Abundant variation in microsatellites of the parasitic nematode

Trichostrongylus tenuis and linkage to a tandem repeat

Paul C. D. Johnson [a, b, *], Lucy M. I. Webster [a, 1], Aileen Adam [a], Robert Buckland [b], Deborah A. Dawson [b], Lukas F. Keller [a, c]

Addresses

[a]Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK

[b]Sheffield Molecular Genetics Facility, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK 11 [c] Zoological Museum, University of Zürich, Winterthurerstrasse 190, CH-8057 12 Zürich, Switzerland 13

*Corresponding author. Tel.: +44 141 330 6625; fax: +33 141 330 5971.

E-mail address: p.johnson@bio.gla.ac.uk

Abbreviations: mtDNA, mitochondrial DNA; PCR, polymerase chain reaction; *He*, expected heterozygosity; PIC, polymorphic information content.

Note: Nucleotide sequence data reported in this paper are available in the EMBL, GenBank™ and DDBJ databases under the accession numbers: **AM167567** – **AM167875**.

[1] Present address: School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, UK

Abstract An understanding of how genes move between and within populations of parasitic nematodes is important in combating the evolution and spread of anthelmintic resistance. Much has been learned by studying mitochondrial DNA markers, but autosomal markers such as microsatellites have been applied to only a few nematode species, despite their many advantages for studying gene flow in eukaryotes. Here we describe the isolation of 307 microsatellites from *Trichostrongylus tenuis*, an intestinal nematode of red grouse. High levels of variation were revealed at sixteen microsatellite loci (including three sex-lined loci) in 111 male *T. tenuis* nematodes collected from four hosts at a single grouse estate in Scotland (average *He* = 0.708; mean number of alleles = 12.2). A population genetic analysis detected no deviation from panmixia either between (*FST* = 0.00) or within hosts (*FIS* = 0.015). We discuss the feasibility of developing microsatellites in parasitic nematodes and the problem of null alleles. We also describe a novel 146-bp repeat element, TteREP1, which is linked to two thirds of the microsatellites sequenced and is associated with marker development failure. The sequence of TteREP1 is related to the TcREP-class of repeats found in several other trichostrongyloid species including *T. colubriformis* and *Haemonchus contortus*.

Keywords: microsatellite; parasitic nematode; population genetics; repetitive elements; sex-linkage; *Trichostronglylus tenuis*

elements; sex-linkage; *Trichostronglylus tenuis*

Introduction

The ability of parasites and pathogens to rapidly evolve drug resistance is a major problem facing human health, agriculture, and animal welfare. Ideally, control strategies should be implemented to prevent or delay the spread of resistance genes to other populations even before drug resistance begins to emerge in a population of hosts [1]. The effectiveness of these measures – such as barriers to gene flow or refugia of susceptible genes [2] – will depend in part on an understanding of parasite population genetics. In other words, we first need to know how parasite genes are distributed between hosts and between populations of hosts and which forces – such as gene flow, genetic drift and selection – are influencing this distribution [3].

Resistance of parasitic nematodes to anthelmintics is a particularly well-studied problem. Anthelmintic resistance is widespread among livestock, and, locally, it has reached levels that threaten livestock production [4]. Several studies suggest that population genetic processes may affect the speed of resistance evolution in parasitic nematodes [2, 5, 6]. The study of parasitic nematode population genetics has been facilitated by the development of a variety of molecular markers, particularly from mitochondrial DNA (mtDNA) [7-9]. Although mtDNA variation has provided valuable information on population genetic structure in nematode parasites of livestock and wild hosts, it represents a single locus and is maternally inherited. In most cases biparentally inherited markers such as microsatellites are required to study the population genetic processes that contribute to the spread of anthelmintic resistance.

Microsatellites are tandemly repeated 1–6 bp DNA motifs that are abundant in eukaryote genomes and can mutate rapidly by loss or gain of repeat units, with the result that most eukaryote species can be expected to have accumulated a wealth of microsatellite length variation [10-12]. This genetic variation can be easily assayed by the polymerase chain reaction (PCR) and subsequent electrophoresis, and consequently

microsatellites have become the most popular resource for studying population genetic variation in eukaryotes.

Despite their potential utility, microsatellite markers have been developed for only a few parasitic nematodes, including parasites of sheep [13], pigs [14], humans [15-17] and rats [18]. The unpopularity of microsatellite as genetic markers for parasitic nematodes may be explained by the unusually high number of nematode microsatellites that fail to produce interpretable PCR banding patterns, possibly as a result of inter-locus flanking sequence homology [13, 16, 18-21].

Here we report the isolation of 307 microsatellite loci from the avian parasitic nematode *Trichostrongylus tenuis*. Inter-locus flanking sequence homology was detected in 196 microsatellites and was strongly associated with marker failure. From the remaining sequences we developed a suite of sixteen highly variable microsatellite markers and demonstrated their utility in an analysis of *T. tenuis* population genetic structure. *T. tenuis* is a caecal nematode of birds that reaches its highest prevalence and abundance in red grouse *Lagopus lagopus scoticus*, an economically important game bird endemic to Great Britain. High burdens of *T. tenuis* are severely detrimental to red grouse both individually and at the population level [22-26]. *T. tenuis* is an excellent candidate for a microsatellite-based study of genetic structure for two principal reasons. Firstly, as a parasite of avian hosts, *T. tenuis* might exhibit unusual population genetic structure: parasitic nematode gene flow among airborne hosts might be expected to be less restricted than among earthbound mammals. Secondly, many red grouse populations (defined here as grouse-shooting estates) have been treated with anthelmintics for up to fifteen years while others have remained untreated [27, 28], creating the opportunity to study the effects on anthelmintic resistance evolution of gene flow between treated and untreated sites.

The objectives of this study were therefore: (1) to isolate microsatellite 2

sequences from *T. tenuis*; (2) to investigate the effects of flanking sequence homology 3 on microsatellite marker development; (3) to develop a suite of polymorphic 4 microsatellite markers; and (4) to demonstrate their utility in a genetic analysis of *T.* 5 *tenuis* population structure. 6

Materials and Methods

Microsatellite isolation

T. tenuis samples were collected at grouse-shooting estates in the UK. Red grouse caeca were removed and preserved at –20°C. *T. tenuis* adults were isolated from thawed caecal contents and stored in 10–15 ml 95% ethanol in 15-ml plastic tubes at 4°C.

