

PRIMER NOTE

Characterization of di-, tri- and tetranucleotide microsatellite markers with perfect repeats for *Trypanosoma brucei* and related species

OLIVER BALMER,* CHRISTOPHER PALMA,* ANNETTE MACLEOD† and ADALGISA CACCONE*

*Yale University, Department of Ecology and Evolutionary Biology, New Haven, USA, †University of Glasgow, Wellcome Centre for Molecular Parasitology, Glasgow, UK

Abstract

Trypanosoma brucei, a unicellular parasite causing human sleeping sickness and animal nagana, has a great impact on the socioeconomic environment of sub-Saharan Africa. The dynamics of the parasite are still poorly understood. We have characterized 14 polymorphic di-, tri- and tetranucleotide microsatellite loci with perfect repeats (only one motif) exhibiting between five and 16 alleles in *T. brucei* isolates from all over Africa and from all described subspecies. The microsatellites will be useful in addressing population genetic questions in *T. brucei* to better understand the population structure and spread of this important parasite.

Keywords: coinfection, cross-amplification, kinetoplastid, microsatellites, population genetics, *Trypanosoma*

Received 1 December 2005; revision accepted 13 January 2006

Trypanosoma brucei, a protozoan parasite transmitted by the bite of tsetse flies (*Glossina* spp.), causes human sleeping sickness and animal nagana in Africa. After near elimination in the 1960s, sleeping sickness has resurged to record heights (Van Nieuwenhove *et al.* 2001) due to declining control efforts caused by neglect and civil unrest.

Microsatellite loci for *T. brucei* have been published previously: five microsatellites by Biteau *et al.* (2000), one by Truc *et al.* (2002) and a further 180 by MacLeod *et al.* (2005), which were used to generate a genetic map of the *T. brucei* genome strain (TREU927/4). Many of these markers, however, do not have perfect repeats (only motif included) and so may not follow a stepwise mutation model, limiting their use for population genetic analyses. And, in the case of the markers used for the genetic map, their variability is unknown. We therefore characterized a set of microsatellite loci with perfect repeats. Loci were identified by extracting all clean di-, tri- and tetranucleotide repeats in the entire

genome (Berriman *et al.* 2005) using TANDEM REPEAT FINDER 3.21 (Benson 1999). Primers were constructed with PRIMER 3 (Rozen & Skaletsky 2000). Where loci identified matched those of MacLeod *et al.* (2005), the same locus names and primers were used (Table 1). All microsatellite loci were amplified using the following polymerase chain reaction (PCR) profile: 4 min at 94 °C, followed by 35 cycles of 45 s at 94 °C, 30 s at 53 °C, 45 s at 72 °C, followed by 7 min at 72 °C. Reactions contained 1× PCR buffer II with MgCl₂ (Applied Biosystems), 0.8 mM of each dNTP (Promega), 0.2 μM of each primer and 0.25 U AmpliTaq polymerase (Applied Biosystems). Allele sizes were determined using an ABI PRISM 3100 Genetic Analyser (Applied Biosystems) and GENEMAPPER version 3.5 software (Applied Biosystems 2002). GENEPOP version 3.3 (Raymond & Rousset 1995) was used to calculate observed heterozygosities. Linkage disequilibrium (LD) between loci was not assessed because the loci are distributed over nine of 11 chromosomes and thus in most combinations not physically linked. Furthermore, samples from all over Africa and from all subspecies were assessed to maximize detected allele diversity. Due to the geographical structure thus introduced, LD is expected without physical linkage.

Correspondence: Oliver Balmer, Yale University, Department of Ecology and Evolutionary Biology, 165 Prospect St., New Haven, USA. Fax: (203) 432 5176; E-mail: oliver.balmer@yale.edu, oliver.balmer@pronet.ch

Table 1 Characteristics of *Trypanosoma brucei* microsatellites

Locus	Chromosome	Primer sequence (5'-3')	Repeat	<i>n</i>	<i>N_a</i>	Range (bp)	<i>H_O</i>	Cross amplification
Tb1/8	1	[FAM]-AGGTTTAGTGCATGTCGGA CCTGTTGTACGGAGGTCA	(CA)	103	11	97-117	0.38	Teq, Tev, Tv, Ts, Tr, Lm, Ld
Tb2/19	2	[HEX]-CTGGTGCCTGTAACCTGTG GAAGTGAGGACATGCACG	(AT)	24	11	84-104	0.71	Teq, Tev, Ld, Lm
Tb2/21	2	[HEX]-CTGTGTGTGCTGTGTCATA AGTTTAAACAGCACTTCCATTT	(AT)	21	12	80-102	0.67	Teq, Tev, Ld, Lm
Tb5/2	5	[HEX]-CAACCGAAAGTAAGGGGAAC TCTGCGCTTCTTTGCCC	(AT)	23	9	83-107	0.57	Teq, Tev, Tc, Ts, Ld, Lm
Tb6/7	6	[HEX]-AAGCTGACAGGTGGTTGA GAACATGCGTGCCTGTG	(AT)	24	12	104-136	0.50	Teq, Tev, Ts, Lm
Tb8/11	8	[FAM]-TGTAGCAGTGGTACGCAC CACCAACGCATGTAAGC	(AT)	22	13	97-127	0.86	Teq, Tev, Tc, Ld, Lm
Tb9/6	9	[HEX]-TGATTCATTGGTTAAGACAGG AATGATAACTGCGGATTACAC	(AC)	24	11	124-158	0.67	Teq, Tev, Lm
Tb9/4	9	[HEX]-CAACTGTCATCTCGTTACTT TTCAATTGCCTATCGTTGTG	(AT)	24	9	101-119	0.50	Tev, Ts, Ld, Lm
Tb10/5	10	[FAM]-AAAGCGATATGTTATATTGA ATTGGGTATACTGTCCCTCA	(TA)	24	13	79-115	0.75	Tev
Tb11/12	11	[FAM]-CCTTCACTCTTAAGTGAAG GTAGCCATTCTGCGTCC	(AT)	23	12	83-109	0.48	Tev, Ts, Ld, Lm
Tb11/13	11	[FAM]-CAAGAACTCTGCATTGAGC ATCTGTTGGCGATGGTGA	(AT)	91	16	125-161	0.78	Teq, Tev
Tr301-1	1	[HEX]-CTCCCCTCACTTCCCTCACG GTTTGCACATGACAATAACACACAGG	(TTA)	25	10	69-108	0.56	Tev, Tc, Tl, Ts, Ld, Lm
Tr401-1	1	[HEX]-GTGAAAAACGAAAGCAACG TGAGTTCAACAATCTTTTATTCC	(GTGA)	106	6	126-150	0.28	Teq, Tev, Ts, Ld, Lm
Tr407-1	7	[HEX]-AACAAATATCTGACAATGAGGATGG GTTTAGATGGGTGGAAAGGGTAGG	(AAAC)	76	5	157-173	0.16	Teq, Tev, Tc, Tl, Ts, Ld, Lm

