CellPress

Special Issue: Human Genetics

Feature Review

Neglected Tropical Diseases in the Post-Genomic Era

Carlos A. Buscaglia,[1] Jessica C. Kissinger,[2] and Fernán Agüero[1,*]

Neglected tropical diseases (NTDs) are a group of viral, bacterial, and eukaryotic parasitic diseases that are especially endemic in low-income populations, with a large health and economic impact on both the developing and developed world. The structure and dynamics of the genomes of the organisms causing these diseases, as well as the modes of expression, exchange, and transmission of their genetic information, often deviate from those found in classical, model organism-centric textbooks. We assess the role of basic and applied genetic research in our understanding of key aspects of their biology and evolution, and discuss the impact of novel high-throughput approaches spawned by the post-genomic era on the development of next-generation drugs, vaccines, molecular epidemiology, and/or diagnostic tools for these important pathogens.

Trends

Neglected tropical diseases have a large health and economic impact on human populations.

The addition of significant effort and resources can move an organism off of the NTD list.

The organisms causing these diseases often have unique genetic features.

We describe the role that genetic research had in our understanding of their biology and evolution.

We discuss the impact of novel high-throughput approaches spawned by the post-genomic era.

Neglected Tropical Diseases and Their Pathogens

NTDs are a heterogeneous group of infectious diseases unified not by pathophysiology, evolution or geography – they can be found outside the tropics – but by their endemicity, especially in low-income populations in developing regions of Africa, Asia, and the Americas, thus perpetuating the poverty of the poorest people in the world [1,2]. In addition to being major pressing public health issues, most NTDs are zoonotic, thus posing an additional economic burden by affecting reliance on livestock for sustainable human development [1,2]. Importantly, the 17 illnesses currently classified as NTDs by the 2012 London Declaration (compiled in Table 1) are preventable and most are treatable. However, they still threaten the lives of more than 1 billion people worldwide [3]. Moreover, NTD outbreaks associated with previously unrecognized biological or epidemiological features were recently reported [4,5].

The causative agents of NTDs (henceforth NTD pathogens) represent a wide phylogenetic sampling of parasitic organisms, which include viruses, bacteria and eukaryotes (both unicellular protozoa and metazoan worms) (Table 1). They often display intricate life cycles, involving multiple developmental forms, both free-living and/or parasitic for either insect vectors or intermediate hosts. NTD pathogens use different strategies to ensure entry, infectivity, survival, and transmission to and from the human host, in which they trigger a wide spectrum of disease-associated pathologies. In accordance with this great biological diversity, the structure and dynamics of their genomes are fairly disparate. At the core of this phenomenon, a myriad of genetic transmission strategies have been unveiled, ranging from true haploid bacteria with a striking capacity for recombination and/or horizontal gene transfer (HGT), to asexual or rarely sexual protozoa, to classical, obligate sexual metazoan helminths (see below). These genetic exchange and/or reproductive strategies, in turn, play a key role in shaping the population structures of NTD pathogens and thus have major implications for the development, evaluation, and application of diagnostic, chemotherapeutic, or vaccine-based control strategies.

[1]Instituto de Investigaciones Biotecnológicas – Instituto Tecnológico de Chascomús (IIB-INTECH), Universidad Nacional de San Martín (UNSAM) – Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), San Martín, B 1650 HMP, Buenos Aires, Argentina
[2]Department of Genetics – Center for Tropical and Emerging Global Diseases – Institute of Bioinformatics, University of Georgia, Athens, GA 30602, USA

*Correspondence: fernan.aguero@gmail.com, fernan@iib.unsam.edu.ar (F. Agüero).

Table 1. The 17 Neglected Tropical Diseases[a]

Disease	Causative Agent	Taxonomy Notes	Genomic Data Available
Buruli ulcer	*Mycobacterium ulcerans*	Actinobacteria (Gram positive)	Complete genome available [144]. https://www.patricbrc.org/ [145].
Chagas disease	*Trypanosoma cruzi*	Kinetoplastid Protozoa	Draft genomes of TcI [146], TcII and TcVI (hybrid) [147] lineages available. Projects to sequence genomes from other lineages are underway. http://tritrypdb.org [148].
Dengue	Dengue virus	Flavivirus	Genomes for all dengue virus serotypes (DENV-1–4) are available (reviewed in [149]), and functional RNA elements of the genome have been described [150]. http://www.viprbrc.org/ [151].
Dracunculiasis (Guinea worm disease)	*Dracunculus medinensis*	Nematoda (roundworms)	Sequencing underway at WTSI.
Echinococcosis	*Echinococcus granulosus, E. multilocularis*	Platyhelminths (flatworms), Cestoda	Complete genomes available [17,152].
Foodborne trematodiases (clonorchiasis, ophisthorciasis, fasciolasis, paragonimiasis)	*Clonorchis* spp., *Opisthorchis* spp., *Fasciola* spp., and *Paragonimus* spp.	Platyhelminths (flatworms), Trematoda	Draft genome for *Fasciola hepatica* available [153].
Human African trypanosomiasis (sleeping sickness)	*Trypanosoma brucei*	Kinetoplastid Protozoa	Complete reference genome available [154]. Genomes from different subspecies and strains/isolates also available [49,155]. http://tritrypdb.org [148].
Leishmaniasis	Several species of *Leishmania*	Kinetoplastid Protozoa	Complete reference genome available [80]. Genomes from different subspecies and strains/isolates also available [82,156,157]. http://tritrypdb.org [148].
Leprosy	*Mycobacterium leprae*	Actinobacteria (Gram positive)	Complete genome sequence of extant and medioeval isolates [158,159]. https://www.patricbrc.org/ [145].
Lymphatic filariasis	*Wuchereria bancrofti, Brugia malayi*, and *Brugia timori*	Filarial worms (Filarioidea); Nematoda (roundworms)	Complete genome sequence of *Brugia malayi* (and its endosymbiont) available [160]. Pre-publication draft genome data for *Wuchereria* and other filarial worms are available at the FWD.
Onchocerciasis (river blindness)	*Onchocerca volvulus*	Filarial worms (Filarioidea); Nematoda (roundworms)	Draft, partially assembled shotgun sequences are available (GenBank CBVM000000000).
Rabies	Rabies virus	Lyssavirus	Complete reference genome available [161]. http://www.viprbrc.org/ [151].
Schistosomiasis	Several species of the `genus *Schistosoma*: *S. haematobium* (urogenital) and any of *S. guineensis*, *S. intercalatum*, *S. mansoni*, *S. japonicum*, and *S. mekongi* (intestinal)	Platyhelminths (flatworms), Trematoda	Three genome sequences are available (*S. mansoni*, *S. japonicum*, *S. haemotobium)*, together with high-quality transcriptome data [16,162,163]. http://schistodb.net [164].

Glossary

Aneuploidy: a condition in which the number of chromosomes in the nucleus of a cell is not an exact multiple of the haploid number of chromosomes for a particular species.

Biovars: variant prokaryotic 'type' that differs physiologically and/or biochemically from other 'types' in a particular species. In the case of *C. trachomatis*, biovars differ both in their antigenic properties and in their pathotype (i.e., in their pathogenicity in a specific host).

Genetic hybridization: interbreeding/sexual recombination between genetically distinct individuals to produce an individual with different alleles of the same gene. Hybridization as used here refers to the genetic offspring produced from interbreeding between individuals of different species or subspecies to produce a stable offspring that carries alleles from each parent.

Introgression: the transfer of genetic material from one species to another as a result of genetic hybridization followed by repeated backcrossing (sexual recombination) with the original 'parent' species as opposed to the donor of the transferred genetic material.

Kinetoplast DNA: the kinetoplast is a network of circular DNA (termed kDNA) inside a large mitochondrion that contains many copies of the mitochondrial genome. It is only found in protozoa of the order Kinetoplastida. The kinetoplast contains circular DNA in two forms, maxicircles and minicircles, which are concatenated to form a network. During cell division, replication of this network requires that these rings are disconnected from the parental kinetoplast and subsequently reconnected in the daughter kinetoplast.

Multi-locus sequence typing (MLST): a DNA amplification and sequencing technique used to characterize pathogen isolates. The technique is based on genotyping DNA fragments from 7–8 independent genomic loci, usually housekeeping genes.

Mosaic aneuploidy: a mixed population of cells having chromosomes displaying extra, or fewer, copies than normal (e.g., haploid, diploid, triploid).

Table 1. (continued)

Disease	Causative Agent	Taxonomy Notes	Genomic Data Available
Soil transmitted helminthiases	*Ascaris lumbricoides* (roundworm), *Trichuris trichiura* (whipworm), *Necator americanus*, and *Ancylostoma duodenale* (hookworms)	Nematoda	http://nematode.net/NN3_frontpage.cgi [165].
Taeniasis/ cysticercosis	*Taenia solium* (pork tapeworm) and *T. saginata* (beef tapeworm)	Platyhelminths (flatworms), Cestoda	http://www.ncbi.nlm.nih.gov/bioproject/PRJNA170813
Trachoma	*Chlamydia trachomatis*	Chlamydiales	Reference genome available [166], and pan-genome from different isolates also available [167]. https://www.patricbrc.org/
Yaws (endemic treponematoses)	*Treponema pallidum* spp.	Spirochaetes	Two genomes available (*T. pallidum pallidum* [168], *T. pallidum pertenue* [169]). https://www.patricbrc.org/ [145].

[a]Current diseases recognized as 'neglected' by the WHO, their causative agents, and available genomic data and bioinformatic resources. Abbreviations: FWD, Filarial Worm Database, Broad Institute; WTSI, Wellcome Trust Sanger Institute.

Advances in our understanding of NTD pathogens have been difficult owing to several financial, technical, and biological obstacles, including a lack of appropriate *in vitro/in vivo* culture systems [6,7], the inherent inaccessibility or scarcity of some of the developmental forms of parasites [8], or the existence of highly structured populations (i.e., populations composed of multiple 'strains' or 'types' showing genome size variation and extensive karyotype polymorphism) within species displaying asexual, or rarely sexual, reproduction [9]. The past decade, however, has witnessed major achievements in NTD pathogen research fueled primarily by recent advances in high-throughput technologies and the introduction of powerful molecular, cellular, and genetic tools. In this review we appraise some of these achievements and discuss the impact that post-genomic developments may bring about in driving innovation and discoveries for much-needed intervention and surveillance tools such as drugs, diagnostics, molecular epidemiology, and vaccines for NTDs.

