Molecular genomics resource for the parasitic nematode *Spirocerca lupi*: Identification of 149 microsatellite loci using FIASCO and Next Generation Sequencing

Kerry Reid[a], Janishtha R Mitha[b], Jaco M Greeff[b] and Pamela J de Waal[b*]

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

[b]Evolution of Interactions Group, Department of Genetics, University of Pretoria, Pretoria 0002, South Africa

*Corresponding author: Dr P J de Waal, Tel.: +27124203949; fax: +27123625327

Postal address: Department of Genetics, Private Bag X20, Hatfield, 0028

E-mail address: pam.dewaal@up.ac.za

Abstract

Understanding genetic diversity and movement patterns in parasitic organisms is paramount to establish control and management strategies. In this study we developed a microsatellite resource as well as a diagnostic multiplex for the cosmopolitan parasitic nematode *Spirocerca lupi*, known to cause spirocercosis in canids. A combination of microsatellite enrichment and 454 sequencing was used to identify 149 unique microsatellite loci in *S. lupi.* Twenty loci were characterized further in two sampling sites in South Africa, with 10 loci identified as polymorphic (allele ranges from 4-17). These loci were designed into a single diagnostic multiplex suitable for species identification and population genetics studies. The markers were also successful in cross-species amplification in *Cylicospirura felineus*, *Philonema oncorhynchi* and *Gongylonema pulchrum*. Our resource provides a large set of candidate loci for a number of nematode studies as well as loci suitable for diversity and population genetics studies of *S. lupi* within the South African context as well as globally.

Keywords: 454 sequencing, canids, cross-species amplification, microsatellites, parasitic nematode, South Africa,

1. Introduction

Spirocerca lupi (order: Spirurida) is a parasitic nematode which is known to cause spirocercosis in canids globally [1]. It is a severe and debilitating disease for canine populations worldwide and shows a higher prevalence in tropical regions [2,3]. Although mostly restricted to domestic canid populations, there have also been cases of spirocercosis occurring in wild canid populations [4].

Several studies have indicated the severity of the disease and the requirements for characterizing genetic diversity in parasitic nematodes for biological and management purposes [5,6,7]. There are to date no nuclear studies available for *S. lupi*, however, two previous studies based on mitochondrial *cox*1 found different results in vastly contrasting geographic ranges in terms of diversity of the species as well as within host diversity [5,6]. Although mitochondrial markers are informative, microsatellites have become the marker of choice for population genetics studies due to their cost effective, quick genotyping and their utility as a diagnostic tool for species identification and genetic diversity assessment [7].

The aim of this study was to develop a resource of potentially polymorphic microsatellite loci for *S. lupi* using a microsatellite enrichment protocol and Roche 454 sequencing. These data were mined for all candidate loci. A subset of loci were further characterized and designed into a single diagnostic multiplex which were tested on samples from the South Africa as well as closely related species.

2. Materials and Methods

Isolation of microsatellite loci using FIASCO and NGS sequencing

Spirocerca lupi nematodes were obtained from post-mortems on dogs conducted by veterinarians and pathologists in Gauteng (85 samples) and the Eastern Cape (35 samples) in South Africa. These nematodes were stored in 70% ethanol at 4°C. The DNeasy Blood and Tissue kit (Qiagen, Hilden)

was used to extract total genomic DNA from a small segment of each nematode following the manufactures specifications. Before the enrichment for microsatellites was performed, seven individual nematode extractions were pooled in order to increase the total amount of DNA and avoid ascertainment bias [8].

The Fast Isolation by AFLP of Sequences COntaining repeats (FIASCO) microsatellite enrichment protocol [9] was performed in three separate reactions which included the following repeat motifs $[(AC)_{15}, (AG)_{15}, (AT)_{15}]_1$, $[(TAT)_{10}, (TGT)_{10}, (AAG)_{10}, (ATC)_{10}]_2$, $[(AAAT)_8, (TTTG)_8]_3$ commonly found in nematode species. Equal concentrations of the three nucleotide enrichments (PCR products) were pooled together and sequenced by Inqaba Biotec (Pretoria, South Africa) on a 454 LifeSciences/ Roche GS-FLX pyrosequencing platform using $1/16^{th}$ of a PicoTiter plate.

