	F: ATACCGTTTCGGTGCCAAG	59	$(AT)_8$	231-249	124	9	0.73	0.67	-0.07	0.944
		R: GCGGCACTCACAGTTGAC									
SL14	PET	F: CCGAGGGTACTCGATGTGG	60	$(AC)_{10}$	221-263	126	18	0.86	0.83	-0.03	0.791
		R: AGCCCGAGCAGTCTTGATG									
SL15	VIC	F: CTGTTGGTGGTCCATTTCGG	60	$(GT)_{14}$	183-209	125	13	0.47	0.56	0.16	0
		R: CAGATGGTCGCAACAGTCC									
SL17	NED	F: TCAACCAATCTGGCGCAAC	60	$(GCT)_8$	210-240	125	9	0.63	0.77	0.19	0
		R: CCGTTCGTCCTTCAACAGC									
SL18	6-FAM	F: TGAGGCGATTGTTGCGTTC	59	$(GAT)_8$	297-318	129	7	0.88	0.78	-0.13	1
		R: AGCGACATCACGTTTCCAG									
SL20	6-FAM	F: GAGCTTTCCGTGGTTCTGC	60	$(GT)_{12}$	184-196	97	-	-	-	-	-
		R: CGGTTGGATGGGCGTAATTC									

2.4.4. Sequencing of homozygotes

Sanger sequencing revealed that the flanking regions in almost all of the markers were conserved and that variation in the length of the alleles was due to variation in the number of repeats (Appendix 2B). There were single nucleotide polymorphisms (SNPs) and occasional indels in the flanking regions of some loci. The repeat region did not sequence in locus SL18, however was not re-sequenced, since all alleles scored unambiguously at this locus. Locus SL20 could not be scored unambiguously and sequencing results confirmed this (Appendix 2B). This locus was excluded from further analysis since the repeat length did not match the allele size for any of the individuals that were sequenced, even though the flanking regions showed no variation.

2.4.5. Cross-species amplification and testing in jackal

Some of the markers successfully amplified DNA from other closely-related species, namely *Cylicospirura felineus* (SL20, SL15, SL14, SL10, SL06 and SL18), *Philonema oncorhynchi* (SL20, SL15, SL17, SL10, SL06 and SL18) and *Gongylonema pulchrum* (SL14 and SL10) (Table 2.4). None of the markers produced amplification in *Cylicospirura subaequalis* and *Toxocara canis*.

Table 2.4: Loci that produced amplification in cross-species tests using markers designed for *S. lupi*

Locus in *Spirocerca lupi*	*Cylicospirura felineus*	*Philonema oncorhynchi*	*Gongylonema pulchrum*	*Cylicospirura subaequalis*	*Toxocara canis*
SL02	-	-	-	-	-
SL04	-	-	-	-	-
SL06	✓	✓	-	-	-
SL10	✓	✓	✓	-	-
SL13	-	-	-	-	-
SL14	✓	-	✓	-	-
SL15	✓	✓	-	-	-
SL17	-	✓	-	-	-
SL18	✓	✓	-	-	-
SL20	✓	✓	-	-	-

All ten loci that amplified *S. lupi* from dogs also produced amplification in worms that were collected from jackal. The same levels of polymorphism were observed as for dogs, where the allele size ranges of the jackal-derived samples fell within the range that was found for dog-derived samples (Figure 2.4) for most loci.

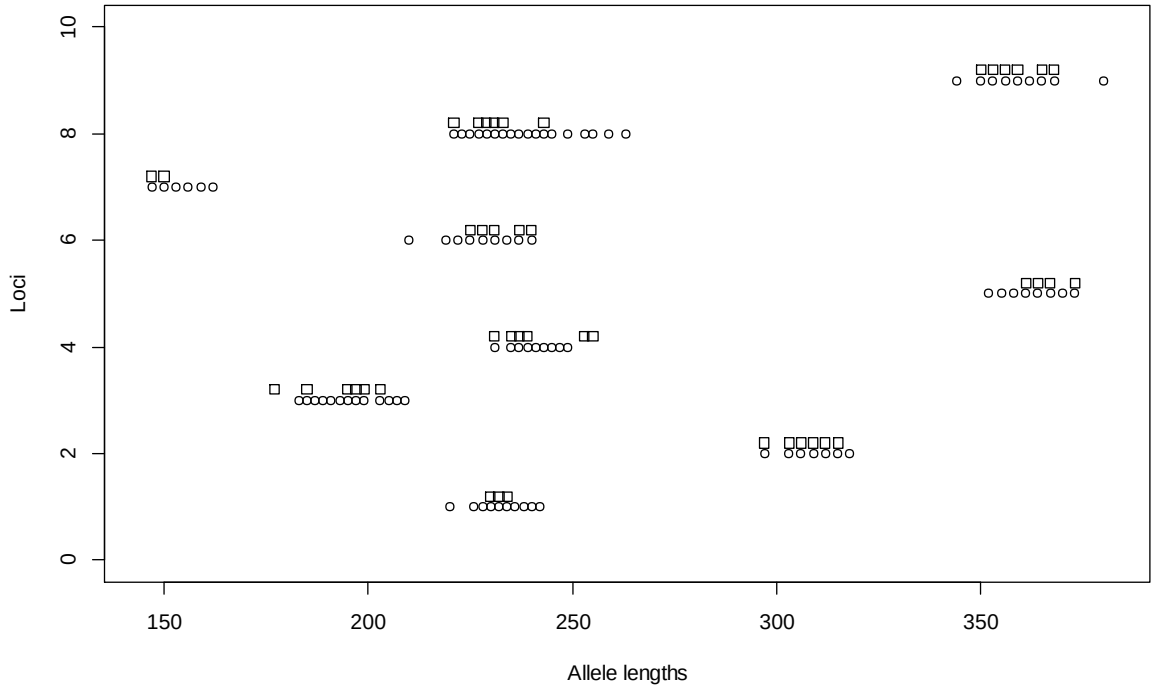


Figure 2.4: Allele size ranges of the nine microsatellite loci with alleles from jackal-derived samples represented by blocks and alleles from dog-derived samples represented by circles.

