Title: An integrated pipeline for the development of novel panels of mapped microsatellite markers for *Leishmania donovani* complex, *Leishmania braziliensis* and *Leishmania major.*

Authors: M. Fakhar[1]; M. H. Motazedian[1]; D. Daly[2]; C.D. Lowe[2]; H.A. Noyes[2]; S.J. Kemp[2]

Addresses:

[1] Department of Medical Parasitology and Mycology, School of Medicine, Shiraz University of Medical Sciences. Shiraz, Iran.

[2] School of Biological Sciences, University of Liverpool, Crown St, Liverpool L69 7ZB

Running title: A pipeline for mapped microsatellites for *Leishmania spp*

Address for correspondence:

Harry A. Noyes

Room 231 Biosciences Building

University of Liverpool

Crown Street

Liverpool

L69 7ZB

0151 795 4512

www.genomics.liv.ac.uk/tryps

harry@liv.ac.uk

Summary:

A panel of microsatellites mapped to the *Leishmania* genome might make it possible to find associations between specific loci and phenotypic traits. To identify such loci, a Perl programme was written that scans the sequence of a genome and writes all loci containing microsatellites to a MySQL database. The programme was applied to the sequences of the *L. braziliensis, L. infantum* and *L. major* genomes. The database is publicly available over the internet: http://www.genomics.liv.ac.uk/tryps/resources.html "Microsatellite Locus Extractor", and allows the selection of mapped microsatellites that meet user defined criteria from a specified region of the selected genome. The website also incorporates a primer design pipeline that will design primers to amplify the selected loci. Using this pipeline 12 out of 17 primer sets generated polymorphic PCR products. A tailed primer protocol was used to label all microsatellite primers with a single set of labelled primers. To avoid the culture of parasites prior to genotyping, sets of nested PCR primers were developed to amplify parasite DNA eluted from microscope slides. The limit of detection was approximately 1.6 parasite equivalents. However only 6/56 DNA from slides stored at ambient temperature for over six months gave positive PCR results.

Key words: *Leishmania donovani* complex, *Leishmania braziliensis* and *Leishmania major,*

Leishmania infantum, Leishmania (Viannia), microsatellite, tailed primers, nested PCR, Iran

Introduction:

The leishmaniases are severe vector borne diseases caused by parasites belonging to the genus *Leishmania*. The diseases are caused by about 30 species of *Leishmania* that are transmitted by the bite of female sand flies. Visceral leishmaniasis (VL) is the most severe form of the disease and is frequently fatal if left untreated. The incidence of VL is approximately 0.5 million cases per annum and it is present in both the Old World (principally North and East Africa, Middle East, Asia and Southern Europe) and the New World (principally Central and South of America; Desjeux 2001). Multilocus enzyme electrophoresis (MLEE) has been the gold standard for taxonomy and strain typing of *Leishmania*, but has several limitations including the relatively small number of characterised loci and alleles and the requirement for bulk parasite culture (Kuhls *et al.* 2007). An alternative reliable and fast genotyping method is Multilocus microsatellite typing (MLMT; Kuhls *et al*. 2007).

MLMT is a powerful tool for population genetic and epidemiological studies of *Leishmania spp* (Kuhls *et al*. 2007). Various panels of microsatellite markers have been reported for *Leishmania* species. Jamjoom *et al*. (2002b) reported a panel of 20 microsatellites that are polymorphic in *L. donovani* and *L. infantum.* Kuhls et *al.* (2007) established 15 microsatellites for *L. donovani* complex. Montoya *et al* (2007) developed 8 microsatellites for *L. infantum* in three regions (ITS, Lm4 and Lm2) and detected 17 alleles. Russell *et al.* (1999) screened all parasites of the subgenus *L. (Viannia)*, except *L. naiffi*, with a panel of 3 loci. Bulle *et al.* (2002) found 10 microsatellites markers for *L. infantum*. The microsatellite panel detailed by Jamjoom *et al*. (2002a) was developed using the genome sequence of *L. major*; however, with this exception, most microsatellites have essentially been selected at random. This is perfectly adequate for population analysis and where it is possible to integrate additional datasets it would be desirable to continue to use these existing panels.

However, a number of additional applications for the genotype data become possible if mapped microsatellites with known positions in the genome are used. For example, it may be possible to undertake association studies to identify correlations between the frequency of marker alleles and different parasite phenotypes. It may also be possible to search for evidence of recombination within a chromosome. For the *Leishmania* genomes that have been sequenced, it is relatively straightforward but laborious to manually extract mapped microsatellites and design primers for PCR (Jamjoom *et al*. 2002a). In this study, to facilitate the selection of microsatellite loci in future research, we have compiled a database of most microsatellite loci that have been mapped to chromosomes in the *L. braziliensis, L. infantum* and *L. major* genomes.

