Barcoding in trypanosomes

RACHEL HUTCHINSON[1] *and* JAMIE R. STEVENS[2]*

[1] *School of Biological Sciences, Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol BS8 1TQ, UK*
[2] *Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK*

SUMMARY

Trypanosomes (genus *Trypanosoma*) are parasites of humans, and wild and domestic mammals, in which they cause several economically and socially important diseases, including sleeping sickness in Africa and Chagas disease in the Americas. Despite the development of numerous molecular diagnostics and increasing awareness of the importance of these neglected parasites, there is currently no universal genetic barcoding marker available for trypanosomes. In this review we provide an overview of the methods used for trypanosome detection and identification, discuss the potential application of different barcoding techniques and examine the requirements of the 'ideal' trypanosome genetic barcode. In addition, we explore potential alternative genetic markers for barcoding *Trypanosoma* species, including an analysis of phylogenetically informative nucleotide changes along the length of the 18S rRNA gene.

Key words: 18S rRNA, Chagas disease, cytochrome oxidase I, molecular methods, sleeping sickness, *Trypanosoma*, Trypanosomatidae.

INTRODUCTION

> DNA barcoding: 'the utilisation of DNA sequences of short standardised gene fragments for quick and accurate determination of the species'
>
> (D'Avila-Levy *et al*. 2015)

Trypanosoma parasites are flagellated protozoa within the class Kinetoplastida, which is characterized by the presence of a kinetoplast: a mass of mitochondrial 'kDNA' (Adl *et al*. 2012). These parasites cause a wide range of diseases in both humans and animals, and are often transmitted between hosts by insect vectors (Fig. 1). Human diseases caused by parasitic trypanosomes carry a combined health burden of 2·2 million daily adjusted life years and primarily affect people from the poorest demographics in tropical and subtropical climates (Stuart *et al*. 2008), while in African animals, trypanosomiasis costs the livestock industry over US$ 4·5 billion every year (Yaro *et al*. 2016). Despite their devastating social and economic impact, these diseases remain widely under-reported; misdiagnosed, unidentified or asymptomatic cases, limited funding and the lack of a universal method for parasite detection and identification make surveillance and monitoring of these parasites difficult (Wastling and Welburn, 2011; Auty *et al*. 2012*a*; Stockdale and Newton, 2013; Franco *et al*. 2014).

Since the development of the first DNA-based identification methods for trypanosomes in the 1980s, the number of molecular detection techniques available (and iterations on these techniques) has increased dramatically; for examples, see the following reviews: Adams and Hamilton (2008); Taberlet *et al*. (2012). Although they constitute a vast improvement in sensitivity and specificity of diagnosis compared with microscopy methods (Gibson, 2007; Enyaru *et al*. 2010), the absence of a 'gold standard' for the detection and classification of trypanosomes has resulted in a distinct lack of comparable data between surveys (Auty *et al*. 2012*b*; Hernández and Ramírez, 2013; D'Avila-Levy *et al*. 2015). Most molecular techniques are too costly or complex for general use in front-line field diagnostics and, while developments in the transport of blood specimens have allowed samples to be analysed at centralized clinical laboratory facilities, the majority of molecular methods are still confined to research laboratories (Deborggraeve and Büscher, 2010).

Nonetheless, in other areas of biology and medicine, standardized, sequence-based barcoding (Hebert *et al*. 2003) has provided a sensitive, reliable method for the identification of species across a vast range of taxa and is now used by thousands of researchers worldwide (Coissac *et al*. 2016). However, despite the growing reference libraries of DNA barcodes for animals, plants and fungi (Ratnasingham and Hebert, 2007), there is currently no universal genetic barcoding marker available for trypanosome species. Accordingly, there is a clear need for a definitive, simple test suitable for the detection of all trypanosomes (Wastling and Welburn, 2011), with sensitivity and specificity

* Corresponding author. Department of Biosciences, College of Life and Environmental Sciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK. E-mail: j.r.stevens@exeter.ac.uk
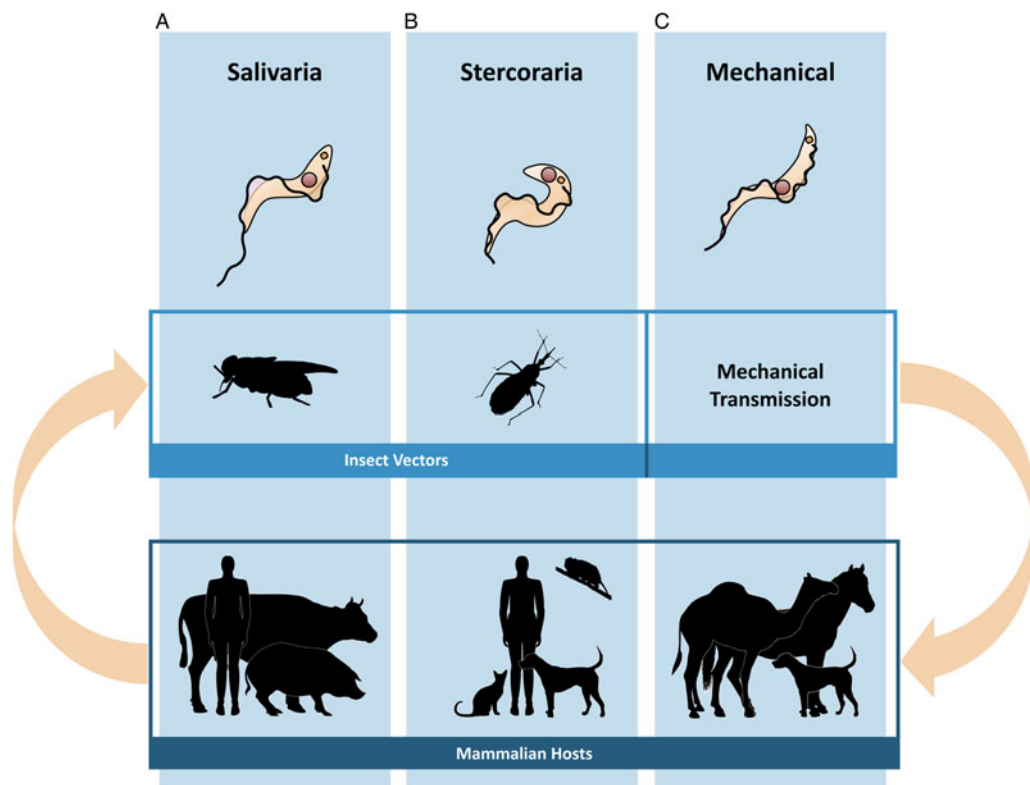
Fig. 1. Pathogenic trypanosomes of mammals. Trypanosomes are responsible for a number of diseases of both humans and animals. Chagas disease and human African trypanosomiasis (HAT) are considered 'neglected tropical diseases' by the World Health Organization and are transmitted between mammalian hosts by blood-feeding insect vectors. (A) Salivarian trypanosomes, characterized by development in the foregut of their insect vector, are confined to sub-Saharan Africa and are spread by the bite of the tsetse fly (*Glossina* spp.). These African trypanosomes, which include the human-infective *T. brucei* spp. and the major livestock pathogen *T. congolense*, cause the wasting diseases sleeping sickness (human African trypanosomiasis, HAT) and Nagana (animal African trypanosomiasis, AAT) across sub-Saharan Africa. (B) Stercorarian trypanosomes, characterized by development in the hindgut of their insect vectors, are mostly non-pathogenic. However, *Trypanosoma cruzi*, transmitted between mammalian hosts by the kissing bug (*Triatoma* spp.), causes Chagas disease, primarily in Latin America. When an infected kissing bug takes a blood meal, *T. cruzi* is passed out in the insect's feces and is typically deposited near the bite wound. The parasite enters the host when infected feces is spread into the wound, the eyes, mouth or breaks in the skin of the unaware host. (C) Three trypanosome species, *T. evansi*, *T. equiperdum* and *T. vivax*, are the major pathogens of livestock and have become adapted to mechanical transmission; they are now transmitted by a range of biting organisms (and, in the case of *T. equiperdum*, sexual contact) and, having lost the need for their ancestral tsetse fly host, they have spread beyond Africa to become disease agents in many parts of Asia and the Americas.

sufficient to differentiate between infections at the subspecies level, and usable for known, unknown and mixed infections. This is particularly pertinent from an epidemiological perspective for organisms that are morphologically identical but which require different treatments, such as the two human-infective trypanosomes that cause sleeping sickness (human African trypanosomiasis, HAT), *Trypanozoma brucei rhodesiense* and *Trypanozoma brucei gambiense* (Adams and Hamilton, 2008).