Genomic Features of NTD Pathogens

Recent advances in sequencing methods have facilitated the completion of genome sequences for several NTD pathogens, providing an invaluable resource to advance our understanding of their fundamental biology and, particularly, of their host–parasite interactions. These efforts were followed closely by re-sequencing and comparative studies which highlighted the evolutionary forces acting on these genomes and allowed the development of maps of single-nucleotide polymorphisms (SNPs) and gene copy-number variation (CNVs). Overall, genomic studies combined with functional genomics profiles of gene expression at both the transcriptome and proteome levels, as well as metabolomic data [10], are revealing variations in the molecular make-up of NTD pathogens. These variations help, in turn, to unravel the biological basis underlying the differential infectivity/pathogenicity of distinct developmental stages, 'types', and/or closely-related species. Table 1 summarizes the data available for NTD pathogens and provides links to mainstream bioinformatics resources holding these data.

A largely unifying evolutionary concept is that a parasitic (and/or endosymbiotic) lifestyle leads to a reduction of genome size either through neutral gene loss or, less likely, through adaptive

Panmixia (or panmixis): random mating. A panmictic population is one where all individuals are potential partners. This assumes that there are no genetic and/or physical mating restrictions.
Parasexual cycle/parasexuality: a nonsexual mechanism for transferring genetic material without meiosis or the development of sexual structures. Originally described in fungi, a parasexual cycle is initiated by fusion of cells and nuclei resulting in the formation of a diploid or polyploid nucleus, which is believed to be unstable and can produce segregants by recombination involving mitotic crossing-over and haploidization. Similarly to a sexual cycle, parasexuality gives the species an opportunity to recombine the genome and produce new genotypes in their offspring. Unlike a sexual cycle, the process lacks coordination and is exclusively mitotic. Parasexual cycles have been hypothesized to have been the source of natural hybrid lineages in *T. cruzi* [37,59].
Population structure/genetic structure: because of geographical (physical) barriers, natural populations rarely interbreed as in a panmictic (theoretical) model. As a result, genetic structuring appears, where groups of geographically closer individuals are more closely related to one another than to others randomly selected from the population.
Promastigote: the most common and replicative developmental stage of *Leishmania* found in the insect vector. It is characterized by a free anterior flagellum, which is not attached to the cell body, and by the presence of the kinetoplast (see above) at the anterior end of the body.
Quantitative trait loci: quantitative traits are phenotypes that display continuous variation and polygenic inheritance. The genes associated with these traits cannot be identified using classical Mendelian methods. Sophisticated statistical techniques have been developed to estimate in a population the most likely DNA locations (loci) that contain the genes that contribute toward the variation observed for the particular trait/characteristic or phenotype.
Synteny: shared gene order and orientation between orthologous chromosomes from different species.
Variant surface glycoprotein (VSG): a family of glycoproteins

genome streamlining [11,12]. The main idea underlying this 'reductive genome evolution' hypothesis is that adaptation to a highly-homeostatic environment results in a long-term trend towards reliance on the host (and/or the endosymbiont) for specific metabolic functions. Among NTD pathogens, this is particularly evident in the protozoa, where the genome sizes are often significantly larger in related free-living taxa (e.g., *Euglena*, ~1300 Mb) relative to their parasitic counterparts (e.g., *Trypanosoma* and *Leishmania*, ~30–60 Mb) [13]. A similar trend is observed in particular NTD bacteria in which adaptation to a new, isolated niche with a narrow ecological spectrum is usually accompanied by a reduction in genome size and ultimately loss of the ability to replicate outside a host. For instance, the spirochete *Treponema pallidum*, the etiological agent of syphilis and yaws, lacks genes related to several key metabolic pathways such as the tricarboxylic acid cycle, the electron transport system, and the synthesis of nucleotides, amino acids, and lipids [14]. Conversely, its genome encodes genes for multiple transport proteins, indicating that it relies heavily on scavenging the required compounds from its human host [15].

anchored to the outer leaflet of the *T. brucei* membrane through a glycosylphosphatidyl inositol moiety with highly variable sequences embedded at the N-terminus. Their mono-allelic and sequential expression on the parasite surface drives the process of 'antigenic variation' which is essential for survival in the bloodstream of mammalian hosts.

Within helminths, the parasitic flatworms (Table 1) also show clear signatures of genome reduction which mirror their morphological regression (i.e., loss of anatomical and light-sensory organs, pigmentation, and all free-living life cycle stages found in related, non-parasitic worms) [16–18]. From a metabolic standpoint, the impact of reductive evolutionary forces on flatworms is more patent in the case of lipids: these organisms lack the ability to synthesize fatty acids and cholesterol *de novo* but display several fatty acid transporters to scavenge essential fats directly from the host [16–18]. Recent comparative genomic studies in nematodes (round and filarial worms) also suggest genome reductions in parasitic (i.e., *Brugia malayi* and *Trichinella spiralis*) as compared to free-living species such as *Caenorhabditis elegans* and *Pristionichus pacificus* [19], although support for this trend awaits future genome projects.

Reductive evolutionary forces in parasites are often counterbalanced by expansion of gene families, particularly those involved in their interaction with the host(s), such as adhesion proteins that mediate host cell invasion, components of immune evasion mechanisms, and/or signaling molecules. Again, protozoan parasites provide an excellent case in point to illustrate this issue [13]. They display significant expansion and diversification of particular gene families, such as secreted rhoptry kinases in *Toxoplasma gondii* [20], red blood cell surface receptors in *Plasmodium* [21], and surface glycoconjugates and/or proteases in the trypanosomatids (trypanosomes and leishmanias) [22,23]. The fact that these expanded gene families are lineage-specific likely underscores the different modes of infectivity exploited by each taxon towards its corresponding insect and/or mammalian host(s). Along these lines, there is also strong bioinformatics evidence that Chlamydiae and Treponemes have undergone several gene duplication events resulting in the creation of several families of outer membrane and secreted proteins important for immune evasion and intracellular survival [14,24].

In addition, mining of complete genomes has unveiled several pathogen-specific genetic signatures that provide the opportunity for development of previously unexplored chemotherapeutic approaches. An *in silico* platform aimed at facilitating these searches has been developed that is specifically focused on NTDs [25], and several target prioritization strategies have been proposed for different NTD pathogens [26]. One particular source of novel and appealing targets is provided by biosynthetic pathways that display several key steps that differ from those described in mammalian hosts. For instance, one relevant and common feature in tapeworms and flukes is the merging of two crucial enzymatic functions for redox homeostasis into a single enzyme, thioredoxin glutathione reductase (TGR). Further studies highlighted the structural basis for the multi-enzymatic function of TGR and showed that these parasites can be readily killed by compounds targeting this molecule [27]. Differences in the detoxification systems between tapeworms/flukes and their mammalian hosts are being currently explored using chemoinformatics and high-throughput screens for drug design [28]. Along these same lines,

the identification of missing or divergent pathways involved in DNA repair and replication facilitated by analysis of the genome sequences of trypanosomatids [13] promises to provide novel drug targets [29].

The completion of several genome sequences of trypanosomatids sheds new light on the unusual organization of their chromosomes. In these parasites, coding genes are arranged in large, directional, gene clusters in which all genes are encoded on the same DNA strand, and there are only a few strand switches along each chromosome [13]. Although resembling the polycistronic operons of bacteria, regulation of gene expression in trypanosomatids is largely post-transcriptional, relying on *cis*-acting motifs located on mRNAs that, in concert with a variety of RNA-binding proteins, guide the fate of individual transcripts in a developmental stage-specific manner [30]. Therefore, in these parasites, the concept of 'RNA regulons' or 'RNA operons' [31] seems to be an appealing way to envisage post-transcriptional regulatory programs [30,32].

A striking feature of some nematode genomes is the presence of a complete bacterial endosymbiont genome. The most recent meta-analysis indicates that 47% of filarial nematodes belonging to the Onchocercidae family are infected with an α-proteobacteria from the genus *Wolbachia* [33]. These endosymbionts encode genes that provide essential metabolic and/or physiologic functions; indeed antibiotic cure of *Wolbachia* halts nematode growth, leading to apoptosis and eventually to death of the worm [34]. This makes *Wolbachia* itself an appealing target for the eradication of river blindness and lymphatic filariasis. Indeed, several *Wolbachia*-encoded biochemical processes, including lipid and lipoprotein biosynthesis, heme biosynthesis, and glycolysis, are currently being further evaluated as potential drug target candidates in high-throughput inhibitor screens [33]. From an evolutionary standpoint, the mutualistic interaction between filarial nematodes and *Wolbachia* is expected to have strengthened the trend of gene loss in the parasitic worm (see above). As shown, *Brugia malayi* lacks enzymes for the *de novo* synthesis of purines, riboflavin, FAD, and heme, all of which are encoded in the *Wolbachia* genome but are missing from related α-proteobacteria (i.e., *Rickettsia*) [33]. Genomic sequencing followed by complete metabolic reconstructions of additional filarial species uninfected by *Wolbachia* (in which ancestral *Wolbachia* infection was followed by secondary loss of the endosymbiont) such as *Loa loa* (the African eyeworm), *Onchocerca flexuosa*, and *Acanthocheilonema viteae* revealed that these worms also lack functional genes to replace most of these pathways [35,36]. Thus, despite the evidence that *Wolbachia*-encoded pathways provide metabolites needed by their filarial hosts, it is probable that the symbiotic role of *Wolbachia* in filarial nematodes either lies outside these pathways or involves more subtle examples of metabolic supplementation.

