Trends in Parasitology

CellPress
REVIEWS

Review

Next-Generation Molecular Surveillance of TriTryp Diseases

Malgorzata Anna Domagalska[1,*] and Jean-Claude Dujardin[1]

Elimination programs targeting TriTryp diseases (Leishmaniasis, Chagas' disease, human African trypanosomiasis) significantly reduced the number of cases. Continued surveillance is crucial to sustain this progress, but parasite molecular surveillance by genotyping is currently lacking. We explain here which epidemiological questions of public health and clinical relevance could be answered by means of molecular surveillance. Whole-genome sequencing (WGS) for molecular surveillance will be an important added value, where we advocate that preference should be given to direct sequencing of the parasite's genome in host tissues instead of analysis of cultivated isolates. The main challenges here, and recent technological advances, are discussed. We conclude with a series of recommendations for implementing whole-genome sequencing for molecular surveillance.

Highlights

Leishmaniasis, Chagas' disease, and human African trypanosomiasis (the TriTryp diseases) are the object of elimination programs. More than ever, 'postelimination' surveillance is required to ensure the long-term success of these programs. Currently, there is no formal molecular surveillance and tracking of TriTryp diseases.

WGS can be used to answer several questions of epidemiological interest in one sequencing run.

WGS analysis of TriTryps is currently done on cultivated isolates; this constitutes a major bias. In the case of *Leishmania*, it was clearly demonstrated that the parasite's genome can be different between clinical samples and derived isolates. New methods are available for direct sequencing of the parasite in host tissues.

Direct WGS in clinical samples should be the method of choice for molecular surveillance of TriTryp diseases and many others.

Advocating for the Integration of Molecular Surveillance in Control Programs of TriTryp Diseases

Despite the fact that trypanosomatids have long been neglected by health authorities, private companies, and media, these parasites are causing major vector-borne diseases in humans, animals, and plants [1]. In contrast, the parasites were never neglected by health professionals and scientists for which they represent challenging targets to control and a unique model for the biology of early branching eukaryotes. Recently, diseases caused by trypanosomatids started to emerge from the shadow: there was a renewed interest in drug companies – with public partnership – to invest in R&D, and health authorities launched a series of major **elimination** (see Glossary) programs. This was guided by the accumulated knowledge and know-how, but there is still a need and a potential for further exploitation of the acquired scientific knowledge and data, especially in this **postgenomic** era. Unique biological features of trypanosomatids, such as their genomic plasticity and adaptive capacity, should be taken into account during and after the elimination programs. In the present review we aim to bridge this gap and advocate for the integration of molecular surveillance in trypanosomatids' control programs, with a special focus on **TriTryp** parasites infecting humans. Firstly, we discuss why molecular surveillance should be implemented and which epidemiological questions of public health relevance it may answer, with, as a test case, three parasites characterized using **targeted genotyping** methods. In a second part, we show the added value and power of WGS for epidemiological monitoring, focusing on the high-throughput analysis of *Leishmania* clinical isolates (Box 1) as a model study. Thirdly, in light of recent reports [2], we show why direct sequencing of the parasites in host tissues should be preferred over the analysis of cultivated isolates, where we highlight a series of technological advances allowing this approach. We conclude by a series of recommendations for the implementation of WGS for molecular surveillance.

TriTryp Elimination Programs

In the peak of epidemics and/or endemicity, millions of (new) cases were reported[i] every year for Chagas' disease (CHD: *Trypanosoma cruzi*), leishmaniases (*Leishmania* sp.), and human African trypanosomiases (HAT: *Trypanosoma brucei gambiense/rhodesiense*). Major regional elimination

[1]Molecular Parasitology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerp, Belgium

*Correspondence: mdomagalska@itg.be (M.A. Domagalska).

Box 1. The Complex and Incomparable Taxonomy of TriTryps

Taxonomy can be considered as the law code for biologists. As highlighted in a provocative opinion paper which is still of actuality [65], the view on the classification of protozoan parasites may differ for evolutionary biologists, biomedical researchers, and decision makers. In the absence of a clear species concept, different approaches are followed, which is well illustrated in the case of TriTryps. *T. cruzi* is considered to be a single species while presenting a similar level of (phylo-)genetic diversity as the genus *Leishmania*, where the latter is divided into tens of species. In between, the subgenus *Trypanozoon* comprises subspecies such as *T. b. gambiense*, *T. b. rhodesiense*, and *T. b. brucei*, but these traditional taxonomic classifications are not fully supported by genetic data [66]. Even if the (sub-)taxonomic categories are not comparable among TriTryps, we refer here to those practically used in the literature. In *T. cruzi*, six discrete typing units (DTUs Tc1–6, concept specifically introduced for classification within *T. cruzi* and originally defined as a set of stocks genetically more similar to each other than to any other, and identifiable by common molecular markers [65]) were for long considered, but recent reports mention two additional ones, Tc1DOM [19] and Tcbat [67]. In *T. b. gambiense*, two groups are considered, Tbg-1 and Tbg-2, albeit the latter one clusters phylogenetically together with the animal-infecting *T. b. brucei* [66]. *T. b. rhodesiense* is paraphyletic with *T. b. brucei*, with some isolates clustering with Tbg-1 [66]. In *Leishmania*, there are two subgenera, *Leishmania* and *Viannia*, and so far 27 species (see last revision in [68]).

Besides these (sub-)taxonomic categories, a series of operational terms are used, but are often confounded, hence we felt it to be convenient to remind the reader of the following agreed definitions (adapted from [69]): (i) *population*, a group of trypanosomatids present at a given time in a given host; (ii) *sample*, part of a trypanosomatid population collected on a single occasion; (iii) *primary isolate*, the viable organisms present in a culture (or an experimental animal) following the introduction of a sample from a naturally infected host; (iv) *stock*, the population derived by serial passage *in vitro* (or *in vivo*) from a primary isolate, without any implication of homogeneity or characterization; (v) *line*, a laboratory derivative of a stock maintained in different physical conditions or different laboratories; (vi) *clone*, trypanosomatids derived from a single individual by binary fission; (vii) *strain*, a clone defined by the possession of one or more designated characteristics.

Glossary

Elimination: reduction to zero of the incidence of disease or infection in a defined geographical area, as a consequence of deliberate efforts; often confounded with (i) eradication, which stands for the permanent reduction to zero of the worldwide incidence of infection, and (ii) control, the reduction in the incidence, prevalence, morbidity, or mortality of an infectious disease to a locally acceptable level [63]. In contrast to eradication, elimination and control require continued interventions to avoid re-emergence of the disease[iv].

Postgenomic: the period following the characterization of reference genomes and corresponding to their exploitation, including high-throughput intraspecies genome diversity studies facilitated by the emergence and spreading of next-generation sequencing methods [64].

Targeted genotyping: molecular characterization of a predefined genetic locus with limited length, for instance the sequence of a given gene or intergenic region.