A key point regarding the continued relevance (or otherwise) of sequence-based barcoding, irrespective of the target locus, is the need for such a test to provide information on unknown taxa. This represents a very different requirement from a binary *yes/no* diagnostic test – generally an antibody-based method, which requires screening against panels of known potential infective agents to establish antibody specificity, levels of cross-reactivity and the

likelihood of scoring false-positives. In this context, a simple sequence-based test continues to offer advantages over an antibody-based diagnostic as, even with an unknown or previously unencountered taxon, such a test will yield a result that allows identification of an unknown organism as being most closely related to an organism of known sequence identity. Having established the continued benefits, a further major requirement is for such a test to work with sub-optimal sample material (and potentially degraded DNA), as is frequently encountered in field and/or clinical situations.

This review provides a critical overview of the development of barcoding techniques from traditional methods of trypanosome detection and identification, and examines the requirements of an 'ideal' barcode. An alternative approach to barcoding, based on the distribution of phylogenetically informative regions along a target gene, is presented

and we discuss whether barcoding can fulfil all the necessary requirements to become a truly universal method of identification. In other words: can barcoding be all things to all people?

Old faithful: microscopy

Despite the development of a variety of molecular methods for the detection and identification of infectious agents, the usual method for diagnosing trypanosome infections in vertebrate hosts remains the most basic: microscopic examination of sample preparations (Mugasa *et al*. 2012; Ricciardi and Ndao, 2015). However, this method is time consuming, dependent on operator expertise, unreliable for mixed infections, fails to detect immature infections and, in the case of African trypanosomes, is only useful for distinguishing between parasites to the level of subgenus (Ouma *et al*. 2000; Gibson, 2009; Enyaru *et al*. 2010; Auty *et al*. 2012b; Mugasa *et al*. 2014).

Early attempts to define the identity of pathogenic trypanosomes relied on a combination of microscopy and the ability, or otherwise, to passage parasites through laboratory host animals. In vertebrate hosts, where bloodstream-form trypanosomes exhibit a variety of distinctive morphological characteristics, this approach worked relatively well. However, the insect stages of trypanosomes from a range of subgenera are morphologically indistinguishable and, prior to the advent of enzymatic and molecular methods, the identification of different trypanosome species relied heavily on the site of infection in the insect vector (Hoare, 1972; Enyaru *et al*. 2010).

For human African trypanosomiasis, microscopic examination of cerebral spinal fluid can be used to determine the stage of disease progression, but the invasive procedure (lumbar puncture) required to collect samples often discourages patients from seeking medical help. A lack of formal training for front-line medical workers, local stigma surrounding diagnosis of sleeping sickness and a delay in patients contacting medical services only exacerbates the problem of surveillance and monitoring of this disease (Mpanya *et al*. 2012; Acup *et al*. 2016).

Not-so-quick kit: isoenzyme analysis

In the late 1960s, Lanham and Godfrey developed a cellulose column-based method utilizing the differential surface charge between trypanosomes and red blood cells to reliably separate parasites from host blood (Lanham and Godfrey, 1970). With this method, they were able to obtain relatively large-scale, pure preparations of live, undisrupted parasites suitable for subsequent biochemical analysis. At around the same time, Godfrey and colleagues developed a method to characterize trypanosomes using isoenzymes (Kilgour and Godfrey, 1973) and the characterization of many trypanosome species, subspecies and strains quickly followed (e.g. Godfrey and Kilgour, 1976; Miles *et al*. 1977). Several major isoenzyme-based studies followed and succeeded in defining the species and groupings of epidemiological significance recognized today (e. g. Gashumba *et al*. 1986; Gibson *et al*. 1988; Godfrey *et al*. 1990). Attempts were made subsequently to both streamline the methodology and to optimize the discriminatory power of the enzymes used (e.g. Stevens and Godfrey, 1992; Abderrazak *et al*. 1993), but ultimately the practical difficulties associated with isolating and preserving parasite enzyme extracts, reproducibility and issues of homoplasy in banding patterns led to the approach being superseded by DNA-based methodologies (e.g. Gibson and Borst, 1986; Hide *et al*. 1990).

Quick kit: serological tests

Antibody-detection tests, such as the card agglutination tests and the direct agglutination test, are widely used for the detection of trypanosomes in human hosts (Ricciardi and Ndao, 2015; Lutumba *et al*. 2016). These tests have excellent field application as they do not require a constant supply of electricity and are cheaper and more rapid than equivalent molecular techniques, although they can vary significantly in their sensitivity and specificity (Ricciardi and Ndao, 2015). Serological tests require relatively large samples and have the potential to yield false-negative results where parasitaemia is low or where antibody production is reduced, such as in immunocompromised patients (Papadopoulos *et al*. 2004; World Health Organisation, 2013). In addition, positive diagnoses obtained using serological tests nearly always require confirmation by microscopy, as these methods cannot distinguish between active infection and residual antigens from past infection or vaccination (Uilenberg and Boyt, 1998; Woods, 2013). Misdiagnosis of trypanosome infections remains a major problem, as treatment often carries a significant inherent risk (Barrett and Croft, 2012; Field *et al*. 2017).

The enzyme-linked immunosorbent assay, offers higher sensitivity than many other serological tests available, but it requires a sophisticated laboratory set-up that has restricted its use for diagnosis in the field (Chappuis *et al*. 2005).

The rise of molecular methods

DNA probes based on non-coding satellite repeats were the first molecular methods sensitive enough for the direct identification of trypanosomes in both host and vector samples without requiring cell cultures (Kukla *et al*. 1987; Gibson *et al*. 1988; McNamara *et al*. 1989). The development of the

polymerase chain reaction (PCR) heralded a major advance in the sensitivity of diagnostic techniques; PCR-based methods can identify trypanosomes at the subspecies level, they are suitable for analysis of mixed infections and can be applied to samples where parasite numbers are vanishingly low (Adams and Hamilton, 2008; Gibson, 2009; Matovu *et al.* 2010). Species-specific PCR-based methods are the most frequently used molecular tests for detection and identification of trypanosomes, but are limited by the number of species for which species-specific primers are available. Critically, these methods only detect known species: they cannot prove an absence of trypanosomes. In addition, screening samples for multiple trypanosome species using species-specific PCR methods requires a panel of probes; this can be expensive, time consuming and limits the number of samples that it is practical to analyse (Gibson, 2009; Adams *et al.* 2010; De Waal, 2012).

GENERIC PCR-BASED METHODS AND THE 'IDEAL' BARCODE

Historically, generic PCR methods have been less sensitive than species-specific PCR methods, but allow for multiple trypanosome species to be identified with a single test (Gibson, 2009). Most generic methods, such as restriction fragment length polymorphism PCR (RFLP-PCR) and ribosomal length-based methods, utilize multipurpose primers that target a semi-conserved region of the genome. Identification of an organism is made based on the length of the amplified regions (Adams and Hamilton, 2008). Although these methods each result in a species-specific 'barcode', none fulfil the requirements for the 'ideal' trypanosome barcode (Box 1).

Target gene

The success of a gene as a DNA barcode depends on a number of attributes, which must be considered when selecting gene targets: Is it a multicopy gene? How conserved is the sequence? How much does it vary across/between taxa/species? Is this level of variation constant across the gene? Some genes have been identified as universal barcodes, and are suitable for vast groups of organisms: the mitochondrial gene cytochrome c oxidase subunit 1 (*cox1*/ COI) is the accepted gold standard for molecular species identification of animals, and equivalents are available for plants and fungi. However, identifying universal barcodes in eukaryotic groups has proved difficult, not least because the level of genetic variability possible within each species is poorly understood (Enyaru *et al.* 2010), and consensus is yet to be reached regarding which genes to target and the criteria for delimiting species groups (Pawlowski *et al.* 2012; Pečnikar and Buzan, 2014).

Box 1. Defining the ideal trypanosome barcode

An ideal trypanosome barcode should be:

1. Optimal length: short enough to be sequenced in a single reaction, but long enough to capture all inter-taxon sequence variation.

2. Conserved: contain regions suitable for targeting with universal primers.

3. Phylogenetically informative: contain enough variability both between and within species to capture the full extent of trypanosome diversity.

4. Utilize a standardized set of universal primers applicable to all trypanosomes.

5. Reliable: the primer binding sites should be highly conserved and/or multicopy, so the primers are still applicable for field samples that may have suffered a degree of DNA degradation.