Pairwise F_{ST} values were not significantly different from zero (0.08359, 0.09248 and 0.02820 with P < 0.009), thus indicating that the jackal-derived samples were not significantly different from the dog-derived samples.

2.5. DISCUSSION

Microsatellites use DNA-based methods such as PCR to amplify the marker of interest from even the smallest amounts of DNA that have been preserved with the simplest tissue preservation methods (Selkoe & Toonen 2006). Since microsatellites are shorter in length (100-300 bp) than sequenced loci (500-1500 bp), they are easily amplified by PCR even when DNA samples are slightly degraded. Formalin is known to cause DNA degradation. In this study, worm samples that were preserved in formalin could still be genotyped, thus providing increased support for the value of using microsatellites as markers.

The method of enrichment and pyrosequencing that was employed in this study to obtain microsatellite markers proved to be very efficient and cost effective. Pyrosequencing produced 36 482 reads which was in the region of that obtained from pyrosequencing of 13 taxa where the highest number of reads was 39 473 after enrichment (Malausa *et al.* 2011). It is expected that, without enrichment, from 1/16th of a

pyrosequencing lane, one can expect to isolate 6000-7000 microsatellites (Abdelkrim *et al.* 2009). With enrichment, we detected 21 390 (58.63%) sequences that contained microsatellites, which is approximately three times more than anticipated without enrichment. This result confirms, as in other studies using pyrosequencing and FIASCO (Santana *et al.* 2009; Zhan *et al.* 2009), that the method is well-optimised to provide high-quality sequences that are rich in microsatellite repeats.

The abundance of repeat motifs in the genome was measured after restrictions were applied to the pyrosequencing data. It should be noted that a potential bias may have been introduced since the probes used to enrich for microsatellites were designed based on their abundance in other nematode genomes. Also, enrichment with di- and tri- nucleotide probes may have resulted in the higher numbers of these repeat sequences being found. However many other repeat types that were not probed for, were detected in the pyrosequencing reads, which indicates that the repeats that were detected were not entirely dependent on the probes that were used. Similarly, this has been observed in attempts to isolate microsatellites from *Haemonchus contortus* where repeat types, different from those that the libraries were probed with, were found (Grillo *et al.* 2006).

It was found that the frequency of repeats decreased as the repeat length increased, which can be explained by the high mutation rate and downward mutation bias that is observed when there is a large number of repeat units (Lai & Sun 2003). Dinucleotides were by far, the most abundant repeat type, which has also been found in other species (Meglecz *et al.* 2012; Techen *et al.* 2010). There are four possible unique dimer combinations: AC, AG, AT and CG. AC-rich repeat units are the most abundant in other eukaryote genomes (Jarne & Lagoda 1996; Schwenkenbecher & Kaplan 2007), whereas CG-rich repeat units are rare (Ellegren 2004; Meglecz *et al.* 2012; Tóth *et al.* 2000). The same was observed for *S. lupi*, where hardly any CG repeats were found and AC/GT was the most frequent repeat type. The underrepresentation of CpG dinucleotides in the genome could be explained by the fact that CpG dinucleotides that are not found in CpG islands will undergo methylation of cytosine, which will then mutate to thymine (Meglecz *et al.* 2012). Very few trimers rich in CG were found, which correlated well with findings from other nematodes (Molnar *et al.* 2012). Tetranucleotide repeats with less than 50% G+C repeats are also more common (Tóth *et al.* 2000). Unlike in other eukaryotes, where it was most abundant, AAT (Meglecz *et al.* 2012) was the least common trinucleotide motif found in *S. lupi*.

There is a high attrition rate when developing markers for invertebrates since more than double the number of sequences are required to obtain the same number of usable loci as plants and vertebrates (Gardner *et al.* 2011). Also in invertebrates there are fewer loci for which suitable primers can be designed, especially for tetra-, penta- and hexa- nucleotides (Gardner *et al.* 2011). The FIASCO protocol that was used to enrich for microsatellites proved to be successful since 233 suitable primer pairs could be designed. Of these, 20 primer sets were tested and ten polymorphic microsatellite markers were developed for *S. lupi*. The choice of

dinucleotide and trinucleotide loci for marker development was appropriate since these repeats often have a small allele range (Gardner *et al.* 2011), which allows for greater numbers of loci to be genotyped in a single reaction. The markers proved to be easily usable in routine conditions and only one multiplex PCR reaction was required to genotype each individual. The advantage of being able to combine all ten microsatellite markers into a single multiplex PCR reaction is that firstly, it significantly reduces the cost of genotyping individuals and secondly, multi-locus genotypes can be obtained even with limited amounts of DNA from an individual. A zero percent error rate was found in individuals that were re-genotyped, thus multiplexes were well-adapted to specified conditions.