A significant part of the cost of a panel of microsatellite markers is the requirement for a labelled primer for each locus. This cost is particularly onerous when developing new marker panels since it is common to find that around half of all loci tested will either not produce a satisfactory PCR product or are not polymorphic between isolates. We have tested a tailed primer method that permits the use of a single set of labelled primers that bind to 'tails' on primers for all loci (Oetting *et al*. 1995; Boutin-Ganache *et al*. 2001).

To date, all MLMT studies of *Leishmania* have required cultured parasites for microsatellites typing. Parasite culture is a well established technique but requires laboratory facilities and is relatively costly, laborious and time consuming. Nested PCR combined with whole genome amplification has recently been developed for genotyping *Trypanosoma brucei* direct from clinical samples (Morrison *et al*. 2007). We have tested this method for genotyping *L. infantum* isolated from Iran.

The aim of this paper is to validate the application of a programme that extracts microsatellite information from existing genome sequences and provides a strategy for generating PCR primers for microsatellite genotyping. In addition, we assess the use of a nested PCR strategy

for the routine genotyping of *Leishmania* parasites from clinical samples, to avoid potentially costly and time consuming parasite culturing.

Materials & Methods:

Parasites – *Leishmania* spp. samples used in this study are detailed in Table 1. Eleven strains were derived from cultured parasites isolated from Brazil, Sudan and Europe (7 *L. donovani* strains, 4 *L. infantum* strains). Six additional cultures (5 *L. infantum* strains, 1 *L. tropica* viscerotropic strain) were isolated from the Fars and East-Azarbaijan Provinces of Iran, which are the main endemic regions in Southern and Northwesten Iran, respectively. (Table 1). A further 56 DNA samples were isolated from bone marrow, spleen and liver tissue smears on microscope slides from humans and dogs also from endemic regions of Iran.

DNA extraction - Genomic DNA was extracted from cultured parasites or slide preparations following a high salt method (Aljanabi & Martinez, 1997): a detailed protocol is available from http://www.genomics.liv.ac.uk/animal/Protocol1.html. For cultured samples, 400-500 µl of media (~10^7 parasites) was pelleted and digested with 20 µl of proteinase K (20 mg/ml) in 200-300 µl of TNES buffer (10 mM of Tris-HCl, pH=7.5; 100 mM of EDTA, pH=8.0; 400 mM NaCL; 0.6% SDS). Cell debris and protein was precipitated with saturated salt solution (6.0 M NaCl). DNA was precipitated with 2.5 volumes of absolute ethanol and subsequently washed twice in cold 100% ethanol and once in 70% ethanol. Samples were air dried and resuspended in 50 µl TE buffer (10.0 mM Tris-HCl, 1.0 mM EDTA, pH7.4). To extract DNA from tissues mounted on microscope slides, sealed slides were first soaked in xylene (Gibco) for one week to remove the cover slip. Samples were left to air dry for ~ 10 minutes and then lysed with 200µl of TNES buffer. After 1-2 mins the tissue smear was completely detached from the slide and transferred to a microcentrifuge tube. To extract DNA from non-sealed microscope slides, 200µl of TNES buffer was applied to the slide and the TNES-sample mixture was aspirated after 1-2 mins and transferred to a microcentrifuge tube. The

subsequent extractions proceeded as detailed above. For all samples, DNA concentrations were measured using a Nanodrop™ spectrophotometer (Labtec co), and were adjusted to ~20 ng/µl for subsequent PCR.

Nested–PCR protocol – All PCR amplifications on clinical samples from microscope slides followed a two step nested PCR strategy using external and internal primer sets (listed in Table 3). For the first round of PCR, each reaction contained 75 mM Tris-HCl (pH 8.8), 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween 20, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, ~20 ng template DNA, 10 pmol of each primer and 0.25 U Taq polymerase (Thermo Life Sciences). PCRs were performed using a PTC-0221 Dyad thermocycler (MJ research) and the following conditions: 95^oC for 3 min, 5 cycles ($95\ ^oC$ 30 s, T$a\ ^oC$ 45 s, $72\ ^oC$ 45 s), 35 cycles ($92\ ^oC$ 30 s, T$a\ ^oC$ 45 s, $72\ ^oC$ 55 s), $72\ ^oC$ 10 mins (where Ta is the locus specific annealing temperature). For the second round of PCR, 1 µl of the amplified product from round one was used as template. Preamplifaction of the DNA template with the Genomiphi™ whole genome amplification kit (GE Healthcare), following the manufacturers protocol but with half volumes of reagents, was tested on a subset of samples.