Two genomic DNA libraries were prepared following a standard enrichment protocol [29]. For the first library genomic DNA was purified from bulked male and female adult nematodes (approximately a 200-µl volume of compressed tissue, probably containing well over 1000 individuals) taken from a single red grouse from a grouse-shooting estate in North Yorkshire, England in September 2001. Ethanol was removed by pipetting followed by vacuum centrifugation at 50°C. The dried nematodes were rinsed in TE (pH 8.0) with 0.5% SDS to remove contaminant DNA (e.g. from the red grouse host) and incubated in 600 µl of extraction buffer (10 mM Tris pH 8.0, 0.1 mM EDTA pH 8.0, 5% Chelex, 2.5 mg/ml proteinase K) for 6 hr at 55°C. Proteins were removed by phenol:chloroform-purification [30] followed by salting-out using 5 M LiCl [31]. Genomic DNA was precipitated using ethanol [30], washed twice in 70% ethanol and resuspended in 20 µl of TE pH 8.0, of which 17 µl were used to prepare the microsatellite-enriched genomic library. Host DNA was detected in 100 × dilutions of *T. tenuis* DNA by PCR amplification of three out of four red grouse microsatellite markers (LLST1, LLSD4 and LLSD8 but not LLSD3 [32]), raising the possibility that the enriched library might contain red grouse DNA fragments.

*Sau*3AI (New England Biolabs) restriction fragments ranging from 100–500 bp were gel-isolated [33], ligated to *Sau*3AI linkers [34] and hybridization-enriched for microsatellites with repeat units AC and AAAG (and their complements). The pre-enrichment hybridization PCR amplification was omitted to minimize the isolation of

duplicate clones [35]. Enrichment hybridizations were performed overnight in 2 × SSC at 55°C (dinucleotides) and 60°C (tetranucleotides) followed by three washes at 55°C and 60°C, respectively, in 2 × SSC with 0.1% SDS.

Four polymorphic microsatellite markers were cloned from the first library (Table 1: Tte002, Tte003, Tte016 and Tte017). Because this number would be insufficient for many analyses (e.g. parentage analysis), a second genomic library was prepared as above [30] but with the following differences. An equivalent quantity of bulked nematodes from the same source was incubated in 400 µl of DNA extraction buffer (100 mM EDTA pH 8.0, 0.5% SDS, 50 µg/ml proteinase K, 12.5 µg/ml RNase A, 1% β-mercaptoethanol) for 2 hr at 50°C, the phenol:chloroform step was omitted and the purified DNA was precipitated with isopropanol. The genomic library was prepared using an isoschizomer of *Sau*3AI, *Mbo*I (New England Biolabs), and enriched for microsatellites with repeat units AC, AG, TTTA, GTAA, GATA and AAAG (and their complements). Both di- and tetranucleotide hybridizations were performed at 60°C.

The enriched fragments were ligated into pUC18-*Bam*HI/BAP (Amersham Pharmacia Biotech) and transformed into XL1-Blue competent cells (Stratagene) according to manufacturers' instructions. The resulting 7700 transformant colonies were screened by hybridization to [α^{32}P]-dCTP-labelled dinucleotide (Pharmacia) and tetranucleotide [29] microsatellite probes (polymers of AC.GT and AAAG.CTTT for the first library and AC.GT, AG.CT, TTTA.TAAA, GTAA.TTAC, GATA.TATC and AAAG.CTTT for the second), leading to the identification of 458 positive colonies. Plasmid vectors were purified from all positive colonies and the insert sequenced in the forward direction by the Department of Genetics Sequencing Facility, University of Cambridge, UK (first library), and Lark Technologies, Saffron Walden, UK (second library). Sequences containing ambiguous base calls were re-sequenced in the reverse direction to allow the creation of a consensus sequence. All sequences were checked for the presence of re-

ligated *Sau*3AI and *Mbo*I restriction sites (GATC) to ensure that each sequence consisted of a single insert. All 458 sequences (length range 78–758 bp) contained microsatellite DNA, and were named Tte001–Tte458. Tte001– Tte019 were isolated from the first library and the remainder from the second library. From this number, the program BLASTALL (ftp://ftp.ncbi.nih.gov/blast/) [36] identified 309 unique sequences (rather than duplicates or alleles), which were submitted to the EMBL database (accession numbers **AM167567**–**AM167875**). Two sequences, Tte008 (**AM167573**) and Tte010 **(AM167575),** were withheld from further analysis because ENSEMBL BLAST (http://www.ensembl.org/Multi/blastview) identified close similarity to chicken genome sequence, suggesting that they may have been cloned from contaminant red grouse DNA. However, neither locus PCR-amplified a product from red grouse DNA. The remaining 307 sequences were tested for similarity with all the DNA sequences in GenBank (17 November 2005) using BLASTN (http://www.ncbi.nlm.nih.gov/BLAST/) [36]. The degree of homology within groups of DNA sequences was gauged by calculating the mean pairwise nucleotide diversity (π) using the p-distance option in MEGA version 3.1 [37].

Ninety-four of the 307 microsatellites were tested for PCR amplification and polymorphism in two stages: (1) initial testing on a few individuals followed by (2) more detailed assessment of the most promising loci in a larger sample. PCR primer pairs, designed using the program Primer3.0 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) [38], were initially tested for amplification and polymorphism in 5–14 adult male nematodes collected from at least two red grouse from at least two sites in Scotland and England. Only male nematodes were genotyped to avoid the possibility of contamination from sperm and fertilized eggs in females [15]. PCR primers were also tested on DNA extractions from two red grouse to guard against the possibility that microsatellites might have been cloned from red grouse DNA. However,

no PCR product was amplified from red grouse DNA. PCRs were carried out in 10-µl volumes containing 1 µl of 20 × diluted DNA extraction [39], 1 × Magnesium-free PCR buffer (Promega), 1.5 mM MgCl2, 0.1 mM dNTPs, 1 µM each primer and 0.25 U Taq (Promega). The PCR program was: 94 °C for 2 min, 12× [92°C for 45s, 55°C for 45s, 72°C for 50s], followed by 25× [89°C for 30s, 55°C for 45s, 72°C for 50s], and finally 5 min at 72°C. PCR products were resolved on 6% denaturing polyacrylamide gels and visualized by silver staining [40].