Linkage disequilibrium was detected between three pairs of loci when the analysis included all the samples as though they were from one population. Even though the relative locations of the markers in the genome are not known, no genetic linkage is suspected since none of the same loci were found to be linked across the different locations when the test for linkage disequilibrium was performed on each of the three locations separately. It is known that when there is genetic structure it can cause linkage disequilibrium between unlinked loci resulting in false-positive linkage signals (Pfaff *et al.* 2001). Thus, the linkage disequilibrium observed is most likely an artefact of grouping separate populations as one. It can be assumed that none of the loci are linked and can be used as independent loci, thus making them informative. The level of polymorphism detected in the sampled individuals was also high, indicating that the markers are suitable for population genetics analyses.

The loci were tested for conformity to Hardy-Weinberg equilibrium by comparing observed genotype frequencies with the frequencies expected for an ideal population (random mating, no mutation, no drift, no migration) (Selkoe & Toonen 2006). The heterozygote deficit that was observed at the SL06, SL10, SL15 and SL17 loci may have been caused by the presence of null alleles, i.e. alleles that are not amplified by PCR due to shorter alleles amplifying more efficiently than larger ones so that only the smaller allele is detected in a heterozygote (Dakin & Avise 2004). Alternatively, a substantial heterozygote deficit (more homozygotes than expected under HWE) may be observed when subdivided populations are studied as one population (Dakin & Avise 2004), also known as the Wahlund effect. It is also important to note that genetic sub-structuring of populations is detected as deviations from Hardy-Weinberg equilibrium for multiple markers for particular populations instead of a single marker in all populations (Grillo *et al.* 2006).

The first suggestion of there being null alleles is unlikely, because even though there were a few missing genotypes at the four loci, it was not because PCR did not amplify the alleles, but rather because the alleles could not be scored due to ambiguous amplification. PCR failure or allelic dropout, as a result of poor template quality (Redman *et al.* 2008) is unlikely because when genotyping was repeated for a selected set of samples, consistent results were obtained. In addition, those samples which failed to amplify for one locus, amplified robustly at other loci. Allelic dropout was also an unlikely explanation since the analysis in

Microchecker version 2.2.3 (Van Oosterhout *et al.* 2004) showed no evidence of short allele dominance. In this study, samples from three different locations were pooled and analysed as one population. Therefore the observed heterozygote deficit is more likely due to the Wahlund effect (Dakin & Avise 2004) and not the presence of null alleles. Evidence for genetic structuring is provided in Chapter 3, when the populations are split according to their locations and analysed separately.

A marker may have a high percentage of null alleles because of variation in the flanking region of the locus, thus affecting its suitability as a marker for population genetics studies. Often, two different alleles may be scored as one when an insertion or deletion close to a microsatellite marker generates size homoplasy (Pompanon *et al.* 2005). In this study, some sequence differences were observed when Sanger and pyrosequencing data were compared, which may have been due to errors in the 454 sequencing data that may be caused by inaccuracies in the sequencing process or errors during the PCR library preparation (Balzer *et al.* 2011). Sequencing of homozygotes did not show mutations in original primer binding site (see Appendix 2B), which further validated the reliability of the markers developed in this study.

A summary of studies developing microsatellite markers for parasitic nematodes of animals is presented in Table 2.5. Many of these studies employ the method of enrichment, however few studies use pyrosequencing to sequence repeats. The number of microsatellite markers developed in this study seem adequate when compared to other studies where ten (Wasimuddin *et al.* 2012) and twelve (Mlonyeni *et al.* 2011) markers were developed using enrichment and pyrosequencing techniques. It is not possible to compare variation of microsatellites among different species because different microsatellite isolation strategies are employed and different restrictions are set on the data when defining optimal markers, which may influence the levels of variation that are observed (Johnson *et al.* 2006). Nevertheless, it was also found that the number of alleles and heterozygosities found for *S. lupi* are within the range detected for other animal parasitic nematodes (Table 2.5).

Table 2.5: Summary of studies to date on microsatellite markers developed for animal parasitic nematodes

Nematode species	Animal host	Method	Markers	N_A	H_O	H_E	Reference
Spirocerca lupi	Canids	Enrichment and pyrosequencing	9	6-18	0.46-0.88	0.47-0.83	This study
Oxyspirura petrowi	Quail and other birds	Genome sequence survey (GSS)	-	-	-	-	(Xiang *et al.* 2013)
Syphacia obvelata	House mouse (*M. musculus*)	Enrichment and pyrosequencing	10	2-6	0-0.57	0-0.72	(Wasimuddin *et al.* 2012)
Trichinella spiralis	Human	Genome screen	7	2-10	0-0.75	0-0.77	(La Rosa *et al.* 2012)
Deladenus siricidicola	Sirex noctilio wasp	FIASCO and pyrosequencing	12	1-2	Very low	-	(Mlonyeni *et al.* 2011)
Neoheligmonella granjoni	Rodents (*M. natalensis* and *M. erythroleucus*)	Enrichment and cloning	13	2-18	0.09-0.95	0.09-0.94	(Brouat *et al.* 2011; Loiseau *et al.* 2010)
Dirofilaria immitis	Mosquito vector, mammals and humans	Enrichment and cloning	11	6-17	0-0.8	0.10-0.88	(Belanger *et al.* 2011)
Trichuris arvicolae	Arvicoline rodents	Enrichment and cloning	12	2-6	0-0.93	0.03-0.67	(Deter *et al.* 2009)
Ancylostoma caninum	Humans and dogs	Enrichment and cloning and genome screen	34	-	-	-	(Schwenkenbecher & Kaplan 2007)
Anguillicola crassus	Eels	Enrichment and cloning	7	13-39	0.49-0.98	0.59-0.95	(Wielgoss *et al.* 2007)
Ascaris lumbricoides	Human	Enrichment and cloning and data mining	35	4-31	0.07-0.89	0.12-0.95	(Criscione *et al.* 2007)
Trichostrongylus tenuis	Red grouse	Enrichment and cloning	16	3-21	0.13-0.2	0.07 -0.91	(Johnson *et al.* 2006)
Teladorsagia (*Ostertagia*) *circumcincta*	Sheep and goats	Enrichment and cloning and data mining	7	4-18	0.03-0.80	0.03-0.91	(Grillo *et al.* 2006)
Haemonchus contortus	Sheep and goats	Genomic library screening and cloning	30	1-7	-	-	(Otsen *et al.* 2000)
Trichuris trichiura	Human	Enrichment and cloning	6	6-10	0.43-0.92	0.57-0.9	(Barker & Bundy 2000)