Labelling of PCR products - To allow fragment detection on a genetic analyser, PCR products were fluorescently labelled using a tailed primer method (Oetting *et al.* 1995; Boutin-Ganache *et al.* 2001). Briefly, four standard primers with unique sequences were separately labelled with one of four fluorescent dyes (forward labelled primers, Table 3: dye labels 6-FAM, NED, PET, VIC; Applied Biosystems). Forward internal primers for microsatellite loci were then designed with a 3' tail (forward tailed primer) complementary to the forward labelled primer. Reaction conditions for the second round of PCRs were identical to the first, except that round two PCRs contained 3 primers (3.0 pmol reverse primer, 3.0 pmols labelled forward primer and 1.0 pmol forward tailed primer). PCR amplification was confirmed via electrophoresis of products on a 2 % agarose gel containing

ethidium bromide and visualisation under UV illumination. To size PCR fragments, products were pooled into one of two genotyping panels along with a GENESCAN-500 LIZ size standard (Applied Biosystems) and separated by capillary electrophoresis through POP7 denaturing polymer on an ABI3730*xl* automated sequencer (Applied Biosystems). Allele sizes were determined using the cubic model of analysis in GENEMAPPER software (Applied Biosystems).

Evaluation of nested PCR - To assess the sensitivity of the nested PCR approach, PCRs were conducted using DNA template prepared from serial dilutions of parasite cultures. Culture medium (Dulbecco's modified eagle's medium, DMEM) containing Chinese hamster ovary cells was seeded with known numbers of cultured promastigotes (*L. infantum* MHOM/BR/76/150406; M4192). Parasites were counted using a Neubauer haemocytometer slide. For each dilution, 80 µl of culture was dispersed on microscopic slides. The final concentrations of the parasites were 10^6, 10^5, 10^4, 10^3, 10^2, 10^1, 10 promastigotes per ml of culture medium corresponding to 8×10^4, 8×10^3, 8×10^2, 80, 8 ,0.08, 0.008 promastigotes per sample on slides. Seeded samples were left at room temperature before DNA extraction as described above.

The RV1 and RV2 primers directed at the conserved region of the kinetoplast minicircle were used to confirm the presence of amplifiable parasite DNA (Ravel *et al*.1995). There are approximately 10,000 copies of the kinetoplast minicircle and consequently the RV1-RV2 PCR is expected to be much more sensitive than even nested PCR for single copy microsatellite loci because multiple minicircles can be present in the sample used for PCR even after many genomic loci have been diluted out.

Results:

Microsatellite primer design - To facilitate the use of the *Leishmania* genome sequences for microsatellite primer design two Web Pages were developed; one is the front end for a

database of microsatellite containing sequences and the second is a front end for the Primer3 programme and designs a single pair of primers against each sequence in a file. *Leishmania* genome sequences were downloaded from ftp://ftp.sanger.ac.uk/pub/databases/. A Perl script was developed to extract all microsatellites in a genome and enter them in a searchable MySQL database. The script crawls along each chromosome looking for a perfect match to 6 repeats of all possible di, tri and tetra-nucleotide repeat motifs. When a match is located, the script identifies the number of perfect repeats and the writes to a database the following information for each locus: the number of perfect repeat units, the repeat unit length, the full repeat sequence, and a 700bp sequence starting 350bp upstream of the microsatellite. To minimise the frequency of overlapping loci in the database, the script then moves 200bp along the chromosome before resuming the search. This strategy means that microsatellite loci that are interrupted by imperfect repeats will be assigned a size that corresponds to the first run of perfect repeats. Many of the loci may therefore have effective lengths greater than that recorded in the database. A web front end was constructed for the microsatellite database and is available at http://www.genomics.liv.ac.uk/tryps/resources.html "Microsatellite Locus Extractor".

All microsatellites on a chromosome or a region of a given chromosome that meet the following criteria set by the user can be extracted: chromosome, start position, end position, repeat unit length, minimum number of repeat units and maximum number of repeat units. The output file from the microsatellite database contains the flanking sequence and is formatted for entry into a second webpage at the same site "Primer Design Programme". This page uses a local copy of the Primer3 engine (Rozen & Skaletsky 2000; obtained from http://primer3.sourceforge.net/releases.php) to design primers against each of the microsatellite containing sequences in the output from the microsatellite database. Most of the usual Primer3 parameters can be set by the user to adjust the length of product and

annealing temperature of primers etc. Primers for this project were designed with an annealing temperature of 60 °C and a 2bp GC clamp. For nested primer design the maximum product length of the inner primer pair was set at 300bp. To design external primers (flanking the internal primers) the output file was modified for resubmission to the primer design website using Excel to specify new target start positions and product lengths. Seventeen microsatellite loci designed against the *L. infantum* genome using this approach were evaluated against 2 species in the *Leishmania donovani* complex.