(Savolainen *et al.* 2005; Ferri *et al.* 2009; Valentini *et al.* 2009; Pečnikar and Buzan, 2014)

Valentini *et al.* (2009) discussed the different requirements of DNA barcodes for different users, and highlighted the differences between DNA barcoding '*sensu stricto*' and '*sensu lato*'. DNA barcoding '*sensu stricto*' is favoured by taxonomists and prioritizes standardization of primers with enough variation to elucidate a high level of phylogenetic information. DNA barcoding '*sensu lato*' is most suited to environmental samples and prioritizes short, robust primer binding sites that are resistant to degradation.

Molecular markers have been developed to target a wide range of trypanosome gene regions (Fig. 2A), but few have been the target of barcoding approaches (Fig. 2B). Fluorescent fragment length barcoding (FFLB) has been used to amplify small target regions in both the 18S small subunit ribosomal RNA (rRNA) and the 28S large subunit rRNA (Hamilton *et al.* 2008; Hamilton *et al.* 2011; Silva-Iturriza *et al.* 2013). This highly sensitive, PCR-based method uses four sets of primers: two target the 18S and are specific to trypanosomes, two target the 28S and are specific to all trypanosomatids (Hamilton *et al.* 2011). The length of the resulting fragments produces a pattern unique to each species, which can be matched to reference pattern profiles for species identification. FFLB can also detect novel trypanosome species and, although further analysis is needed to identify these novel species, the fragment patterns may provide an indication of phylogenetic relationships (Hamilton *et al.* 2008). However, there are a limited number
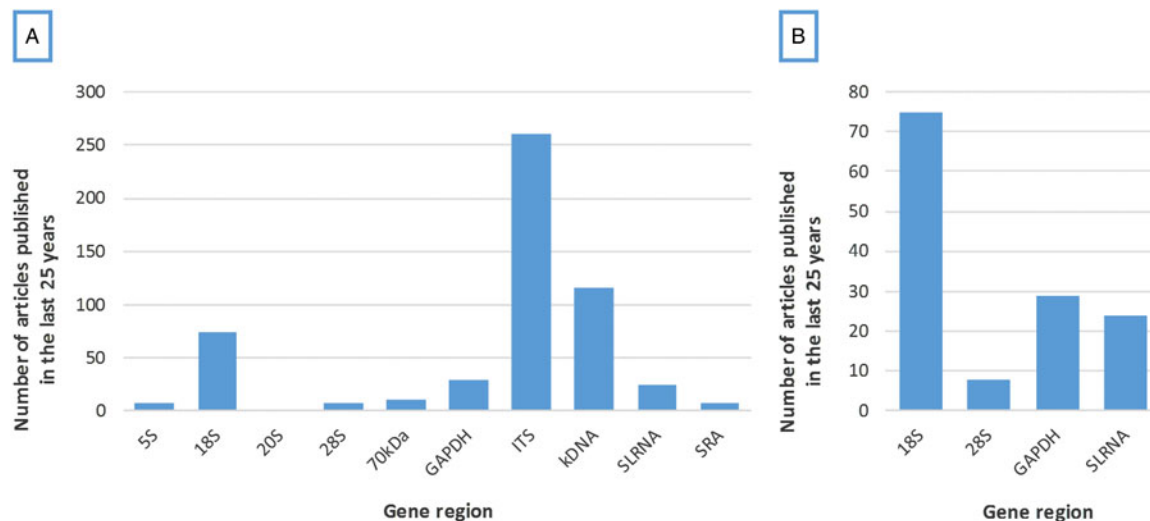
Fig. 2. (A) Gene regions commonly used for the identification of trypanosomes. (B) Gene regions that have also been the target of barcoding techniques for trypanosomes. The 18S and 28S rRNA regions have both been targeted using fluorescent fragment length barcoding (FFLB) (Hamilton *et al.* 2008, 2011; Silva-Iturriza *et al.* 2013), whilst cluster analysis has been used to delimit species when targeting the V7–V8 regions of the 18S rRNA gene (Lima *et al.* 2015) and regions of SL RNA (Votýpka *et al.* 2010). *Definitions:* 5S, 5S rRNA gene insertions; 18S, 18S small subunit ribosomal RNA (rRNA); 20S, editosome protein complex; 28S, 28S large subunit rRNA; 70 kDa, 70 kDa heat shock protein; GAPDH, glyceraldehyde phosphate dehydrogenase; ITS, internal transcribed spacer; SL RNA, spliced leader RNA; SRA, serum resistance associated gene. *Meta analysis method:* A literature search of the Web of Science (Clarivate Statistics, 2017) database was conducted in September 2017 to identify gene regions used for the identification or classification of trypanosomatids in the past 25 years. Subsequent searches were conducted to quantify the number of papers published between 1993 and 2017 that utilized each identified gene region (full Boolean search criteria available in Supplementary Information).

of reference profiles available for FFLB, which restricts its use as a trypanosome identification tool at this time (Hamilton *et al.* 2011; Silva-Iturriza *et al.* 2013), and this method cannot be used to discriminate between *T. brucei* subspecies (Hamilton *et al.* 2008).

The 18S rRNA gene has long been a popular target for molecular detection methods in protists (D'Avila-Levy *et al.* 2015). It is a highly expressed multicopy gene, present in all eukaryotes, with an assortment of conserved and variable nucleotide sequences that offer targets for universal primers, whilst still providing a wealth of taxonomic information. As sequence-based molecular methods gained popularity, the 18S rRNA gene succeeded protein-coding genes (e.g. Fernandes *et al.* 1993; Hashimoto *et al.* 1995; Adjé *et al.* 1998) to become the gene of choice for nearly all trypanosome evolutionary analysis (Maslov *et al.* 1996; Lukes *et al.* 1997; Haag *et al.* 1998; Stevens *et al.* 1998, 1999) and, as a result, has formed the basis of all modern trypanosome taxonomic frameworks (e.g. Hamilton *et al.* 2007; Lima *et al.* 2015; Dario *et al.* 2017). However, while nearly all trypanosome phylogenies have been constructed using 18S rRNA sequences, inadequate signals at certain depths of phylogenetic reconstruction have necessitated the use of additional trypanosome gene markers such as the glyceraldehyde phosphate dehydrogenase (GAPDH) gene. Nonetheless, the framework described using the 18S

rRNA has proven robust: other gene markers have complemented and strengthened this framework without fundamentally changing the nature of the basic relationships described based on 18S rRNA data; ultimately, this framework has also been fully supported by whole-genome phylogenetic comparisons (Leonard *et al.* 2011). In addition, as the 18S rRNA is one of the most widely used markers for trypanosomes, it is well represented in sequence databases such as GenBank (D'Avila-Levy *et al.* 2015).

Another popular molecular marker is the GAPDH gene. Few if any gaps are required for alignment of trypanosome GAPDH sequences, and sequences are shorter than those of 18S rRNA; sequencing this 'housekeeping gene' can be more economical, but provides a complementary depth of phylogenetic information in trypanosomes (Hamilton *et al.* 2004; Adams and Hamilton, 2008). GAPDH genes are relatively conserved and are therefore useful for resolving deep phylogenetic relationships (Hamilton *et al.* 2004). However, in order to determine close relationships, GAPDH must be used in conjunction with another barcoding marker; GAPDH has been used successfully with the 18S rRNA for trypanosome identification, and has proven suitable for novel species and mixed infections (e.g. Hamilton *et al.* 2008; Barbosa *et al.* 2016).

Internal transcribed spacer (ITS) regions have been widely used for barcoding in some organisms, e.g. fungi (Pawlowski *et al.* 2012); however, while

they have long been utilized for the detection of trypanosomes (Desquesnes and Davila, 2002; Adams *et al*. 2008; Desquesnes *et al*. 2011; Hernández and Ramírez, 2013), they have not yet been used specifically for the barcoding of different trypanosome species. Identification of species depends on the length of the amplified fragments of ribosomal RNA produced *via* PCR using primers complementary to conserved regions of the 18S, 28S and 5·8S rRNA genes matching all species of interest. This means species determination is possible for mixed infections, except in cases where the amplicon length is similar between species or there is intra-species variation (Adams and Hamilton, 2008; Hamilton *et al*. 2008; Gibson, 2009). Another constraint of the ITS region, as with all mitochondrial genes as targets for barcoding, is its relatively low copy number (100–200 repeats), compared with that of satellite DNA (10 000–20 000 repeats), which can limit the sensitivity of tests (Desquesnes and Davila, 2002).