Technical difficulties with *de novo* isolation and PCR amplification makes cross-species testing an effective alternative (Temperley *et al.* 2009) for marker development. Since flanking regions of microsatellites may be conserved across some taxa, testing transferability of loci to other organisms is beneficial since it can significantly reduce costs of marker development. Some of the markers developed for *S. lupi* successfully amplified loci from *Cylicospirura felineus*, *Philonema oncorhynchi* and *Gongylonema pulchrum*. Although levels of polymorphism must still be determined, this result indicates that many other loci can be tested from the remaining set of 233 primer pairs to conduct studies in these organisms.

No amplification was observed in *Cylicospirura subaequalis* and *Toxocara canis* and since microsatellites cannot be applied, in most cases, across distantly-related species (Barbará *et al.* 2007), this may indicate that these genera are very divergent from *S. lupi*. This may also reflect the evolutionary distance between these species – Table 2.6 indicates that *T. canis* is in the order Ascaridida whereas the other worms that were tested belong to the order Spirurida. It was found previously that *Cylicospirura felineus* and *Cylicospirura subaequalis* grouped closely with and formed a clade with *S. lupi* within the order Spirurida, based on the sequence of the *cox1* gene (Ferguson *et al.* 2011). However, since the markers developed in this study amplified in one species of C*ylicospirura* and not the other, this supports the conclusion reached by Ferguson *et al.* (2011) that C*ylicospirura felineus* and *Cylicospirura subaequalis* are in fact separate species.

Table 2.6: Summary of the classification of species tested for cross-species amplification, based on information obtained from the NCBI Taxonomy Browser

Organism	Phylum	Class	Order	Superfamily	Family	Genus
Spirocerca lupi	Nematoda	Chromadorea	Spirurida	Thelazioidea	Thelaziidae	Spirocerca
Philonema oncorhynchi	Nematoda	Chromadorea	Spirurida	Dracunculoidea	Philometridae	Philonema
Toxocara canis	Nematoda	Chromadorea	Ascaridida	Ascaridoidea	Toxocaridae	Toxocara
Gongylonema pulchrum	Nematoda	Chromadorea	Spirurida	Spiruroidea	Gongylonematidae	Gongylonema
Cylicospirura felineus	Nematoda	Chromadorea	Spirurida	Spiruroidea	Spirocercidae	Cylicospirura
Cylicospirura subaequalis	Nematoda	Chromadorea	Spirurida	Spiruroidea	Spirocercidae	Cylicospirura

Parasites are usually identified based on their morphological features, type of host they infect, effects on the host, pathological effects, transmission patterns and geographical origin (Gasser 2006). However, these cues may sometimes be insufficient for proper identification. In this study, nematodes obtained from jackal were

identified morphologically as *S. lupi*. Testing with the microsatellite markers developed in this study provided additional support that they were correctly identified. All the loci designed for *S. lupi* in dogs amplified in putative *S. lupi* adults isolated from jackal, with similar allele ranges. This result indicates that *S. lupi* is in fact found in jackal, as has been reported in very early studies on *S. lupi* (Bailey 1972), thus making it a concern for wild populations. In Iran, jackal were previously found to be infested with *T. canis*, *A. caninum*, *D. repens* and other helminth parasites (Sadighian 1969) and only later was *S. lupi* detected in jackal (Meshgi *et al.* 2009). A better understanding of the genetic structure of the parasite population in domestic dogs will allow us to predict the implications of infestation in wildlife populations. This finding also opens new avenues for future work surrounding parasite gene flow between wild and domesticated populations.

2.6. CONCLUSION

The low cost and minimal time spent in developing microsatellites using the FIASCO and 454 sequencing method compared to cloning and traditional sequencing methods, has proven this method to be a powerful tool for population genetics studies in non-model species. The technique employed to isolate microsatellite sequences allowed the identification of a large number of potentially polymorphic loci for the nematode *S. lupi*. Although only 20 microsatellite loci were analysed, there are still 213 microsatellite-containing sequences available for testing, without having to perform any cloning or sequencing. Very stringent criteria were used to obtain the final list of 233 primer sets, thus there is a possibility that if the criteria are relaxed slightly, there may even be additional primers. Only a single multiplex reaction is required to genotype individuals across all nine loci, making genotyping very efficient and cost-effective. Since cross-species utility testing revealed that some markers are transferable between species, it is likely that more primers for related species can be obtained from the data. This study has also provided molecular confirmation of the occurrence of *S. lupi* in jackal from South Africa. These markers can further be used to perform population genetic analyses on adult nematodes from this species to gain more information on disease spread.

2.7. ACKNOWLEDGEMENTS

This work is based on the research supported in part by a number of grants from the National Research Foundation of South Africa: **UID 78566** (NRF RISP grant for the ABI3500) and **UID 80678** (NRF grant to Dr Pamela de Waal). The Grant holders acknowledge that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research are that of the author(s), and that the NRF accepts no liability whatsoever in this regard.