The programmes used to populate the database are avialbale on request from the authors. A modified version of the programme that writes microsatellite loci to a file instead of a database, and that can be used on any sequence, are available for download from the website http://www.genomics.liv.ac.uk/tryps/resources.html. The output file can be loaded into the Primer Design webpage after removal of metadata from the top of the file.

Microsatellite loci - The abundance of different microsatellite repeats, identified by the Perl programme, in the *Leishmania* species for which sequence data are available is shown in Table 2. Microsatellite repeat classes that were represented by fewer than 10 loci are not shown. CA repeats were the most abundant repeat class in all genomes tested at ~78% of all loci detected. Over twice as many microsatellite loci were identified in the *L. major* genome as in the *L. infantum or L. braziliensis* genomes; however, this may be a consequence of the higher finished quality of this genome rather than a reflection of any fundamental difference in number of microsatellite loci.

Primer testing - Primers were designed for 17 microsatellite loci for the *L. donovani* complex using the *L. infantum* sequence data. Two principles guided the selection of loci for testing. Firstly four and five markers respectively were selected for the two largest chromosomes (35 and 36) so that the data could be tested for the presence of recombination within those chromosomes. Secondly eight markers were selected from eight different chromosomes so

that any apparent linkage disequilibrium would be evidence for a clonal population structure. Primers were tested against a panel of control DNA: 7 *L. donovani* from Sudan and 4 *L. infantum* from Sudan, Brazil and the UK. Primers designed against 12 out of the 17 loci generated PCR products suitable for microsatellite genotyping by standard-PCR using internal primers. To allow fragment detection on a genetic analyser (i.e. ABI3130*xl*), PCR products were fluorescently labelled using a tailed primer strategy. Briefly, in the initial PCR cycles, amplification is primed by a forward tailed primer (5' end complementary to the target sequence) which generates extension products bearing a 'tail' complementary to the labelled primer, subsequent rounds of PCR are then primed by the forward labelled primer, which produces a pool of fluorescently labelled products. The ratio of labelled forward primer to forward tailed primer is critical to produce a majority of labelled PCR products for fragment detection: if the relative concentration of the unlabelled tailed primer is too high, the proportion of labelled products will be low and subsequent detection will be poor. Labelled to tailed primer ratios of 1:1, 3:1, and 10:1 were tested to optimise the labelling protocol. Here, a forward-labelled to forward-tailed primer ratio of ≥3:1 was found to produce a sufficient pool of labelled PCR products for fragment detection on the ABI3130*xl*. At a ratio of 1:1, PCR products were detected by agarose gel electrophoresis, but did not produce a detectable signal on the ABI3130*xl*, indicating that an insufficient pool of fragments were labelled during PCR. The observed heterozygosity and number of alleles observed at each locus for the control DNA samples is shown in Table 4.

Five *L. infantum* samples from Iran were also tested and gave PCR products at 8 out of the 12 loci tested. The loss of alleles in the Iranian samples may be a consequence of DNA polymorphisms in the primer binding sites or may be because the parasite cultures were shipped from Iran to the UK at ambient temperature in culture medium and the DNA may

have degraded. A single cultured strain of *L. tropica* was also analysed, but only amplified for 4 of the 12 microsatellite loci (data not shown).

Sensitivity of nested PCR - To estimate the sensitivity of the nested PCR a series of dilutions of parasites in hamster fibroblast cells were spotted onto microscope slides from which DNA was subsequently prepared. DNA precipitated from microscope slides was resuspended in 50µl of TE and diluted to a concentration of 20ng/µl for PCR, 1µl of this solution was used for the nested PCR reaction. The PCR detected down to 800 parasites per slide. After dilution this corresponds to 1.6 parasite equivalents in the PCR reaction. The RV1 and RV2 primers for kinetoplast DNA generated a positive PCR product at all parasite dilutions, indicating the presence of at least one parasite on all slides.

Nested PCR on clinical samples on microscope slides from Iran - DNA prepared from 6 out of 56 microscope slides bearing Iranian *L. infantum* parasites generated PCR products by nested-PCR at 8 out of the 12 microsatellite loci tested. Preamplifcation of DNA with Genomiphi™ whole genome amplifcation protocol did not increase sensitivity and the results presented are from samples that did not receive a Genomiphi™ pretreatment. However RV1 and RV2 primers for the conserved region of the kinetoplast minicircle generated PCR products from all slides, confirming the presence of parasites on these slides.