The kinetoplast is a modified mitochondrion unique to kinetoplast protists and kinetoplast DNA (kDNA) minicircles have been successfully used in PCR assays for the identification of a number of *Trypanosoma* species. The high copy number of these minicircles – several thousand per cell – lends itself to highly sensitive diagnostics. However, high levels of nucleotide polymorphism between repeats of kDNA fragments make these genes unsuitable for sequence alignment (De Oliveira Ramos Pereira and Brandão, 2013). Only very short regions (100–200 base pairs) of kDNA minicircles are conserved and for some trypanosomes, such as *T. brucei*, there is only one of these regions per minicircle (Jensen and Englund, 2012). Low levels of conserved sequences in kDNA make it difficult to develop universal primers and limit the depth of phylogenetic information that can be elucidated from these sequences.

Spliced leader RNA (SL RNA) or 'mini-exon donor RNA' is another feature unique to kinetoplastid protists and has also been used as a target for barcoding (Rodrigues *et al*. 2010; Lima *et al*. 2015). The SL RNA genes are arranged as tandem repeats, with each repeat comprising many repeat units with regions of differing variability (Rodrigues *et al*. 2010). The conserved regions are convenient for primer targeting, whilst the more variable intergenic regions permit distinction between closely related trypanosomes (Westenberger *et al*. 2004). However, there are no primers currently available that are applicable to all trypanosomes (D'Avila-Levy *et al*. 2015), and the high mutation rate of intergenic regions makes it difficult to compare sequences across the full spectrum of trypanosomes or to define any meaningful phylogeny beyond closely related taxa (Gibson *et al*. 2000). Previous attempts to use SL RNA barcodes for trypanosomes delimited species using an arbitrary level of sequence similarity (90%) (Votýpka *et al*. 2010). However, this threshold is insufficient for discriminating between closely related *Trypanosoma* species that share up to 98% similarity in their SL transcripts (Gibson *et al*. 2000).

A significant (and pragmatic) consideration when choosing a target gene for barcoding is the availability of sequences. Protists are poorly represented in sequence libraries and comprise just over 2% of the sequences currently in GenBank (National Center for Biotechnology Information (NCBI), 2017), despite constituting the majority of samples in environmental surveys (Del Campo *et al*. 2015). In addition, the sequence availability of *Trypanosoma* species is further skewed towards human-infective species and those infecting important agricultural species, such as cattle, which are over-represented in sequence databases relative to other trypanosomes, including parasites of insects and plants (D'Avila-Levy *et al*. 2015).

Whilst a bias towards medically important parasites is understandable, the paucity of genomic data from other *Trypanosoma* is a continuing impediment to our understanding of the evolutionary history and intricate phylogenetic relationships within this diverse group of parasites.

Gene or genes?

As the number of genes scrutinized for their barcoding potential has increased, it has become apparent that no test amplifying a single fragment has the differential power necessary to fully and reliably resolve the phylogeny of all trypanosomes (Hamilton *et al*. 2007; Adams and Hamilton, 2008; Pompanon and Samadi, 2015). Barcoding methods that utilize multiple loci have the advantage of additional power and accuracy (Mallo and Posada, 2016), and nested strategies that utilize 'a universal pre-barcode' and a 'group specific' barcode have been proposed by the Protist Working Group (ProWG) as alternative methods to resolve interspecies relationships (Pawlowski *et al*. 2012). In addition, we anticipate that the increasing ease and ever reducing costs of genome-wide SNP discovery in non-model organisms will lead to major advances in the use of SNP chip-based diagnostics in the near future.

Optimizing fragment length

In the past, target fragment length has been limited by the technology available. When molecular methods were first introduced, sequencing was only possible up to a few hundred base pairs. However, with the growth of Next-Generation Sequencing, the cost of sequencing has decreased by a factor of 10^4 in the last 10 years (Hayden, 2014; Van Nimwegen *et al*. 2016).

But is bigger always better?.   Should we strive for barcode fragments with a length at the ever-increasing limit of our sequencing ability? Here, there is a

significant trade-off to consider; optimal sequence length of the target region is highly dependant on the user's requirements. Shorter fragments result in higher sensitivity tests, favourable for analysis of degraded DNA from field samples. In diagnostic or clinical situations, for example, where the objective is to discriminate between the two human-infective subspecies of *T. brucei,* a shorter fragment is likely to provide all the required information. However, it is only with longer fragments that we can infer robust phylogenetic information at the subspecies level (Pompanon and Samadi, 2015); recreating the evolutionary history of a collection of poorly known or newly discovered species is likely to call for a very long target region, though this is, of course, a very different task than routine, high-throughput barcoding of large numbers of specimens.

An alternative future for diagnostics? Isothermal techniques

The use of isothermal amplification molecular methods, such as loop-mediated isothermal amplification and nucleic acid sequence-based amplification are becoming increasingly popular for the detection of trypanosomes as they offer simple, rapid and cheap alternatives to traditional PCR-based methods (Mugasa *et al*. 2014; Besuschio *et al*. 2017; Rivero *et al*. 2017). Isothermal tests involve a single reaction in a single tube incubated at a constant temperature; therefore, these techniques do not require the expensive thermocycling equipment that is

necessary for PCR (Matovu *et al*. 2010; Wastling and Welburn, 2011). The simplicity, sensitivity and low cost of isothermal techniques make them strong candidates for the application of molecular methods in field diagnostics in resource-poor areas (Laohasinnarong, 2011; Ricciardi and Ndao, 2015). However, a number of additional costs must be considered when evaluating the suitability of these methods for field diagnostics, including: the need for six primers, heating and maintaining samples at 65 °C, and expensive dyes for visualization of results (Enyaru *et al*. 2010; Wastling and Welburn, 2011). In addition, the ability of these tests to amplify extremely small amounts of DNA mean that they are highly prone to contamination. Developing simplified 'kit' forms of these techniques, and refining those already available, may yield promising alternatives to sequence-based barcoding for clinical purposes (Mugasa *et al*. 2014).

DISCUSSION

Towards a spectrum of similarity: an alternative approach to barcoding

Rather than identifying species by the length of their amplified fragments, we propose the adoption of a technique that identifies species, within a defined group, by the level of concordance across a selected gene, e.g. 18S rRNA, or partial gene (Fig. 3). Sequence differences between a cohort of species are tracked along a specified gene, highlighting
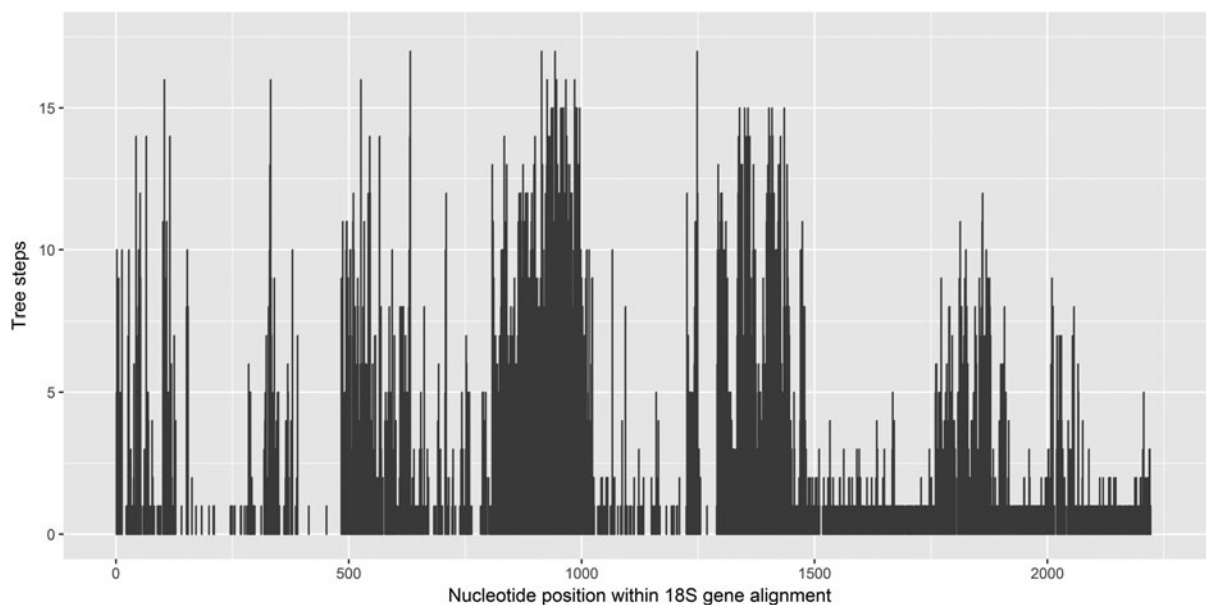
Fig. 3. Plot of phylogenetically informative nucleotide changes (based on the sequence alignment file and phylogeny presented by Hamilton *et al*. 2007) along the length of the 18S rRNA gene. Phylogenetic analysis –bootstrapped maximum parsimony analysis of 129 18S ssu rRNA sequences– was performed using the program PAUP* Ver 4·0a152 (Swofford, 2002). The default options of PAUP* were used: initial upper bound computed stepwise; only minimal trees kept; addition sequence = furthest; zero length branches collapsed. For further details of methodology, see Stevens and Wall (2001).

regions rich with phylogenetic variety. Species can then be identified by the degree of similarity across the selected region(s), for example, see Stevens and Wall (2001). The resulting *spectrum of similarity* can provide a valuable tool for understanding the relative level of sequence differentiation of any putative species, as their place in the spectrum will provide clues as to their phylogenetic placement.