Locus, TB1-TB11 as in MacLeod *et al.* (2005), all sequences based on the *T. brucei* genome (Berriman *et al.* 2005); *n*, number of isolates assessed; *N_a*, number of alleles detected in *T. brucei*; *H_O*, observed heterozygosity; Cross-amplification, further species that amplified: Teq, *Trypanosoma equiperdum*; Tev, *Trypanosoma evansi*; Tc, *Trypanosoma congolense*; Tl, *Trypanosoma lewisi*; Ts, *Trypanosoma simiae*; Tr, *Trypanosoma rangeli*; Tv, *Trypanosoma vivax*; Ld, *Leishmania donovani*; Lm, *Leishmania major*.

We screened between 21 and 106 *T. brucei* isolates at 41 loci to determine variability. Fourteen loci were highly variable and reliably produced allele patterns consistent with the repeat motif (Table 1). We assessed the potential of our markers to study related species by analysing one isolate each of *Trypanosoma evansi*, *Trypanosoma equiperdum*, *Trypanosoma vivax*, *Trypanosoma congolense*, *Trypanosoma lewisi*, *Trypanosoma simiae*, *Trypanosoma rangeli*, *Leishmania donovani* and *Leishmania major*. Most markers amplified *T. evansi* and *T. equiperdum* (the closest relatives of *T. brucei*) well, and the *Leishmania* species very weakly. *Trypanosoma simiae* was amplified by eight markers, the remaining *Trypanosoma* species by one to four markers (Table 1). Cross-amplification could potentially pose problems in those five markers that amplify *T. vivax* or *T. congolense* as these species have geographic and host ranges overlapping with *T. brucei*. The other species do not occur sympatrically or have very restricted host ranges (Mulligan & Potts 1970). Genotyping of cultured parasites is unproblematic with all markers

because all species survive under different culture conditions. Also, all species are distinguishable morphologically or with species-specific primers (Desquesnes & Davila 2002) if necessary. The markers did not amplify tsetse flies (*Glossina* spp.), or a vertebrate host, wildebeest (*Connochaetes taurinus*).

The presented markers will help reveal the population structure and spread of this important parasite and greatly facilitate the determination of different strains directly in field samples.

Acknowledgements

We thank R. Brun (Swiss Tropical Institute, Basel), P. Grébaut (CIRAD-IRD/LRCT, Montpellier) and W. Gibson (University of Bristol) for providing trypanosome DNA, and the National Science Foundation (DEB-0408083 to OB), the Yale Institute of Biospheric Studies (GC), the Wellcome Trust (grant no. 074732/z/04/z to AML), The Royal Society of Edinburgh (AML) and Tenovus Scotland (AML) for funding.

References

- Applied Biosystems (2002) *GENEMAPPER Software*, Version 3.0. Applied Biosystems, Foster City, CA.
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, **27**, 573–580.
- Berriman M, Ghedin E, Hertz-Fowler C *et al.* (2005) The Genome of the African Trypanosome *Trypanosoma brucei*. *Science*, **309**, 416–422.
- Biteau N, Bringaud F, Gibson W, Truc P, Baltz T (2000) Characterization of Trypanozoon isolates using a repeated coding sequence and microsatellite markers. *Molecular and Biochemical Parasitology*, **105**, 187–202.
- Desquesnes M, Davila AMR (2002) Applications of PCR-based tools for detection and identification of animal trypanosomes: a review and perspectives. *Veterinary Parasitology*, **109**, 213–231.
- MacLeod A, Tweedie A, McLellan S *et al.* (2005) The genetic map and comparative analysis with the physical map of *Trypanosoma brucei*. *Nucleic Acids Research*, **33**, 6688–6693.
- Mulligan HW, Potts WH (1970) *The African Trypanosomiases*. Wiley-Interscience, New York, NY.
- Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Rozen S, Skaletsky HJ (2000) PRIMER 3 on WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, NJ.
- Truc P, Ravel S, Jamonneau V, N'Guessan P, Cuny G (2002) Genetic variability within *Trypanosoma brucei gambiense*: evidence for the circulation of different genotypes in human African trypanosomiasis patients in Cote d'Ivoire. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**, 52–55.
- Van Nieuwenhove S, Betu-Ku-Mesu VK, Diabakana PM, Declercq J, Bilenge CM (2001) Sleeping sickness resurgence in the DRC: the past decade. *Tropical Medicine and International Health*, **6**, 335–341.