From Genetics to Genomics of Pathogen Populations

Genetic Exchange in NTD Pathogens

Classical genetic exchange is one way to increase genetic diversity in a population. This, in turn can impinge on the phenotypic diversity of parasites, affecting the appearance and expression of virulence factors, antigens, host-specificity molecules, and disease severity determinants. Current models suggest that the ability of sexual reproduction to promote genetic variation is important for lineage survival because it can promote adaptation to fluctuating environments while limiting the accumulation of deleterious alleles [37]. However, sexual reproduction is associated with increased cellular 'costs' and with the risk of genetic conflicts and/or the breakdown of well-adapted genetic combinations. In general, pathogens exploit a wide array of strategies to generate genetic and phenotypic diversity, such as sexual or **parasexual** reproduction (see Glossary) [37]. Despite this, a unifying theme surfaces among NTD pathogens: perhaps with the exception of metazoan helminths, which readily enter classic sexual reproduction, almost all these pathogens rely at some point in their life cycles on the generation of clonal populations that are well adapted to the host [9].

Studies of the genetic diversity of populations have revealed the main modes of reproduction utilized by pathogen species. A highly structured (e.g., clonal) population indicates that the main mode of reproduction for such a species lacks genetic exchange (is primarily asexual or sex occurs only rarely; see below for some examples); alternatively, and disregarding sampling issues, a highly structured population may be the consequence of geographical or ecological barriers [9]. Highly structured populations have been demonstrated for several viruses, bacteria, protozoa, and fungi [37], including *Mycobacterium tuberculosis* [38], *Mycobacterium leprae* [39], *Mycobacterium ulcerans* [40], most medically relevant trypanosomatids [9], and to a lesser extent *Toxoplasma gondii* [41]. The **population structure** of most metazoan NTDs and many other significant human pathogens (e.g., *Cryptosporidium* and *Pneumocystis carinii*) in general remains unknown [42].

Trypanosoma brucei (the agent of human African trypanosomiasis, HAT [43]) is a complex of three subspecies, two of which cause human disease (*T. brucei rhodesiense* and *T. brucei gambiense*). All are strict diploid organisms, with sexual recombination occurring in the salivary glands of tsetse fly vectors [8]. Sexual exchange proceeds under a classical Mendelian segregation of alleles, supporting both cross- and self-fertilization [44]. As a consequence, the parasite has the capacity for both clonal and sexual propagation, with varying degrees of inbreeding or outcrossing. Indeed, all these variants have been observed in the field, with *T. brucei gambiense* showing strict clonality [45], whereas different subpopulations of *T. brucei rhodesiense* display both clonality and epidemic or close to **panmictic** behavior [46]. Interestingly, the capacity of these two subspecies to cause human infections is due to independently evolved mechanisms of resistance to a lytic factor present in human serum [47,48]. Recent reports of **introgression** between *T. brucei* subspecies [49], and the demonstration of direct genetic recombination between human- and livestock-infective strains in coinfected tsetse flies [50], warn about the possibility that important traits, such as human serum resistance or drug resistance, may be transferred between parasite subspecies.

The two other trypanosomatids of medical importance, *Trypanosoma cruzi* (the agent of Chagas disease or American trypanosomiasis [43]) and the genus *Leishmania* (a complex of species causing various clinical forms of leishmaniasis [43]) show striking similarity in their unusual genetic features. Both originally appeared to be diploids with high karyotypic diversity [51,52], an observation that seemed to rule out the possibility of sexual exchange as observed in *T. brucei.* Indeed, pioneer molecular epidemiology studies reported strong linkage disequilibrium in several genes and/or genetic markers, leading to the proposal that these parasites were essentially clonal [53]. However, the existence of natural hybrids [54–58] suggested that, although infrequent, genetic exchange occurred in the wild. Later, genetic exchange was demonstrated in the laboratory for *Leishmania* and *T. cruzi*. In both cases, it was suggested to be mediated by a process that includes fusion of parental cells, homologous recombination, loss of alleles, and uniparental inheritance of **kinetoplast DNA** [54,59]. As for *T. brucei*, mating in *Leishmania* must be considered a non-obligatory part of the life cycle, because the majority of coinfected insect vectors failed to yield a hybrid but were still permissive to the development of mature, human-infectious forms [8,60].

Meiosis-specific genes have been identified in *Trypanosoma* and *Leishmania* genomes by phylogenomic analysis [13]. Recent studies of insect-derived parasites during the window of peak expression of these meiosis-specific genes allowed the identification of *T. brucei* 'gametes'; in other words, **promastigote**-like cells with haploid DNA content that interact with each other via their flagella and undergo cytoplasmic fusion [8]. Although these developmental forms have not yet been found in *Leishmania*, dynamic studies of the timing of hybrid formation within coinfected flies suggested nectomonad-related forms as the most likely sexually competent stage, with hybrids emerging well before the first appearance of metacyclic promastigotes, the

developmental forms that bring the infection into mammals [60]. The fact that these nectomonad-related forms are unique to *Leishmania* stage differentiation *in vivo*, and are not observed amongst the pleomorphic promastigotes that appear during *in vitro* growth, which have so far remained mating incompetent, further support this hypothesis [60]. Interestingly, while *Leishmania* hybrids were recovered at similar efficiencies in all pairwise crosses tested [60], recent observations of fly-derived stages of *T. brucei ex vivo* showed that cell fusion was rare unless the gametes were of different genotypes [61], indicative of mating-compatibility phenomena controlled by the haploid gametes. Although the underlying molecular basis remains to be addressed, this mating-compatibility mechanism may have been selected for to maximize the genetic variability of the species [37,62].

Historically, genetic diversity analyses were carried out using several informative genome markers, but genomic approaches that capture the genetic diversity of a species at the scale of **MLST (multi-locus sequence typing)** or the whole genome are increasingly being used for this purpose. Indeed, bacterial pathogen genomics has recently moved towards large-scale, population-based genome sequencing projects. In the case of *Chlamydia trachomatis*, the causative agent of Trachoma, there are now more than 100 completed genomes that in concert cover a large temporal and geographic diversity. Most importantly, these genotypes include *C. trachomatis* strains belonging to each of the three disease **biovars** [urogenital, lymphogranuloma venereum (LGV), and ocular] [63,64]. These phylogenomic studies confirmed the high similarity and almost identical **synteny** between *Chlamydia* genomes, with only a few thousand SNPs separating urogenital and LGV strains [63]. The most surprising finding was the large extent of apparent recombination between strains of *C. trachomatis*, either between virulent and non-virulent strains, or between biovars associated with different anatomical sites, a finding which has major implications for our understanding of chlamydial biology and epidemiology [63,64]. This pervasive capacity for genetic recombination has now been demonstrated *in vivo* and in the laboratory using mixed-infection models [65].

In the case of *Mycobacterium ulcerans* (responsible for Buruli ulcer), a whole-genome sequencing (WGS) approach, followed by a phylogenomic reconstruction of more than 35 isolates covering a wide geographical range and known genetic diversity, indicated that it constitutes a single evolutionary lineage whose divergence from *Mycobacterium marinum* (a free-living species) was characterized by acquisition of the pMUM plasmid [66]. This plasmid harbors three genes required for the synthesis of the macrocyclic polyketide mycolactones which play a key role in the pathogenesis and in the neuro-immunomodulatory properties of *M. ulcerans* [67].

Syphilis is one of the oldest recognized sexually transmitted infections and, despite the availability of inexpensive and effective therapy, the incidence is increasing in many parts of the world [68]. *T. pallidum* is a challenging infectious agent to study because of its inability to be cultured or genetically manipulated, its physical fragility, and its outbred animal model [6]. Despite these challenges, sequencing of the entire *T. pallidum* Nichols genome in 1998 [14] started an era of whole-genome analyses of pathogenic treponemes, which shed light on novel chromosomal targets for molecular diagnostics of treponemal infections and on treponemal evolution [69–71]. Indeed, these studies revealed striking similarity among isolates as well as a high degree of similarity to the rabbit pathogen, *Treponeme paraluiscuniculi*, which is not infectious to humans [70]. Further genome comparisons between *pallidum* and non-*pallidum* treponemes revealed genes with potential involvement in human infectivity [70]. Analysis of the genome also revealed that *T. pallidum* lacks genetic elements (e.g., plasmids, bacteriophages, and transposons) commonly associated with HGT mechanisms (i.e., transformation, transduction, and conjugation) that are a major means for acquiring antibiotic resistance. These features, coupled with the unique β-lactamase and penicillin-binding activities of the Tp47 protein [72], are perhaps the

explanation of why penicillin resistance has never developed in these organisms, even after sustained use of this antibiotic for more than 60 years [15].

Although it poses considerable challenges for interpreting the data in terms of population genetics and evolutionary genetics, the trend for sequencing complete genomes of clinical and field isolates is now being followed for most other viral [73,74] and bacterial NTD parasites. For those organisms with larger genome sizes such as protozoa and helminths, however, this is still a costly endeavor.

Non-Sexual Mechanisms of Generating Diversity and Variation in NTD Pathogens

In the absence (or near-absence) of genetic exchange between individuals, pathogens have evolved other strategies to generate variation. In Treponemes, for instance, molecular studies of the highly polymorphic *TprK* locus showed that new variants arise by segmental gene conversion, with the new sequences arising from a large repertoire of 'donor sites' located elsewhere on the chromosome [68]. The resulting changes in exposed variable regions of the *TprK* protein enable the organism to evade antibody binding and opsonophagocytosis, allowing *TprK* variant treponemes to survive clearance and persist during chronic latent infection [68]. Very similar mechanisms for creating genetic variability involving DNA recombination and repair have evolved in Trypanosomes. *T. brucei*, which dwells in the bloodstream of infected mammals, evades the host immune system by undergoing antigenic variation (i.e., periodically changing the expression of a group of **variant surface glycoproteins**, VSGs) [75]. In this parasite, a large repository of antigenically distinct 'donor' VSGs, which are scattered along silent sites in the genome, are moved by recombination into an active VSG expression site [75]. By contrast, the intracellular parasite *T. cruzi* does not use this strategy of antigenic variation for host immune evasion but instead relies on the simultaneous display of multiple polymorphic glycoproteins (e.g., mucins, *trans*-sialidases, mucin-associated proteins) on its surface during its mammal-dwelling stages [22,76,77]. The presence of a mosaic antigenic coat could modulate the mammalian host immune response by impairing the development of highly-specific lymphocytes and/or facilitate the recognition of the broad spectrum of cell types and mammalian hosts infected by this parasite [22,76]. Cumulative functional data (reviewed in [23]) and the recent finding of a significantly smaller number of gene copies encoding mucins, *trans*-sialidases, and mucin-associated proteins in the genome of a closely related species, the avirulent human-infective *Trypanosoma rangeli* further support a role for these molecules in *T. cruzi* pathogenesis [78]. As in *T. brucei*, the gene families encoding *T. cruzi* polymorphic glycoproteins constitute a huge genetic repertoire, which includes functional genes and pseudogenes, and can provide an additional pool of genetic variability for the generation of new glycoprotein variants through recombination and gene conversion mechanisms [13,79]. In this regard, it is worth noting that the vast majority of *T. cruzi* polymorphic glycoproteins are commonly found in unstable, subtelomeric chromosomal regions enriched in retrotransposon-like elements, and hence prone to frequent rearrangements [13].