TriTryp: the term initially used in genomics referring to the three trypanosomatids for which reference genomes were first established, that is *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania major*; by extension, the term was used here to name the three principal diseases (TriTryp diseases) caused by these parasites in humans: human African trypanosomiasis, Chagas' disease, and leishmaniasis.

programs were launched in Latin America (CHD), the Indian subcontinent (ISC; anthroponotic visceral leishmaniasis, AVL) and Africa (HAT). These programs were essentially based on diagnosis, treatment, and/or vector control, and demonstrated a major impact, resulting in significant reduction in the number of cases: (i) from 700 000 new CHD cases per year in 1980–1985 to 38 593 in 2010 [3], (ii) from 77 000 AVL cases in the ISC in 1992 to less than 7000 in 2016 [4], and (iii) from 37 385 reported HAT cases in 1998 to less than 2200 in 2016 [5]. Despite major successes, these programs remain vulnerable to decrease and shift in political attention and priority and they are confronted with several biological and ecoepidemiological challenges such as changes in drug susceptibility and virulence, the occurrence of untreated asymptomatic individuals, the presence of animals as reservoirs (sometimes not suspected), the existence of superspreaders, the natural fluctuations of the disease, the rise of insecticide resistance hampering vector control, or the adaptation of nondomiciliated vectors [4–6]. Accordingly, the probability of new outbreaks is high, and this could jeopardize the performance of ongoing elimination programs. Next to this, in the postelimination phase, there is a need for improving the sensitivity and specificity of diagnostic tools, development of new drugs, and implementation of better vector-control measures; however, such activity must always be combined with continued surveillance to ensure the long-term success of the elimination programs [4]. Here, one of the most neglected aspects is systematic molecular surveillance by parasite genotyping, an approach well established for other infectious diseases such as tuberculosis [7], but not yet for TriTryps. A battery of genotyping methods is available to achieve this surveillance, with a discriminative power ranging from genus to species, subspecies, phenotypic variants, or strains (Box 1). The required resolution of the molecular methods obviously depends on the epidemiological question being addressed: for instance, a simple species-specific PCR is sufficient to identify a new focus of a given disease, while highly discriminatory tools like WGS are needed to track strains of the studied pathogen between different vertebrate and invertebrate hosts and to identify transmission chains.

Why Molecular Surveillance in the (Post-)Elimination Context?

Molecular surveillance by parasite genotyping is relevant for addressing at least six overlapping questions of major public health importance (subsequently called Q1 to Q6 in this review) and

this is illustrated for TriTryps further below. We deliberately exclude WGS from this section to highlight its applications better in the next section and concentrate here on targeted genotyping methods (Box 2).

Following the Evolution of Epidemics in Time and Space (Q1)

Multilocus microsatellite typing (MLMT) of *Leishmania* samples from new foci of visceral leishmaniasis (VL) in Northern Italy showed a different population of *L. infantum* in humans and dogs, suggesting a cycle not involving a canine reservoir, as is the case in other parts of the country [8]. New foci of HAT have been encountered in Central Uganda, and MLMT analysis suggested that expansion of HAT was associated with *T. b. rhodesiense* lineages spreading from old to new foci [9]. MLMT was also used to follow the population dynamics of *T. cruzi* in the Central Ecuadorian coast and revealed a gene flow between sylvatic and (peri-) domestic transmission cycles, highlighting the risk of reinvasion after control measures targeting domiciliary environments [10].

Characterizing (New) Transmission Cycles (Q2)

Using a PCR-RFLP method for *Leishmania* species identification, Torrellas *et al.* [11] showed *Lutzomyia migonei* to be the putative vector in two distinct epidemiological cycles of cutaneous leishmaniasis (CL) in the Andean region of Venezuela, involving *L. guyanensis* and *L. mexicana*, respectively. MLMT allowed the detection of *T. b. gambiense* group 1 (Box 1) in both clinical and aparasitemic/seropositive subjects, suggesting that the latter individuals might play a role as a reservoir in HAT foci [12]. A discrete typing unit (DTU)-specific (Box 1) multiplex qPCR was used to study the epidemiology of canine *T. cruzi* infections in kennels in Texas; the study identified two distinct DTUs, each associated with a different triatomine vector species, and it demonstrated that canine kennels constituted a high-risk environment for the transmission of the parasites in that region [13].

Box 2. 'Targeted' Genotyping Methods Used to Track TriTryps

This concerns only methods cited in the text; for a complete review see also [70].

MLMT, multilocus microsatellite typing. Microsatellites are short repeats of two or three nucleotides with high variability in their copy numbers. The genomic region flanking them is amplified and the length measured; because of the high risk of convergent evolution (homoplasy), multiple loci are analyzed. This method is used for evolutionary studies, population genetics, and strain tracking.

MLST, multilocus sequence typing. This method involves the same principle as PCR-sequencing, but it targets multiple variable regions. MLST is used for evolutionary studies, population genetics, and strain tracking.

PCR-RFLP, PCR-restriction fragment-length polymorphism. In this approach, a given genomic region is PCR-amplified and the product is cleaved with restriction enzymes; it is less discriminatory than PCR-sequencing, but it can be applied in the absence of sequencers. PCR-RFLP is generally used for identification at species and intraspecific levels.

PCR-sequencing. In this method, a given fragment of genome is PCR-amplified and sequenced; depending on the variability of the amplified region, the method can be used for the identification of species, subspecies or phenotypic variants. Targeting minicircles of kinetoplast DNA (kDNA) provides a discriminatory fingerprinting tool – given the high variability of the repeated minicircles – but the method is difficult to standardize.

Quantitative PCR (qPCR). This method is generally used to quantify a product; it provides the copy number of a given gene or – if combined with reverse transcription – the abundance of a given transcript. Combined with hybridization with specific probes, the method can also be used to identify point mutations in a given amplicon and, multiplexed for increased resolution, it can generally be used for identification at species and intraspecific levels.

Specific PCR. In this method, genetic discrimination is provided by the primers themselves, which are designed to anneal correctly only with the genetic variant to be detected; the method is often used for the identification of one given species (species-specific PCR) or intraspecific variant (allele-specific PCR).

Outbreak Studies and Source Identification (Q3)

An outbreak of leishmaniasis occurred in 2009 in a suburb of Madrid, and a *Leishmania* sp.-specific PCR demonstrated the presence of the parasite in rabbits; while dogs are the usual suspects, this PCR highlighted the role of rabbits as a potential reservoir of *L. infantum* and guided further control of the disease through ecological interventions in the reservoir's habitat [14]. PCR analysis of trypanosomes in two foci of HAT (classically considered as an anthroponosis) in Ivory Coast, identified pigs and cattle as potential reservoirs of *T. brucei* s.l., suggesting the need for a 'one health' approach to achieve a complete elimination of the disease [15]. DTU identification of *T. cruzi* allowed identification of guava juice as a common source of infection in a school-related oral CHD outbreak in Venezuela [16].