Discussion:

For population studies it is recommended to use at least 12-15 polymorphic loci (Kuhls *et al.* 2007). In this study 12 out of 17 newly designed loci generated polymorphic products, which is a higher yield of functioning polymorphic loci than is commonly obtained by *de novo* isolation of microsatellite loci by cloning and sequencing genomic DNA. However, while we have developed a new panel of microsatellites, the markers *per se* are not the major output of this study. Indeed, numerous microsatellite panels exist for *Leishmania* species and are in extensive use for population analyses (See Introduction). The objective of this paper is the

validation of a rapid and cost effective strategy for the development and application of mapped microsatellite markers from existing genomic sequence. We anticipate that the main application of this 'pipeline' will be the development of mapped markers for specific applications such as association studies and the search for recombination within chromosomes.

In addition to the development of the microsatellite extractor programme we have also assessed the use of clinical samples and nested-PCR to avoid the need for culturing in the genotyping of *Leishmania* parasites. We are not aware of any previous reports of the use of MLMT directly on clinical samples from leishmaniasis cases without prior cultivation of the parasite. However, it has been successfully applied using DNA isolated from FTA cards spotted with blood from Human African Trypanosomiasis cases (Morrison *et al.* 2007). We have confirmed that microsatellite genotyping can be performed directly from biological material recovered from microscope slides, sufficient DNA could even be recovered from archival Giemsa stained slides that were protected by a cover slip and sealed by a mountant. However, the number of slides from which amplifiable DNA could be recovered was too low to make this a routine sample collection strategy for microsatellite based studies although it may be appropriate where no other sample is available. A more appropriate sample collection and storage medium such as FTA cards may give an increased success rate. *Trypanosoma brucei gambiense* infected human blood samples on FTA cards have been used as a source of DNA for microsatellite PCR using a similar protocol and 11/11 samples carrying 250 -1100 parasites per ml of blood were positive (Morrison *et al.* 2007).

Further to the low success rate achieved here for genotyping directly from clinical samples on microscope slides, 2 additional technical issues should be noted. Firstly, when only very small amounts of parasite DNA are present in the PCR there is a substantial risk of amplifying only one of the two alleles present giving rise to a false excess of apparent

homozygotes and therefore it is necessary to use the products of at least three independent genome amplification reactions as template for the PCR (Morrison *et al.* 2007). This was not attempted here since no downstream analysis was undertaken but would be necessary before any population genetic analysis was attempted. Secondly, it is common practice to dilute the product of the first round PCR before the second round to reduce the potential occurrence of PCR artefacts, which may occur if external primers (from the first PCR round) continue to prime fragment synthesis in the second round. Such artefacts may be erroneously scored as higher molecular weight alleles after electrophoresis. Dilution was not used in this study in an attempt to maximise sensitivity. At most loci the observed allele size range was small, which suggests that all products were from the internal primers. However at 2079709_205_450 the size range was 48bp which could include PCR artefacts. The large size range for this locus was attributable to a single isolate MHOM/SD/XX/9011 that had an exceptionally small size (222bp) whereas the remaining 4 alleles were between 256 and 269bp. It is possible that this anomalously small allele is an artefact, but the small size suggests that it is not a consequence of carry-over of first round primers.

The main causative agent of VL in Iran is *L. infantum*, (Mohebali *et al.* 2007) but recently *L. tropica* (Viscerotropic strain) has been reported as another aetiological agent for immunocompetent patients infected with VL in Southern Iran (Geramizadeh *et al.* 2006, Alborzi *et al.* 2006). In the present study a viscerotropic strain of *L. tropica* isolated from southern Iran, was included. Four loci gave a PCR product for this strain. In a previous study we showed that markers designed against the *L. major* genome would amplify homologous loci in the *L. donovani* genome but that the loci were not polymorphic in *L. donovani* (Jamjoom *et al.* 2002a). It remains to be seen whether loci that are polymorphic in *L. infantum* are also polymorphic in *L. tropica,* however since some *L. infantum* loci were

amplified in *L. tropica* it may at least be possible to identify this species of parasite when it is circulating in the same focus even if it may not be possible to classify it using these primers.

