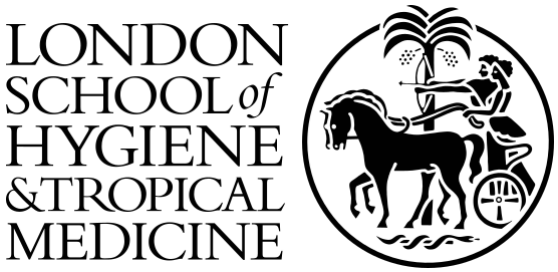


Diagnostics development for the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis

TEGWEN MARLAIS

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the University of London

July 2019

Department of Pathogen Molecular Biology

Faculty of Infectious and Tropical Diseases

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE,

UNIVERSITY OF LONDON

Funded by The Sir Halley Stewart Trust and the John Henry Memorial Fund

Research group affiliation:
Professor Michael A. Miles

Dedicated to Helen Fullerton

Declaration

I, Tegwen Marlais, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Tegwen Marlais

July 2019

Abstract

Strongyloidiasis, caused by the gut nematode *Strongyloides stercoralis*, and visceral leishmaniasis (VL), caused by the protozoa *Leishmania donovani* and *L. infantum*, are two potentially fatal parasitic diseases with wide global distribution and close association with poverty. Although both infections are treatable, it is imperative to validate cure after treatment. Available diagnostics for both infections have reasonable to high sensitivity for current infection but cannot easily distinguish cure. There is a need for diagnostic tests that are rapid, simple to use and deployable in field conditions to diagnose infection and validate cure.

This project aimed to identify candidate coproantigens of *S. stercoralis* and urine antigens of *L. donovani*, and to investigate the utility of IgG1 serology for determining cure versus relapse after treatment for VL. For *Strongyloides* coproantigen discovery, open access 'omic' data were analysed using computational tools. For *Leishmania* urine antigen discovery, antibodies were used to capture parasite antigens from VL urine, which were then identified by mass spectrometry. For *Leishmania* IgG1 serology, paired sera from pre- and post-treatment (cured) VL were compared with relapse sera in ELISA and with novel IgG1 specific rapid diagnostic tests (RDTs).

This work identified over 40 candidate coproantigens of *Strongyloides* that satisfied the required characteristics. Seven *L. donovani* proteins were identified in VL urine, within which 22 protein sequences were indicated as having high epitope potential and specificity to *L. donovani*. In VL serology, IgG1 was able to differentiate between cure and relapse of VL in both ELISA and RDT assays. With development and optimisation, the candidate antigens and IgG1 assays presented here have potential to contribute to disease control for these parasitic infections. The computational method of antigen selection used here for *S. stercoralis* can be applied to multiple parasitic helminth infections using the wealth of open access 'omic' data.

Acknowledgements – General

My sincere thanks to:

My parents who have encouraged and supported me to pursue my dreams, and all my family who made this PhD possible.

My husband Reinaldo for his belief in me, his patience and inspirational support.

All of the Miles Group whom I have had the joy and privilege to work alongside.

All the study participants whose samples have contributed to the quest for better diagnostics.

At LSHTM

Michael Miles, my supervisor, for his guidance, support and encouragement.

Tapan Bhattacharyya, for his constant support and guidance throughout my time at LSHTM.

Anna Last, for giving me the invaluable opportunity to join her fieldwork team in Guinea Bissau.

Olivia Farrant, for her outstanding organisation, teamwork and positivity during fieldwork in Guinea Bissau.

Safiyyah Marhoon, for assistance with using 3D protein modelling software.

Claire Rogers, for her invaluable support on all things parasitological. Specifically- for providing slides of *Strongyloides* gut biopsies and *Leishmania* splenic impression smears and for assisting me to take microscopic photographs of these. Also, for introducing me to Olivia Farrant, which ultimately led to my field work in Guinea Bissau, and for assisting with the preparations for field work and sample transport.

Stephanie Airs, for her excellent work on western blots during the project in India.

Marianne de la Roche, for unwavering kindness during our time in India.

Sue Passarelli, for helping me to prepare for doing Kato Katz stool microscopy in the field.

Debbie Nolder, for detailed advice on planning the *Strongyloides* qPCR of stool samples.

Hai Le, for a very thorough approach to primer design and persistence in PCR development.