This work is also based on the research supported by the University of Pretoria, South Africa. The authors would like to thank all the veterinarians and researchers involved in collecting samples for this study. Special thanks to Jayde Ferguson (Oregon, USA) and Soraya Naem (Urmia, Iran) for providing samples for cross-species tests.

2.8. REFERENCES

Abdelkrim J, Robertson BC, Stanton JL, Gemmell NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* **46**, 185-191.

Bailey WS (1972) *Spirocerca lupi*: A Continuing Inquiry. *The Journal of Parasitology* **58**, 3-22.

Balzer S, Malde K, Jonassen I (2011) Systematic exploration of error sources in pyrosequencing flowgram data. *Bioinformatics* **27**, i304-i309.

Barbará T, Palma-Silva C, Paggi GM, *et al.* (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Molecular Ecology* **16**, 3759-3767.

Barker GC, Bundy DAP (2000) Isolation and characterization of microsatellite loci from the human whipworm Trichuris trichiura. *Molecular Ecology* **9**, 1181-1183.

Belanger DH, Parkins SL, Rockwell RF (2011) Inference of Population Structure and Patterns of Gene Flow in Canine Heartworm (*Dirofilaria Immitis*). *Journal of Parasitology* **97**, 602-609.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Brouat C, Tatard C, Machin A, *et al.* (2011) Comparative population genetics of a parasitic nematode and its host community: The trichostrongylid *Neoheligmonella granjoni* and Mastomys rodents in southeastern Senegal. *International Journal for Parasitology* **41**, 1301-1309.

Castoe TA, Poole AW, Gu W, *et al.* (2010) Rapid identification of thousands of copperhead snake (Agkistrodon contortrix) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Molecular Ecology Resources* **10**, 341-347.

Chandrasekharon KP, Sastry GA, Menon MN (1958) Canine spirocercosis with special reference to incidence and lesions. *The British Veterinary Journal* **114**, 388-395.

Criscione CD, Anderson JD, Raby K, *et al.* (2007) Microsatellite Markers for the Human Nematode Parasite *Ascaris lumbricoides*: Development and Assessment of Utility. *Journal of Parasitology* **93**, 704-708.

Criscione CD, Poulin R, Blouin MS (2005) Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology* **14**, 2247-2257.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* **93**, 504-509.

Deter J, Galan M, Charbonnel N, Cosson JF (2009) Characterization and PCR multiplexing of polymorphic microsatellite loci in the whipworm Trichuris arvicolae, parasite of arvicoline rodents and their

cross-species utilization in T. muris, parasite of murines. *Molecular and Biochemical Parasitology* **167**, 144-146.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* **107**, 1-15.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* **5**, 435-445.

Excoffier L, G. L, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

Faircloth BC (2008) MSATCOMMANDER : detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**, 92-94.

Ferguson JA, Woodberry K, Gillin CM, *et al.* (2011) *Cylicospirura* species (Nematoda: *Spirocercidae*) and stomach nodules in cougars (*Puma concolor*) and bobcats (*Lynx rufus*) in Oregon. *Journal of Wildlife Diseases* **47**, 140-153.

Gardner MG, Fitch AJ, Bertozzi T, Lowe AJ (2011) Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources* **11**, 1093-1101.

Gasser RB (2006) Molecular tools—advances, opportunities and prospects. *Veterinary Parasitology* **136**, 69-89.

Glenn TC, Schable NA, Elizabeth AZ, Eric HR (2005) Isolating Microsatellite DNA Loci. In: *Methods in Enzymology*, pp. 202-222. Academic Press.

Grillo V, Jackson F, Gilleard JS (2006) Characterisation of *Teladorsagia circumcincta* microsatellites and their development as population genetic markers. *Molecular and Biochemical Parasitology* **148**, 181-189.

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series* **41**, 95-98.

Hoekstra R, Criado-Fornelio A, Fakkeldij J, Bergman J, Roos MH (1997) Microsatellites of the parasitic nematode *Haemonchus contortus*: polymorphism and linkage with a direct repeat. *Molecular and Biochemical Parasitology* **89**, 97-107.

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

Johnson PCD, Webster LMI, Adam A, *et al.* (2006) Abundant variation in microsatellites of the parasitic nematode *Trichostrongylus tenuis* and linkage to a tandem repeat. *Molecular and Biochemical Parasitology* **148**, 210-218.

Kok DJ, Schenker R, Archer NJ, Horak IG, Swart P (2011) The efficacy of milbemycin oxime against pre-adult Spirocerca lupi in experimentally infected dogs. *Veterinary Parasitology* **177**, 111-118.

La Rosa G, Marucci G, Rosenthal BM, Pozio E (2012) Development of a single larva microsatellite analysis to investigate the population structure of Trichinella spiralis. *Infection, Genetics and Evolution* **12**, 369-376.

Lai Y, Sun F (2003) The Relationship Between Microsatellite Slippage Mutation Rate and the Number of Repeat Units. *Molecular Biology and Evolution* **20**, 2123-2131.

Leach MM, Agudelo PA, Lawton-Rauh A (2012) Genetic Variability of *Rotylenchulus reniformis*. *Plant Diseases* **96**, 30-36.

Loiseau A, Duplantier JM, Kane M, Brouat C (2010) Isolation and characterisation of microsatellites in a rodent nematode parasite, *Neoheligmonella granjoni*. In: *Molecular Ecology Resources*, pp. 1-8.