Since the publication of a genome sequence for *Leishmania major* [80], complete genome sequences were also obtained for several Leishmanias, including 17 *L. donovani* lines derived from clinical samples [81,82]. Chromosomal analysis of pulse-field gel electrophoresis and WGS indicated that gene content and synteny were remarkably well conserved across the species, despite an estimated divergence time of 20–100 million years [80]. Despite this conservation, the sizes of homologous chromosomes were highly polymorphic; likely as the result of DNA amplification and deletion events, mainly in the subtelomeric repeat regions [83]. Strikingly, what appeared to be high karyotypic diversity turned out to be a constitutive and universal **mosaic aneuploidy**, which was previously suggested based on several lines of evidence collected in laboratory strains. This aneuploidy was definitively demonstrated using a single cell resolution technique performed on multiple isolates of at least four species of *Leishmania* [84,85]. Several independent observations using cloned isolates also indicated that mosaic

aneuploidy may be a shared feature with *T. cruzi* [86,87]. Moreover, small-scale studies [88,89] as well as comparative genomic hybridization of different *T. cruzi* strains have illustrated frequent variation in gene dosage as well as in chromosome copy numbers [90].

Aneuploidy is usually associated with severe abnormalities and decreased cell fitness because it leads to cell proliferation defects and to stoichiometric imbalances in essential proteins resulting from changes in gene dosage [91]. However, under specific circumstances some organisms appear to rely on aneuploidy for rapid adaptation to changing environments. Significantly, aneuploidy is predominantly found in clonal populations such as cancer cells and pathogenic fungi [92]. In the case of *Leishmania* and *T. cruzi*, DNA amplification and deletion events as well as aneuploidy can be also regarded as a gene dosage adjustment system in the absence of transcriptional regulation [30]. In terms of public health, the potential ability of trypanosomatids to acquire novel phenotypes (i.e., drug resistance in *Leishmania* isolates) by adjusting their chromosome copy numbers poses significant challenges for basic and applied research.

Among NTD pathogens, karyotype plasticity may not be restricted to trypanosomatids. Interestingly, genome sequences of several isolates of *Echinococcus multilocularis* revealed tetraploidy in one isolate, and transient trisomy of chromosome 9 (the smallest chromosome, and possibly the only one for which trisomy is tolerated) in different developmental stages from two different isolates [17], consistent with previous observations of karyotype plasticity in flatworms [93].

Another strategy to increase localized genetic variability is through the horizontal acquisition of foreign DNA sequences. Although this phenomenon is frequent in prokaryotes, it was thought to be rare among eukaryotes with sexual reproduction [94]. However, HGT into the genomes of nematodes is found at a much greater frequency than in other metazoans [19]. These transferred DNA sequences were acquired primarily from bacteria (particularly the endosymbiont *Wolbachia*), but also from fungi and amoebozoa, and may have contributed to the phenotypic plasticity and evolution of nematodes [94]. The possibility of HGT from different bacteria (i.e., *Chlamydiae*, *Salmonella*, and *Shigella*) to the *Trypanosoma/Leishmania* group has also been hypothesized based on genome-wide comparisons [95,96].

The Future Looks (Genome-) Wide

Although scientific and genetic breakthroughs for many NTD pathogens required discovery of the full life cycle and *in vivo* or *in vitro* culture (for an illustrative example, see Box 1), recent advances in DNA sequencing methods have contributed a wealth of information about NTD pathogens, particularly for those whose study is intractable by other means. Considering that genome sequencing capacity has increased while decreasing in cost, and single cell sequencing and genome amplification are possible, it is expected that these forward genetic approaches as well as functional genomics of gene expression will dominate the landscape of NTD pathogen research in the coming years. The integration of these genome-wide data, together with genetic manipulation of individual genes by means of highly-robust DNA and/or RNAi strategies when possible [97,98], will provide a global insight into the molecular architecture of the biology, pathogenesis, and host–parasite interactions of these organisms. In this section we discuss some of the genome-wide functional interrogations of genes that have already started to pay off for particular NTD parasites.

Genome-wide transposon mutagenesis in *Mycobacterium* has led to the identification of essential genes for growth *in vitro* [99,100] and *in vivo* [101]. In *T. brucei*, one of the few trypanosomes with an intact machinery for double-stranded (ds) RNA interference [102], a very elegant genome-scale RNAi knockdown screen allowed the identification of loss- and gain-of-fitness phenotypes in different life cycle stages and during differentiation [103,104]. This study,

Box 1. Building a Genetic System From Scratch: Cloning the Chloroquine Resistance Gene in the Malaria Parasite *Plasmodium falciparum*

The road from NTD to biological understanding, genetic system, and intervention is a long one that is far from over. All genetic systems start somewhere, and the malaria parasite had its first breakthrough with the discovery of its complex life cycle that involves both a mosquito vector, where sexual recombination takes place, and a vertebrate host in which asexual replication occurs. This breakthrough earned Ronald Ross a Nobel Prize in 1902 [170]. Once the life cycle was described, experimental *in vivo* systems, primarily in birds, rodents, and non-human primates, were established and, with time, some parasite strains from the field could be passaged in these animal models, and some eventually adapted to *in vitro* blood culture [171]. It is against this backdrop that experiments to identify the genetic basis of chloroquine (CQ) resistance in *Plasmodium falciparum* began. CQ was the most important drug in the battle against malaria. However, once resistance developed it spread quickly across the planet, erasing many of the gains made towards malaria eradication [172]. The origin of the resistance needed to be determined and new therapeutics and/or vaccines developed. The *Plasmodium* karyotype was determined to be 14 chromosomes, and genetic markers to differentiate strains were developed [173]. In 1987 a study reported a genetic cross in *Aedes freeborni* and in chimpanzees between *P. falciparum* strains showing many differences including pyrimethamine sensitivity [174]. This successful cross demonstrated Mendelian inheritance. It was followed in 1990 by a cross between CQ-resistant and CQ-sensitive parasites. Studies of the recombinant progeny revealed that a single genetic locus controlled the majority of the resistance phenotype [175]. Ten years of developing more refined maps and markers among the recombinant progeny narrowed down the location of CQ resistance [176–178]. Massive manual sequencing of the region and bioinformatics eventually identified the PfCRT (CQ resistance transporter) locus [179]. It was subsequently shown that mutation of this gene leads to the resistance phenotype [180]. This major accomplishment led to identification of a marker for CQ resistance [181] and numerous studies were undertaken to determine the origin and study how it swept around the globe [182]. The genetic and molecular genetic tools developed here and elsewhere, combined with advances in genomics, opened the door to major research on the *Plasmodium* genome, its global population structure [183] and, sadly, the mapping of additional resistance loci to antimalarials [184–187].

the first of its type, revealed that ~50% of the protein-coding genes in the genome are essential in at least one of the stages or conditions tested. The same technology was applied to identify novel determinants of human-serum sensitivity in *T. brucei* [105] and in a chemogenomic screen to identify genes associated to resistance or susceptibility to five drugs historically used for the treatment of HAT [106,107]. As in the former example, the genome-wide scale allowed the authors to identify a network of proteins involved in the mechanism of action of each drug.

In the case of transposon mutagenesis in *M. tuberculosis* it was clear that even essential genes could tolerate some transposon insertions (e.g., near the N- or C-termini, in linker regions between domains, etc.). Therefore new statistical methods for inferring gene essentiality from transposon insertions [108] had to be developed to take advantage of the rich data produced by the deep-sequencing experiments used to characterize the composition of mutant libraries [100]. A similar observation can be made regarding the *T. brucei* RNAi screens. All published experiments relying on deep sequencing of collections of transgenic parasites carrying different siRNA-producing constructs show that not all siRNAs directed against the same gene produce the same fitness phenotype. Therefore, inference of gene essentiality must be made after careful analysis of all available data [104]. It is expected that these types of experiments using deep sequencing to characterize complex collections of modified cells will continue to provide clues to the function of genes and their role in the biology of these important parasites.

RNAi in parasitic nematodes has been reported to result in transcript knockdown of some target genes, but not others, thus limiting its use as a potential functional genomics tool [98]. Recent work in *Haemonchus contortus*, a livestock parasite of worldwide significance and a suitable experimental model for the strongylid nematode group, however, suggests that this most likely reflects limited uptake of dsRNA from the environment, a phenomenon also observed in other free-living nematodes. Novel technologies to improve dsRNA delivery (such as nanoparticles) and methods to monitor RNAi effects are currently underway, and will be important in furthering the application of RNAi to identify essential gene function in parasitic nematodes [98].

Unfortunately, gene knockdown by RNAi in NTD protozoans is limited to *T. brucei*, *L. braziliensis*, and other New World *Leishmania* species. As a complementary approach, several transposon toolkits for use *in vivo* or *in vitro* have been developed for a variety of applications [109]. However, perhaps because of their relative low frequency of transposition, and the fact that in these mostly asexual diploid organisms a loss-of-function mutation requires two genetic events, no genome-wide transposon mutagenesis experiments have been successful thus far [109]. Novel and powerful tools for the analysis of gene function such as CRISPR/Cas9 have been recently established in *T. cruzi* [110] and other apicomplexan protist pathogens [111,112]. These studies are only the beginning of the exploitation of the CRISPR/Cas9 system in pathogens, and there are high hopes that this molecular toolkit can finally facilitate much-needed genome-wide phenotypic and chemogenomic screenings in these parasites.