Matthew Yeo, for his invaluable advice and support throughout my PhD process.

Richard Stabler, for assistance with setting up a custom blast database.

Jody Phelan, for assistance with annotating phylogenetic trees on iTOL.

Others providing support

The Sir Halley Stewart Trust, my funders, who not only made this work possible, but showed kindness and encouragement by their genuine interest in my progress. I truly appreciate being a recipient of their award which inspired me throughout the project.

The John Henry Memorial Fund for funding towards my PhD fees which enabled me to continue my studies.

Carlos Talavera-López for sharing his time and expertise so generously with me at the beginning of the *Strongyloides* work and for making me welcome at the Karolinska Institutet in Stockholm.

Lisette van Lieshout, Leiden University Medical Centre, for advice on PCR techniques for both strongyloidiasis and leishmaniasis, and for feedback and encouragement on my conference presentations.

Jürg Utzinger, Swiss TPH, for encouragement, discussion and collaboration on my *Strongyloides* work.

Björn Andersson, Karolinska Institutet. For taking time to understand my work and for his considered advice. Also for kindly hosting me in his group during two visits to the KI and for welcoming me at the group's Julbord (Christmas dinner).

Norbert Brattig, BNITM, Hamburg. For helpful discussions and encouragement as part of our collaboration.

James Cotton, WTSI, for useful discussions on helminth genomics and for providing a list of *S. ratti* unique genes.

Cornelis Hokke, Leiden University Medical Centre, for a very helpful discussion and advice on how to deal with glycosylation of my candidate protein antigens.

The British Society for Parasitology (BSP) for their very student-friendly events which led to many interesting discussions and excellent opportunities.

The International Society for Neglected Tropical Diseases (ISNTD) for its diverse and inspiring programme of events, especially to **Marianne Comparet** and **Kamran Rafiq** for their support.

The London Centre for Neglected Tropical Disease Research (LCNTDR) for the networking, presenting and funding opportunities throughout my PhD.

Contents

Additional items

A CD containing supplementary files for each of the enclosed publications is also included.

Abbreviations

aa	amino acid
AChE	acetylcholinesterase
AIDS	acquired immunodeficiency syndrome
BCG	Bacille Calmett-Guérin
BLAST	Basic logical alignment search tool
bwa	Burrows-Wheeler aligner
CAA	circulating anodic antigen
CAP	cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 protein superfamily
CC BY	Creative Commons Attribution licence
CCA	circulating cathodic antigen
CDS	coding sequence
CD	cluster of differentiation
CI	confidence interval
CrAg	crude antigen
CRS	composite reference standard
DAT	direct agglutination test
DDT	dichlorodiphenyltrichloroethane
DE	differentially expressed
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
ELISA	enzyme linked immunosorbent assay
E/S	Excreted/secreted
FL	free-living
HASPB	hydrophilic acylated surface protein B
HIV	human immunodeficiency virus
HTLV-1	human T-cell lymphotropic virus 1

IFAT	immunofluorescent antibody test
Ig	immunoglobulin
IGV	Integrative Genome Viewer
IL	interleukin
INDEL	insertion or deletion
INF	interferon
IRS	indoor residual spraying
KA	kala azar
LAM	lipoarabinomannan
LCA	latent class analysis
MDA	mass drug administration
MSA	multiple sequence alignment
LAMP	loop-mediated isothermal amplification
LC-MS/MS	liquid chromatography tandem mass spectrometry
LIPS	luciferase immunoprecipitation system
MHC	major histocompatibility complex
MSF	Médecins Sans Frontières
NCBI	National Centre for Biotechnology Information
NIE	Undefined meaning. A serological antigen of *Strongyloides stercoralis*.
nr	non-redundant
P	parasitological
PCR	polymerase chain reaction
PDB	Protein Data Bank
PFL	post free-living
PKDL	post kala-azar dermal leishmaniasis
PM	paromomycin
POP	prolyl oligopeptidase
PP	post parasitic
qPCR	quantitative (real-time) polymerase chain reaction