Twenty-one loci were polymorphic (Table 1). Five of these were excluded from further analysis because of low variability (Tte025, Tte027), suspected null alleles (Tte114, Tte199) and relative difficulty in scoring due to artefactual peaks (Tte201). However, given their high variability, Tte114, Tte199 and Tte201 are promising candidates for future optimization. Population genetic analyses were conducted at the population and infrapopulation levels using the remaining sixteen loci. (Here a population is defined as all the *T. tenuis* infecting red grouse on single grouse estate, and an infrapopulation as all the *T. tenuis* infecting a single host individual [41].) We genotyped 111 adult male nematodes from four male red grouse (numbered 1–4) harvested on a single estate in Inverness-shire, Scotland on 18 August 2003. Respectively, the numbers of males genotyped from each host were 31, 28, 26 and 26, and the nematode burdens of the hosts were estimated as 3380, 2450, 1780 and 2360 [23]. The sixteen loci were divided into three multiplex panels for automated genotyping using fluorescently labelled primers (Applied Biosystems; see Table 1). One primer of each pair was 5′ -labelled with a fluorescent dye, with the exception of locus Tte134 for which both primers were labelled to boost its signal.

Each panel was amplified by multiplex PCR using the Multiplex PCR Kit (Qiagen) with the exception of Tte017, which failed to amplify in multiplex PCR and was therefore amplified singly (also using the Multiplex Kit) and mixed with panel 2 following PCR.

Multiplex PCR was carried out on 1 µl of 20 × diluted DNA extraction [39] following the manufacturer's instructions except that the reaction volume was 10 µl, 35 amplification cycles were used and the annealing temperatures were 50°C for panel 1, 54°C for panels 2 and 3 and 60°C for Tte017. Allele lengths were measured using an ABI3730 DNA Analyzer (Applied Biosystems) by The Sequencing Service (University of Dundee, UK) and analyzed using GeneMapper (Applied Biosystems).

Population genetic analyses The level of genetic variability among the 111 *T. tenuis* males was gauged by calculating expected heterozygosity (*He*) and its standard error [42]. To allow comparison with *H. contortus*, for which polymorphic information content (PIC) rather than *He* has been reported, PIC was calculated using MolKin [43, 44]. We tested for the presence of null alleles, short allele dominance and scoring of stutter peaks using Micro-Checker [45]. Null allele frequencies were estimated according to van Oosterhout's method [45]. Linkage disequilibrium between each pair of loci was tested in Arlequin using 1000 permutations [46]. Deviation from Hardy-Weinberg equilibrium was tested using exact tests in Genepop version 3.4 [47]. Population genetic structure was investigated by performing an analysis of molecular variance (AMOVA) in Arlequin [46]. AMOVA was used to quantify the partitioning of genetic variance at three levels: among infrapopulations (equivalent to *FST*, a measure of interinfrapopulation structure); among individuals within infrapopulations (analogous to the inbreeding coefficient, *FIS*); and within individuals (equivalent to 1 – *FIT*).

AMOVA was performed assuming the infinite alleles mutation model. Three sex-linked loci and two loci showing evidence of null alleles were excluded from the AMOVA to avoid inflating the estimate of *FIS*. Deviation from the null hypothesis that all of the variance is distributed within individuals and none between individuals or between

populations was tested using 10,000 permutations. Genetic structure among the four infrapopulations was further examined by estimating total and pairwise *FST* with 95% confidence intervals in FSTAT [48]. For all analyses except for those investigating population structure genotypes were pooled across infrapopulations.

Results

Microsatellite isolation We isolated 307 microsatellites from *T. tenuis* with an average of 10.4 repeats (range 3–73; Table 2). We detected flanking sequence homology between 196 of the 294 (AC)n microsatellites and microsatellites of at least one of two trichostrongyloid nematodes, *Trichostrongylus colubriformis* and *Haemonchus contortus*. In each case the source of sequence similarity was a class of conserved repetitive elements associated with (AC)n microsatellites called TcREP in *T. colubriformis* [49] and HcREP1 in *H. contortus* [13, 19]. More than half of all known trichostrongyloid microsatellites (excluding those found within genes) are linked to repeats of this class, which are found tandemly repeated downstream of (GT)n microsatellites [13, 19, 49]. For brevity we refer to trichostrongyloid microsatellites that share flanking sequence homology with TcREP- and HcREP1-linked microsatellites as "REP$^+$" and the remainder as "REP", and for consistency with previous reports [13, 19, 49] we describe REP$^+$ microsatellites in the GT rather than the AC orientation. The 196 REP$^+$ *T. tenuis* microsatellites are distinguished by the suffix "REP", e.g. Tte001REP.

Multiple alignment of the REP$^+$ sequences showed that homology begins about 35 bp upstream of (GT)n microsatellites (Fig. 1a) and continues for at least 242 bp downstream (Fig. 1b). Multiple alignment of the 60 longest downstream sequences (at least 150 bp) with each other and with the full lengths of TcREP and HcREP1 divided them into two distinct groups of homologous sequences, which are exemplified by Tte323REP (51 sequences, π = 13.1%) and Tte036REP (nine sequences; π = 5.0%). The downstream regions of the Tte323REP-type sequences consist of 146-bp tandem repeats sequences closely homologous to TcREP and HcREP1 (Fig. 1b). The consensus of the Tte323REP-type sequences, which we have designated TteREP1,

shares 79% identity with TcREP and 54% with HcREP1, in concordance with the expected phylogenetic relationship [50]. Homology between the TteREP1 element adjacent to the microsatellite and elements further downstream declines with distance from the (GT)n microsatellite (Fig. 1b), a pattern previously observed in HcREP1 [13]. The consensus of the Tte036REP-type downstream region is closely homologous to a sequence just downstream of TcREP but not to TcREP, HcREP1, or TteREP1 (Fig. 1c).

Prediction of RNA secondary structure of conserved sequences from Tte009REP, Tte036REP, Hcms21 and Tc15 (consisting of the upstream conserved 35 bp, the microsatellite and one REP element downstream) using the MFold server (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi) [51] showed no consistent structural motif that might suggest that they could form similar functional RNA molecules (e.g. transposable elements). Six-way translated BLAST searching of the same sequences revealed no matches to indicate that these repeats could have been translated into a known class of protein.

Microsatellite polymorphism High levels of variability were detected among the 111 *T. tenuis* males using the sixteen microsatellite loci (Table 1; mean *He* = 0.708, mean PIC = 0.682, mean alleles

per locus = 12.2). Polymorphism is restricted to REP microsatellites: REP^{+} loci were invariably monomorphic as well as being relatively refractory to PCR amplification (Table 3). No evidence was found for short allele dominance, spurious scoring of stutter peaks or linkage disequilibrium. However, both Tte003 and Tte030 were identified as harbouring null alleles at estimated frequencies of 0.074 and 0.174 respectively. No heterozygotes were observed among the 111 males at three of the loci (Tte016, Tte218 and Tte378) despite high *He*. We have observed high frequencies of heterozygotes in genotypes of females at these three loci. Given that males are the heterogametic sex (XO) in

Trichostrongylus nematodes [52], we conclude that Tte016, Tte218 and Tte378 are sex-linked.