Other WGS approaches, together with powerful mass spectrometry methods, are also gaining traction for strain typing in diagnostic and epidemiology applications [113]. In particular, these methods have proved to be particularly useful for typing recombinogenic bacteria, such as *C. trachomatis*, in which the use of diagnostic and/or typing schemes that rely on a single, or limited number of loci, compromised the ability to detect infections [63]. Most strains of *C. trachomatis* harbor a plasmid of ~7–8 kb which is highly conserved and essential for pathogenesis [114]. Because of their multi-copy nature, and available data suggesting that the plasmids have not been freely exchanged but have remained closely linked to their cognate bacterial chromosome [115], genes encoded in the plasmid have been preferred as markers for diagnostic purposes. Nevertheless, this did not prevent the emergence and spread of *C. trachomatis* variants that went undetected when diagnostic tests were used that relied only on molecular amplification of plasmid-encoded genes [63,116]. As a consequence, WGS is becoming the ultimate typing tool, allowing epidemiologists to obtain fine-grained snapshots in clinical settings in real-time [117,118]. To facilitate the task of analyzing large WGS datasets for the typing of bacterial isolates, new genome-wide MLST schemes and tools have recently been developed [119].

Several new post-genomic applications driven by the availability of genomic and transcriptomic data are starting to emerge. In the case of helminths, several alternative methods based on sequencing of miRNAs are changing the landscape of diagnostic applications for these diseases. Deep sequencing of RNAs isolated from the blood of hosts infected with filarial nematodes allowed the identification of parasite-specific miRNAs for the human pathogen *Onchocerca volvulus* [120]. A similar strategy may be used in the case of *Brugia malayi* now that several miRNAs have been characterized [121]. It is worth noting that the lack of a reliable typing method is a major limiting factor in the fight for global elimination of lymphatic filariasis [7].

Large-scale transcriptome/proteome information can also be utilized for high-throughput discovery of diagnostic biomarkers in NTD pathogens. Using *T. cruzi* as an example, we recently designed high-density tiling peptide chips spanning the full-length of ~500 parasite proteins from the deduced proteome [122]. These peptide arrays have been screened with sera from Chagasic patients (and healthy controls), and this led to the simultaneous identification and mapping of hundreds of novel parasite B cell antigens/epitopes. The same platform may be employed to explore other vacant areas in Chagas disease research, such as the identification of serology-based biomarkers suitable for the evaluation of drug efficacy, for disease prognosis, and for early diagnosis of congenital transmission. Moreover, immunological screening of peptides showing polymorphisms among *T. cruzi* 'types' is expected to improve the robustness of currently-available immunotyping methods [123,124]. Similar genome-wide strategies using alternative platforms/screening methods are being employed to identify immune signatures in other pathogens [125].

Host Genetics: The Neglected Side of the Host–Pathogen Interaction

Pathogen-related factors contributing to the outcome of infection have been widely recognized. However, host genetic factors are also a major determinant in susceptibility and resistance to disease, but historically these have only been identified serendipitously or searched for tepidly. Perhaps the best-known example is the classical textbook case of the HbS hemoglobin allele responsible for sickle cell anemia, which is maintained by balancing selection at ~10% in regions where *Plasmodium falciparum* is endemic because heterozygotes have a greatly reduced risk of severe malaria ([126] for review). In the case of other NTDs, only a few association studies in humans have been conducted. In one such study it was found that a specific haplotype (termed H-RISK) was associated with high risk of trichiasis and scarring trachoma, the leading cause of infectious preventable blindness worldwide [63]. This haplotype consists of four SNPs spanning upstream and downstream regions of the interleukin 10 (*IL10*) gene [127]. Recent studies have also shown different frequencies in genetic variants of the human gene coding for apolipoprotein-1 in Africa, which are thought to correlate with protection against *T. brucei* rhodesiense [128]. Other studies that highlight the role that protozoan parasites had in shaping the genetics of human populations were recently reviewed [129]. These studies suggest that **quantitative trait loci** and/or human polymorphisms play a role in genetic susceptibility to parasite burden and/or associated pathology for different pathogens of human [130–133] and livestock [134,135]. These types of investigations, however, require large cohorts of individuals (patients or animals), generate large sequence data sets, and usually lack causal conclusions. For these reasons the use of sets of genetically diverse advanced recombinant inbred mice have proved to be valuable tools in forward genetic studies to help relate host susceptibility, resistance, and disease severity phenotypes to specific genetic loci. Alternatively, and despite some limitations (such as compensating mechanisms to overcome the targeted deficiency and the bias in host factors specifically selected by the investigator), the use of knockout mice has also provided novel insights concerning the importance of immune-regulated cytokines (e.g., interferon γ), a variety of inflammatory proteins (e.g., matrix metalloproteases), chemokines (e.g., CXC motif chemokine receptors), and elements of the adaptive response that are important for resolving chlamydial infections [65]. For diseases such as malaria that have been removed from the NTD list, extensive studies are underway in both humans and non-human primates to better understand the role of the host and of host–pathogen interactions [126,136–138].

Classical genetic strategies relying on crosses between genetically defined mouse strains allowed the mapping of chromosomal regions involved in conferring resistance to infection to *T. brucei* [139] and *Leishmania tropica* [140]. The combined use of genetically modified mice strains and interleukin treatment demonstrated a redundant and essential role for nucleic acid-sensing Toll-like receptors TLR3, TLR7, and TLR9 in conferring resistance to *L. major* infection [141].

Overall, although there is strong evidence that both host and parasite polymorphisms have a role in the diversity of response to infection, much research is needed in this area. The identification of signature pathogen genetic variants, and of key host polymorphisms involved in susceptibility/resistance to infections (and their combinations), is likely to provide highly specific information for tailoring the application of therapies and interventions for effective elimination of NTDs.

Concluding Remarks

The number of diseases that are considered as 'neglected' has changed over time, reflecting to some extent the degree to which cell culture and genetic systems could be established and the level of attention and funding provided (for a discussion, see [1]). Thus, diseases that were historically considered to be 'neglected', such as malaria and tuberculosis, no longer carry this status because the combined funding for these diseases has grown significantly in recent years, in part because of our increased genetic understanding of the organisms. Public and private partners, including drug companies, donors, and governments, committed to the 2012 London

Box 2. Our Battle Against NTD Pathogens: The Road Ahead

We are 5 years away from the 2012 London Declaration to control, eliminate, or eradicate by 2020 12 of the 17 currently recognized NTDs. Sequencing technology breakthroughs have permitted tremendous advances and have allowed researchers to obtain genome sequences directly from field isolates of many NTD pathogens. This vast amount of information, in turn, is making possible more comprehensive and accurate approaches for drug target identification, the generation of diagnostic/therapy-efficacy markers, and the elucidation of parasite population structures. However, they do not solve all problems. For many species, a very large number of genes are hypothetical and tests of function are not possible in the absence of culture or powerful molecular genetic systems. Likewise, genome sequences may point to choke-points in metabolic processes but, in the absence of *in vivo* or *in vitro* assay systems, the testing and screening of potential therapeutic compounds is currently not possible.

We outline below some of the pending challenges that apply to most NTD pathogens.

- Establishment of appropriate *in vitro* culture systems or *in vivo* animal models (e.g., Treponemes). In this regard, the development of 'humanized' chimeric animal models may hold promise. Such a model has been recently applied to investigate liver-stage development and hypnozoites (dormant stages) in *Plasmodium vivax*, a parasite for which no easily tractable animal model is yet available [188].
- Establishment of powerful molecular genetic systems for difficult pathogens, including the application of CRISPR/Cas9 (e.g., for *T. cruzi* and helminths) and dsRNA interference (in some helminths).
- Insight into the role of genetic diversity (in both the host and pathogen) in the pathogenesis and ecoepidemiological dynamics of pathogens. Functional characterization and mechanistic elucidation of these genetic variations (SNPs, CNVs, or other genomic differences found in the population of NTD pathogens or in the host) represent a major challenge, particularly in the case of protozoa with a natural tendency towards chromosomal mosaic aneuploidy, such as Leishmanias and *T. cruzi*.
- Big data integration. The next big step forward for our understanding of disease biology will certainly come from the 'big data' revolution that is already underway. From population genomics, to high-resolution metabolic reconstructions of pathogen cells in different states, to the analysis of the host immunome in response to infection [122], these (and other) high-tech endeavors will have a significant impact on our understanding of the biology of NTD pathogens. For these data-analysis and integration efforts to succeed, a significant push is needed to attract and nurture new generations of bioinformaticians in the NTD field [189,190].
- Finally, we would like to stress the idea that an NTD is a dynamic concept that changes over time owing to multiple parameters. Several human-restricted and/or zoonotic diseases need the urgent support of the global community and would benefit from obtaining 'NTD status'. In particular, several invasive fungal infections such as Cryptococciasis and *Pneumocystis* pneumonia are rapidly emerging as a global threat to health, primarily owing to an increasing population of immunocompromised individuals and our dynamic interface with natural ecosystems [191].

Outstanding Questions

How important are both human and pathogen genetic diversity in determining the clinical outcome and development of disease? A better understanding of the host–pathogen interaction should provide a clearer strategy towards drug and vaccine design as well as other strategies aimed at pharmacological manipulation of host functions to avoid nonproductive and/or detrimental responses to infection.

What are the consequences, in terms of drug and/or vaccine development, of different genetic traits of NTD pathogens such as their adaptive genome streamlining and karyotype plasticity?

Unfortunately, many of the NTDs discussed here do not occur in isolation but in conjunction with a multitude of additional factors ranging from malnutrition to concurrent infections. How does the interplay of these factors facilitate and/or contribute to research and treatment of NTDs?

Declaration to control, eliminate, or eradicate by 2020 12 of the 17 currently recognized NTDs (lymphatic filariasis, blinding trachoma, soil-transmitted helminthiases, onchocerciasis, schistosomiasis, leprosy, dracunculiasis, yaws, visceral leishmaniasis, Chagas disease, and HAT). To achieve this goal, promises made were to ensure the supply of drugs, to enhance collaboration and coordination at national and international levels, to enable adequate funding, and to monitor R&D programs to lay a foundation on which to develop effective vaccines as well as to identify drug targets and diagnostic and epidemiology markers for treatment and control of NTDs (see Box 2 for a discussion of the challenges in meeting this goal). Still, many open questions remain (see Outstanding Questions).