RDT	rapid diagnostic test
RNA	ribonucleic acid
RPA	recombinase polymerase assay
RSEM	RNA-Seq by Expectation Maximization
S	serology
SCP/TAPS	sperm-coating protein/Tpx-1/Ag5/PR-1/Sc7
SDS/PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
SRA	Sequence Read Archive
SSG	sodium stibogluconate
SsIR	*Strongyloides stercoralis* immunoreactive antigen
STH	soil transmitted helminths
TB	tuberculosis
TGF	transforming growth factor
T_H	T helper cell
TNF	tissue necrosis factor
TTL	transthyretin-like
UCSF	University of California, San Francisco
UK	United Kingdom
USA	United States of America
UPD	University of Pennsylvania, dog
VL	visceral leishmaniasis
WBPS	WormBase ParaSite
WHO	World Health Organization
w/o	without
WTSI	Wellcome Trust Sanger Institute

List of Figures

CHAPTER 1: Introduction, Literature Review, Aim, Objectives

1.1 General Introduction

1.1.1 Diagnostics development

Diagnosis is fundamental to appropriate treatment of individuals as well as to wider disease control for public health. Diagnostic methods for infectious agents, including parasites, can detect either the disease agent itself or the host immune response to it. In both cases, development of the test relies on knowledge of the target organism- morphologically, molecularly and ecologically.

Diagnostic methods detecting the disease agent itself include nucleic acid detection by polymerase chain reaction (PCR), quantitative PCR (qPCR), loop-mediated amplification (LAMP) or recombinase polymerase assay (RPA), and immunoassays capturing antigens shed by the organism. Diagnostics detecting the host immune response include serological immunoassays for anti-pathogen antibodies and assays for host immune factors such as cytokines. This thesis focuses on immunoassays targeting both parasite antigens and host antibody responses (Figure 1).

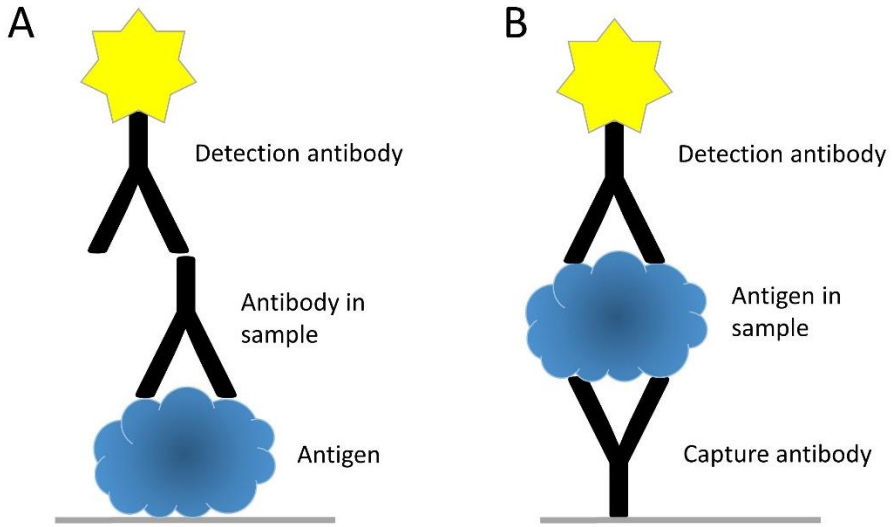


Figure 1. The different formats of immunoassays detecting (A) antibodies (serology) and (B) pathogen antigens in biological samples.

Immunoassays can be developed in several formats including: enzyme-linked immunosorbent assay (ELISA) which is ideal for screening large numbers of samples under laboratory conditions and can provide a quantitative result; rapid diagnostic tests (RDT)

which, with minimal sample preparation, can usually be carried out in the field in less than 30 minutes and most often provide a qualitative result. Additional formats include the luciferase immunoprecipitation system (LIPS), which involves expression of a synthetic protein comprising the antigen of interest fused with luciferase[1], and the Luminex platform, both of which provide high dynamic range quantification and assays that can be multiplexed. Technological developments in the field of diagnostics are making rapid test devices smaller, with cheaper materials such as paper and with the potential for multiplexing. In addition, there is a drive to associate diagnostic tests with mobile phone technology to quantify, record and share results[2-4].

Developing any immunoassay requires access to the antigen for optimisation of the assay, either immobilised on a surface in the case of serology, or to raise antibodies and assess their performance in the case of an antigen capture assay. Following initial development and determination of the lower limit of detection, well-characterised biological samples, both negative and positive for the target infection, are needed in order to assess background reactivity, cross-reactivity, and to determine positivity thresholds that give appropriate sensitivity and specificity.