Such an approach offers several benefits, including (as with any barcoding approach) the adoption of a standardized marker (or set of markers) and the ability to compare findings across studies, together with the practical benefits of being able to utilize a limited number of standardized primers. In the 'sliding window' approach proposed by Stevens and Wall (2001), the use of a given molecular marker in conjunction with a particular group of taxa allows the gene region (to be adopted for subsequent barcoding) to be selected based on the degree of phylogenetic resolution delivered by the particular sequence positions used within the target gene. More recently, Hadziavdic *et al.* (2014) undertook a much broader study along similar lines, screening for variation across more than 500 000 eukaryote 18S rRNA sequences (see also Pawlowski *et al.* (2012) for a review of the potential role of the V4 region of 18S rRNA as a candidate universal barcoding marker). Such approaches go a long way towards fulfilling the requirements for marker selection as set out in Box 1. To date, however, while several studies have focused on the use of the V7–V8 sub-region of 18S rRNA (e.g. Smith *et al.* 2008; Averis *et al.* 2009), citing its phylogenetic informativeness (but, see Hamilton and Stevens, 2011), this approach remains to be systematically applied across the full 18S rRNA gene in trypanosomes.

Can barcoding be all things to all people?

The ideal barcode from a gene region that yields enough sequence variation to capture the vast diversity of trypanosomes may provide a level of discrimination sufficient for diagnostic and identification purposes. However, it is questionable whether the same barcode could also provide enough variation to fully capture the phylogenetic relationships or complex evolutionary history of such a diverse group of organisms. In cases where genetic functionality is the key interest, barcoding is likely to be of little use. In the field, adequate preservation methods would be required to maintain the integrity of DNA from samples in order to apply any barcoding method successfully (Reeves *et al.* 2016).

The development of a perfect and truly universal barcode, based on a single primer pair, may be not only unattainable but also impractical. Different avenues of research have different requirements, in terms of both the techniques they use and the information required/acquired. A geneticist studying the evolution of trypanosomes needs a way to detect

intricate relationships over a range of evolutionary timescales (from, for example, the (putatively) most ancient to most recent: Simpson *et al.* 2004; Flegontov *et al.* 2013; Hamilton *et al.* 2004; Stevens & Rambaut, 2001; Haag *et al.* 1998; Lima *et al.* 2015; Balmer *et al.* 2011; Messenger *et al.* 2012), and it may be that a suite of gene markers is required to provide sufficient detail at all levels of phylogenetic depth. Conversely, for a clinician diagnosing patients in a resource-poor community, the nuances of an organism's evolutionary history are all but irrelevant. Identification of the parasite often determines treatment, so in this case the sensitivity and specificity of a diagnostic test becomes the overriding priority.

The range of requirements for the detection and identification of trypanosomes must be considered when selecting gene targets for barcoding, and the benefits of each molecular marker weighed against its limitations. For example, SL RNA is an ideal marker for detection of parasites in field samples, as this region is not present in either insect or vertebrate hosts (Westenberger *et al.* 2004). However, 18S rRNA may be preferable for field samples with potentially poor quality template DNA, as this region is relatively well protected against degradation (Basiye *et al.* 2011). Moreover, if a sample is for clinical diagnosis, diagnostic sensitivity is likely to be a priority – especially if parasitaemia is low. Therefore, a target marker would ideally be one with a high copy number (Hernández and Ramírez, 2013).

To date, there has been limited investigation into the comparative efficacy of different target regions for barcoding in trypanosomes. The barcoding technique presented in Fig. 3 can be applied to existing barcoding markers, as well as identifying the most phylogenetically informative regions, guiding the development of new primer targets. Rather than striving for a single, universal trypanosome barcode it may be advisable to adopt a multi-locus barcoding approach, similar to that suggested by Pawlowski *et al.* (2012) that can be adapted depending on the user's particular circumstances and requirements.

Concluding remarks and future directions

At present, molecular methods are mostly used only in sophisticated research laboratories, and there is a concern that new techniques are 'merely another addition to an ever-expanding toolbox of molecular assays for research' (Wastling and Welburn, 2011), rather than having any clinical diagnostic utility. And, whilst there has been a drive to develop and refine new molecular diagnostics, the sensitivity of existing techniques may be greatly improved if more research was conducted on initial stages, such as sample preparation and DNA extraction (Dunlop

et al. 2014). However, recent developments in molecular methods for trypanosome identification have succeeded in unveiling a number of previously unidentified species (Adams *et al.* 2010; Hutchinson and Gibson, 2015) and may offer new opportunities for the identification of novel hybrids (Koffi *et al.* 2015; Tihon *et al.* 2017) and the epidemiological tracking of trypanosome strains spread by the movement of host cattle (Févre *et al.* 2005). Nonetheless, a lack of comparable data between parasite surveys makes it difficult to draw any firm conclusions regarding species prevalence, and the full extent of trypanosome diversity remains unknown at this time (Adams *et al.* 2010; D'Avila-Levy *et al.* 2015). Priority should be given to the establishment of a standardized barcoding protocol for the detection and identification of trypanosomes (matching as close as possible the criteria given in Box 1). A standard barcoding protocol with requirement-dependant refinements is likely to be the closest we can ever come to obtaining a truly universal barcode for trypanosomes.

DATA STATEMENT

The research materials supporting this publication can be publicly accessed at the *Parasitology* journal website as Supplementary Information and/or by contacting the corresponding author.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at https://doi.org/10.1017/S0031182017002049.

AUTHOR ORCIDS

R. Hutchinson: orcid ID 0000-0003-0336-9594; J. R. Stevens: orcid ID: 0000-0002-1317-6721.

FINANCIAL SUPPORT

REFERENCES

Abderrazak, S., Guerrini, F., Mathieu-Daudé, F., Truc, P., Neubauer, K., Lewicka, K., Barnabé, C. and Tibayrenc, M. (1993). Isoenzyme electrophoresis for parasite characterisation. In Hyde, J. (ed.). Methods in Molecular Biology, Vol 21. Protocols in Molecular Parasitology. Totowa, New Jersey: Humana Press Inc., pp. 361–382.

Acup, C., Bardosh, K. L., Picozzi, K., Waiswa, C. and Welburn, S. C. (2016). Factors influencing passive surveillance for *T. b. rhodesiense* human African trypanosomiasis in Uganda. *Acta Tropica* **165**, 230–239.

Adams, E. R. and Hamilton, P. B. (2008). New molecular tools for the identification of trypanosome species. *Future Microbiology* **3**, 167–176.

Adams, E. R., Hamilton, P. B., Malele, I. I. and Gibson, W. C. (2008). The identification, diversity and prevalence of trypanosomes in field caught tsetse in Tanzania using ITS-1 primers and fluorescent fragment length barcoding. *Infection, Genetics and Evolution* **8**, 439–444.

Adams, E. R., Hamilton, P. B. and Gibson, W. C. (2010). African trypanosomes: celebrating diversity. *Trends in Parasitology* **26**, 324–328.

Adjé, C. A., Opperdoes, F. R. and Michels, P. A. M. (1998). Molecular analysis of phosphoglycerate kinase in *Trypanosoma borrelia* and the evolution of this enzyme in Kinetoplastida. *Gene* **217**, 91–99.