Analysis of population genetic structure

The AMOVA provided no evidence for genetic structure between infrapopulations or deviation from random mating within infrapopulations (Table 4). Had we not screened the loci for null alleles, and consequently included Tte003 and Tte030 in the AMOVA, we would have detected a moderate but highly significant deficit of heterozygotes within infrapopulations (*FIS* = 0.060, *P* < 0.0001). For all six pairwise *FST* estimates between the four infrapopulations, the 95% confidence interval comfortably overlaps zero, confirming the lack of support in the data for genetic structure between infrapopulations (Table 5). The overall *FST* estimate among the four infrapopulations was 0.001 (95% confidence interval: –0.003, 0.005).

Discussion

This study reports the isolation of 307 microsatellite sequences from the nematode *T. tenuis*, and the development of sixteen highly variable microsatellite markers. The utility of these markers for population genetic analysis was demonstrated by examining the partitioning of genetic variation between and within infrapopulations from four host birds from a single location. We also describe a new repeat element linked to *T. tenuis* AC/GT microsatellites, TteREP1, which is related to TcREP [49] and HcREP1 [13], and whose presence appears to thwart marker development.

Similarly high levels of microsatellite variation have been found in other animal parasitic nematodes [15, 16], including the trichostrongyloids *H. contortus* [13, 19] and *Teladorsagia circumcincta* [53], as well as in plant parasitic nematodes [20, 21] (Table 6). However, too few species have been studied to allow any correlates of microsatellite variation to be identified. Inter-species comparison of levels of microsatellite variation is also confounded by the fact that the level of variation reported may differ between studies due to differences in microsatellite isolation methods, such as differences in hybridization temperature which influences the number and sequence of microsatellites isolated. In addition, the threshold for a marker to be ascertained as polymorphic may differ between studies, so that in some studies markers with low polymorphism may be either undiscovered or unreported. Nevertheless, given that 22 of the 48 *He* values summarized in Table 6 exceed 0.75, it is clear that microsatellite markers are potentially powerful tools for uncovering genetic variation in a wide range of parasitic nematode species.

Nevertheless, the utility of microsatellite markers must be balanced against the effort required to develop them, which for parasitic nematodes appears to be unusually high. Only sixteen of the 307 *T. tenuis* microsatellites were successfully developed into markers, a quarter of the success rate achieved using exactly the same methods in the

common buzzard [54]. The trouble was not that microsatellite sequences were difficult to isolate, nor that an unusually high proportion was monomorphic (even setting aside the REP$^+$ sequences), but that so many of them failed to yield an interpretable PCR product (Table 3). The PCR amplification failure rate was 58% for REPmicrosatellites compared with 11% in the common buzzard, which is a typical rate in birds and mammals (PCDJ, LFK, DAD, pers. obs.). Similar problems have been experienced in other parasitic nematodes. Five out of six loci failed to amplify in *Strongyloides ratti*, a strongyloid parasite of rats [18]; 30 out of 69 in *H. contortus* [13, 19]; six out of 17 in the human whipworm *Trichuris trichiura* [16]; five out of 19 in the potato cyst nematode *Globodera pallida* [20]; and five out of ten in the sugar beet cyst nematode *Heterodera schachtii* [21]. In contrast, no problems were encountered in amplifying seven microsatellites of a parasite of grapevine and fig, the dagger nematode *Xiphinema index* [55].

Why are nematode microsatellites so difficult to amplify? Fisher and Viney proposed that the amplification of multiple products might be explained by the location of microsatellites in repetitive sequences leading to multiple priming sites [18]. This theory would explain the 82% amplification failure rate of REP$^+$ microsatellites in this study, and undetected repetitive features may also explain the low success of REP loci. However, flanking sequence homology did not hinder the development of polymorphic microsatellites in *Xiphinema index* [55], while in *H. contortus* HcREP1-linked microsatellites were successfully developed into variable markers, with HcREP1 even being exploited as the site for a generic primer [19] (a similar approach using generic primers in *T. tenuis* was unsuccessful; N. Temperley, pers. comm.). In the absence of any pattern to indicate whether a particular species is amenable to microsatellite development, developing microsatellite markers in nematodes will likely remain a challenging venture with unpredictable results.

The utility of the microsatellites developed here was demonstrated by showing that the nematodes sampled from four male red grouse could have been drawn from a single panmictic population. At the inter-infrapopulation level, this result could be explained by a high degree of nematode gene flow between infrapopulations within a single estate, which would indicate that male red grouse do not predominantly reinfect themselves despite their strongly territorial behaviour from autumn to spring [56]. However, because *FST* conflates the effects of gene flow and effective population size, the lack of population structure observed could also be explained by high effective population size (and correspondingly slow genetic drift) and low levels of gene flow. Nevertheless, the 100% prevalence and high intensity of *T. tenuis* infections in adult grouse [23] also suggest that inter-host transmission is frequent. This scenario contrasts with *Ascaris* spp., where small infrapopulations, frequent autoinfection and recruitment of groups of siblings appear to have caused strong partitioning of genetic variance between infrapopulations [15, 57-59].

The absence of genetic structure within infrapopulations of *T. tenuis* is perhaps also unsurprising, given the high nematode abundance found in the hosts. High levels of within-infrapopulation structure (the Wahlund effect [60]), as have been detected in *Ascaris suum*, could occur through the infection of a single host with multiple genetically distinct groups of larvae [58]. However, this effect requires the prior existence of inter-infrapopulation structure, which we have shown to be absent in *T. tenuis*.

The discovery and exclusion from the analysis of two loci with suspected null alleles that would otherwise have caused a significant positive *FIS* highlights the dangers of trying to detect biologically meaningful heterozygote deficits using microsatellites. Inferences of non-random mating within infrapopulations based on high *FIS* estimates should be viewed with scepticism when they are dependent on heterozygote deficits in one or a few loci, or where *FIS* varies widely between loci. Screening loci for null alleles is

therefore imperative. Although unsuitable for some purposes, loci affected by a null allele can often be accommodated in analyses [45].