Assays may be developed to detect a specific known antigen or an uncharacterised antigen or antigens. Equally, serology may detect antibodies against a known antigen, or an unprocessed mixture of antigens in a whole organism lysate for example. Both methods have advantages and disadvantages: a crude mixture contains a wide range of epitopes, potentially conferring higher sensitivity, however, this can also increase the chance of cross-reactivity of the assay with other parasite species. Availability of native parasite material must also be considered and can be a limitation to standardising and scaling-up production of an assay. Therefore, an optimal immunoassay uses a defined antigen/epitope which is both specific to the target parasite and can be optimised to give a highly sensitive assay.

In this thesis, these aspects of diagnostic development are considered in the context of the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis.

1.1.2 *Strongyloides stercoralis* and strongyloidiasis

Strongyloides stercoralis is a parasitic nematode (roundworm) primarily of humans and occasionally of dogs and great apes[5, 6]. It has a complex life cycle involving two reproductive

phases: parthenogenesis, the laying of unfertilized yet viable eggs, which occurs inside the host gut; and a facultative sexual phase in the environment which increases the number of infective larvae in the soil[7, 8]. Infection is acquired when infective stage larvae penetrate the skin. Inside the host, they migrate, coming to reside in the duodenum where the adult female digests host tissue and lays eggs which hatch before leaving the host, therefore larvae are passed in stool. Larvae may also reach infective stage prior to being excreted and will penetrate the host gut or perianal skin to restart the parasitic life cycle, termed autoinfection, which leads to the potential for lifelong infection. This unusual feature among human nematodes means that infection may persist for decades after a person has left an endemic area.

Symptoms of strongyloidiasis may be mild and non-specific. However, in cases of immunosuppression the autoinfective cycle may amplify to become hyperinfection, or dissemination if larvae spread to organs beyond their normal migration route, often accompanied by bacterial infection. Diagnosis of uncomplicated *S. stercoralis* infection can be made by: microscopy, using *Strongyloides*-specific methods to concentrate the larvae from a stool sample; DNA detection in stool; or serology.

1.1.3 *Leishmania donovani* and visceral leishmaniasis

Leishmania donovani is an intracellular protozoan parasite, transmitted by the phlebotomine sand fly. It causes the disease visceral leishmaniasis (VL) (also known as kala-azar (KA), meaning 'black fever' in Hindi) which occurs focally in parts of Africa and Asia. The closely-related species *L. infantum* is the disease agent in parts of Europe, the Middle East and South America. *L. infantum* and African *L. donovani* have animal reservoirs but in Asia the parasite is considered anthroponotic. Without successful treatment, VL is considered fatal. It has historically caused large epidemics in both India and eastern Africa which have had high fatality rates[9, 10].

In the vast majority of cases, VL is a disease of poverty due to poor housing conditions which allow proximity to sand flies, and poor nutrition[11, 12]. In turn, the productive time lost due to sickness and medical treatment place additional financial burden on affected families due to lost income and associated costs of medical care[11]. However, the first-line drug in India, liposomal amphotericin B (AmBisome), is donated free-of-charge to the World Health Organization by the manufacturer Gilead[13].

Disease control is based on control of the vector by indoor residual spraying of insecticide, early case detection and completed treatment as well as surveillance and health system capacity strengthening[14]. As yet there is no vaccine in use. Factors complicating the control of VL are the large proportion of infected people who remain asymptomatic[15], the possibility of disease relapse in inadequately treated individuals[16], post kala-azar dermal leishmaniasis (PKDL)[17] and the atypical presentation of VL in human immunodeficiency virus (HIV)-coinfected individuals[18].

1.2 Literature Review

1.2.1 *Strongyloides stercoralis*

Life cycle

Strongyloides stercoralis is a parasitic gut nematode (roundworm) of humans (Figure 2).