Studies of NTD parasites have contributed substantially to the field of genetics. Research in pathogenic protozoa was instrumental for the elucidation of mechanisms leading to programmed DNA rearrangements [75], polycistronic transcription and *trans*-splicing to generate mature mRNAs [30], templated RNA-editing [142], and epigenetic control of gene expression [21,143]. Therefore, it is expected that the tools and resources that will emerge in our fight against NTD pathogens will continue to boost the discovery and the study of fundamental genetic processes.

Acknowledgments

We apologize to those authors whose work could not be cited owing to space restrictions. Research carried out in our laboratories is supported by grants and contracts from the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, Argentina) (to F.A. and C.A.B.), the Fundación Bunge y Born (Argentina) (to C.A.B.), and in part by the US National Institutes of Health (contracts HHSN272201400030C, HHSN272201200031C, and grant 1R03AI115339) and by the Ciência sem Fronteiras (CNPQ, Brazil; grant 400278/2014-6) to J.C.K. and collaborators. F.A. and C.A.B. are career investigators of the National Research Council of Argentina (CONICET).

References

1. Bhopal, A. *et al.* (2013) Strength in numbers? Grouping, fund allocation and coordination amongst the neglected tropical diseases. *J. Global Health* 3, 020302
2. Holmes, P. (2014) Neglected tropical diseases in the post-2015 health agenda. *Lancet* 383, 1803
3. Kappagoda, S. and Ioannidis, J.P.A. (2014) Prevention and control of neglected tropical diseases: overview of randomized trials, systematic reviews and meta-analyses. *Bull. World Health Organ.* 92, 356–366C
4. Eberhard, M.L. *et al.* (2014) The peculiar epidemiology of dracunculiasis in Chad. *Am. J. Trop. Med. Hyg.* 90, 61–70
5. Ramírez, J.D. *et al.* (2013) Molecular epidemiology of human oral Chagas disease outbreaks in Colombia. *PLoS Negl. Trop. Dis.* 7, e2041
6. Singh, S. *et al.* (2013) Axenic culture of fastidious and intracellular bacteria. *Trends Microbiol.* 21, 92–99
7. McNulty, S.N. *et al.* (2013) Inter and intra-specific diversity of parasites that cause lymphatic filariasis. *Infect. Genet. Evol.* 14, 137–146
8. Peacock, L. *et al.* (2014) Meiosis and haploid gametes in the pathogen *Trypanosoma brucei*. *Curr. Biol.* 24, 181–186
9. Tibayrenc, M. and Ayala, F.J. (2012) Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proc. Natl. Acad. Sci. U.S.A.* 109, E3305–E3313
10. Preidis, G.A. and Hotez, P.J. (2015) The newest 'omics' – metagenomics and metabolomics – enter the battle against the neglected tropical diseases. *PLoS Negl. Trop. Dis.* 9, e0003382
11. Wolf, Y.I. and Koonin, E.V. (2013) Genome reduction as the dominant mode of evolution. *Bioessays* 35, 829–837
12. Keeling, P.J. and Slamovits, C.H. (2005) Causes and effects of nuclear genome reduction. *Curr. Opin. Genet. Dev.* 15, 601–608
13. El-Sayed, N.M. *et al.* (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404–409
14. Fraser, C.M. *et al.* (1998) Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* 281, 375–388
15. Stamm, L.V. (2010) Global challenge of antibiotic-resistant *Treponema pallidum*. *Antimicrob. Agents Chemother.* 54, 583–589
16. Berriman, M. *et al.* (2009) The genome of the blood fluke *Schistosoma mansoni*. *Nature* 460, 352–358
17. Tsai, I.J. *et al.* (2013) The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* 496, 57–63
18. Zarowiecki, M. and Berriman, M. (2015) What helminth genomes have taught us about parasite evolution. *Parasitology* 142 (Suppl. 1), S85–S97
19. Sommer, R.J. and Streit, A. (2011) Comparative genetics and genomics of nematodes: genome structure, development, and lifestyle. *Annu. Rev. Genet.* 45, 1–20
20. Peixoto, L. *et al.* (2010) Resource integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. *Cell Host Microbe* 8, 208–218
21. Guizetti, J. and Scherf, A. (2013) Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell. Microbiol.* 15, 718–726
22. Buscaglia, C.A. *et al.* (2006) *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. *Nat. Rev. Microbiol.* 4, 229–236
23. De Pablos, L.M. and Osuna, A. (2012) Multigene families in *Trypanosoma cruzi* and their role in infectivity. *Infect. Immun.* 80, 2258–2264
24. Voigt, A. *et al.* (2012) The *Chlamydia psittaci* genome: a comparative analysis of intracellular pathogens. *PLoS ONE* 7, e35097
25. Agüero, F. *et al.* (2008) Genomic-scale prioritization of drug targets: the TDR Targets database. *Nat. Rev. Drug Discov.* 7, 900–907
26. Crowther, G.J. *et al.* (2010) Identification of attractive drug targets in neglected-disease pathogens using an *in silico* approach. *PLoS Negl. Trop. Dis.* 4, e804
27. Cioli, D. *et al.* (2008) Will new antischistosomal drugs finally emerge? *Trends Parasitol.* 24, 379–382
28. Gaba, S. *et al.* (2014) Cheminformatics models for inhibitors of *Schistosoma mansoni* thioredoxin glutathione reductase. *ScientificWorldJournal* 2014, 957107
29. Machado, C.R. *et al.* (2014) Nucleotide excision repair in *Trypanosoma brucei*: specialization of transcription-coupled repair due to multigenic transcription. *Mol. Microbiol.* 92, 756–776
30. De Gaudenzi, J.G. *et al.* (2011) Gene expression regulation in trypanosomatids. *Essays Biochem.* 51, 31–46
31. Keene, J.D. (2007) RNA regulons: coordination of post-transcriptional events. *Nat. Rev. Genet.* 8, 533–543
32. Najafabadi, H.S. *et al.* (2013) Global identification of conserved post-transcriptional regulatory programs in trypanosomatids. *Nucleic Acids Res.* 41, 8591–8600
33. Scott, A.L. *et al.* (2012) Filarial and *Wolbachia* genomics. *Parasite Immunol.* 34, 121–129
34. Landmann, F. *et al.* (2011) Anti-filarial activity of antibiotic therapy is due to extensive apoptosis after *Wolbachia* depletion from filarial nematodes. *PLoS Pathog.* 7, e1002351
35. Desjardins, C.A. *et al.* (2013) Genomics of *Loa loa*, a *Wolbachia*-free filarial parasite of humans. *Nat. Genet.* 45, 495–500
36. McNulty, S.N. *et al.* (2010) Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *PLoS ONE* 5, e11029
37. Heitman, J. (2010) Evolution of eukaryotic microbial pathogens via covert sexual reproduction. *Cell Host Microbe* 8, 86–99
38. Dos Vultos, T. *et al.* (2008) Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. *PLoS ONE* 3, e1538
39. Monot, M. *et al.* (2005) On the origin of leprosy. *Science* 308, 1040–1042
40. Röltgen, K. *et al.* (2012) The genome, evolution and diversity of *Mycobacterium ulcerans*. *Infect. Genet. Evol.* 12, 522–529
41. Sibley, L.D. and Ajioka, J.W. (2008) Population structure of *Toxoplasma gondii*: clonal expansion driven by infrequent recombination and selective sweeps. *Annu. Rev. Microbiol.* 62, 329–351
42. Tibayrenc, M. and Ayala, F.J. (2014) *Cryptosporidium*, *Giardia*, *Cryptococcus*, *Pneumocystis* genetic variability: cryptic biological species or clonal near-clades? *PLoS Pathog.* 10, e1003908
43. Stuart, K. *et al.* (2008) Kinetoplastids: related protozoan pathogens, different diseases. *J. Clin. Invest.* 118, 1301–1310
44. Peacock, L. *et al.* (2009) Intraclonal mating occurs during tsetse transmission of *Trypanosoma brucei*. *Parasit. Vectors* 2, 43
45. Koffi, M. *et al.* (2009) Population genetics of *Trypanosoma brucei* gambiense, the agent of sleeping sickness in Western Africa. *Proc. Natl. Acad. Sci. U.S.A.* 106, 209–214
46. Duffy, C.W. *et al.* (2013) Population genetics of *Trypanosoma brucei rhodesiense*: clonality and diversity within and between foci. *PLoS Negl. Trop. Dis.* 7, e2526
47. Capewell, P. *et al.* (2013) The TgsGP gene is essential for resistance to human serum in *Trypanosoma brucei gambiense*. *PLoS Pathog.* 9, e1003686
48. Uzureau, P. *et al.* (2013) Mechanism of *Trypanosoma brucei gambiense* resistance to human serum. *Nature* 501, 430–434
49. Goodhead, I. *et al.* (2013) Whole-genome sequencing of *Trypanosoma brucei* reveals introgression between subspecies that is associated with virulence. *Mbio* 4, e00197–e213
50. Gibson, W. *et al.* (2015) Genetic recombination between human and animal parasites creates novel strains of human pathogen. *PLoS Negl. Trop. Dis.* 9, e0003665
51. Engman, D.M. *et al.* (1987) *Trypanosoma cruzi* exhibits inter- and intra-strain heterogeneity in molecular karyotype and chromosomal gene location. *Mol. Biochem. Parasitol.* 22, 115–123
52. Bastien, P. *et al.* (1992) Molecular karyotype analysis in *Leishmania*. *Subcell. Biochem.* 18, 131–187
53. Tibayrenc, M. and Ayala, F.J. (2002) The clonal theory of parasitic protozoa: 12 years on. *Trends Parasitol.* 18, 405–410
54. Akopyants, N.S. *et al.* (2009) Demonstration of genetic exchange during cyclical development of *Leishmania* in the sand fly vector. *Science* 324, 265–268