The discovery of TteREP1 in *T. tenuis* brings to four the number of species from which TcREP-class repeats have been sequenced: they have also been found in *T. colubriformis* [49], *H. contortus* [13] and *Teladorsagia circumcincta* [53], pers. comm.). PCR amplification using TcREP-specific primers has shown that their distribution is widespread throughout the Trichostrongyloidea, and in the cases of *Oesophagostomum radiatum* and *Nematodirus spathiger* extends beyond this superfamily [61]. Although the sequence similarity between TteREP1, TcREP and HcREP1 makes phylogenetic sense, the origin of this class of repeats is a mystery, as does its mode of replication. Might it spread by retrotransposition? Various retrotransposable elements have been suspected of spreading microsatellites throughout genomes [62-64]. However, the lack of homology of TcREP-class repeats to any known retrotransposable element and the failure to form a consistent RNA secondary structure do not support this view.

In conclusion, this study illustrates the feasibility, as well as some pitfalls, of developing microsatellite markers with sufficient power to investigate parasitic nematode population genetics at an unprecedented level of detail.

Acknowledgements

We thank Andrew Leviston for help with BLAST analysis, Stuart Piertney for providing grouse primers and DNA, Peter Hudson, Victoria Grillo, John Gilleard and Pedro Pedro for valuable discussions and insights, and Barbara Mable, Terry Burke and two anonymous referees for commenting on the manuscript. We are also grateful to the Dave Newborn, the Game Conservancy Trust, UK, and to many gamekeepers and estate owners for access to sampling sites and assistance with sampling. This study was funded by the Natural Environment Research Council (NERC), UK and The Leverhulme Trust, UK.

References

[1] Taylor SM. Resistanceto chemical treatment of endoparasites. Meat & Livestock Commission, Milton Keynes, UK, 1999.

[2] Van Wyk JA. Refugia – overlooked as perhaps the most potent factor concerning the development of anthelmintic resistance. Onderstepoort J Vet Res 2001;68:55-67.

[3] Lenormand T, Raymond M. Resistance management: the stable zone strategy. Proc R Soc Lond B Biol Sci 1998;265:1985-90.

[4] Wolstenholme AJ, Fairweather I, Prichard R, von Samson-Himmelstjerna G, Sangster NC. Drug resistance in veterinary helminths. Trends Parasitol 2004;20:469-76.

[5] Lenormand T, Bourguet D, Guillemaud T, Raymond M. Tracking the evolution of insecticide resistance in the mosquito Culex pipiens. Nature 1999;400:861-4.

[6] Cornell SJ, Isham VS, Smith G, Grenfell BT. Spatial parasite transmission, drug resistance, and the spread of rare genes. Proc Natl Acad Sci U S A 2003;100:7401-5.

[7] Blouin MS, Liu J, Berry RE. Life cycle variation and the genetic structure of nematode populations. Heredity 1999;83:253-9.

[8] Brant SV, Orti G. Evidence for gene flow in parasitic nematodes between two host species of shrews. Mol Ecol 2003;12:2853-9.

[9] Braisher TL, Gemmell NJ, Grenfell BT, Amos W. Host isolation and patterns of genetic variability in three populations of Teladorsagia from sheep. Int J Parasitol 2004;34:1197-204.

[10] Tautz D, Renz M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 1984;12:4127-38.

[11] Jarne P, Lagoda PJL. Microsatellites, from molecules to populations and back. Trends Ecol Evol 1996;11:424-9.

[12] Toth G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: Survey and analysis. Genome Res 2000;10:967-81.

[13] Hoekstra R, Criado-Fornelio A, Fakkeldij J, Bergman J, Roos MH. Microsatellites of the parasitic nematode Haemonchus contortus: polymorphism and linkage with a direct repeat. Mol Biochem Parasitol 1997;89:97-107.

[14] Conole JC, Chilton NB, Jarvis T, Gasser RB. Mutation scanning analysis of microsatellite variability in the second internal transcribed spacer (precursor ribosomal RNA) forthree species of Metastrongylus (Strongylida: Metastrongyloidea). Parasitology 2001;122:195-206.

[15] Anderson JD, Williams-Blangero S, Anderson TJC. Spurious genotypes in female nematodes resulting from contamination with male DNA. J Parasitol 2003;89:1232-4.

[16] Barker GC, Bundy DAP. Isolation and characterization of microsatellite loci from the human whipworm Trichuris trichiura. Mol Ecol 2000;9:1181-3.

[17] Underwood AP, Bianco AE. Identification of a molecular marker for the Y chromosome of Brugia malayi. Mol Biochem Parasitol 1999;99:1-10.

[18] Fisher MC, Viney ME. Microsatellites of the parasitic nematode Strongyloides ratti. Mol Biochem Parasitol 1996;80:221-4.

[19] Otsen M, Plas ME, Lenstra JA, Roos MH, Hoekstra R. Microsatellite diversity of isolates of the parasitic nematode Haemonchus contortus. Mol Biochem Parasitol 2000;110:69-77.

[20] Thiéry M, Mugniéry D. Microsatellite loci in the phytoparasitic nematode Globodera. Genome 2000;43:160-5.

[21] Plantard O, Porte C. Isolation and characterization of microsatellite loci in the sugar beet cyst nematode Heterodera schachtii. Mol Ecol Notes 2003;3:139-41.

[22] Watson H, Lee DL, Hudson PJ. The effect of Trichostrongylus tenuis on the caecal mucosa of young, old and anthelmintic-treated wild red grouse, Lagopus lagopus scoticus. Parasitology 1987;94:405-11.

[23] Hudson PJ. The effect of a parasitic nematode on the breeding production of red grouse. J Anim Ecol 1986;55:85-94.

[24] Shaw JL, Moss R. Effects of the cecal nematode Trichostrongylus tenuis on egg-laying by captive red grouse. Res Vet Sci 1990;48:59-63.

[25] Delahay RJ, Moss R. Food intake, weight changes and egg production in captive red grouse before and during laying: Effects of the parasitic nematode Trichostrongylus tenuis. Condor 1996;98:501-11.

[26] Hudson PJ, Dobson AP, Newborn D. Prevention of population cycles by parasite removal. Science 1999;282:2256-8.

[27] Hudson PJ, Newborn D. A Manual of Red Grouse and Moorland Management. Fordingbridge, UK: The Game Conservancy Trust, 1995.

[28] Newborn D, Foster R. Control of parasite burdens in wild red grouse Lagopus lagopus scoticus through the indirect application of anthelmintics. J Appl Ecol 2002;39:909-14.