BIOEDIT version 5.0.6. [10] was used to detect and discard all sequences shorter than 60 bp in length and to remove adaptors. Sequences were reverse complemented and aligned with CLUSTAL X version 1.81 [11] to identify and remove duplicated repeat fragments. Sequences containing repeats (SCRs) were identified with MSATCOMMANDER [12] using the following search parameters: minimum number of repeats for di- (8), tri- (8), tetra- (6), penta- (5), hexa- (5) and a maximum for all repeat types of 30. Primers were designed using default parameters of PRIMER3 version 1.1.1 [13]. Primers that were not immediately adjacent to the repeat unit and that had a length of expected amplicon size between 100-400 bp were selected.

Identification of polymorphic microsatellites in S. lupi

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 amplifications of a final volume of 10 µl 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. The PCR conditions were 95 °C for 2 min followed by 30 cycles of 95°C for 1 min, 50-63°C for 30s and 72°C for 30s and a final elongation step at 72°C for 5 min. Primer pairs that produced consistent amplification of the expected fragment sizes were tested on six *S. lupi* individuals to determine whether loci were polymorphic. These individuals were re-amplified with PCR reactions containing 0.02 pmol ChromaTide Alexa Fluor® 488-5-dUTP (Invitrogen), which are fluorescently labelled dUTPs that incorporate during PCR cycles allowing the fragments to be visualised on an automated sequencer [14]. Genotyping 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).

The G5 dye set (6-FAM, VIC, NED and PET; Applied Biosystems) was used to fluorescently label the forward primers of loci that were found to be polymorphic (Table 1). All loci were combined into a single multiplex reaction and amplified using the Quantitect Multiplex PCR kit (Qiagen) following the protocol described by the manufacturer. One hundred and twenty *S. lupi* individuals were genotyped, including an additional random subset of 22 individuals to assess allele migration and genotyping error. The genotypes for each individual were determined by analysing chromatograms in GENEMARKER™ version 2.4.0. (SoftGenetics, State College, Pennsylvania, USA).

Levels of null alleles, scoring errors due to stuttering and large-allele dropout were assessed in MICROCHECKER version 2.2.3 [15]. A test for linkage disequilibrium was conducted in GENEPOP version 1.2. [16]. Summary statistics including observed heterozygosity (H_O), expected heterozygosity (H_E) and Weir and Cockerham's estimate of F_{IS} were computed using GENETIX version 4.05.2. [17]. Deviations from Hardy-Weinberg equilibrium for each locus and globally, were calculated with ARLEQUIN version 3.5.1[18] with 1 000 000 Markov chain steps. Finally

Table 1: 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, values in bold indicate significant p-values)

Locus	**Dye**	**Primer sequence (5'-3')**	**Tm**	**Repeat**	**Allele**	**Gauteng (N = 85)**				**Eastern Cape (N = 35)**			
			(°C)	**motif**	**range (bp)**	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}
SL02	PET	F: CCATCCTCTTGCGTTGCAC R: TCCTGTGCAGGCCATTACC	60	$(ATC)_{10}$	344-380	9	0.717	0.683	0.053	5	0.567	0.559	0.029
SL04	PET	F: GAACGGTTTCCGCGAACTC R: CCGAAAGTCTGAACGTTGTC	59	$(AGC)_8$	147-162	6	0.447	0.434	0.035	4	0.506	0.543	-0.058
SL06	VIC	F: GAGATTGGCCGGAAAGGTG R: TACAGCATTGCCCGAAAGC	59	$(AAC)_9$	352-373	8	0.793	0.655	**0.180**	7	0.741	0.618	**0.180**
SL10	6-FAM	F: CTCCCTGGAATCTATTTGCCC R: AGCACTGTTAGGGATCAGC	58	$(AG)_{10}$	220-242	9	0.761	0.583	**0.239**	6	0.663	0.706	-0.050
SL13	VIC	F: ATACCGTTTCGGTGCCAAG R: GCGGCACTCACAGTTGAC	59	$(AT)_8$	231-249	9	0.673	0.735	-0.086	5	0.673	0.697	-0.021
SL14	PET	F: CCGAGGGTACTCGATGTGG R: AGCCCGAGCAGTCTTGATG	60	$(AC)_{10}$	221-263	17	0.831	0.831	0.006	7	0.723	0.882	-0.206
SL15	VIC	F: CTGTTGGTGGTCCATTTCGG R: CAGATGGTCGCAACAGTCC	60	$(GT)_{14}$	183-209	12	0.668	0.565	**0.160**	5	0.209	0.226	-0.063
SL17	NED	F: TCAACCAATCTGGCGCAAC R: CCGTTCGTCCTTCAACAGC	60	$(GCT)_8$	210-240	9	0.772	0.691	0.111	4	0.719	0.516	**0.297**
SL18	6-FAM	F: TGAGGCGATTGTTGCGTTC	59	$(GAT)_8$	297-318	7	0.783	0.869	-0.104	6	0.749	0.914	-0.206