Detection of New Variants, Possibly with New Clinical Features (Q4)

In Sri Lanka, new variants of *L. donovani* were identified by PCR sequencing of kDNA minicircles. They were genetically clearly distinct from the main parasite population causing AVL in the ISC and were associated with an outbreak of CL on the island [17]. In HAT, a coevolutionary arms race has been described between trypanosomes and the host's lytic factor APOL1, with (i) parasite molecules SRA or *TgsGP* protecting *T. b. rhodesiense* and *T. b. gambiense* from lysis, (ii) human APOL1 variants protecting hosts against SRA-positive *T. b. rhodesiense*, and (iii) *T. b. rhodesiense* variants lacking SRA and group 2 *T. b. gambiense*, lacking *TgsGP,* being able to cope with APOL1 by unknown mechanisms [18]. In Colombia, nuclear multilocus sequence typing (MLST) allowed the identification of a new domestic *T. cruzi* genotype (TcIDOM) that has been associated with severe chronic cardiomyopathy in contrast to the disease caused by sylvatic DTU I lineages [19].

Detection of Sexual Recombination (Q5)

Besides clonal reproduction, TriTryps may undergo hybridization and sexual recombination. As a consequence, this phenomenon might bring together in the same organism different traits of clinical importance such as resistance to different drugs, thereby possibly jeopardizing combination therapy with the respective drugs. In Ethiopia, intraspecific hybrids were detected by MLMT and MLST in *L. donovani* [20]. In East Africa, it was reported that new strains of *T. b. rhodesiense* are constantly generated by recombination with animal-infective trypanosomes, carrying the risk of future outbreaks of HAT [21]. In *T. cruzi*, hybridization led to two DTUs that successfully colonized several regions of Latin America [22].

Genetic Markers of Clinically and Epidemiologically Relevant Traits (Q6)

This topic is also relevant in the context of virulence; however, we focus here only on the detection of drug-resistance markers. Resistance to antimonials in clinical isolates of *L. tropica* was found to be associated with changes in the expression of several genes [23]. In *T. b. gambiense*, mutations in Aquaporin 2 were found to correlate with a decreased susceptibility to pentamidine and melarsoprol [24]. In *T. cruzi*, MLMT could discriminate Mexican strains according to their susceptibility to benznidazole [25].

The examples described above illustrate well the role that molecular surveillance might play in the support of control activities of leishmaniasis, HAT, and CHD. Moreover, this approach could easily be extrapolated to other diseases caused by trypanosomatids. MLMT for long appeared to be the method of choice for fine tracking of the parasites (Box 2). However, these targeted genotyping methods present several limitations. First, even if they are suited for standardization, the corresponding molecular assays are far from being standardized, with most laboratories using their favorite target or home-made protocol [26]. Second, the discriminatory power of fingerprinting tools depends on the evolutionary history of the population under study: for instance, most ISC isolates of *L. donovani* (recent population reported to have emerged around

1850 [27]) were indistinguishable by MLST and MLMT [28]. Third, these methods are targeted genotyping systems, thereby covering only a small and predefined fraction of the genome, resulting in serious limitations in the detection of both quantitative and qualitative differences at the genomic level. This is best illustrated by the quest for drug-resistance markers: as explained in a previous review [29], different and/or unknown genes can have been modified in drug-resistant parasites, and for each of them, different alteration mechanisms can be operating: different SNPs in the coding sequences, local gene copy number variation, gene deletion, aneuploidy… a phenomenon that we called 'the many roads to drug resistance' [29]. Accordingly, targeted methods like an allele-specific PCR might miss these alternative mutations. If possible, untargeted approaches should be preferred, and currently the only method that allows the different limitations mentioned above to be overcome is WGS.

Why Should We Use WGS for Molecular Surveillance?

WGS is increasingly used for diagnosis and molecular surveillance of diseases such as tuberculosis [7] or foodborne diseases [30], around the six questions mentioned above and with impact. For instance, in tuberculosis, high-resolution information on emergence, spread, genetic makeup, and evolution of highly resistant or virulent clones allows the implementation of targeted measures [7]. The size of the TriTryp genome (reference haploid genomes ranging between 25 and 55 Mb [31]) represents a bigger sequencing challenge for high-throughput studies than in bacteria and viruses, but improved reference genomes are available and the multiplication of next-generation sequencing platforms has democratized genome sequencing beyond the big sequencing centers, allowing a strong reduction in costs. However, well-resourced centers with good computing facilities and bioinformatics are still needed [32]. Up to now, very few large-scale WGS studies have been undertaken in a TriTryp molecular epidemiology context; we will focus here on *Leishmania* as most data are available for this organism. Noteworthy, the use of 'whole' in WGS should be taken with care, as most published reports (i) considered only the nuclear genome, where 100% coverage is not achieved, and (ii) kept aside the kinetoplast genome (kDNA) because of the bioinformatic challenge of aligning highly repeated circular DNA molecules, like the maxicircles and minicircles. Recently, first attempts at integrating kDNA data were published [33–35], indicating that this neglected part of the trypanosomatid genome will become more accessible in the foreseeable future, mainly thanks to recently developed analytical pipelines[ii]. Further below, we focus only on studies of nuclear genomes.

So far, the most exhaustive published study concerned 204 isolates collected during two clinical studies in the ISC a decade ago [27], most of them originating from the historical focus of AVL epidemics in Bihar (India) and the neighboring Eastern Terai (Nepal). This pioneer study answered the six epidemiological questions mentioned above, thereby demonstrating the power of WGS for molecular surveillance, particularly in recent populations. First, it highlighted the 'ultimate' discriminatory power of WGS: as mentioned above, most *L. donovani* isolates in the ISC were indistinguishable by MLMT, while WGS could distinguish each isolate and revealed the so far hidden genetic structure, with a core group (CG) of nine subpopulations (called ISC groups, Q1). Second, Bayesian phylogenetic models provided an unprecedented insight into the evolution of epidemics in the ISC (Q1 and Q3), estimating the emergence of the CG around 1850 (first reported epidemics of AVL in the ISC), evidencing major bottlenecks around 1960 (possibly linked to the malaria-eradication campaign by insecticide spraying, also affecting sand flies), and assessing several recent radiations during the outbreaks following that campaign (Figure 1, Key Figure). If a new sampling was performed now it would probably give evidence of similar bottlenecks as a result of the current kala-azar elimination program (KAEP) and it would possibly reveal new radiations associated with current outbreaks or evidence of new genotypes undetected so far (Figure 1, Q3 and Q4). Third, genome analysis allowed the identification of a new genetic variant

Key Figure

The Importance of Molecular Surveillance with Whole-genome Sequencing (WGS) in the Context of Elimination Programs