[29] Armour JAL, Neumann R, Gobert S, Jeffreys AJ. Isolation of human simple repeat loci by hybridization selection. Hum Mol Genet 1994;3:599-605.

[30] Sambrook J, Fritsh EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press, 1989.

[31] Gemmell NJ, Akiyama S. An efficient method for the extraction of DNA from vertebrate tissue. Trends Genet 1996;12:338-9.

[32] PiertneySB, Dallas JF. Isolation and characterization of hypervariable microsatellites in the red grouse Lagopus lagopus scoticus. Mol Ecol 1997;6:93-5.

[33] Glenn TC, Glenn SJ. Rapid elution of DNA from agarose gels using polyester plug spin inserts (PEPSIs). Trends Genet 1994;10:344.

[34] Royle NJ, Hill MC, Jeffreys AJ. Isolation of telomere junction fragments by anchored polymerase chain reaction. Proc R Soc Lond B Biol Sci 1992;247:57-61.

[35] Gibbs M, Dawson DA, McCamley C, Wardle AF, Armour JAL, Burke T. Chicken microsatellite markers isolated from libraries enriched for simple tandem repeats. Anim Genet 1997;28:401-17.

[36] Altschul SF, Madden TL, Schaffer AA, et al. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389-402.

[37] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 2004;5:150-63.

[38] Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, eds. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa: Humana Press, 2000:365-86.

[39] Floyd R, Abebe E, Papert A, Blaxter M. Molecular barcodes for soil nematode identification. Mol Ecol 2002;11:839-50.

[40] Bassam BJ, Caetano-Anolles G, Gresshoff PM. Fast and sensitive silver staining of polyacrylamide gels. Anal Biochem 1991;196:80.

[41] Bush AO, Lafferty KD, Lotz JM, Shostak AW. Parasitology meets ecology on its own terms: Margolis et al revisited. J Parasitol 1997;83:575-83.

[42] Nei M, Roychoudhury AK. Sampling variances of heterozygosity and genetic distance. Genetics 1974;76:379-90.

[43] Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment polymorphisms. Am J Hum Genet 1980;32:314-31.

[44] Gutiérrez JP, Royo LJ, Álvarez I, Goyache F. MolKin (version 2.0) A Computer Program for Genetic Analysis of Populations Using Molecular Coancestry Information. J Hered 2005;96:718-21.

[45] van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. MICRO CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 2004;4:535-8.

[46] Schneider S, Roessli D, Excoffier L. Arlequin version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of

Geneva, Switzerland, 2000.

[47] Raymond M, Rousset F. GENEPOP (version1.2): population genetics software for exact test and ecumenicism. J Hered 1995;86:248-9.

[48] Goudet J. FSTAT (vers. 1.2): a computer program to calculate F-statistics. J Hered 1995;86:485-6.

[49] Callaghan MJ, Beh KJ. A middle-repetitive DNA sequence element in the sheep parasitic nematode, Trichostrongylus colubriformis. Parasitology 1994;109:345-50.

[50] Blaxter ML. Molecular Analysis of Nematode Evolution. In: Kennedy MW, Harnett W, eds. Parasitic Nematodes: Molecular Biology, Biochemistry and Immunology. Wallingford. UK: CABI Publishing, 2001:139-66.

[51] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406-15.

[52] John B. The chromosomes of zooparasites. Chromosoma 1957;9:61-8.

[53] Grillo V, Jackson F, Gilleard JS. Characterisation of microsatellites from Teladorsagia circumcincta and the development of a panel suitablefor population genetic studies (in press). Mol Biochem Parasitol 2006.

[54] Johnson PCD, Fowlie MK, Amos W. Isolation of microsatellite loci from the common buzzard (Aves, Accipitridae, Buteo buteo). Mol Ecol Notes 2004;5:208-11.

[55] He Y, Li HM, Brown DJF, Lamberti F, Moens M. Isolation and characterisation of Microsatellites for Xiphinema index using degenerate oligonucleotide primed PCR. Nematology 2003;5:809-19.

[56] Watson A. Social-class, socially-induced loss, recruitment and breeding of red grouse. Oecologia 1985;67:493-8.

[57] Anderson TJC, Romeroabal ME, Jaenike J. Mitochondrial DNA and Ascaris microepidemiology: the composition of parasite populations from individual hosts, families and villages. Parasitology 1995;110:221-9.

[58] Nadler SA, Lindquist RL, Near TJ. Genetic structure of midwestern Ascaris

suum populations: a comparison of isoenzyme and RAPD markers. J Parasitol 1995;81:385-94.

[59] Criscione CD, Poulin R, Blouin MS. Molecular ecology of parasites: elucidating ecological and microevolutionary processes.Mol Ecol 2005;14:2247-57.

[60] Wahlund S. Zusammensetzung von Population und Korrelationserscheinung vom Standpunkt der Vererbungslehre aus Betrachtet. Hereditas 1928;11:65-106.

[61] Gasser RB, Nansen P, Bogh HO. Specific fingerprinting of nematodes by PCR with single primers to defined repetitive elements (vol 60, pg 127, 1995). Acta Trop 1996;61:273.

[62] Beckman JS, Weber JL. Survey of human and rat microsatellites. Genomics 1992;12:627-31.

[63] Arcot SS, Wang Z, Weber JL. Alu repeats: a source for the genesis of primate microsatellites. Genomics 1995;29:136-44.

[64] Wilder J, Hollocher H. Mobile elements and the genesis of microsatellites in dipterans. Mol Biol Evol 2001;18:384-92.

[65] Rice WR. Analyzing tables of statistical tests. Evolution 1989;43:223-5.

Table 1

Primer sequences, diversity statistics and other characteristics of 21 polymorphic microsatellite marker loci developed for the parasitic nematode *Trichostrongylus tenuis*. T_a: annealing temperature; N: number of alleles; H_o: observed heterozygosity; H_e: expected heterozygosity (± SE); PIC: polymorphic information content; P1–3: multiplex panels 1–3.