Adl, S. M., Simpson, A. G. B., Lane, C. E., Lukes, J., Bass, D., Bowser, S. S., Brown, M. W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., La Gall, L., Lynn, D. H., McManus, H., Mitchell, E. A. D., Mozley-Stanridge, S. E., Parfrey, L. W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C. L. and Smirnov, F. A. (2012). The revised classification of eukaryotes. *Journal of Eukaryotic Microbiology* **59**, 429–493.

Auty, H., Anderson, N. E., Picozzi, K., Lembo, T., Mubanga, J., Hoare, R., Fyumagwa, R. D., Mable, B., Hamill, L., Cleaveland, S. and Welburn, S. C. (2012a). Trypanosome diversity in wildlife species from the Serengeti and Luangwa Valley ecosystems. *PLoS Neglected Tropical Diseases* **6**, e1828.

Auty, H. K., Picozzi, K., Malele, I., Torr, S. J., Cleaveland, S. and Welburn, S. (2012b). Using molecular data for epidemiological inference: assessing the prevalence of *Trypanosoma brucei rhodesiense* in tsetse in Serengeti, Tanzania. *PLoS Neglected Tropical Diseases* **6**, e1501.

Averis, S., Thompson, R. C. A., Lymbery, A. J., Wayne, A. F., Morris, K. D. and Smith, A. (2009). The diversity, distribution and host-parasite associations of trypanosomes in Western Australian wildlife. *Parasitology* **136**, 1269–1279.

Balmer, O., Beadell, J. S., Gibson, W. and Caccone, A. (2011). Phylogeography and taxonomy of *Trypanosoma brucei*. *PLoS Neglected Tropical Diseases* **5**, e961.

Barbosa, A. D., Mackie, J. T., Stenner, R., Gillett, A., Irwin, P. and Ryan, U. (2016). *Trypanosoma teixeirae*: a new species belonging to the *T. cruzi* clade causing trypanosomiasis in an Australian little red flying fox (*Pteropus scapulatus*). *Veterinary Parasitology* **223**, 214–221.

Barrett, M. P. and Croft, S. L. (2012). Management of trypanosomiasis and leishmaniasis. *British Medical Bulletin* **104**, 175–196.

Basiye, F. L., Schoone, G. J., Beld, M., Minnaar, R., Ngeranwa, J. N., Wasunna, M. K. and Schallig, H. D. F. H. (2011). Comparison of short-term and long-term protocols for stabilization and preservation of RNA and DNA of *Leishmania*, *Trypanosoma*, and *Plasmodium*. *Diagnostic Microbiology and Infectious Disease* **69**, 66–73.

Besuschio, S. A., Murcia, M. L., Benatar, A. F., Monnerat, S., Cruz, I., Picado, A., Curto, M. D. L. A., Kubota, Y., Wehrendt, D. P., Pavia, P., Mori, Y., Puerta, C., Ndung'u, J. M. and Schijman, A. G. (2017). Analytical sensitivity and specificity of a loop-mediated isothermal amplification kit prototype for detection of *Trypanosoma cruzi* DNA in human blood samples. *PLoS Neglected Tropical Diseases* **11**, e0005779.

Chappuis, F., Loutan, L., Simarro, P., Lejon, V. and Büscher, P. (2005). Options for field diagnosis of human African trypanosomiasis. *Clinical Microbiology Reviews* **18**, 133–146.

Clarivate Statistics (2017). Web of Science [v.5.24] – Web of Science Core Collection Home. Available at http://webofknowledge.com (Accessed September 2017).

Coissac, E., Hollingsworth, P. M., Lavergne, S. and Taberlet, P. (2016). From barcodes to genomes: extending the concept of DNA barcoding. *Molecular Ecology* **25**, 1423–1428.

Dario, M. A., Moratelli, R., Schwabi, P., Jansen, A. M. and Llewellyn, M. S. (2017). Small subunit ribosomal metabarcoding reveals extraordinary trypanosomatid diversity in Brazilian bats. *PLoS Neglected Tropical Diseases* **11**, e0005790.

D'Avila-Levy, C. M., Boucinha, C., Kostygov, A., Santos, H. L. C., Morelli, K. A., Grybchuk-Ieremenko, A., Duval, L., Votýpka, J., Yurchenko, V., Grellier, P. and Lukeš, J. (2015). Exploring the environmental diversity of kinetoplastid flagellates in the high-throughput DNA sequencing era. *Memórias Do Instituto Oswaldo Cruz* **110**, 956–965.

Del Campo, J., Sieracki, M. E., Molestina, R., Keeling, P., Massana, R. and Ruiz-Trillo, I. (2015). The others: our biased perspective of eukaryotic genomes. *Trends in Evolution & Ecology* **29**, 252–259.

De Oliveira Ramos Pereira, L. and Brandão, A. (2013). An analysis of trypanosomatids kDNA minicircle by absolute dinucleotide frequency. *Parasitology International* **62**, 397–403.

De Waal, T. (2012). Advances in diagnosis of protozoan diseases. *Veterinary Parasitology* **189**, 65–74.

Deborggraeve, S. and Büscher, P. (2010). Molecular diagnostics for sleeping sickness: what is the benefit for the patient? *The Lancet Infectious Diseases* **10**, 433–439.

Desquesnes, M. and Davila, A. (2002). Applications of PCR-based tools for detection and identification of animal trypanosomes: a review and perspectives. *Veterinary Parasitology* **109**, 213–231.

Desquesnes, M., Kamyingkird, K., Yangtara, S., Milocco, C., Ravel, S., Wang, M. H., Lun, Z. R., Morand, S. and Jittapalapong, S. (2011). Specific primers for PCR amplification of the ITS1 (ribosomal DNA) of *Trypanosoma lewisi*. *Infection, Genetics and Evolution* **11**, 1361–1367.

Dunlop, J., Thompson, C. K., Godfrey, S. S. and Thompson, R. C. A. (2014). Sensitivity testing of trypanosome detection by PCR from whole blood samples using manual and automated DNA extraction methods. *Experimental Parasitology* **146**, 20–24.

Enyaru, J. C., Ouma, J. O., Malele, I. I., Matovu, E. and Masiga, D. K. (2010). Landmarks in the evolution of technologies for identifying trypanosomes in tsetse flies. *Trends in Parasitology* **26**, 388–394.

Fernandes, A. P., Nelson, K. and Beverley, S. M. (1993). Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. *Proceedings of the National Academy of Sciences of the USA* **90**, 11608–11612.

Ferri, G., Alù, M., Corradini, B., Licata, M. and Beduschi, G. (2009). Species identification through DNA 'barcodes'. *Genetic Testing and Molecular Biomarkers* **13**, 1–6.

Févre, E. M., Picozzi, K., Fyfe, J., Waiswa, C., Odiit, M., Coleman, P. G. and Welburn, S. C. (2005). A burgeoning epidemic of sleeping sickness in Uganda. *Lancet* **366**, 745–747.

Field, M. C., Horn, D., Fairlamb, A. H., Ferguson, M. A. J., Gray, D. W., Read, K. D., De Rycker, M., Torrie, L. S., Wyatt, P. G., Wyllie, S. and Gilbert, I. H. (2017). Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nature Reviews Microbiology* **15**, 217–231.

Flegontov, P., Votýpka, J., Skalický, T., Logacheva, M. D., Penin, A. A., Tanifuji, G., Onodera, N. T., Kondrashov, A. S., Volf, P., Archibald, J. M. and Lukeš, J. (2013). *Paratrypanosoma* is a novel early-branching *Trypanosomatid*. *Current Biology* **23**, 1787–1793.

Franco, J. R., Simarro, P. P., Diarra, A., Ruiz-Postigo, J. A. and Jannin, J. G. (2014). The journey towards the elimination of gambiense human African trypanosomiasis: not far, nor easy. *Parasitology* **141**, 748–760.

Gashumba, J. K., Gibson, W. C. and Opiyo, E. A. (1986). A preliminary comparison of *Trypanosoma simiae* and *T. congolense* by isoenzyme electrophoresis. *Acta Tropica* **43**, 15–19.

Gibson, W. (2007). Resolution of the species problem in African trypanosomes. *International Journal for Parasitology* **37**, 829–838.

Gibson, W. (2009). Species-specific probes for the identification of the African tsetse-transmitted trypanosomes. *Parasitology* **136**, 1501–1507.

Gibson, W. and Borst, P. (1986). Size-fractionation of the small chromosomes of *Trypanozoon* and *Nannomonas* trypanosomes by pulsed field gradient gel electrophoresis. *Molecular and Biochemical Parasitology* **18**, 127–140.