Despite the absence of polymorphisms in *L. donovani* when using *L. major* microsatellites, it has been reported that loci informative in *L. braziliensis* are also informative in other members of the *L. braziliensis* complex (Russell *et al*, 1999). Consequently the loci that were identified in the *L. braziliensis* genome in this study may also be amplifiable and polymorphic in *L. guyanensis, L. peruviana* and *L. panamensis*. The programme is currently being used to generate a whole genome mapping panel of tri- and tetra-nucleotide microsatellites in the bovine genome. Additional parasite genomes can be added to the microsatellite database on request to the authors. A standalone version of the Perl programme that extracts the microsatellite loci is available for download from the website http://www.genomics.liv.ac.uk/tryps/resources.html. This programme outputs loci to a file rather than a database, the file is formatted for upload to the primer design programme after removal of metadata.

In summary, we have presented a programme that offers a rapid strategy for the extraction of microsatellite loci from existing genome sequences and the design of PCR primers for genotyping. Here we illustrate the use of the strategy for the genotyping of 2 *Leishmania* species in the first report of the use of the pipeline.

References:

Alborzi, A., Rasouli, M., Shamsizadeh, A. (2006). *Leishmania tropica*-isolated patient with visceral leishmaniasis in southern Iran. *American Journal of Tropical Medicine and Hygiene* **74** ,306–307.

Aljanabi, S. & Martinez, I.. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25**, 4692-4693.

Botstein, D., White, R. L., Skolnick, M., Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32** (3), 314-31.

Boutin-Ganache, I., Raposo, M., Raymond, M., Christian F., Deschepper, C. F. (2001) M13-Tailed Primers Improve the Readability and Usability of Microsatellite Analyses Performed with Two Different Allele-Sizing Methods. BioTechniques 31 (1), 24-27

Bulle, B., Millon, L., Bart, J.M., Gállego, M., Gambarelli, F., Portús, M., Schnur, L., Jaffe, CL., Fernandez- Barredo, S., Alunda J.M., Piarroux, R. (2002). Practical approach for typing strains of Leishmania infantum by microsatellite analysis. *Journal of Clinical Microbiology* **40** (9), 3391-7.

Desjeux P. (2001).The increase in risk factors for leishmaniasis world wide. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **95** (3), 239-43.

Geramizadeh, B., Fakhar, M., Motazedian M. H. (2006). Visceral leishmaniasis with duodenal involvement: three immunocompetent cases from southern Iran. *Annals of Tropical Medicine & Parasitology* **100** (7), 637–640.

Jamjoom, M.B., Ashford R.W., Bates P.A ., Kemp S. J. & Noyes H. A. (2002a) Polymorphic microsatellite repeats are not conserved between *Leishmania donovani* and *Leishmania major*. *Molecular Ecology Notes* **2**, 104-106.

Jamjoom, M.B., Ashford R.W., Bates P.A., Kemp S. J. & Noyes H. A. (2002b) Towards a standard battery of microsatellite markers for the analysis of *Leishmania donovani* complex. *Annals of Tropical Medicine and Parasitology* **96**, 265-270.

Kuhls, K., Keilonat, L., Ochsenreither, S., Schaar, M., Schweynoch, C., Presber, W., Schonian,G. (2007). Multilocus microsatellite typing (MLMT) reveals genetically isolated populations between and within the main endemic regions of visceral leishmaniasis. *Microbes and Infection* **9**, 334-343.

Mohebali, M., Edrissian, Gh.H., Nadim, A., Hajjaran, H., Akhoundi, B., Hooshmand, B., Zarei, Z., Arshi, Sh., Mirsamadi, N., Manouchehri Naeini, K., Mamishi, S., Sanati, AA., Moshfe, AA., Charehdar S., Fakhar, M. (2007). Application of direct agglutination test (DAT) for the diagnosis and seroepidemiological studies of visceral leishmaniasis in Iran. *Iranian Journal of Parasitology* **1**, 15-25.

Montoya, L., Gállego, M., Gavignet, B., Piarroux, R., Rioux, J.A., Portús, M., Fisa, R. (2007).Application of microsatellite genotyping to the study of a restricted *Leishmania infantum* focus: different genotype compositions in isolates from dogs and sand flies. *American Journal of Tropical Medicine and Hygiene* **76** (5), 888-95.

Morrison, L.J., McCormack, G., Sweeney, L., Likeufack, AC., Truc, P., Turner, CM., Tait, A., MacLeod, A.(2007).Use of multiple displacement amplification to increase the detection andgenotyping of *trypanosoma* species samples immobilized on FTA filters. *American Journal of Tropical Medicine and Hygiene* **76** (6), 1132-1137.

Oetting,W.S., Lee,H.K., Flanders,D.J., Wiesner,G.L., Sellers,T.A., King,R.A. (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics, **30**, 450-458.