Locus	EMBL Accession number	Longest repeat motif	Primer sequences (5′–3′)[e]	Panel[f] or T_a (°C)	Allele size range	N/ adults typed	H_o	H_e (± SE)	PIC
Tte002	**AM167568**	$(AC)_8$	GCTCTCAGGCGTCATATTCC fam-AGCCGTCTGTGAACATTCG	P1	139–151	10/105	0.762	0.760 (± 0.017)	0.727
Tte016[a]	**AM167580**	$(AC)_{10}$	fam-GCCGGGGTTGATTGAATG CTTCACCCTGTTTCGGGTTC	P1	171–184	13/104	n/a	0.695 (± 0.046)	0.681
Tte102	**AM167659**	$(AC)_7$	TCAATTCATGGTGAGATGGC fam-GCGTTTTAATGGCAACCAAG	P1	80–108	12/111	0.757	0.805 (± 0.010)	0.778
Tte211	**AM167742**	$(AC)_6$	ned-CTAGCACGGGTGAAGACTCC TGCCATCGAAATTAGAAAACTG	P1	130–148	8/106	0.651	0.660 (± 0.026)	0.619
Tte218[a]	**AM167748**	$(AC)_{13}$	ACACCACCCCCATCCTCTAC hex-CCTGTAGTGCACAAAGCTACG	P1	102–120	16/111	n/a	0.829 (± 0.020)	0.810
Tte254	**AM167772**	$(AC)_6$	hex-TGATTCGTTCTACATAATTTGG AACGGGGTAACTTGTTTTG	P1	131–166	18/110	0.836	0.818 (± 0.015)	0.798
Tte335	**AM167821**	$(AC)_{13}$	ned-CAGGAGCAGTTAAGACAATG GATCACCGCTCTTCCAC	P1	77–89	8/111	0.793	0.788 (± 0.015)	0.760
Tte003	**AM167569**	$(AC)_{17}$	CGAGCATGTTGGCGAAATC hex-GATCAAAGCAAGCGAAGG	P2	133–156	14/109	0.606	0.713 (± 0.028)	0.690
Tte030	**AM167594**	$(AG)_{12}$	AGAGCAAAAGGGGATGGAAC fam-GATCCGTTTTGTGTGTGTGTG	P2	118–159	22/107	0.589[b]	0.912 (± 0.008)	0.906
Tte303	**AM167799**	$(AC)_{11}$	GATCCAGAAAGGTGTGTG ned-ACTATTCATTTGACAAGGTG	P2	97–113	9/111	0.568	0.562 (± 0.031)	0.519

Tte378[a]	**AM167841**	$(AC)_8$	ned-ATCCCCGTAGAAGCGCTATG AGATAACGCGCGACATGG	P2	133–162	10/109	n/a	0.530 (± 0.054)	0.510
Tte057	**AM167620**	$(AC)_{14}$	GTTCATCCTTAAAACTCCTC ned-GATCGAATTTCTCTGTTC	P3	78–116	17/110	0.764	0.791 (± 0.020)	0.770
Tte134	**AM167688**	$(AC)_{12}$	hex-GCTGTTTGCTCTGGGAGTTC hex-GATCGAAGTCGTCGTGG	P3	111–137	14/109	0.835	0.870 (± 0.010)	0.857
Tte331	**AM167818**	$(AC)_{11}$	fam-GATCTCAGAGAATGCATTTC GATCTTGCTGGTGATTCG	P3	68–92	14/110	0.755	0.767 (± 0.024)	0.748
Tte365	**AM167832**	$(AC)_{10}$	GGCATGGTGGTGTCTTTTG fam-GCCTCGCCATGCTGTG	P3	126–136	7/110	0.573	0.577 (± 0.024)	0.508
Tte017	**AM167581**	$(AC)_5$	CGGGTGTTCCTTACCTTCAG ned-TTGCTGAGACGCGAGATATG	60[d]	175–181	4/99	0.273	0.245 (± 0.039)	0.231
Tte025	**AM167589**	$(AC)_5$	TACGATATCAGCACGGCGTA TATGTGCAGGGCGAGTACAG	60	102–106	2/14	0.125[c]	0.069 (± 0.063)	0.067
Tte027	**AM167591**	$(AG)_6$	TACGATATCAGCACGGCGTA TATGTGCAGGGCGAGTACAG	60	143–157	2/14	0.125[c]	0.069 (± 0.063)	0.067
Tte114	**AM167671**	$(AC)_{10}$	TACGATATCAGCACGGCGTA TATGTGCAGGGCGAGTACAG	60	147–163	3/6	0.167[c]	0.569 (± 0.079)	0.477
Tte199	**AM167735**	$(AC)_9$	GATCTAGGGTAGGCAATCCAGGAG CAAGGTTTGAACTGGGACCAA	63	100–124	7/12	0.545[c]	0.777 (± 0.047)	0.744
Tte201	**AM167737**	$(AG)_{14}$	GATCATCCACCGTTAGTGTCC CGGCGTAAGATATGAATTTCG	60	264–289	11/12	0.917[c]	0.854 (± 0.043)	0.841

[a] Sex-linked.

[b] Null hypothesis of Hardy-Weinberg equilibrium (HWE) rejected ($P < 0.05$ after sequential Bonferroni adjustment for multiple tests [65]).

[c] Deviation from HWE was not tested due to low sample size.

[d] Tte017 was amplified singly but was added to multiplex panel 2 after PCR for automated genotyping.

[e] The fluorescent label (FAM, HEX or NED) is indicated at the 5′-end of the primer sequence for loci that were automatically genotyped.

[f] The panel number (P1–3) is given for loci that were amplified by multiplex PCR, while for loci that were amplified singly the annealing temperature is shown (see Materials and Methods).

Table 2

Frequencies of seven repeat motifs among 307 *Trichostrongylus tenuis* microsatellite loci isolated from *T. tenuis*. *n*: mean number of repeat units.

Repeat unit	Frequency	*n* (range)
AC	294	10.5 (3–73)
AG	7	11.1 (6–15)
CAT	1	5
AAAC	1	6
AAAG	2	4 (3–5)
AAGT	1	5
GTAT	1	10

Table 3

PCR amplification success and polymorphism of TteREP1-linked (REP^+) and -unlinked (REP^-) microsatellite loci. REP^+ microsatellites amplified less frequently than REP^- loci ($\chi^2 = 5.2$, $P = 0.02$), and those REP^+ loci that did amplify were less likely to be polymorphic ($\chi^2 = 10.3$, $P = 0.001$). Amplification failure is defined as the absence of a PCR product or the appearance of a smear on a polyacrylamide gel.

	Total loci tested	Amplification failed	Monomorphic	Polymorphic
REP^+ loci	28	23	5	0
REP^- loci	66	38	7	21

Table 4

Hierarchical analysis of molecular variance (AMOVA) between and within infrapopulations (IPs) of *T. tenuis* calculated using data from eleven loci. The *P*-value is the probability of observing a higher variance component and associated *F*-statistic under the null hypothesis of no partitioning of variance among hosts or individuals.