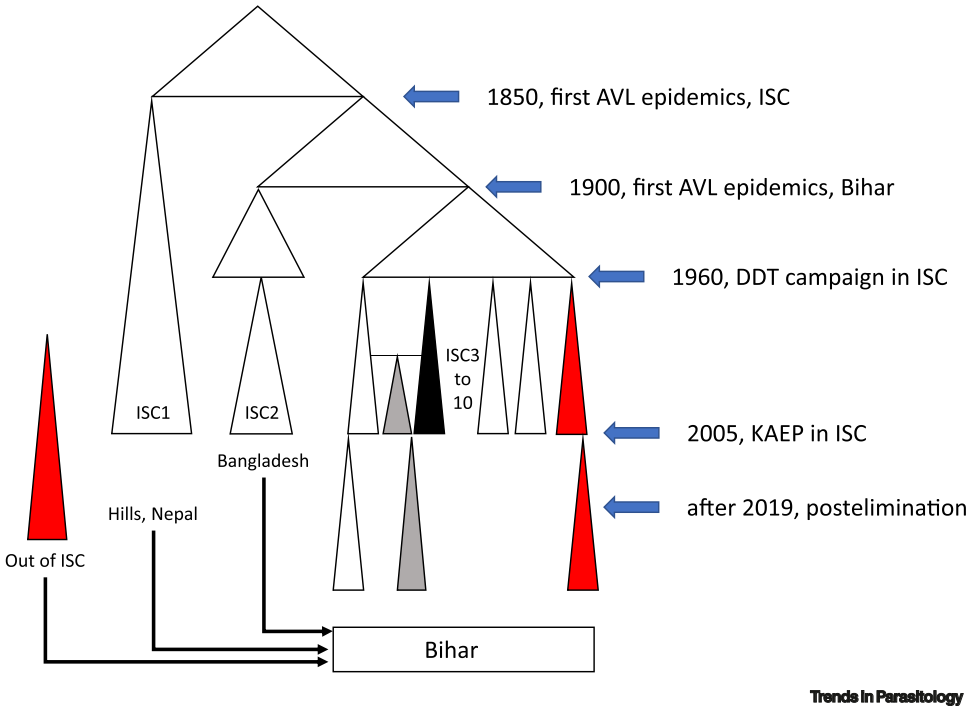

Figure 1. The figure shows the evolution of anthroponotic visceral leishmaniasis (AVL) outbreaks in the Indian subcontinent (ISC) since the first reported epidemics around 1850 (adapted from [27], based on the combination of WGS data with historical and epidemiological data). Each triangle stands for radiation of a parasite subpopulation, the upper tip pointing at the bottleneck at the origin of the population. ISC genetic groups detected between the 2002–2011 survey are uncolored, black, or gray (hybrids). Red triangles stand for (yet) undetected genetic groups. Two control programs are referred to. First, the DDT campaign in the 1960s, which is followed by re-emergence and radiation of several ISC groups. The second one is the Kala-Azar Elimination Programme (KAEP) which started in 2005. The lower part, after 2019, postelimination is virtual. It shows that new postelimination outbreaks in Bihar could originate from different sources: (i) genotypes already detected in 2002–2011 (white and gray), (ii) genotypes already present in Bihar but not yet detected (red), for instance as a consequence of culture-based typing biases, (iii) invasion by ISC1 or ISC2 genotypes from hilly districts in Nepal and from Bangladesh, respectively, and (iv) invasion by new genotypes originating from outside the ISC. This review shows that the tools are available for monitoring these potential re-emergences. DDT, dichlorodiphenyltrichloroethane.

(Q4), significantly different from the CG parasites associated with the main epidemics in the ISC. ISC1 (nickname, 'Yeti') was found in Nepal, especially in hilly districts where AVL had not been reported before (Q2), and these parasites presented several molecular signatures of different virulence [36] and lower drug susceptibility [37] (Q6). Later analysis with an ISC1-specific assay showed an increase in the prevalence of that genotype in the ISC lowlands (Q3) [38], and a recent report demonstrated that this new genetic variant was fully transmissible by *Phlebotomus argentipes*, the vector of AVL in the lowlands [39]. Altogether, these results evidenced the spreading of a new genetic and phenotypic variant. A similar phenomenon was observed in the south of the ISC, where a new genotype of *L. donovani* was first encountered in Sri Lanka in association with CL [40], now also in southern India: WGS of six isolates revealed several genomic signatures likely to explain differences in pathogenicity and response

to treatment [41]. Fourth, WGS has the greatest potential to detect known genetic markers of drug resistance and identify new ones (Q6). In the ISC, we found that all the isolates of the CG presented an intrachromosomal amplificon that includes the MRPA gene, one of the major drivers of antimonial resistance. This amplificon was already present at the origin of the CG (around 1850), much earlier than the discovery of the anti-*Leishmania* activity of antimonials. Possibly resulting from a cross-resistance to other compounds (like arsenicals, very common in water in the ISC [42]), this amplification preadapted CG parasites to antimonials [37]. When this drug was massively used to treat AVL, a series of additional mechanisms, such as the inactivation of AQP1, rendered CG *L. donovani* insensitive to antimonials, leading to the dramatic loss of efficacy of that drug in the ISC [27]. Importantly, WGS of different isolates allowed the identification of many types of molecular change linked to the inactivation, or a decrease in the activity, of AQP1, such as multiple mutations affecting key residues, a 2-nt insertion causing a premature stop codon, or a subtelomeric deletion in chromosome 31, including the loss of the gene [29]. This 'multiple' identification would never have been possible with targeted genotyping methods. Fifth, the high resolution of WGS makes it ideal to study recombination even among parasites with high genetic similarity (Q5). In the ISC, we found evidence that a 2-nt insertion in the AQP1 gene had spread by recombination between ISC groups, thereby resulting in hybrids showing an intermediate resistance (Q5 and Q6).

Why Should Direct Genome Sequencing in Host Tissues Be Preferred to Analysis of Isolates for Epidemiological Monitoring and Clinical Studies?

Current genomic studies assume that the parasite's genome remains the same regardless of the source from which the DNA is isolated, and as such strongly rely on the availability of isolates derived from clinical human samples (Box 1), vectors or animals, and obtained by culturing parasites *in vitro* [27,43,44] or passaging in experimental animals [45,46]. This may introduce a series of important biases. First, isolation is generally done from clinical cases only, while asymptomatic infections are not considered; however, this type of infection may represent the major part of the parasite's population (up to 90% of the cases in AVL [47], about 70% in CHD [48], reported but not quantified in HAT [49]). As such, we get access to the genetic information of only a small part of the parasite's population, clearly the tip of the iceberg. The occurrence of a (hidden) animal reservoir will further increase this representation bias. Secondly, isolation and *in vitro*/*in vivo* maintenance introduces a selection bias. Isolation itself succeeds in only a fraction of the cases (due to contamination, too low parasitemia…etc.), and when successful it may create a bottleneck as only a few parasites present in the initial sample will be able to survive and multiply in the medium or experimental animal. In addition, *in vitro*/*in vivo* maintenance will select parasites which are fitter for this 'artificial' experimental environment, and in case of sample heterogeneity, a genotype/species can emerge in the isolate used for sequence analysis, while not being dominant in the initial clinical sample. This was, for example, clearly documented in cases of *L. donovani*/*Leptomonas* coinfections in AVL patients, for which *Leptomonas* rapidly overgrew *L. donovani* in culturing conditions [50].

Recent experimental evidence demonstrated that the genome of a strain of parasite (Box 1) could be different depending on the environment. Highly aneuploid *L. donovani* strains cultivated *in vitro* were passed to sand fly vectors and hamsters, and subsequently the respective parasite genomes were compared. Much lower levels of aneuploidy were observed in the *Leishmania* genome in the mammalian host as compared to *in vitro* grown parasites [51,52]. The aneuploidy pattern was shown to be very dynamic during early adaptation to culture conditions after isolation from a mammalian host environment: at passage 2 (around 20 generations) already four new aneuploidies arose [52]. Such fluctuations in chromosome numbers during the life cycle are likely due to the selection of better adapted variants in a mosaic background [51,52] and may have an

impact on SNP diversity, especially in cases of heterozygosity [52]. Aneuploidy was also reported in *T. cruzi* [53], but not in *T. brucei* subspecies [54]. Genomic differences might thus also be expected among life stages of *T. cruzi*, but this needs to be further explored.