Ravel, S., Cuny, G., Reynes, J., Veas, F. (1995). A highly sensitive and rapid procedure for direct PCR detection of *Leishmania infantum* within human peripheral blood mononuclear cells. *Acta Tropica* **59** (3),187-96.

Rozen, S., & Skaletsky, H. (2000). Primer 3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* **132**, 365-86.

Russell, R., Iribar, M.P., Lambson, B., Brewster, S., Blackwell, J.M., Dye, C., Ajioka J.W.(1999).Intra and inter-specific microsatellite variation in the *Leishmania* subgenus Viannia. *Molecular and Biochemical Parasitology*. **103**(1), 71-77.

Table 1. Strain, and country of origin data for the *Leshmania* spp. cultures used in this study.

Species	Country of Origin	WHO code	*Zymodeme*
L. infantum	Brazil	MHOM/BR/76/150406;M4192	MON-1
L. donovani	Sudan	MHOM/SD/97/LEM34 41	MON-30
L. donovani	Sudan	MHOM/SD/97/LEM 3431	MON-30
L. infantum	Great Britain	MCAN/GB/96/LV755	MON-1
L. donovani	Sudan	MHOM/SD/90/D75;LEM2134	MON-82
L. donovani	Sudan	MHOM/SD/90/2655;LEM2135	MON-82
L. donovani	Sudan	MHOM/SD/90/D99;LEM2137	MON-82
L. donovani	Sudan	MHOM/SD/91/D1783;LEM2211	MON-82
L. donovani	Sudan	MHOM/SD/90/D100;LEM2140	MON-18
L. donovani	Sudan	MHOM/SD/90/2828;LEM2139	MON-18
L. donovani	Sudan	MHOM/SD/XX/9011	
L. infantum	Iran	MCAN/IR/96/LON 49	MON-1
L. infantum	Iran	MHOM/IR/05/SHZ 5	MON-1
L. infantum	Iran	MHOM/IR/05/SHZ 1	MON-1
L. infantum	Iran	MCAN/IR/05/SHZ 2	MON-1
L. infantum	Iran	MCAN/IR/06/SHZ 4	MON-1
L. tropica	Iran	MHOM/IR/06/SHZ 6	

Table2. The numbers of microsatellites identified in the *L. major*, *L. donovani*, and *L. braziliensis* genomes. Only microsatellite motifs that were found more than ten times in the genome are shown. The ac repeat is by far the most abundant in all three genomes. It is likely that the larger number of loci observed in *L. major* is a consequence of the higher quality assembly and finishing of this genome.

Genome	Repeat motif	No. loci	Mean no. repeats
L. infantum	ac	993	14.6
L. infantum	tgc	61	8.6
L. infantum	acc	25	9.4
L. infantum	ggga	24	6.7
L. infantum	ta	22	12.5
L. infantum	cgg	13	8.7
L. infantum	gga	13	9.8
L. major	ac	2268	15.6
L. major	ta	316	14.3
L. major	ct	155	13.5
L. major	acc	93	8.8
L. major	tgc	72	8.5
L. major	ggga	52	7.1
L. major	gga	42	8.8
L. major	gggt	21	6.6
L. major	caga	15	6.3
L. major	gtt	13	8.5

L. major	ggca	10	6.5
L. major	cacg	10	7.9
L. braziliensis	ac	1131	15.7
L. braziliensis	ta	150	15.8
L. braziliensis	ct	94	13.6
L. braziliensis	tgc	84	8.8
L. braziliensis	ggga	41	6.51
L. braziliensis	acc	31	9.3
L. braziliensis	gga	17	8.3