Variance component	Variance	% Total	*F*-statistic	*P*-value
Among IPs	0.0005	0.01	$F_{ST} = 0.000$	1.00
Among individuals within IPs	0.0564	1.47	$F_{IS} = 0.015$	0.16
Within individuals	3.7824	98.5	$F_{IT} = 0.015$	0.17

Table 5

Pairwise F_{ST} with 95% confidence intervals between four *T. tenuis* infrapopulations (hosts 1–4) estimated using 111 sixteen-locus microsatellite genotypes. Confidence intervals are given in brackets and were estimated by bootstrapping over loci.

	Host 1	Host 2	Host 3
Host 2	–0.001 (–0.010, 0.011)		
Host 3	0.002 (–0.006, 0.013)	–0.001 (–0.007, 0.006)	
Host 4	0.001 (–0.005, 0.008)	–0.002 (–0.009, 0.005)	0.007 (–0.002, 0.019)

Table 6

Levels of microsatellite variation in six species of animal and plant parasitic nematodes, gauged by expected heterozygosity (H_e) or polymorphic information content (PIC) averaged over *n* marker loci.

Nematode species	Host	*n*	H_e	PIC	Source
Trichostrongylus tenuis	Red grouse	21	0.650	0.624	This study
Haemonchus contortus	Sheep	31	–	0.535	[13, 19]
Teladorsagia circumcincta	Sheep	7	0.611	–	[53]
Trichuris trichiura	Human	6	0.787	–	[16]
Globodera pallida	Potato	9	0.526	–	[20]
Heterodera schachtii	Sugar beet	5	0.562	–	[21]

Figure legends

Fig. 1. Intra- and inter-specific alignment of sequences upstream (a) and downstream (b, c) of *T. tenuis* $(GT)_n$ microsatellites. (a) Alignment of the two consensus sequences (exemplified by Tte323REP and Tte036REP) of the highly conserved 35-bp upstream sequence with homologous sequences upstream of *T. colubriformis* (Tc15) and *H. contortus* (Hcms3, Hcms21) microsatellites. (b) Alignment of the consensus TteREP1 sequence with three complete and two incomplete TteREP1 repeat elements from Tte009REP and Tte323REP, the TcREP element of Tc15, and the HcREP1 element of Hcms21. Adjacent TteREP1 repeats are numbered (e.g. Tte009REP#3) to indicate order downstream of the $(GT)_n$ microsatellite. (c) Alignment of the downstream sequence of Tte036REP with the homologous sequence from Tc15, located immediately downstream of TcREP. The first 75 bp of the Tc15 sequence (underlined) closely match the first half of TcREP (not shown) but not Tte036REP, while the remainder is highly homologous to Tte036REP. Nucleotide identity (ignoring insertions and deletions) is given at the end of the first row of each aligned sequence.

Figure 1 MOLBIO-S-06-00050

(a)

```
Tte323REP     ATCCTACTAT ATATAATCCG CTAAGTATGT CTGTC  35
Tte036REP     .......... .......T.. ..T.......  .....       94.3%
Tc15          .......... ........G. ....A.T...  .....       91.4%
Hcms3         .......... ......CG.. ......G...  ....T       88.6%
Hcms21        .......... .......T.. ..T.......  ....T       91.4%
```

(b)

```
TteREP1       GTCACGAAAT TCATATCTCG CGTCACAGCA ATGATGGTGA CCTGAATTTT ACAAGTGATA  60
Tc15          .......... .......... ....G...T. GA.C...GAC .......G.. TA........        78.8%
Hcms21        .A.CG..... .....A...C T...C.C.TG GACC.A.A.C GT.C..A... GGT..GAG..        53.6%
Tte009REP#1   .A.CG..... .......... .A..GA.... .A.....AAT .........C TAG...A...        78.1%
Tte323REP#1   .......... .......... ..G....... .......... .......... ..........        97.9%
Tte009REP#2   A..GG..... .......... ....GA.... TAA....AAT .......... TAG...T...        76.6%
Tte323REP#2   .....A...G ..G....G.C TT..G..TTG G.CC...A.C GT.C.CAC.C ...G...C..        63.5%
Tte009REP#3   A..GG.G... ......T... A..TG.T... .A.....-.. T.                          73.2%

TteREP1       TAGGGGCTGG GACGGAGACC CTAGCGGTGG TGGACCGAAA GTGATAGCAG CTGATAAGGG  120
Tc15          G.....G... T.A..GA..A TC.AT..... G...A..G.. .........A ..........
Hcms21        AG....GCA. ..T..GA..A .CTAT..... GT.G.G.G.C C--------- -------..A
Tte009REP#1   .....AG.A. ..G..GA..A TC.A.....A G.C....... .........A ..........
Tte323REP#1   .......... ......C... .......... .......... .......... .......C..
Tte009REP#2   G.A..-G.A. T.G..GA..A TC..T....A G....A.... .......T.A ..........
Tte323REP#2   .GT...G... .....G...A T..A...... GT..T.

TteREP1       CTGATAAGGG GCGTCGAAAG GCGAAT  146
Tc15          .........A ..C....GT. ..T...
Hcms21        ..CGGGT..C .G.A.C---- .-..C.
Tte009REP#1   .....G...A ..C....... ......
Tte323REP#1   .......... .......... ......
Tte009REP#2   .........A ..C....... ......
```

(c)

```
Tte036REP     GTCACGAAAT TCGGTATTAT ATGAAAAACG ACCTGTGGAG ATGGGGAC-- -AAGTTTGGG  60
Tc15          .CGGGT.... ..------.. ..CTCGCCGT CG.A...... C....AC.CT G..T...T.A        75.4%

Tte036REP     ATGGGTAGGG TATGGGAGAA GAAGTAAGGA GCAGACATGG ACTA-AGCTA AGCTCCGCCC  120
Tc15          .-----GCA. .....---- ------T..T .GTAT.C.-- ...GT..A.. ..........

Tte036REP     ACAGTGGGCG GGGCGTGCAC TAGAATGTT- CCGTTTTCAG ACCGAAAATT GTTAATTGGG  180
Tc15          ..G....... .AA....... .G-..-...T ..T....-.. .....G.... ..........

Tte036REP     GTCCAATATG TCCCAGACGT GGGGAAAAGG TGGAT-C  217
Tc15          T......G.. -....A.... ....G..-.A .....T.
```