For a first or general description of circulating parasite populations in specific foci or basic phylogenomic studies, genomic analysis of maintained isolates should be sufficient. However, direct sequencing without isolating and maintaining parasites *in vitro* or *in vivo* will increase the genomic representativity of the parasites present in the infected tissue samples. Furthermore, given the genomic differences encountered during experimental evolution studies throughout the life cycle as described above, we might expect (major) differences between samples and isolates. This may have a high impact when addressing specific questions, in particular the relationship between the parasite's genome and clinical parameters.

How Can a Parasite's Genome Be Sequenced without Isolation and Maintenance *in vitro* or *in vivo*?

Direct sequencing of a parasite's genome is prone to specific challenges, essentially the high abundance of host DNA (particularly problematic for intracellular stages as in *Leishmania* and *T. cruzi*) and the low parasite loads in some clinical samples or in asymptomatic individuals [55,56]. When the proportion of pathogen DNA (vs host DNA) is at least 5%, WGS sequencing can be done without further processing of the samples. The disadvantage of this approach is a relatively high sequencing cost, as samples need to be sequenced deeply enough to obtain enough reads which map to the pathogen's genome. However, in most cases parasite loads in collected samples are too low [55,56] (also depending on the organ [57]) to allow sequencing without applying some enrichment. Generally, clinical samples can be enriched in a pathogen's DNA during two stages: when the cells are still intact, and after the DNA extraction from a collected sample.

Pre-DNA-extraction enrichment is based on different chemical, physical, biochemical, or biological properties of host and microbial cells. For instance, immunomagnetic separation, or differential cell lysis with commercial kits, is used for bacterial genomic studies [58]. Similarly, a filtration method that removes leukocytes and platelets was developed and applied for WGS analysis of *Plasmodium vivax*-infected blood samples [59]. In strains of *T. brucei* (extracellular parasites), anion-exchange chromatography is used to obtain a pure suspension of trypanosomes from infected blood [60], and, depending on parasitemia, sequencing could be done with or without whole-genome amplification. Alternatively, buffy coat collection can be used to concentrate trypanosomes, but these cells will still be mixed with leukocytes.

Enrichment strategies after total DNA extraction would be required for buffy coat analyses, for trypanosomatids living in nucleated cells, and for any parasite collected in remote sites, often with a basic laboratory infrastructure. A series of methods have been developed for other pathogens (Box 3), and we adapted and evaluated one of them, that is, target enrichment with SureSelect, for direct sequencing of *L. donovani* [2]. The method was applied to 63 clinical samples (bone marrows and splenic aspirates) from AVL patients and allowed (i) to assign 97% of samples to previously defined ISC genotypes using a set of diagnostic SNPs, (ii) to get access to comprehensive genome-wide information in 83% of the samples, and (iii) to get a mean coverage of 61.1% (minimal coverage with five reads) to 78.4% (minimal coverage with one read). Phylogeny was consistent with previous analyses [27], but a new ISC group was identified [2]. However, the most striking result appeared when comparing 12 clinical samples and the corresponding cultured isolates: all 12 isolates showed a different genome (SNPs, aneuploidy and indel) compared to the parasites present in the paired bone-marrow samples. This was likely explained (and partly

Box 3. Methods for Sample Enrichment of Pathogen DNA

A series of approaches have been used for pathogens, including (i) removal of methylated host DNA [71], (ii) selective whole-genome amplification (SWGA) [72], and (iii) targeted whole-genome capture [73]. The first method takes advantage of the differential levels of methylation between the mammalian host and pathogenic DNA. In particular, it has been applied to enrich samples in the DNA of bacterial or *Plasmodium* spp. through differential restriction digestion or commercial kits available to facilitate this approach [71,74]. Although promising, this method requires large amounts of total DNA (1–2 μg) and a high content of pathogen DNA (10% for *Plasmodium*-containing samples), precluding its application on many clinical samples. The second method is based on amplification with phi29 using short oligonucleotides that preferentially bind to several sequences across the whole genome of a parasite [72,75,76]. As a result, the pathogen's genome is preferentially amplified, allowing the enrichment. The advantage of this method is that it can be applied to relatively low amounts of total DNA (as low as 5 ng) and to samples with much lower levels of the parasite's DNA content (in the range of 0.03– 0.1%, depending on the species and the sample used). However, as the enrichment is based on whole-genome amplification, it can be a challenge to obtain even genome coverage [77]; the method is not cost-effective for small studies, but is good for large ones [77]. As it uses multiple displacement amplification (MDA) it is not directly applicable for aneuploidy measurements [78]. In the third method, targeted whole-genome capture, a large set of specific probes are used to capture the genome of interest. Several commercial systems exist [79]. The SureSelect system (Agilent) was initially developed for human exome capture [80] but can be customized to capture any genome in a complex mixture of DNA. The method uses biotinylated 120-base RNA probes which hybridize with target DNA. Captured sequences are enriched with streptavidin-conjugated paramagnetic beads, amplified, and submitted to high-throughput sequencing. For information, in the *L. donovani* study presented in this review, the final design of the array contained 218 904 probes, covering 26 Mbp of the 32 Mbp haploid genome of that species [2]. The protocol required a minimum of 100 ng of DNA, and a percentage of *Leishmania* DNA of 0.006% was found to be the lowest limit so far for suitable analysis of genome diversity [2].

Outstanding Questions

Genome differences between samples and derived isolates are well documented for *L. donovani* in the ISC. Does this phenomenon also occur in other parasites?

How could genome capture methods be improved, especially in terms of analytical sensitivity? Is it possible to allow applications in samples with low parasite loads such as blood or skin?

Parasite diversity studies by direct genome sequencing, or other methods, allow academic research on the biology and microecology of trypanosomatids. How could the acquired information be used to fine-tune disease-transmission models and realign control strategies?

How should we implement a true molecular surveillance platform for the TriTryps?

demonstrated) by the occurrence of polyclonal infections, with different dominating clones in the clinical samples and the derived isolates, likely because of clone fitness differences and subsequent selection during isolation. Among these, one clinical sample presented a mutated form of AQP1, associated with resistance to antimonials, while the wild-type allele was detected in the derived isolate. Probably, when in competition with parasites containing the wild-type allele during the isolation process, parasites with the mutated AQP1 were growing slower and progressively escaped from the WGS radar. Altogether, this pioneer study demonstrated the risk of bias when sequencing isolates, illustrating the clear need for direct sequencing when aiming to establish a link between parasite's genome variation and clinical phenotypes such as treatment failure and drug resistance. To our knowledge, this is the only study published on direct sequencing of TriTryps and it should be extended by similar studies in strains of *T. brucei* and *T. cruzi*.