Gibson, W. C., Dukes, P. and Gashumba, J. K. (1988). Species-specific DNA probes for the identification of African trypanosomes in tsetse flies. *Parasitology* **97**, 63–73.

Gibson, W., Bingle, L., Blendeman, W., Brown, J., Wood, J. and Stevens, J. (2000). Structure and sequence variation of the trypanosome spliced leader transcript. *Molecular and Biochemical Parasitology* **107**, 269–277.

Godfrey, D. G. and Kilgour, V. (1976). Enzyme electrophoresis in characterizing the causative organism of Gambian trypanosomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **70**, 219–224.

Godfrey, D. G., Baker, R. D., Rickman, L. R. and Mehlitz, D. (1990). The distribution, relationships and identification of enzymic variants within the subgenus *Trypanozoon*. *Advances in Parasitology* **29**, 1–74.

Haag, J., O'hUigin, C. and Overath, P. (1998). The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. *Molecular and Biochemical Parasitology* **91**, 37–49.

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M. and Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS ONE* **9**, eB7624.

Hamilton, P. B. and Stevens, J. R. (2011). Resolving relationships between Australian trypanosomes using DNA barcoding data. *Trends in Parasitology* **27**, 99.

Hamilton, P. B., Stevens, J. R., Gaunt, M. W., Gidley, J. and Gibson, W. C. (2004). Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *International Journal for Parasitology* **34**, 1393–1404.

Hamilton, P. B., Gibson, W. C. and Stevens, J. R. (2007). Patterns of co-evolution between trypanosomes and their hosts deduced from ribosomal RNA and protein-coding gene phylogenies. *Molecular Phylogenetics and Evolution* **44**, 15–25.

Hamilton, P. B., Adams, E. R., Malele, I. I. and Gibson, W. C. (2008). A novel, high-throughput technique for species identification reveals a new species of tsetse-transmitted trypanosome related to the *Trypanosoma brucei* subgenus, *Trypanozoon*. *Infection, Genetics and Evolution* **8**, 26–33.

Hamilton, P. B., Lewis, M. D., Cruickshank, C., Gaunt, M. W., Yeo, M., Llewellyn, M. S., Valente, S. A., Da Silva, F. M., Stevens, J. R., Miles, M. A. and Teixeira, M. M. G. (2011). Identification and lineage genotyping of South American trypanosomes using fluorescent fragment length barcoding. *Infection, Genetics and Evolution* **11**, 44–51.

Hashimoto, T., Nakamura, Y., Kamaishi, T., Adachi, J., Nakamura, F., Okamoto, K. and Hasegawa, M. (1995). Phylogenetic place of kinetoplastid protozoa inferred from a protein phylogeny of elongation factor 1α. *Molecular and Biochemical Parasitology* **70**, 181–185.

Hayden, E. (2014). The $1,000 genome. *Nature* **507**, 294–295.

Hebert, P. D. N., Cywinska, A., Ball, S. L. and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* **270**, 313–321.

Hernández, C. and Ramírez, J. D. (2013). Molecular diagnosis of vector-borne parasitic diseases. *Air & Water Borne Diseases* **2**, 1–10.

Hide, G., Cattand, P., Le Ray, D., Barry, J. D. and Tait, A. (1990). The identification of *T. brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology* **39**, 213–226.

Hoare, C. (1972). *The Trypanosomes of Mammals*. Oxford: Blackwell Scientific Publications.

Hutchinson, R. and Gibson, W. (2015). Rediscovery of *Trypanosoma* (*Pycnomonas*) *suis*, a tsetse-transmitted trypanosome closely related to *T. brucei*. *Infection, Genetics and Evolution* **36**, 381–388.

Jensen, R. and Englund, P. (2012). Network news: the replication of kinetoplast DNA. *Annual Review of Microbiology* **66**, 473–491.

Kilgour, V. and Godfrey, D. G. (1973). Isoenzymes of alanine aminotransferase as possible specific characters of trypanosomes. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **67**, 11.

Koffi, M., De Meeûs, T., Séré, M., Bucheton, B., Simo, G., Njiokou, F., Salim, B., Kaboré, J., MacLeod, A., Camara, M., Solano, P., Belem, A. M. G. and Jamonneau, V. (2015). Population genetics and reproductive strategies of African trypanosomes: revisiting available published data. *PLoS Neglected Tropical Diseases* **9**, e0003985.

Kukla, B. A., Majiwa, P. A. O., Young, J. R., Moloo, S. K. and Ole-Moiyoi, O. K. (1987). Use of species-specific DNA probes for detection and identification of trypanosome infection in tsetse flies. *Parasitology* **95**, 1–16.

Lanham, S. M. and Godfrey, D. G. (1970). Isolation of salivarian trypanosomes from man and other mammals using DEAE-cellulose. *Experimental Parasitology* **28**, 521–534.

Laohasinnarong, D. (2011). Loop-mediated isothermal amplification (LAMP): an alternative molecular diagnosis. *Journal of Applied Animal Science* **4**, 9–19.

Leonard, G., Soanes, D. M. and Stevens, J. R. (2011). Resolving the question of trypanosome monophyly: a comparative genomics approach using whole genome data sets with low taxon sampling. *Infection, Genetics and Evolution* **11**, 955–959.

Lima, L., Espinosa-Álvarez, O., Pinto, C. M., Cavazzana, M., Jr, Pavan, A. C., Carranza, J. C., Lim, B. K., Campaner, M., Takata, C. S. A., Camargo, E. P., Hamilton, P. B. and Teixeira, M. M. G. (2015). New insights into the evolution of the *Trypanosoma cruzi* clade provided by a new trypanosome species tightly linked to Neotropical *Pteronotus* bats and related to an Australian lineage of trypanosomes. *Parasites & Vectors* **8**, 657.

Lukes, J., Jirku, M., Dolezel, D., Kral'ova, I., Hollar, L. and Maslov, D. A. (1997). Analysis of the ribosomal RNA gene suggests that trypanosomes are monophyletic. *Journal of Molecular Evolution* **44**, 521–527.

Lutumba, P., Matovu, E. and Boelaert, M. (2016). Human African Trypanosomiasis (HAT). In Gyapong, J and Boatin, B (eds). Neglected Tropical Diseases - Sub-Saharan Africa. Switzerland: Springer International Publishing, pp. 63–85.

Mallo, D. and Posada, D. (2016). Multilocus inference of species trees and DNA barcoding. *Philosophical Transactions of the Royal Society of London B* **371**, 20150335.

Maslov, D. A., Lukes, J., Jirku, M. and Simpson, L. (1996). Phylogeny of trypanosomes as inferred from the small and large subunit rRNAs: implications for the evolution of parasitism in the trypanosomatid protozoa. *Molecular and Biochemical Parasitology* **75**, 197–205.

Matovu, E., Mugasa, C. M., Ekangu, R. A., Deborggraeve, S., Lubega, G. W., Laurent, T., Schoone, G. J., Schallig, H. D. and Büscher, P. (2010). Phase II evaluation of sensitivity and specificity of PCR and NASBA followed by oligochromatography for diagnosis of human African trypanosomiasis in clinical samples from D.R. Congo and Uganda. *PLoS Neglected Tropical Diseases* **4**, 3–8.

McNamara, J., Dukes, P., Snow, W. F. and Gibson, W. C. (1989). Use of DNA probes to identify *Trypanosoma congolense* and *T. simiae* in tsetse flies from The Gambia. *Acta Tropica* **46**, 55–61.

Messenger, L. A., Llewellyn, M. S., Bhattacharyya, T., Franzén, O., Lewis, M. D., Ramírez, J. D., Carrasco, H. J., Andersson, B. and Miles, M. A. (2012). Multiple mitochondrial introgression events and heteroplasmy in *Trypanosoma cruzi* revealed by maxicircle MLST and next generation sequencing. *PLoS Neglected Tropical Diseases* 6, e1584.

Miles, M. A., Toye, P. J., Oswald, S. C. and Godfrey, D. G. (1977). The identification by isoenzyme patterns of two distinct strain-groups of *Trypanosoma cruzi* circulating independently in a rural area of Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 71, 217–225.

Mpanya, A., Hendrickx, D., Vuna, M., Kanyinda, A., Lumbala, C., Tshilombo, V., Mitashi, P., Luboya, O., Kande, V., Boelaert, M., Lefèvre, P. and Lutumba, P. (2012). Should I get screened for sleeping sickness? A qualitative study in Kasai Province, Democratic Republic of Congo. *PLoS Neglected Tropical Diseases* 6, e1467.