Malausa T, Gilles A, Meglecz E, *et al.* (2011) High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Molecular Ecology Resources*, 1-7.

Mazaki-Tovi M, Baneth G, Aroch I, *et al.* (2002) Canine spirocercosis: clinical, diagnostic, pathologic, and epidemiologic characteristics. *Veterinary Parasitology* **107**, 235-250.

Meglecz E, Neve G, Biffin E, Gardner MG (2012) Breakdown of Phylogenetic Signal: A Survey of Microsatellite Densities in 454 Shotgun Sequences from 154 Non Model Eukaryote Species. *PloS one* **7**, e40861.

Meshgi B, Eslami A, Bahonar AR, Kharrazian-Moghadam M, Gerami-Sadeghian A (2009) Prevalence of parasitic infections in the red fox (*Vulpes vulpes*) and golden Jackal (*Canis aureus*) in Iran. *Iranian Journal of Veterinary Research* **10**, 387-391.

Mlonyeni XO, Wingfield BD, Wingfield MJ, *et al.* (2011) Extreme homozygosity in Southern Hemisphere populations of *Deladenus siricidicola*, a biological control agent of *Sirex noctilio*. *Biological Control* **59**, 348-353.

Molnar RI, Witte H, Dinkelacker I, Villate L, Sommer RJ (2012) Tandem-Repeat Patterns and Mutation Rates in Microsatellites of the Nematode Model Organism Pristionchus pacificus. *G3: Genes|Genomes|Genetics* **2**, 1027-1034.

Nadler SA (1995) Microevolution and the Genetic Structure of Parasite Populations. *Journal of Parasitology* **81**, 395-403.

Otsen M, Plas ME, Lenstra JA, Roos MH, Hoekstra R (2000) Microsatellite diversity of isolates of the parasitic nematode *Haemonchus contortus*. *Molecular and Biochemical Parasitology* **110**, 69-77.

Perry JC, Rowe L (2011) Rapid Microsatellite Development for Water Striders by Next-Generation Sequencing. *Journal of Heredity* **102**, 125-129.

Pfaff CL, Parra EJ, Bonilla C, *et al.* (2001) Population Structure in Admixed Populations: Effect of Admixture Dynamics on the Pattern of Linkage Disequilibrium. *The American Journal of Human Genetics* **68**, 198-207.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-846.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity* **86**, 248-249.

Redman E, Packard E, Grillo V, *et al.* (2008) Microsatellite analysis reveals marked genetic differentiation between *Haemonchus contortus* laboratory isolates and provides a rapid system of genetic fingerprinting. *International Journal for Parasitology* **38**, 111-122.

Reid K, Hoareau TB, Bloomer P (2012) High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. *Molecular Ecology Resources* **12**, 740-752.

Rosen S, Skaletsky H (2000) Primer3 on the WWW for General Users and for Biologist Programmers. *Methods in Molecular Biology: Bioinformatics Methods and Protocols* **132**, 365-386.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Sadighian A (1969) Helminth Parasites of Stray Dogs and Jackals in Shahsavar Area, Caspian Region, Iran. *The Journal of Parasitology* **55**, 372-374.

Santana QC, Coetzee MPA, Steenkamp ET, *et al.* (2009) Microsatellite discovery by deep sequencing of enriched genomic libraries. *BioTechniques* **46**, 217-223.

Schwenkenbecher JM, Kaplan RM (2007) Development and characterization of microsatellite markers for the canine hookworm, *Ancylostoma caninum*. *Parasitology Research* **100**, 1015-1021.

Scoble J, Lowe A, Gardner M (2011) Isolation via 454 sequencing, and characterisation of microsatellites for *Drymodes brunneopygia*, southern scrub-robin (Aves: Petroicidae): a species at risk due to substantial habitat loss and climate change. *Conservation Genetics Resources*, 1-3.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615–629.

Techen N, Arias RS, Glynn NC, *et al.* (2010) Optimized construction of microsatellite-enriched libraries. *Molecular Ecology Resources* **10**, 508-515.

Temperley ND, Webster LMI, Adam A, Keller LF, Johnson PCD (2009) Cross-Species Utility of Microsatellite Markers in Trichostrongyloid Nematodes. *Journal of Parasitology* **95**, 487-489.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. *Nucleic Acids Research* **25**, 4876-4882.

Tóth G, Gáspári Z, Jurka J (2000) Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Research* **10**, 967-981.

Van der Merwe LL, Kirberger RM, Clift SJ, *et al.* (2008) *Spirocerca lupi* infection in the dog: A review. *The Veterinary Journal* **176**, 294-309.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

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

Wasimuddin, Cizcova D, Ribas A, *et al.* (2012) Development and characterization of multiplex panels of microsatellite markers for *Syphacia obvelata*, a parasite of the house mouse (*Mus musculus*), using a high throughput DNA sequencing approach. *Molecular and Biochemical Parasitology*.

Wielgoss S, Sanetra M, Meyer A, Wirth T (2007) Isolation and characterization of short tandem repeats in an invasive swimbladder nematode, parasitic in Atlantic freshwater eels, Anguillicola crassus. *Molecular Ecology Notes* **7**, 1051-1053.

Xiang L, Guo F, Zhang H, *et al.* (2013) Gene discovery, evolutionary affinity and molecular detection of *Oxyspirura petrowi*, an eye worm parasite of game birds. *BMC Microbiology* **13**, 233.

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

Zhan A, Bao Z, X. H, W. L, J. H (2009) Methods Comparison for Microsatellite Marker Development: Different Isolation Methods, Different Yield Efficiency. *J. Ocean Univ. China (Oceanic and Coastal Sea Research)* **8**, 161-165.