Concluding Remarks

In this review, we highlighted the neglect of molecular surveillance despite its demonstrated relevance, the added value of WGS for it, and the relevance of direct sequencing to answer clinical and epidemiological questions. The concepts and tools here discussed should pave the way towards platforms for molecular surveillance of TriTryp diseases, but further work is needed to achieve this objective (see Outstanding Questions). First, the relevance of direct sequencing in host tissues (vs isolates) should be documented in other species of *Leishmania* and trypanosomatids, and polyclonality should be further explored. Second, clinical samples to be analyzed in a surveillance context should be carefully chosen, where tissues presenting a higher parasite load and allowing a better sampling feasibility could be the priority (Table 1). However, the clinical and epidemiological relevance of the tissues could also be considered: for instance, blood and skin might be more adequate to study transmission of TriTryp parasites, while other tissues could be more relevant for monitoring pathogenic features or drug resistance (Table 1). Asymptomatic individuals should be included as they constitute the majority of infections for many trypanosomatids. Third, depending on the targeted samples and parasite load, the analytical sensitivity of direct sequencing should be optimized. If large amounts of total DNA are available, (partial) removal of methylated human DNA could be introduced as the first enrichment step, because no DNA methylation could be found in the genomes of *Leishmania* and *T. brucei* [61] (still to be verified in *T. cruzi*). Depending on the remaining amount of DNA, the second step

Table 1. Applicability and Relevance of TriTryp Direct Sequencing in Different Tissues[a]

Type of sample	*Trypanosoma brucei* strains (HAT)	*Trypanosoma cruzi* (CHD)	*Leishmania donovani* (AVL)	Sampling feasibility (routine)	Pre-DNA-extraction enrichment	Enrichment after DNA extraction	Relevance (transmission)	Relevance (pathogenic features)
Vector	x	x	x	x	(x)[b]	x	x	–
Animal reservoir	(x)[c]	x	?	x	(x)[d]	x	x	–
Skin	x	x	x	(x)[e]	–	(x)[f]	x	–
Blood	x	x	x	x	x	(x)[g]	x	–
Lymph node	x	x	?	–	–	–	–	x
Central nervous system	x	–	–	–	–	–	–	x
Adipose tissue	x	x	?	–	–	(x)[f]	–	x
Heart	–	x	–	–	–	(x)[f]	–	x
Digestive track	–	x	–	–	–	(x)[f]	–	x
Bone marrow	–	–	x	(x)[h]	–	x	–	x
Spleen	–	–	x	(x)[h]	–	x	–	x
Cutaneous lesion	–	–	x	x	–	x	x	x

[a]Each column represents the tissue tropism of TriTryps responsible for HAT, CHD, or AVL (partially after [57]), sampling feasibility in routine conditions, possibility of enrichment in pathogen DNA before whole-genome sequencing (WGS) (with the sensitivity of currently developed methods, still to be improved, performance will depend essentially on parasite load, which can differ between species, and according to the clinical features), and relevance of the samples (in the context of transmission and/or pathogenic features). x, yes; –, no; ?, unknown, (x), yes with limitations.
[b]Parasites can be directly purified from insects, but infection rates and amounts of the parasite per insect are rather low.
[c]Well established for *T. brucei rhodesiense*; preliminary evidence for *T. brucei gambiense* [5].
[d]Possible with blood.
[e]Experimental proof of concept [81], not yet used routinely.
[f]Theoretically possible but parasite loads probably too low.
[g]Possible in clinical cases, less in asymptomatic individuals.
[h]Very sensitive but not recommended any more in several countries.

could be based on whole-genome capture or (selective) whole-genome amplification (Box 3). Fourth, strategies should be developed to allow sampling in less-accessible foci or when outbreaks arise. To improve sampling, reagents preserving DNA integrity in field conditions, such as DNA Shield, could be applied. This would allow sampling in several, often remote, places, but DNA extraction could then take place in a centralized place, allowing a standardization of this and further downstream processes. Fifth, other criteria should be considered to develop molecular surveillance platforms (see also [62]), specifically in terms of data analysis and management. Bioinformatic pipelines should be developed and (semi-)automated to allow analysis by different types of user, including non-bioinformaticians. A central computational infrastructure, training, and standard operating procedures should be available. It should be possible for scientists to cumulatively feed the database with results of newly sequenced samples and to compare them with those already present in the database. The molecular surveillance platform should be integrated in existing clinical and epidemiological platforms, in partnership with stakeholders in the endemic countries and international health authorities. Data should be visualized and displayed in an interactive interface (such as Nextstrain), with geolocation, simple analytical tools, and aggregated metadata (such as the Molecular Surveyors used in *Plasmodium falciparum*[iii]). The data visualization should equip managers of control programs, researchers, and clinicians to understand the genetic landscape of the corresponding diseases, answer specific questions such as those mentioned in the third section of this review, and plan adequate counter measures or address further research questions. A pilot study could be done on one disease in a specific region (we would recommend

AVL in the ISC, where most preliminary data are available), but exploration should be done to transfer the concept to other diseases in other regions of the world.

Acknowledgments

We thank Jan Van den Abbeele and Pieter Monsieurs for reading and commenting on this paper which is based on studies that received financial support from (i) EWI, Flemish Agency for Science and Innovation (SOFI-projects GEMINI and SINGLE), and (ii) EC-FP7 (KALADRUG-R, grant 222895).

Resources

[i]http://faculty.ucmerced.edu/kjensen5/index.php/research/global-burden-of-parasitic-disease/