55. Ravel, C. *et al.* (2006) First report of genetic hybrids between two very divergent *Leishmania* species: *Leishmania infantum* and *Leishmania major*. *Int. J. Parasitol.* 36, 1383–1388
56. Westenberger, S.J. *et al.* (2005) Two hybridization events define the population structure of *Trypanosoma cruzi*. *Genetics* 171, 527–543
57. De Freitas, J.M. *et al.* (2006) Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. *PLoS Pathog.* 2, e24
58. Rogers, M.B. *et al.* (2014) Genomic confirmation of hybridisation and recent inbreeding in a vector-isolated *Leishmania* population. *PLoS Genet.* 10, e1004092
59. Gaunt, M.W. *et al.* (2003) Mechanism of genetic exchange in American trypanosomes. *Nature* 421, 936–939
60. Inbar, E. *et al.* (2013) The mating competence of geographically diverse *Leishmania major* strains in their natural and unnatural sand fly vectors. *PLoS Genet.* 9, e1003672
61. Peacock, L. *et al.* (2014) Mating compatibility in the parasitic protist *Trypanosoma brucei*. *Parasit. Vectors* 7, 78
62. Perrin, N. (2012) What uses are mating types? The 'developmental switch' model. *Evolution* 66, 947–956
63. Harris, S.R. *et al.* (2012) Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat. Genet.* 44, 413–419
64. Joseph, S.J. *et al.* (2012) Population genomics of *Chlamydia trachomatis*: insights on drift, selection, recombination, and population structure. *Mol. Biol. Evol.* 29, 3933–3946
65. Abdelsamed, H. *et al.* (2013) Genetic variation in *Chlamydia trachomatis* and their hosts: impact on disease severity and tissue tropism. *Future Microbiol.* 8, 1129–1146
66. Doig, K.D. *et al.* (2012) On the origin of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *BMC Genomics* 13, 258
67. Marion, E. *et al.* (2014) Mycobacterial toxin induces analgesia in buruli ulcer by targeting the angiotensin pathways. *Cell* 157, 1565–1576
68. Ho, E.L. and Lukehart, S.A. (2011) Syphilis: using modern approaches to understand an old disease. *J. Clin. Invest.* 121, 4584–4592
69. Gray, R.R. *et al.* (2006) Molecular evolution of the tprC, D, I, K, G, and J genes in the pathogenic genus *Treponema*. *Mol. Biol. Evol.* 23, 2220–2233
70. Harper, K.N. *et al.* (2008) On the origin of the treponematoses: a phylogenetic approach. *PLoS Negl. Trop. Dis.* 2, e148
71. Knauf, S. *et al.* (2012) *Treponema* infection associated with genital ulceration in wild baboons. *Vet. Pathol.* 49, 292–303
72. Cha, J.Y. *et al.* (2004) A novel beta-lactamase activity from a penicillin-binding protein of *Treponema pallidum* and why syphilis is still treatable with penicillin. *J. Biol. Chem.* 279, 14917–14921
73. Talbi, C. *et al.* (2009) Evolutionary history and dynamics of dog rabies virus in western and central Africa. *J. Gen. Virol.* 90, 783–791
74. Parameswaran, P. *et al.* (2012) Genome-wide patterns of intra-human dengue virus diversity reveal associations with viral phylogenetic clade and interhost diversity. *J. Virol.* 86, 8546–8558
75. Deitsch, K.W. *et al.* (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat. Rev. Microbiol.* 7, 493–503
76. Dos Santos, S. *et al.* (2012) The MASP family of *Trypanosoma cruzi*: changes in gene expression and antigenic profile during the acute phase of experimental infection. *PLoS Negl. Trop. Dis.* 6, e1779
77. Kahn, S.J. *et al.* (1999) *Trypanosoma cruzi*: monoclonal antibodies to the surface glycoprotein superfamily differentiate subsets of the 85-kDa surface glycoproteins and confirm simultaneous expression of variant 85-kDa surface glycoproteins. *Exp. Parasitol.* 92, 48–56
78. Stoco, P.H. *et al.* (2014) Genome of the avirulent human-infective trypanosome – *Trypanosoma rangeli*. *PLoS Negl. Trop. Dis.* 8, e3176
79. Campo, V. *et al.* (2004) Differential accumulation of mutations localized in particular domains of the mucin genes expressed in the vertebrate host stage of *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* 133, 81–91
80. Ivens, A.C. *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436–442
81. Peacock, C.S. *et al.* (2007) Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat. Genet.* 39, 839–847
82. Downing, T. *et al.* (2011) Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. *Genome Res.* 21, 2143–2156
83. Rogers, M.B. *et al.* (2011) Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. *Genome Res.* 21, 2129–2142
84. Lachaud, L. *et al.* (2014) Constitutive mosaic aneuploidy is a unique genetic feature widespread in the *Leishmania* genus. *Microbes Infect.* 16, 61–66
85. Sterkers, Y. *et al.* (2011) FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in *Leishmania major*. *Cell. Microbiol.* 13, 274–283
86. McDaniel, J.P. and Dvorak, J.A. (1993) Identification, isolation, and characterization of naturally-occurring *Trypanosoma cruzi* variants. *Mol. Biochem. Parasitol.* 57, 213–222
87. Lima, F.M. *et al.* (2013) Interclonal variations in the molecular karyotype of *Trypanosoma cruzi*: chromosome rearrangements in a single cell-derived clone of the G strain. *PLoS ONE* 8, e63738
88. Urban, I. *et al.* (2011) Molecular diversity of the *Trypanosoma cruzi* TcSMUG family of mucin genes and proteins. *Biochem. J.* 438, 303–313
89. Burgos, J.M. *et al.* (2013) Differential distribution of genes encoding the virulence factor trans-sialidase along *Trypanosoma cruzi* Discrete typing units. *PLoS ONE* 8, e58967
90. Minning, T.A. *et al.* (2011) Widespread, focal copy number variations (CNV) and whole chromosome aneuploidies in *Trypanosoma cruzi* strains revealed by array comparative genomic hybridization. *BMC Genomics* 12, 139
91. Torres, E.M. *et al.* (2008) Aneuploidy: cells losing their balance. *Genetics* 179, 737–746
92. Kwon-Chung, K.J. and Chang, Y.C. (2012) Aneuploidy and drug resistance in pathogenic fungi. *PLoS Pathog.* 8, e1003022
93. Spakulová, M. *et al.* (2011) Cytogenetics and chromosomes of tapeworms (Platyhelminthes, Cestoda). *Adv. Parasitol.* 74, 177–230
94. Hirt, R.P. *et al.* (2014) Lateral gene transfers and the origins of the eukaryote proteome: a view from microbial parasites. *Curr. Opin. Microbiol.* 23C, 155–162
95. Gupta, R.S. and Griffiths, E. (2006) Chlamydiae-specific proteins and indels: novel tools for studies. *Trends Microbiol.* 14, 527–535
96. Silva, D.C.F. *et al.* (2013) Examining marginal sequence similarities between bacterial type III secretion system components and *Trypanosoma cruzi* surface proteins: horizontal gene transfer or convergent evolution? *Front. Genet.* 4, 143
97. Beverley, S.M. (2003) Protozomics: trypanosomatid parasite genetics comes of age. *Nat. Rev. Genet.* 4, 11–19
98. Britton, C. *et al.* (2012) Ups and downs of RNA interference in parasitic nematodes. *Exp. Parasitol.* 132, 56–61
99. Sassetti, C.M. *et al.* (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48, 77–84
100. Griffin, J.E. *et al.* (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* 7, e1002251
101. Rengarajan, J. *et al.* (2005) Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8327–8332
102. Kolev, N.G. *et al.* (2011) RNA interference in protozoan parasites: achievements and challenges. *Eukaryot. Cell* 10, 1156–1163
103. Alsford, S. *et al.* (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.* 21, 915–924
104. Glover, L. *et al.* (2015) Genome-scale RNAi screens for high-throughput phenotyping in bloodstream-form African trypanosomes. *Nat. Protoc.* 10, 106–133

105. Alsford, S. *et al.* (2014) Cathepsin-L can resist lysis by human serum in *Trypanosoma brucei brucei*. *PLoS Pathog.* 10, e1004130

106. Baker, N. *et al.* (2011) Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6. *Mol. Biochem. Parasitol.* 176, 55–57

107. Alsford, S. *et al.* (2012) High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature* 482, 232–236

108. DeJesus, M.A. *et al.* (2013) Bayesian analysis of gene essentiality based on sequencing of transposon insertion libraries. *Bioinformatics* 29, 695–703

109. Damasceno, J.D. *et al.* (2010) A transposon toolkit for gene transfer and mutagenesis in protozoan parasites. *Genetica* 138, 301–311

110. Peng, D. *et al.* (2014) CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *Mbio* 6, e02097–e2114

111. Wagner, J.C. *et al.* (2014) Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*. *Nat. Methods* 11, 915–918

112. Sidik, S.M. *et al.* (2014) Efficient genome engineering of *Toxoplasma gondii* using CRISPR/Cas9. *PLoS ONE* 9, e100450

113. Dubourg, G. *et al.* (2014) Culturomics and pyrosequencing evidence of the reduction in gut microbiota diversity in patients with broad-spectrum antibiotics. *Int. J. Antimicrob. Agents* 44, 117–124

114. O'Connell, C.M. *et al.* (2007) Plasmid-deficient *Chlamydia muridarum* fail to induce immune pathology and protect against oviduct disease. *J. Immunol.* 179, 4027–4034

115. Seth-Smith, H.M.B. *et al.* (2009) Co-evolution of genomes and plasmids within *Chlamydia trachomatis* and the emergence in Sweden of a new variant strain. *BMC Genomics* 10, 239

116. Unemo, M. *et al.* (2010) The Swedish new variant of *Chlamydia trachomatis*: genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* 156, 1394–1404

117. Reuter, S. *et al.* (2013) Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern. Med.* 173, 1397–1404

118. 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

119. Larsen, M.V. *et al.* (2012) Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361

120. Tritten, L. *et al.* (2014) Detection of circulating parasite-derived microRNAs in filarial infections. *PLoS Negl. Trop. Dis.* 8, e2971

121. Poole, C.B. *et al.* (2014) Diversity and expression of microRNAs in the filarial parasite, *Brugia malayi*. *PLoS ONE* 9, e96498

122. Carmona, S.J. *et al.* (2015) Towards high-throughput immunomics for infectious diseases: use of next-generation peptide microarrays for rapid discovery and mapping of antigenic determinants. *Mol. Cell. Proteomics* 14, 1871–1884