2.9. APPENDICES

Appendix 2A Sequences of microsatellite loci obtained from 454 sequencing of microsatellite-enriched *S. lupi* genome. Forward and reverse primer binding sites and repeat motifs are indicated in bold text.

SL02

CCATCCTCTTGCGTTGCACCATGTGATTGTTTCCCAAATATTTCCTGTTCTTCCTGATCTC**ATCATCATCATCATCATCATCATCATCA**
TCGTCGTCTTTACTTCACAACGAAACTTCTTGCATCAATAATAACTCGATGAGACGAAGCGTGTCCTGTATTCTGTTATATAATTTCGC
CTTCCTTCCCCCAGCCGTGGATTCAGTGTCATGGAGATGGGGTGACAGTTGCTGGTGGGCAGTATTGGGGACATACGTCTACTGCTTCG
AAGTCGAGACGGCCCAAATCGCTGACAGCTTGGCCGTCGGATGGATTCATCATCGGTACGCTTATTCCCGGG**GGTAATGGCCTGCACAG**
GA

SL04

GAACGGTTTCCGCGAACTCTGCGCCCGGTCGCAGAATTATCGGACGGGATGAGATCGGAAATTCCAGGAATCAGCGTACC**AGCAGCAGC**
AGCAGCAGCAGCAGCAGTCACTTTTACGAGGTCAGGTTACCTGGCTCAGTTGACT**GACAACGTTCAGACTTTCGG**

SL06

GAGATTGGCCGGAAAGGTGCAGTACGTAGCTACGGTTATCATACAACCGGAAGATAATTGTTTCGGAATGTGCAGATGACATGTTCAGG
CTATTGGCCACCAGAACAGCC**AACAACAACAACAACAACAACAACAAC**ACGCAGACGGTGATGGGCTATCCGTCATTTGGCTGGCTCCC
TTTCAGCGTTTACGATTATTGTCGCCTGTCCCTGTCCCATTAGCGAACCACAAAAGAAATGAACATGATGATGCACTGGGCCTCAGTTA
CTCCGTCGTGTACAAACTGACGTATTTGCTCCTCCGGTGTAAGTGCTTCTGGAAAAGGGGGTCAGGCTGTAATAGCATGTG**GCTTTCGG**
GCAATGCTGTA

SL10

CTCCCTGGAATCTATTTGCCCTCAAGAGACAGAGTATTCTGTGTAGCTTACCCGAGACAGATACTCAACATTTCCCAGAAGTCTGCCCT
CC**AGAGAGAGAGAGAGAGAGAG**AAGAGTATTTTGTGTAGCTTATCTGAGAGAGATATCCGAAATTTCCCTGAATTCTGCCTTCCAAAGA
GAGTATTCTATGTAAGCTTATTACAAAATACCCATAGTA**GCTGATCCCTAACAGTGCT**

SL13

ATACCGTTTCGGTGCCAAGTTGCACATCCTTTCGTCCCAAAACG**ATATATATATATATATAT**TGCATACGGTTGTTCAGCGTTGGTATACT
TTGTAGCGAGGCAACGCGCTACGCGACTTCCACATATTTCACGCTAATGCCAGCCGTGGTCGCGCTGTGCAATTCCTAAGTGTTACACA
TCACAATGTGCAGACCGTGGTCCACCTCGCGCTATAGCAAA**GTCAACTGTGAGTGCCGC**

SL14

CCGAGGGTACTCGATGTGGTTAGGCGGGATTGAAAACGTGTTTATTTCTTGATTGTACTAACAAAATCCATCATTCA**ACACACACACAC**
ACACACACATGGTGCTTTGGGAAGAGGGAGGAGTAAAAGTTAGGTTGATGAACGAGAACACGATAACGGACTACCACCACCATCAACAT
CATGACATCCTGTGACATTATGAAATT**CATCAAGACTGCTCGGGCT**

SL15

CTGTTGGTGGTCCATTTCGGTACTTTTCCTGTCGTCAACGTCTTCTGTGTTTGTGTGTATGTGC**GTGTGTGTGTGTGTGTGTGTGTGTG**
TGTCAGTATCAGTATCAGTGGGGTCAGTCGCCTTGCTACTTAATTTCTGTAGTGCCTAGTTGCGGTATTACCGAAGCTATTTCCGTTAT
T**GGACTGTTGCGACCATCTG**

SL17

TCAACCAATCTGGCGCAACATTTTTTCAGTCTCCTCTCGGCATTTGATAGTGCTGGTTGAGCGACGACCGGCTGAAAACTGTTCCACAG
TGGGAAAAGTTGATCCGGTTGTCCGAAACCTGGAGCAGTGCTACCCGGAACCGGAACTCCTCCTCCT**GCTGCTGCTGCTGCTGCTGCTG**
CTCCCTGATTGACGACGGGCAGTGATTGTTGTGGAGCA**GCTGTTGAAGGACGAACGG**

SL18

GAGGCGATTGTTGCGTTCTTTTCATCCTTG**GATGATGATGATGATGATGATGAT**GAAATTTGTTTGAAGGTAGCCGATATGTCCGTACC
ATCCGGTTTTGCCGACGACGCAGCTAGCTTCGACAAATCATCAAATTCCGGATATTCACCAATCATTGCCTGAAATAAACTTAGTTGTA
TCATTTATTACACTGTAGTACGGTCTCTAATTATAACTATTTTCCTAAGATCTGAATACTAGAAATTTACAATATTGGTGTGCTGGTGC
TCCACCCTCCCGATCTTTTGAGAATT**CTGGAAACGTGATGTCGCT**