[ii]https://frebio.github.io/komics/

[iii]www.wwarn.org/molecular-surveyor-k13

[iv]www.who.int/bulletin/volumes/84/2/editorial10206html/en/

References

1. Schwelm, A. *et al.* (2018) Not in your usual Top 10: protists that infect plants and algae. *Mol. Plant Pathol.* 19, 1029–1044
2. Domagalska, M.A. *et al.* (2019) Genomes of *Leishmania* parasites directly sequenced from patients with visceral leishmaniasis in the Indian subcontinent. *PLoS Negl. Trop. Dis.* 13, e0007900
3. Pérez-Molina, J.A. and Molina, I. (2018) Seminar Chagas disease. *Lancet* 391, 82–94
4. Rijal, S. *et al.* (2019) Eliminating visceral leishmaniasis in South Asia: The road ahead. *BMJ* 364, k5224
5. Büscher, P. *et al.* (2018) Do cryptic reservoirs threaten gambiense-sleeping sickness elimination? *Trends Parasitol.* 34, 197–207
6. Flores-Ferrer, A. *et al.* (2018) Evolutionary ecology of Chagas disease; what do we know and what do we need? *Evol. Appl.* 11, 470–487
7. Meehan, C.J. *et al.* (2019) Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat. Rev. Microbiol.* 17, 533–545
8. Rugna, G. *et al.* (2018) Multilocus microsatellite typing (MLMT) reveals host-related population structure in *Leishmania infantum* from northeastern Italy. *PLoS Negl. Trop. Dis.* 12, e0006595
9. Echodu, R. *et al.* (2015) Genetic diversity and population structure of *Trypanosoma brucei* in Uganda: implications for the epidemiology of sleeping sickness and nagana. *PLoS Negl. Trop. Dis.* 9, e0003353
10. Costales, J.A. *et al.* (2015) *Trypanosoma cruzi* population dynamics in the Central Ecuadorian Coast. *Acta Trop.* 151, 88–93
11. Torrellas, A. *et al.* (2018) Molecular typing reveals the coexistence of two transmission cycles of American cutaneous leishmaniasis in the Andean Region of Venezuela with *Lutzomyia migonei* as the vector. *Mem. Inst. Oswaldo Cruz* 113, e180323
12. Kaboré, J. *et al.* (2011) First evidence that parasite infecting apparent aparasitemic serological suspects in human African trypanosomiasis are *Trypanosoma brucei gambiense* and are similar to those found in patients. *Infect. Genet. Evol.* 11, 1250–1255
13. Curtis-Robles, R. *et al.* (2017) Epidemiology and molecular typing of *Trypanosoma cruzi* in naturally-infected hound dogs and associated triatomine vectors in Texas, USA. *PLoS Negl. Trop. Dis.* 11, e0005298
14. García, N. *et al.* (2014) Evidence of *Leishmania infantum* infection in rabbits (*Oryctolagus cuniculus*) in a natural area in Madrid, Spain. *Biomed. Res. Int.* 2014, 318254
15. N'Djetchi, M.K. *et al.* (2017) The study of trypanosome species circulating in domestic animals in two human African trypanosomiasis foci of Côte d'Ivoire identifies pigs and cattle as potential reservoirs of *Trypanosoma brucei gambiense*. *PLoS Negl. Trop. Dis.* 11, e0005993
16. Segovia, M. *et al.* (2013) Molecular epidemiologic source tracking of orally transmitted Chagas disease, Venezuela. *Emerg. Infect. Dis.* 19, 1098–1101
17. Kariyawasam, U.L. *et al.* (2017) Genetic diversity of *Leishmania donovani* that causes cutaneous leishmaniasis in Sri Lanka: A cross sectional study with regional comparisons. *BMC Infect. Dis.* 17, 791
18. Capewell, P. *et al.* (2015) A co-evolutionary arms race: Trypanosomes shaping the human genome, humans shaping the trypanosome genome. *Parasitology* 142, S108–S119
19. Ramírez, J.D. *et al.* (2013) Genetic structure of *Trypanosoma cruzi* in Colombia revealed by a high-throughput nuclear multilocus sequence typing (nMLST) approach. *BMC Genet.* 14, 96
20. Gelanew, T. *et al.* (2014) Multilocus sequence and microsatellite identification of intra-specific hybrids and ancestor-like donors among natural ethiopian isolates of *Leishmania donovani*. *Int. J. Parasitol.* 44, 751–757
21. Gibson, W. *et al.* (2015) Genetic recombination between human and animal parasites creates novel strains of human pathogen. *PLoS Negl. Trop. Dis.* 9, e0003665
22. Izeta-Alberdi, A. *et al.* (2016) Geographical, landscape and host associations of *Trypanosoma cruzi* DTUs and lineages. *Parasit. Vectors* 9, 631
23. Mohebali, M. *et al.* (2019) Gene expression analysis of antimony resistance in *Leishmania tropica* using quantitative real-time PCR focused on genes involved in trypanothione metabolism and drug transport. *Arch. Dermatol. Res.* 311, 9–17
24. Graf, F.E. *et al.* (2013) Aquaporin 2 mutations in *Trypanosoma brucei gambiense* field isolates correlate with decreased susceptibility to pentamidine and melarsoprol. *PLoS Negl. Trop. Dis.* 7, e2475
25. Martínez, I. *et al.* (2013) Microsatellite and Mini-Exon analysis of Mexican human DTU I *Trypanosoma cruzi* strains and their susceptibility to nifurtimox and benznidazole. *Vector-Borne Zoonotic Dis.* 13, 181–187
26. Van der Auwera, G. *et al.* (2016) Comparison of *Leishmania* typing results obtained from 16 European clinical laboratories in 2014. *Eurosurveillance* 21 30418
27. Imamura, H. *et al.* (2016) Evolutionary genomics of epidemic visceral leishmaniasis in the Indian subcontinent. *eLife* 5, e12613
28. Downing, T. *et al.* (2012) Genome-wide SNP and microsatellite variation illuminate population-level epidemiology in the *Leishmania donovani* species complex. *Infect. Genet. Evol.* 12, 149–159
29. Hefnawy, A. *et al.* (2017) Exploiting knowledge on *Leishmania* drug resistance to support the quest for new drugs. *Trends Parasitol.* 33, 162–174
30. Gerner-Smidt, P. *et al.* (2019) Whole genome sequencing: bridging one-health surveillance of foodborne diseases. *Front. Pub. Health* 7, 172
31. El-Sayed, N.M. *et al.* (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404–409
32. Imamura, H. and Dujardin, J.C. (2019) A guide to next generation sequence analysis of *Leishmania* genomes. *Methods Mol. Biol.* 1971, 69–94
33. Van den Broeck, F. *et al.* (2018) Mitonuclear genomics challenges the theory of clonality in *Trypanosoma congolense*: Reply to Tibayrenc and Ayala. *Mol. Ecol.* 27, 3425–3431