		R: AGCGACATCACGTTTCCAG												
SL20	6-FAM	F: GAGCTTTCCGTGGTTCTGC R: CGGTTGGATGGGCGTAATTC	60	$(GT)_{12}$	184-196	-	-	-	-	-	-	-	-

POWSIM [19] was used to assess the statistical power of the loci when detecting various genetic divergence values (F_{ST} ranging from 0.0025-0.04).

Cross-species amplification

Microsatellite loci that were identified as polymorphic in *S. lupi* were used to evaluate their success in amplifying DNA from other species within the order Spirurida. One worm from each of the following species (hosts in brackets) was tested: *Cylicospirura felineus* (bobcats, Oregon,USA), *Philonema onco oncorhynchi* (sockeye salmon, Alaska, USA) and *Gongylonema pulchrum* (sheep, Urmia, Iran).

3. Results and discussion

Microsatellite identification and characterization

The 454 sequencing yielded 36,482 sequence reads of which 21,390 sequences (58.63%) contained repeat motifs. After applying restrictions 7,835 (21.48%) sequences contained repeats with 149 primers sets being designed for unique loci (Table S1). These included 109 di-, 31 tri, 3 tetra, 1 penta and 5 hexanucleotide loci. Of the twenty loci tested, ten loci were found to be polymorphic and could be combined into a single multiplex reaction, with loci at least 30 bp apart.

One locus (SL 20) had ambiguous scoring which resulted in missing data and was therefore excluded from further analyses. No evidence was found for scoring errors due to stuttering or large allele dropout and only low levels of potential null alleles were detected in one locus (SL17, 13%). No loci showed significant deviation from Hardy-Weinberg equilibrium and F_{IS} values were generally negligible (ranging from 0.206 to 0.239) as indicated in Table 1. Summary statistics showed alleles ranging from 4-17 per sampling site, with H_O ranging from 0.447 – 0.831 and H_E from 0.226 – 0.831 (Table 1). The POWERSIM analysis showed a hundred percent significance call rate (for chi_2 and Fisher's exact test) when genetic differentiation of more than 0.0075 was tested, indicating that these loci are suitable for detecting even subtle genetic structure (Figure 1).

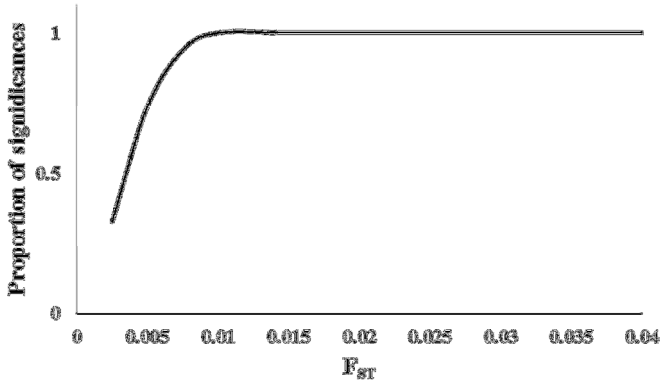

Fig. 1. POWSIM analyses showing simulated estimates of power for various F_{ST} values of the nine microsatellite loci (based on the allele frequencies and sample sizes of this study).

Cross-species amplification

Technical difficulties with *de novo* isolation and PCR amplification makes cross-species testing an effective alternative [20] for marker development. Since flanking regions of microsatellites may be conserved across some taxa, testing transferability of loci to other organisms is beneficial as it can significantly reduce costs of marker development. 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). Although levels of polymorphism must still be determined, this result indicates that many other loci can be tested from the remaining set of loci to conduct studies in these organisms.

Utility of microsatellite markers in S. lupi

This study showed the efficiency of using a combination of FIASCO and 454 in a parasitic nematode to identify a large number of potentially polymorphic loci. The low cost and minimal time spent in developing microsatellites using this 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 [14, 21]. The results obtained in this study (58% SCRs, 149 unique loci) are comparable to other studies using NGS sequencing to develop microsatellites [summarised, 12] and specifically to other parasitic nematodes (e.g. *Deladenus siricidicola* (40% SCRs, 269 unique loci) using the same method [21].