Locus	Dye label	Pair	Forward Sequence	Reverse Sequence	Repeat Motif	Chr	Position	Expected Size
2079372_72_448		Ex	gaccattccaacatcctccc	tgcctctccaccattattcg	$(tgc)_{14}$	2	126036	376
2079372_119_394	NED	In	CTCATTTCATGTGAACTTTCCAa cacccgtgctttctgacg	agtagagcagcagacgtgcg				275
2079447_148_646		Ex	cacaacaagcagagaacagagg	ttgatgagcaccactccg	$(cacg)_{11}$	4	161662	498
2079447_193_429	6-FAM	In	GCGCATTGAAGCGGTTAC accatccg ttactctcgcc	tgtgtgtgtgtgtgtgtctgg				236
2079619_25_438		Ex	ctctctcttcggcttgttgg	ggatacggcaatacaatggg	$(acc)_{10}$	10	225980	413
2079619_166_373	PET	In	GAAGACTGCTCAGAGGAGCC atg gtagagaacgacgcacc	cacggagaggacaacaacg				207
2079709_131_590		Ex	tcgacaacaacaagaggagg	ttacccgcttcatcagcc	$(tgc)_{11}$	13	252447	459
2079709_205_450	NED	In	CTCATTTCATGTGAACTTTCCAa ggcaagagaacagggagg	gtctgtgtgcggacttgtgc				245
2079734_89_538		Ex	gcagagaggggagaaagaggg	agagagagggaggtaagggcg	$(gga)_{12}$	14	354668	449
2079734_196_471	6-FAM	In	GCGCATTGAAGCGGTTAC agacaca cacagacacacgc	ctaacccgattgcgaagg				275
2079764_93_547		Ex	cagcatcctctcagctactcc	cgtactataaaggcagacggg	$(tgc)_{11}$	15	264557	454
2079764_207_434	PET	In	GAAGACTGCTCAGAGGAGCC cag cagcaccaacatcagc	acttgagtaccgcagcaacg				227
2079862_64_549		Ex	aggaggtgtctgtgtattcgg	ggagagagaaagagagcgagc	$(gga)_{10}$	19	395073	485
2079862_235_376	VIC	In	CGTTCTTCGTCTTCTGGGATcgtc ctcttgttgtttgttcg	agtgggagaaagcgtcaacc				141
2080398_188_553		Ex	ctacttctgctctcttgccg	catggcttgctttagtcacc	$(ac)_{19}$	35	781548	365
2080398_243_507	NED	In	CTCATTTCATGTGAACTTTCCAc ctctcttctagggttcgcc	caggtgattcgcaggtacg				264
2080476_49_610		Ex	cccttctgtctgcttctcc	gtggtgtctctgtgctttcc	$(ac)_{12}$	36	421783	561
2080476_227_539	VIC	In	CGTTCTTCGTCTTCTGGGATtttcc tctttccctcggc	gatgcacagtatcgcaaacg				312
2080483_70_595		Ex	cacccttgctgtatttcacg	cccagctccaacattacttacc	$(ac)_{16}$	36	965053	525
2080483_178_461	NED	In	CTCATTTCATGTGAACTTTCCA gaagagcaagaagaagaggcg	gagaacggcacgagaaagg				283
2080455_230_522		Ex	tgtcggtctctctcttgtttgc	aactggaggaaagcagtcgg	$(ac)_{12}$	36	1425639	292
2080455_268_439	6-FAM	In	GCGCATTGAAGCGGTTAC cttcctcct cctcacagtcg	caaacaaactagccagccg				171
2080492_169_582		Ex	gtattcgcagtcgatgagcc	cgtcagtctctcaattcacgc	$(ac)_{12}$	36	1988273	413
2080492_192_516	PET	In	GAAGACTGCTCAGAGGAGCC cact acagcgacacacgagc	agggtcttctctgcctctcc				324
Labeled F primers	VIC		atcccagaagacgaagaacg					
	NED		tggaaagttcacatgaaatgag					
	6-FAM		gtaaccgcttcaatgcgc					
	PET		ggctcctctgagcagtcttc					

Table3. Twelve microsatellite loci successfully amplified in *Leishmania donovani* complex. Internal (In) and external (Ex) primer pairs for nested PCR are indicated. Each internal forward primer also included a 3- 'tail' (in upper case) complimentary to the forward labelled primer with the required fluorescent dye. The primer name is made up of three components, the locus id from the database and the start and end positions within the 700bp sequence that are associated with the locus in the database. The microsatellite that was identified by the Perl script starts at position 350 within the 700bp sequence.

Table 4. The number of alleles observed at each locus (N_a), observed heterozygosity (Ho), allele size range, and PIC (polymorphism information content, calculated according to Botstein *et al.* 1980) at 12 polymorphic microsatellite loci for *L. donovani* and *L. infantum* strains. For *L. donovani* and *L. infantum* 7 and 4 strains were genotyped respectively.

Loci/Markers	N_a	Size range (bp)	H_O	PIC
2079372_119_394	3	283-300	0.11	0.40
2079447_193_429	3	252-266	0.0	0.49
2079619_25_438	3	232-236	0.50	0.51
2079709_205_450	5	222-269	0.40	0.68
2079734_196_471	3	278-290	0.0	0.44
2079764_207_434	2	241-248	0.0	0.20
2079862_235_376	3	147-166	0.0	0.44
2080398_243_507	2	264-286	0.0	0.18
2080455_268_439	2	187-190	0.0	0.16
2080476_227_539	2	338-340	0.0	0.35
2080483_178_461	4	304-324	0.20	0.49
2080492_192_516	2	348-364	0.22	0.34

L. donovani