34. Camacho, E. *et al.* (2019) *Leishmania* mitochondrial genomes: maxicircle structure and heterogeneity of minicircles. *Genes (Basel)* 10, 758
35. Van den Broeck, F. *et al.* (2019) Ecological divergence and hybridization of Neotropical *Leishmania* parasites. *bioRxiv*. Published online November 22, 2019. https://doi.org/10.1101/824912
36. Cuypers, B. *et al.* (2018) Integrated genomic and metabolomic profiling of ISC1, an emerging *Leishmania donovani* population in the Indian subcontinent. *Infect. Genet. Evol.* 62, 170–178
37. Dumetz, F. *et al.* (2018) Molecular preadaptation to antimony resistance in *Leishmania donovani* on the Indian subcontinent. *mSphere* 3, e00548-17
38. Rai, K. *et al.* (2017) Single locus genotyping to track *Leishmania donovani* in the Indian subcontinent: Application in Nepal. *PLoS Negl. Trop. Dis.* 11, e0005420
39. Seblova, V. *et al.* (2019) ISC1, a new *Leishmania donovani* population emerging in the Indian sub-continent: Vector competence of *Phlebotomus argentipes*. *Infect. Genet. Evol.* 76, 104073
40. Siriwardana, H.V.Y.D. *et al.* (2007) *Leishmania donovani* and cutaneous leishmaniasis, Sri Lanka. *Emerg. Infect. Dis.* 13, 476–478
41. Samarasinghe, S.R. *et al.* (2018) Genomic insights into virulence mechanisms of *Leishmania donovani*: Evidence from an atypical strain. *BMC Genomics* 19, 843
42. Perry, M. *et al.* (2015) Arsenic, antimony, and *Leishmania*: has arsenic contamination of drinking water in India led to treatment-resistant kala-azar? *Lancet* 385, S80
43. Rogers, M.B. *et al.* (2014) Genomic confirmation of hybridisation and recent inbreeding in a vector-isolated *Leishmania* population. *PLoS Genet.* 10, e1004092
44. Schwabl, P. *et al.* (2019) Meiotic sex in Chagas disease parasite *Trypanosoma cruzi*. *Nat. Commun.* 10, 3972
45. Goodhead, I. *et al.* (2013) Whole-genome sequencing of *Trypanosoma brucei* reveals introgression between subspecies that is associated with virulence. *mBio* 4, e00197-13
46. Weir, W. *et al.* (2016) Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *eLife* 5, e11473
47. Dorlo, T.P.C. *et al.* (2016) Treatment of visceral leishmaniasis: Pitfalls and stewardship. *Lancet Infect. Dis.* 16, 777–778
48. Murray, D.L. and Burke, A. (2012) Chagas disease often asymptomatic but can be life-threatening if untreated. *AAP News* 33, 12
49. Capewell, P. *et al.* (2019) Resolving the apparent transmission paradox of African sleeping sickness. *PLoS Biol.* 17, e3000105
50. Srivastava, P. *et al.* (2010) Detection of *Leptomonas* sp. parasites in clinical isolates of Kala-azar patients from India. *Infect. Genet. Evol.* 10, 1145–1150
51. Dumetz, F. *et al.* (2017) Modulation of aneuploidy in *Leishmania donovani* during adaptation to different *in vitro* and *in vivo* environments and its impact on gene expression. *mBio* 8, e00599-17
52. Barja, P.P. *et al.* (2017) Haplotype selection as an adaptive mechanism in the protozoan pathogen *Leishmania donovani*. *Nat. Ecol. Evol.* 1, 1961–1969
53. Reis-Cunha, J.L. *et al.* (2018) Whole genome sequencing of *Trypanosoma cruzi* field isolates reveals extensive genomic variability and complex aneuploidy patterns within TcII DTU. *BMC Genomics* 19, 816
54. Almeida, L.V. *et al.* (2018) Chromosomal copy number variation analysis by next generation sequencing confirms ploidy stability in *Trypanosoma brucei* subspecies. *Microb. Genom.* Published online September 27, 2018. https://doi.org/10.1099/mgen.0.000223
55. Verma, S. *et al.* (2010) Quantification of parasite load in clinical samples of leishmaniasis patients: Il-10 level correlates with parasite load in visceral leishmaniasis. *PLoS One* 5, e10107
56. Macchiaverna, N.P. *et al.* (2020) Human infectiousness and parasite load in chronic patients seropositive for *Trypanosoma cruzi* in a rural area of the Argentine Chaco. *Infect. Genet. Evol.* 78, 104062
57. Pereira, S.S. *et al.* (2019) Tissue tropism in parasitic diseases. *Open Biol.* 9, 190124
58. Seth-Smith, H.M.B. *et al.* (2013) Generating whole bacterial genome sequences of low-abundance species from complex samples with IMS-MDA. *Nat. Protoc.* 8, 2404–2412
59. Sriprawat, K. *et al.* (2009) Effective and cheap removal of leukocytes and platelets from *Plasmodium vivax* infected blood. *Malar. J.* 8, 115
60. Lejon, V. *et al.* (2019) The separation of trypanosomes from blood by anion exchange chromatography: From Sheila Lanham's discovery 50 years ago to a gold standard for sleeping sickness diagnosis. *PLoS Negl. Trop. Dis.* 13, e0007051
61. Cuypers, B. *et al.* (2019) C-5 DNA methyltransferase 6 does not generate detectable DNA methylation in *Leishmania*. *bioRxiv*. Published online August 28, 2019. https://doi.org/10.1101/747063
62. Mintzer, V. *et al.* (2019) Operational models and criteria for incorporating microbial whole genome sequencing in hospital microbiology – A systematic literature review. *Clin. Microbiol. Infect.* 25, 1086–1095
63. Dowdle, W.R. (1998) The principles of disease elimination and eradication. *Bull. World Health Organ.* 76, 22–25
64. Meystre, S.M. *et al.* (2012) Clinical research in the postgenomic era. In *Clinical Research Informatics* (Richesson, R. and Andrews, J., eds), pp. 113, Springer
65. Tibayrenc, M. (2006) The species concept in parasites and other pathogens: a pragmatic approach? *Trends Parasitol.* 22, 66–70
66. Richardson, J.B. *et al.* (2017) Genomic analyses of African Trypanozoon strains to assess evolutionary relationships and identify markers for strain identification. *PLoS Negl. Trop. Dis.* 11, e0005949
67. Zingales, B. (2018) *Trypanosoma cruzi* genetic diversity: Something new for something known about Chagas disease manifestations, serodiagnosis and drug sensitivity. *Acta Trop.* 184, 38–52
68. Akhoundi, M. *et al.* (2016) A historical overview of the classification, evolution, and dispersion of *Leishmania* parasites and sandflies. *PLoS Negl. Trop. Dis.* 10, e0004349
69. Baker, J.R. *et al.* (1978) Proposals for the nomenclature of salivarian trypanosomes and for the maintenance of reference collections. *Bull. World Health Organ.* 56, 467–480
70. van der Auwera, G. and Dujardina, J.C. (2015) Species typing in dermal leishmaniasis. *Clin. Microbiol. Rev.* 28, 265–294
71. Feehery, G.R. *et al.* (2013) A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One* 8, e76096
72. Oyola, S.O. *et al.* (2016) Whole genome sequencing of *Plasmodium falciparum* from dried blood spots using selective whole genome amplification. *Malar. J.* 15, 597
73. Doyle, R.M. *et al.* (2018) Direct whole-genome sequencing of sputum accurately identifies drug-resistant *Mycobacterium tuberculosis* faster than MGIT culture sequencing. *J. Clin. Microbiol.* 56, e00666-18
74. Oyola, S.O. *et al.* (2013) Efficient depletion of host DNA contamination in malaria clinical sequencing. *J. Clin. Microbiol.* 51, 745–751
75. Leichty, A.R. and Brisson, D. (2014) Selective whole genome amplification for resequencing target microbial species from complex natural samples. *Genetics* 198, 473–481
76. Cowell, A.N. *et al.* (2017) Selective whole-genome amplification is a robust method that enables scalable whole-genome sequencing of *Plasmodium vivax* from unprocessed clinical samples. *mBio* 8, e02257-16
77. Clarke, E.L. *et al.* (2017) swga: a primer design toolkit for selective whole genome amplification. *Bioinformatics* 33, 2071–2077
78. Macaulay, I.C. and Voet, T. (2014) Single cell genomics: advances and future perspectives. *PLoS Genet.* 10, e1004126
79. Gaudin, M. and Desnues, C. (2018) Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Front. Microbiol.* 9, 2924
80. Chen, R. *et al.* (2015) Whole-exome enrichment with the Agilent Sureselect Human All Exon Platform. *Cold Spring Harb. Protoc.* 2015, 626–633
81. Doehl, J.S.P. *et al.* (2017) Skin parasite landscape determines host infectiousness in visceral leishmaniasis. *Nat. Commun.* 8, 57