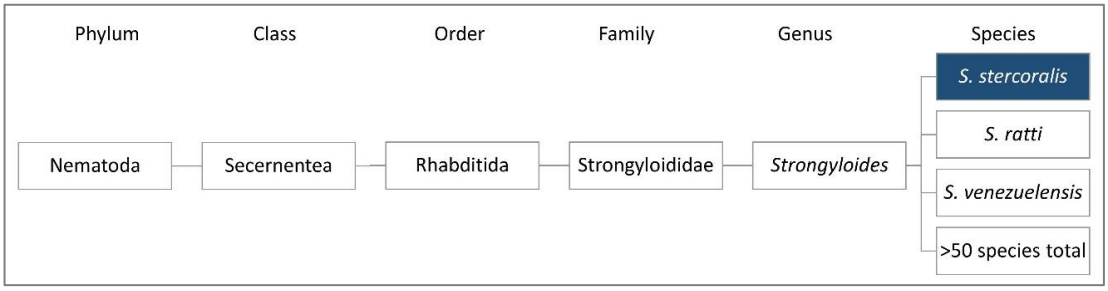


Figure 2. *Strongylodies stercoralis* taxonomy.

After penetrating the skin, infective stage (iL3) larvae progress through two further moults as they migrate inside the host via the bloodstream, penetrating into the lungs and ascending the airways to be swallowed into the digestive tract where they become parasitic adult females (Figure 3).

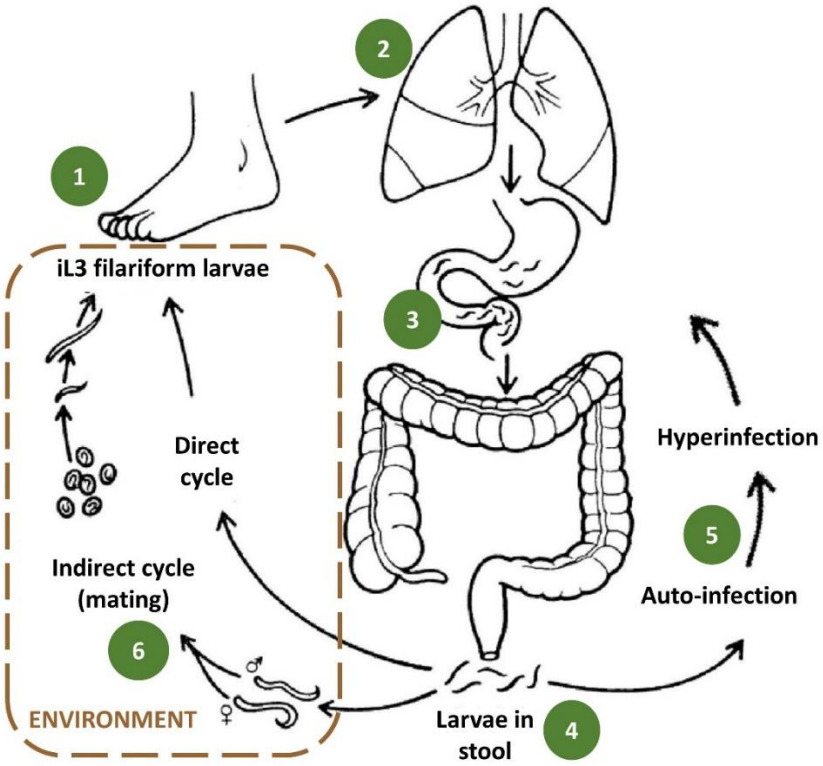


Figure 3. *Strongyloides stercoralis* life cycle. Adapted from Tefé-Silva *et al.* (2012)[19] under a Creative Commons license. (1) Infection is acquired when infective stage iL3 filariform larvae penetrate the skin. (2) Within the host, larvae migrate via the bloodstream to the airways where they ascend the trachea and are swallowed into the digestive system, during their migration they pass from iL3 to L4 larvae. (3) In the duodenum, adult parasitic female worms burrow into the epithelium where they digest host tissue and lay eggs by parthenogenesis. (4) Eggs hatch before leaving the host gut and larvae are passed in stool. (5) Some larvae become iL3 before leaving the host and penetrate the gut to re-commence their life cycle. If autoinfection is deregulated by host immunosuppression, it may result in hyperinfection. (6) Larvae that enter the environment may follow the indirect cycle of becoming adult male and female worms that undergo sexual reproduction and lay eggs which hatch through L1 to iL3, or alternatively, larvae in stool may develop directly to iL3 to infect a new host.