123. De Marchi, C.R. *et al.* (2011) Evaluation of a recombinant *Trypanosoma cruzi* mucin-like antigen for serodiagnosis of Chagas' disease. *Clin. Vaccine Immunol.* 18, 1850–1855

124. Risso, M.G. *et al.* (2011) Immunological identification of *Trypanosoma cruzi* lineages in human infection along the endemic area. *Am. J. Trop. Med. Hyg.* 84, 78–84

125. Legutki, J.B. *et al.* (2014) Scalable high-density peptide arrays for comprehensive health monitoring. *Nat. Commun.* 5, 4785

126. Kwiatkowski, D.P. (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* 77, 171–192

127. Natividad, A. *et al.* (2008) Susceptibility to sequelae of human ocular chlamydial infection associated with allelic variation in IL10 cis-regulation. *Hum. Mol. Genet.* 17, 323–329

128. Thomson, R. *et al.* (2014) Evolution of the primate trypanolytic factor APOL1. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2130–E2139

129. Pozzoli, U. *et al.* (2010) The role of protozoa-driven selection in shaping human genetic variability. *Trends Genet.* 26, 95–99

130. Frade, A.F. *et al.* (2013) Genetic susceptibility to Chagas disease cardiomyopathy: involvement of several genes of the innate immunity and chemokine-dependent migration pathways. *BMC Infect. Dis.* 13, 587

131. Williams-Blangero, S. *et al.* (2012) Host genetics and population structure effects on parasitic disease. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 367, 887–894

132. Quinnell, R.J. (2003) Genetics of susceptibility to human helminth infection. *Int. J. Parasitol.* 33, 1219–1231

133. Nogueira, L.G. *et al.* (2015) Functional IL18 polymorphism and susceptibility to chronic Chagas disease. *Cytokine* 73, 79–83

134. Skallerup, P. *et al.* (2012) Detection of a quantitative trait locus associated with resistance to *Ascaris suum* infection in pigs. *Int. J. Parasitol.* 42, 383–3891

135. McManus, C. *et al.* (2009) Genetic factors of sheep affecting gastrointestinal parasite infections in the Distrito Federal, Brazil. *Vet. Parasitol.* 166, 308–313

136. Auburn, S. *et al.* (2008) Association of the GNAS locus with severe malaria. *Hum. Genet.* 124, 499–506

137. Jallow, M. *et al.* (2009) Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nat. Genet.* 41, 657–665

138. Salinas, J.L. *et al.* (2014) Metabolomics in the fight against malaria. *Mem. Inst. Oswaldo Cruz* 109, 589–597

139. Sima, M. *et al.* (2011) Genetic control of resistance to *Trypanosoma brucei brucei* infection in mice. *PLoS Negl. Trop. Dis.* 5, e1173

140. Sohrabi, Y. *et al.* (2013) Mapping the genes for susceptibility and response to *Leishmania tropica* in mouse. *PLoS Negl. Trop. Dis.* 7, e2282

141. Schamber-Reis, B.L.F. *et al.* (2013) UNC93B1 and nucleic acid-sensing Toll-like receptors mediate host resistance to infection with *Leishmania major*. *J. Biol. Chem.* 288, 7127–7136

142. Simpson, L. *et al.* (2000) Evolution of RNA editing in trypanosome mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6986–6993

143. Borst, P. and Sabatini, R. (2008) Base J: discovery, biosynthesis, and possible functions. *Annu. Rev. Microbiol.* 62, 235–251

144. Stinear, T.P. *et al.* (2007) Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Res.* 17, 192–200

145. Wattam, A.R. *et al.* (2014) PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* 42, D581–D591

146. Franzén, O. *et al.* (2011) Shotgun sequencing analysis of *Trypanosoma cruzi* I Sylvio X10/1 and comparison with *T. cruzi* VI CL Brener. *PLoS Negl. Trop. Dis.* 5, e984

147. El-Sayed, N.M. *et al.* (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309, 409–415

148. Aslett, M. *et al.* (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* 38, D457–D462

149. Guzman, M.G. *et al.* (2015) Dengue: a continuing global threat. *Nat. Rev. Microbiol.* 8 (Suppl.), S7–S16

150. Gebhard, L.G. *et al.* (2011) Functional RNA elements in the dengue virus genome. *Viruses* 3, 1739–1756

151. Pickett, B.E. *et al.* (2012) ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* 40, D593–D598

152. Zheng, H. *et al.* (2013) The genome of the hydatid tapeworm *Echinococcus granulosus*. *Nat. Genet.* 45, 1168–1175

153. Cwiklinski, K. *et al.* (2015) The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol.* 16, 71

154. Berriman, M. *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416–422

155. Jackson, A.P. *et al.* (2010) The genome sequence of *Trypanosoma brucei gambiense*, causative agent of chronic human African trypanosomiasis. *PLoS Negl. Trop. Dis.* 4, e658

156. Raymond, F. *et al.* (2012) Genome sequencing of the lizard *parasite Leishmania tarentolae* reveals loss of genes associated

to the intracellular stage of human pathogenic species. *Nucleic Acids Res.* 40, 1131–1147

157. Real, F. *et al.* (2013) The genome sequence of *Leishmania* (*Leishmania*) *amazonensis*: functional annotation and extended analysis of gene models. *DNA Res.* 20, 567–581
158. Cole, S.T. *et al.* (2001) Massive gene decay in the leprosy bacillus. *Nature* 409, 1007–1011
159. Schuenemann, V.J. *et al.* (2013) Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* 341, 179–183
160. Ghedin, E. *et al.* (2007) Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* 317, 1756–1760
161. Tordo, N. *et al.* (1988) Completion of the rabies virus genome sequence determination: highly conserved domains among the L (polymerase) proteins of unsegmented negative-strand RNA viruses. *Virology* 165, 565–576
162. Young, N.D. *et al.* (2012) Whole-genome sequence of *Schistosoma haematobium*. *Nat. Genet.* 44, 221–225
163. Protasio, A.V. *et al.* (2012) A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. *PLoS Negl. Trop. Dis.* 6, e1455
164. Zerlotini, A. *et al.* (2013) SchistoDB: an updated genome resource for the three key schistosomes of humans. *Nucleic Acids Res.* 41, D728–D731
165. Martin, J. *et al.* (2012) Nematode.net update 2011: addition of data sets and tools featuring next-generation sequencing data. *Nucleic Acids Res.* 40, D720–D728
166. Thomson, N.R. *et al.* (2008) *Chlamydia trachomatis*: genome sequence analysis of lymphogranuloma venereum isolates. *Genome Res.* 18, 161–171
167. Seth-Smith, H.M.B. *et al.* (2013) Whole-genome sequences of *Chlamydia trachomatis* directly from clinical samples without culture. *Genome Res.* 23, 855–866
168. Carlton, J.M. *et al.* (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207–212
169. Pětrošová, H. *et al.* (2012) Whole genome sequence of *Treponema pallidum* ssp. *pallidum*, strain Mexico A, suggests recombination between yaws and syphilis strains. *PLoS Negl. Trop. Dis.* 6, e1832
170. Cox, F.E. (2010) History of the discovery of the malaria parasites and their vectors. *Parasit. Vectors* 3, 5
171. Trager, W. and Jensen, J.B. (1976) Human malaria parasites in continuous culture. *Science* 193, 673–675
172. WHO Malaria Action Programme (1988) Development of recommendations for the protection of short-stay travellers to malaria endemic areas: memorandum from two WHO Meetings. *Bull. World Health Organ.* 66, 177–196
173. Sanderson, A. *et al.* (1981) Enzyme typing of *Plasmodium falciparum* from African and some other Old World countries. *Trans. R. Soc. Trop. Med. Hyg.* 75, 263–267
174. Walliker, D. *et al.* (1987) Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* 236, 1661–1666
175. Wellems, T.E. *et al.* (1990) Chloroquine resistance not linked to mdr-like genes in a *Plasmodium falciparum* cross. *Nature* 345, 253–255
176. Su, X. *et al.* (1999) A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science* 286, 1351–1353
177. Su, X-Z. and Wellems, T.E. (1996) Toward a high-resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. *Genomics* 33, 430–444
178. Wellems, T.E. *et al.* (1999) Genome projects, genetic analysis, and the changing landscape of malaria research. *Curr. Opin. Microbiol.* 2, 415–419
179. Fidock, D.A. *et al.* (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol. Cell* 6, 861–871
180. Sidhu, A.B.S. *et al.* (2002) Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by pfcrt mutations. *Science* 298, 210–213
181. Djimdé, A. *et al.* (2001) A molecular marker for chloroquine-resistant falciparum malaria. *N. Engl. J. Med.* 344, 257–263
182. Wootton, J.C. *et al.* (2002) Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* 418, 320–323
183. Volkman, S.K. *et al.* (2007) A genome-wide map of diversity in *Plasmodium falciparum*. *Nat. Genet.* 39, 113–119
184. Miotto, O. *et al.* (2015) Genetic architecture of artemisinin-resistant *Plasmodium falciparum*. *Nat. Genet.* 47, 226–234
185. Roper, C. *et al.* (2004) Intercontinental spread of pyrimethamine-resistant malaria. *Science* 305, 1124
186. Maïga, O. *et al.* (2007) A shared Asian origin of the triple-mutant dhfr allele in *Plasmodium falciparum* from sites across Africa. *J. Infect. Dis.* 196, 165–172
187. Plowe, C.V. *et al.* (1997) Mutations in *Plasmodium falciparum* dihydrofolate reductase and dihydropteroate synthase and epidemiologic patterns of pyrimethamine–sulfadoxine use and resistance. *J. Infect. Dis.* 176, 1590–1596
188. Mikolajczak, S.A. *et al.* (2015) *Plasmodium vivax* liver stage development and hypnozoite persistence in human liver-chimeric mice. *Cell Host Microbe* 17, 526–535
189. Chang, J. (2015) Core services: reward bioinformaticians. *Nature* 520, 151–152
190. Savage, N. (2014) Bioinformatics: big data versus the big C. *Nature* 509, S66–S67
191. Armstrong-James, D. *et al.* (2014) A neglected epidemic: fungal infections in HIV/AIDS. *Trends Microbiol.* 22, 120–127