A large number of polymorphic microsatellite loci were identified with stringent search criteria (Table S1) and there are likely to be additional loci if the parameters are relaxed. Although

only 20 microsatellite loci were analysed, there are still 129 repeat-containing (Table S1) sequences available for testing, without having to perform any cloning or sequencing.

Only a single multiplex reaction was required to genotype individuals across all nine loci, making genotyping very efficient due to the low cost and small amount of DNA required. The loci developed in this study are suitable for advanced population genetics analyses, to understand dispersal potential, mating strategies within and between host organisms and spread of *S. lupi* to new hosts [7]. Finally, these loci will be applicable for studies of this species in other regions globally where this parasitic nematode is a threat to canine populations as well as closely related nematodes through cross-species transferability studies.

Competing interests

The author(s) declare that they have no competing interests.

Acknowledgements

All the veterinarians and parasitologists involved in sample collection for this study are thanked for their valuable contributions - specifically, Dr Sarah Clift, at the Onderstepoort Veterinary Institute for *S. lupi* nematode samples from the Tshwane Metropole and Prof Dawie Kok of ClinVet International Research Organization for the nematode samples from the Grahamstown area. Special thanks to Jayde Ferguson (Oregon, USA) and Soraya Naem (Urmia, Iran) for providing samples for cross-species tests. The Animal and Zoonotic Diseases (AZD) Institutional Research Theme (IRT) at the University of Pretoria (UP) is thanked for financial support. 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 authors, and

that the NRF accepts no liability whatsoever in this regard. This work is also based on the research supported from the University of Pretoria, South Africa.

Supplementary material

Table S1 129 additional loci with amplicon size, primers and sequence

Data deposited

DRYAD: All raw sequence reads

Microsatellite sequences will be submitted to Genbank

4. References

1. Chandrasekharon KP, Sastry GA, Menon MN: **Canine spirocercosis with special reference to incidence and lesions.** *The British Veterinary Journal* 1958, **114:**388-395.
2. Kok DJ, Schenker R, Archer NJ, Horak IG, Swart P: **The efficacy of milbemycin oxime against pre-adult Spirocerca lupi in experimentally infected dogs.** *Veterinary Parasitology* 2011, **177:**111-118.
3. Van der Merwe LL, Kirberger RM, Clift SJ, Williams M, Keller N, Naidoo V: ***Spirocerca lupi* infection in the dog: A review.** *The Veterinary Journal* 2008, **176:**294-309.
4. Al-Sabi MNS, Hansen MS, Chriel M, Holm E, Larsen G, Enemart HL: **Genetically distinct isolates of Spirocerca sp. from a naturally infected red fox (*Vulpes vulpes*) from Denmark.** *Veterinary Parisitology* 2014, *in press*.
5. de Waal PJ, Gous A, Clift SJ, Greeff JM. **High within-host genetic variation of the nematode Spirocerca lupi in a high-density urban dog population**. Veterinary Parisitology 2012, **187:** 259-266
6. Traversa D, Costanzo F, Iorio R, Aroch I, Lavy E: **Mitochondrial cytochrome c oxidase**

subunit 1 (*cox*1) gene sequence of *Spirocerca lupi* (Nematoda, Spirurida):Avenues for potential implications. *Veterinary Parisitology* 2007, **146:** 263-270.

7. Gilabert A, Wasmuth JD: **Unravelling parasitic nematode natural history using population genetics**. *Trends in Parasitology* 2013, **29:**438-448.

8. Glenn TC, Schable NA, Elizabeth AZ, Eric HR: **Isolating Microsatellite DNA Loci.** *Methods in Enzymology. Volume* Volume 395: Academic Press; 2005: 202-222.

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

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

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

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

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

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

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

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

17. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F: **GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.** In *Book 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); 1996-2004.

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

19. Ryman N, Palm S. **POWSIM: a computer program for assessing statistical power when testing for genetic differentiation**. *Molecular Ecology* 2006, **6**:600-602.

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

21. Santana QC, Coetzee MPA, Steenkamp ET, Mlonyeni OX, Hammond GNA, Wingfield MJ, Wingfield BD: **Microsatellite discovery by deep sequencing of enriched genomic libraries.** *BioTechniques* 2009, **46:**217-223.