Only the adult female is parasitic. In the duodenum, it embeds itself in the epithelium, digesting host tissue and laying viable eggs by parthenogenesis, which hatch while still within the host gut (Figure 4). Larvae are passed in stool and may be at L1, L2 or L3/iL3 stages by this point. Larvae reaching iL3 may penetrate the gut or perianal area and continue development within the host or remain developmentally arrested in the environment until they are able to infect a new host. Alternatively, larvae ending up in the environment may continue to develop from L3, to L4 and into free-living dioecious adults

which undergo sexual reproduction. In the case of *S. stercoralis*, eggs laid following sexual reproduction mature into the developmentally arrested iL3 (Figure 3)[20].

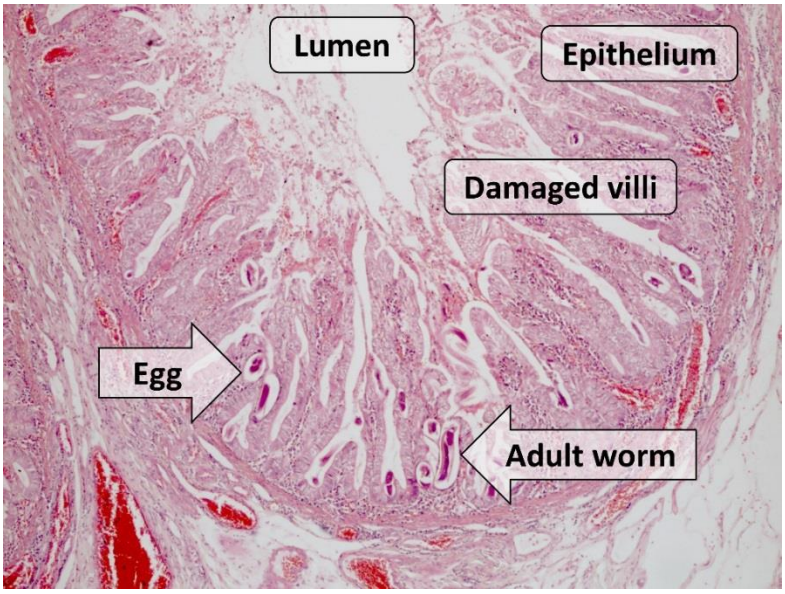


Figure 4. *Strongyloides stercoralis* parasitic adult females and eggs embedded in the epithelium of the duodenum where they cause damage to the microvilli. Photo by the author, slide of infected gut section kindly provided by Claire Rogers.

The unusual ability of *S. stercoralis* to complete its life cycle without leaving the host, termed 'autoinfection', means that once infected, a person may remain so for years or decades unless treated. This is reflected in the epidemiology where prevalence increases with age[21-23].

The free-living phase of the life cycle enables the parasite to amplify its numbers in the environment where there are humid and warm soil conditions. These conditions occur from tropical and sub-tropical to temperate regions throughout the world, enabling focal transmission[22, 24].

There are about 50 known species of *Strongyloides*, infecting a wide range of vertebrate hosts and with high host specificity[25]. Other species relevant to this work are *S. ratti*, a natural parasite of rats which is often used as a model for *S. stercoralis* due to the ability to maintain it in laboratory animals; *S. venezuelensis*, also a parasite of rats; *S. papillosus* which has a wider host range including sheep, cattle, goats and other ungulates; *Parastrongyloides trichosuri*, a parasite of certain possum species in Australia[26]. *S. stercoralis* has also been found in non-human primates and dogs, which may overlap with human infection[5, 27]. In addition to *S. stercoralis*, the species *S. fuelleborni fuelleborni* and *S. fuelleborni kelleyi* infect humans, the former is a potential zoonotic infection from non-

human primates[5, 6, 28] and the latter a severe infection in Papua New Guinea[29, 30] which has been proposed to form a separate species (i.e. *S. kelleyi*), rather than a subspecies[31].

Nematode anatomy has been well studied in *Caenorhabditis elegans*, the soil-dwelling model organism which is also in the order Rhabditida, along with *Strongyloides* and various bacteria-eating nematodes, both free-living and parasitic. Rhabditida are classified by their double-bulbed (rhabditoid) pharynx which in *Strongyloides* is a feature of stage 1 and 2 larvae (L1 and L2) whereas L3/iL3 larvae have a filariform (straight) pharynx (Figure 5). Features of the diagnostic stage larvae are the shallow buccal cavity (mouthparts) and tri-forked tail which distinguish them from hookworm larvae in stool microscopy (Figure 5).