Mugasa, C. M., Adams, E. R., Boer, K. R., Dyserinck, H. C., Büscher, P., Schallig, H. D. H. F. and Leeflang, M. M. G. (2012). Diagnostic accuracy of molecular amplification tests for human African trypanosomiasis-systematic review. *PLoS Neglected Tropical Diseases* 6, e1438.

Mugasa, C. M., Katiti, D., Boobo, A., Lubega, G. W., Schallig, H. D. F. H. and Matovu, E. (2014). Comparison of nucleic acid sequence-based amplification and loop-mediated isothermal amplification for diagnosis of human African trypanosomiasis. *Diagnostic Microbiology and Infectious Disease* 78, 144–148.

National Center for Biotechnology Information (NCBI) (2017). Bethesda, MD: National Library of Medicine (US), National Center for Biotechnology Information website. Available at http://www.ncbi.nlm.nih.gov (Accessed April 2017).

Ouma, J. O., Masake, R. A., Masiga, D. K., Moloo, S. K., Njuguna, J. T. and Ndung'u, J. M. (2000). Comparative sensitivity of dot-ELISA, PCR and dissection method for the detection of trypanosome infections in tsetse flies (Diptera: *Glossinidae*). *Acta Tropica* 75, 315–321.

Papadopoulos, M. C., Abel, P. M., Agranoff, D., Stich, A., Tarelli, E., Bell, B. A., Planche, T., Loosemore, A., Saadoun, S., Wilkins, P. and Krishna, S. (2004). A novel and accurate diagnostic test for human African trypanosomiasis. *Lancet* 363, 1358–1363.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S. S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A. M., Gile, G. H., Holzmann, M., Jahn, R., Jirků, M., Keeling, P. J., Kostka, M., Kudryavtsev, A., Lara, E., Lukeš, J., Mann, D. G., Mitchell, E. A. D., Nitsche, F., Romeralo, M., Saunders, G. W., Simpson, A. G. B., Smirnov, A. V., Spouge, J. L., Stern, R. F., Stoeck, T., Zimmermann, J., Schindel, D. and De Vargas, C. (2012). CBOL protist Working Group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biology* 10, e1001419.

Pečnikar, Z. F. and Buzan, E. V. (2014). 20 years since the introduction of DNA barcoding: from theory to application. *Journal of Applied Genetics* 55, 43–52.

Pompanon, F. and Samadi, S. (2015). Next generation sequencing for characterizing biodiversity: promises and challenges. *Genetica* 143, 133–138.

Ratnasingham, S. and Hebert, P. D. N. (2007). BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes* 7, 355–364.

Reeves, L. E., Holderman, C. J., Gillett-Kaufman, J. L., Kawahara, A. Y. and Kaufman, P. E. (2016). Maintenance of host DNA integrity in field-preserved mosquito (Diptera: Culicidae) blood meals for identification by DNA barcoding. *Parasites & Vectors* 9, 503.

Ricciardi, A. and Ndao, M. (2015). Diagnosis of parasitic infections: what's going on? *Journal of Biomolecular Screening* 20, 6–21.

Rivero, R., Bisio, M., Velázquez, E. B., Esteva, M. I., Scollo, K., González, N. L., Altcheh, J. and Ruiz, A. M. (2017). Rapid detection of *Trypanosoma cruzi* by colorimetric loop-mediated isothermal amplification (LAMP): A potential novel tool for the detection of congenital Chagas infection. *Diagnostic Microbiology and Infectious Disease* 89, 26–28.

Rodrigues, A. C., Garcia, H. A., Batista, J. S., Minervino, A. H. H., Góes-Cavalcante, G., Da Silva, F. M., Ferreira, R. C., Campaner, M., Paiva, F. and Teixeira, M. M. G. (2010). Characterisation of spliced leader genes of *Trypanosoma (Megatrypanum) theileri*: phylogeographical analysis of Brazilian isolates from cattle supports spatial clustering of genotypes and parity with ribosomal markers. *Parasitology* 137, 111–122.

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K. and Lane, R. (2005). Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B* 360, 1805–1811.

Silva-Iturriza, A., Nassar, J. M., Garcia-Rawlins, A. M., Rosales, R. and Mijares, A. (2013). *Trypanosoma evansi* kDNA minicircle found in the Venezuelan nectar-feeding bat *Leptonycteris curasoae* (Glossophaginae), supports the hypothesis of multiple origins of that parasite in South America. *Parasitology International* 62, 95–99.

Simpson, A. G. B., Gill, E. E., Callahan, H. A., Litaker, R. W. and Roger, A. J. (2004). Early evolution within kinetoplastids (Euglenozoa), and the late emergence of trypanosomatids. *Protist* 155, 407–422.

Smith, A., Clark, P., Averis, S., Lymbery, A. J., Wayne, A. F., Morris, K. D. and Thompson, R. C. A. (2008). Trypanosomes in a declining species of threatened Australian marsupial, the brush-tailed bettong *Bettongia penicillata* (Marsupialia: Potoroidae). *Parasitology* 135, 1329–1335.

Stevens, J. and Rambaut, A. (2001). Evolutionary rate differences in trypanosomes. *Infection, Genetics and Evolution* 1, 143–150.

Stevens, J. and Wall, R. (2001). Genetic relationships between blowflies (Calliphoridae) of forensic importance. *Forensic Science International* 120, 116–123.

Stevens, J., Noyes, H. and Gibson, W. (1998). The evolution of trypanosomes infecting humans and primates. *Memórias Do Instituto Oswaldo Cruz* 93, 669–676.

Stevens, J. R. and Godfrey, D. G. (1992). Numerical taxonomy of *Trypanozoon* based on polymorphisms in a reduced range of enzymes. *Parasitology* 104, 75–86.

Stevens, J. R., Noyes, H. A., Dover, G. A. and Gibson, W. C. (1999). The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *T. cruzi*. *Parasitology* 118, 107–116.

Stockdale, L. and Newton, R. (2013). A review of preventative methods against human leishmaniasis infection. *PLOS Neglected Tropical Disease* 7, e2278.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S. and Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. *The Journal of Clinical Investigation* 118, 1301–1310.

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4.0. Sunderland, Massachusetts: Sinauer Associates.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. and Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* 21, 2045–2050.

Tihon, E., Imamura, H., Dujardin, J. C., Van Den Abbeele, J. and Van den Broeck, F. (2017). Discovery and genomic analysis of hybridization between divergent lineages of *Trypanosoma congolense*, causative agent of animal African trypanosomiasis. *Molecular Ecology*, 1–15, doi: 10.1111/mec.14271.

Uilenberg, G. and Boyt, W. P. (1998). *A Field Guide for the Diagnosis, Treatment and Prevention of African Animal Trypanosomosis*. Rome, Italy: Food and Agriculture Organisation of the United Nations.

Valentini, A., Pompanon, F. and Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution* 24, 110–117.

Van Nimwegen, K. J. M., Van Soest, R. A., Veltman, J. A., Nelen, M. R., Van Der Wilt, G. J., Vissers, L. E. L. M. and Grutters, J. P. C. (2016). Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. *Clinical Chemistry* 62, 1458–1464.

Votýpka, J., Maslov, D. A., Yurchenko, V., Jirků, M., Kment, P., Lun, Z. and Lukeš, J. (2010). Probing into the diversity of trypanosomatid flagellates parasitizing insect hosts in South-West China reveals both endemism and global dispersal. *Molecular Phylogenetics and Evolution* 54, 243–253.

Wastling, S. L. and Welburn, S. C. (2011). Diagnosis of human sleeping sickness: sense and sensitivity. *Trends in Parasitology* 27, 394–402.

Westenberger, S. J., Sturm, N. R., Yanega, D., Podlipaev, S. A., Zeledon, R., Campbell, D. A. and Maslov, D. A. (2004). Trypanosomatid biodiversity in Costa Rica: genotyping of parasites from Heteroptera using the spliced leader RNA gene. *Parasitology* 129, 537–547.

Woods, C. R. (2013). False-positive results for immunoglobulin M serologic results: explanations and examples. *Journal of the Pediatric Infectious Diseases Society* 2, 87–90.

World Health Organisation (2013). *Control and Surveillance of Human African Trypanosomiasis*. WHO Technical Report Series. No 984. Geneva, Switzerland: World Health Organisation.

Yaro, M., Munyard, K. A., Stear, M. J. and Groth, D. M. (2016). Combatting African animal trypanosomiasis (AAT) in livestock: the potential role of trypanotolerance. *Veterinary Parasitology* 225, 43–52.