SL20

GAGCTTTCCGTGGTTCTGCAAAAAAAAGTGCACGATTTCGAAAGATATTTTTCACGATCAGTGAATATGATGAGCATAAGTTT**GTGTGT**
GTGTGTGTGTGTGTGTGTTCGTTCGTTCGTTCGGTTAGAACTGATAACGATAACATATACAAACATTATTTTGTTTACTGATCCATAC**G**
AATTACGCCCATCCAACCG

Appendix 2B Homozygotes sequenced to determine flanking sequence variability and to compare observed allele size with sequenced repeat length. Allele length shown in parentheses on left of sequence and number of repeats in each sequenced allele shown on the right. Note that no sequences were obtained for marker SL18 and allele lengths did not correspond to the number of repeats for marker SL20.

SL02

(356) TTCCTGATCTC**ATCATCATCATCATCATCGTCGTCGTC**---TTTTACTTCACAACGAAACTTCTTGCATCAA (9)
(359) TTCCTGATCTC**ATCATCATCATCATCATCATCGTCGTCGTC**TTTTACTTCACAACGAAACTTCTTGCATCAA (10)
(359) TTCCTGATCTC**ATCATCATCATCATCATCATCATCGTCGTC**TTTTACTTCACAACGAAACTTCTTGCATCAA (10)

SL04

(162) AGATCGGAAATTCCAGGAATCAGCGTACC**AGCAGCAGCAGCAGCAGCAGCAGC**ACTTTTACAGGCAGTT (9)

SL06

(361) TGGCCACCAGAACAGCC**AACAACAACAACAACAACAACAAC**----------ACGCAGACGGTGATGGGCT (8)
(355) TGGCCACCAGAACAGCC**AACAACAACAACAACAAC**----------ACGCAGACGGTGATGGGCT (6)
(373) TGGCCACCAGAACAGCC**AACAACAACAACAACAACAACAACAACAACAAC**ACGCAGACGGTGATGGGCT (12)
(364) TGGCCACCAGAACAGCC**AACAACAACAACAACAACAACAACAAC**---------ACGCAGACGGTGATGGGCT (9)

SL10

(230) CATTTCCCAGAAGTCTGCCCTCC**AGAGAGAGAGAGAGAGAGAG**----TATTCTGTGTAGCTTATCTGAGAGA (10)
(228) CATTTCCCAGAAGTCTGCCCTCC**AGAGAGAGAGAGAGAGAG**------TATTTTGTGTAGCTTATCTGAGAGA (9)
(234) CATTTCCCAGAAGTCTGCCCTCC**AGAGAGAGAGAGAGAGAGAGAGAG**TATTTTGTGTAGCTTATCTGAGAGA (12)

SL13

(231) NNNNNNNN**ATATATATAT**----------TGCATACGGTTGTTCAGCGTTGGTATAGT (5)
(239) NNNNNNNN**ATATATATATATATATAT**----------TGCATACGGTTGTTCAGCGTTGGTATAGT (9)

SL14

(229) ACTAACAAAATCCATCATTCA**ACACACACACACACACACACAC**ATGGTGCTTTGGGAAGAGGGAGGAGTAAA (11)

(227) ACTAACAAAATCCATCATTCA**ACACACACACACACACACACA—**TGGTGCTTTGGGAAGAGGGAGGAGTAAAA (10)

(227) ACTAACAAAATCCATCATTCA**ACACACACACACACACACACA—**TGGTGCTTTGGGAAGAGGGAGGAGTAAAA (10)

SL15

(177) GTGTGTATGTGC**GTGTGTGTGT**----------CAGTATCAGTATTATGGGGGTCAGTG (5)

(185) GTGTGTATGTGC**GTGTGTGTGTGTGTGTGT**----------CAGTATCAGTATTATGGGGGTCAGTG (9)

(195) GTGTGTATGTGC**GTGTGTGTGTGTGTGTGTGTGTGTGT**------CAGTATCAGTATTATGGGGGTCAGTG (14)

SL17

(237) GAACTCCTCCTCCT**GCTGCTGCTGCTGCTGCTGCTGCT**---CCCTGATTGACGACGGGCAGTGGTTGTTGTG (8)

(231) GAACTCCTCCTCCT**GCTGCTGCTGCTGCTGCT**---------CCCTGATTGACGACGGGCAGTGGTTGTTGTG (6)

SL20

(---) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGTGTGTGT**--TCGTTCGTTCGTTCGGTTAGACACTGAT (10)

(187) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGT**----------TCGTTCGTTCGTTCGGTTAGACACTGAT (6)

(---) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGTGTGTGT**--TCGTTCGTTCGTTCGGTTAGACACTGAT (10)

(---) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGTGTGTGT**--TCGTTCGTTCGTTCGGTTAGACACTGAT (10)

(187) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGT**----------TCGTTCGTTCGTTCGGTTAGACACTGAT (6)

(197) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGTGTGTGTGT**TCGTTCGTTCGTTCGGTTAGACACTGAT (11)

(187) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGT**---------TCGTTCGTTCGTTCGGTTAGACACTGAT (7)

(189) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGTGT**------TCGTTCGTTCGTTCGGTTAGACACTGAT (8)

(185) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGT**----------TCGTTCGTTCGTTCGGTTAGACACTGAT (6)

(185) TGAATATGATGAGCATAAGTTT**GTGTGTGTGTGT**----------TCGTTCGTTCGTTCGGTTAGACACTGAT (6)