S. stercoralis ranges from 0.5 to 2 mm in length at different stages of its life cycle, with L1 larvae being the smallest and parasitic females the longest. Free-living life stages are shorter and wider than their parasitic counterparts.

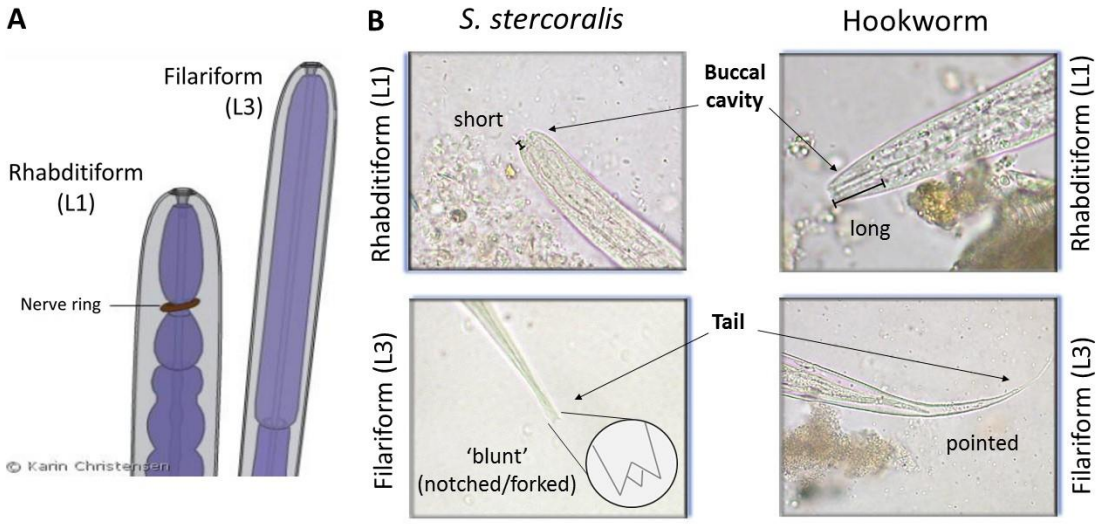


Figure 5. *Strongyloides stercoralis* morphology. A) Bulbed rhabditiform and straight filariform pharynx. Image copyright Karin Christensen, reproduced with permission. B) Diagnostic features to differentiate *S. stercoralis* from hookworm by microscopy: rhabditoid (L1) (top) and filariform (L3) (bottom) morphologies, showing the short buccal cavity and blunt tail of *S. stercoralis* (left), compared to the long buccal cavity and pointed tail of hookworm (right). Photographs from www.tropicalmed.eu[32] with permission.

Genomics

The genomes of 4 *Strongyloides* species were published in 2016[33]. These revealed that particular gene families had expanded in parasitic species when compared to non-parasitic Rhabditoid nematodes and therefore that certain proteins are closely linked to survival of the nematode in the host[34]. The currently available genomic data is based on analysis of *S. stercoralis* reference strain PV0001 (also known as PV001). This was derived from strain UPD (University of Pennsylvania, dog)[35] which originated in a naturally-infected dog from Pennsylvania, USA ([36] and Thomas Nolan, personal communication).

The closely related *S. stercoralis* and *S. ratti* each have 2 pairs of autosomes and XX/XO sex chromosomes where diploid males have only one X chromosome, and thus only 5 chromosomes in total (Figure 6). The ratio of male to female offspring of parthenogenesis is controlled by environmental factors so that more males are produced if the host immune response is strong[37]. This would favour environmental (indirect) development and sexual reproduction in order to maintain the species. Conversely, all progeny of free-living adults are female, which may arise from defective X-deficient sperm and mortality among male embryos, among other causes, so that all larvae become infective[37] (Figure 6). This genomic feature accounts for the limited number of free-living generations seen in *S. stercoralis* and the obligate parasitism compared with *P. trichosuri* which differs in life cycle and can maintain free-living generations indefinitely[26].

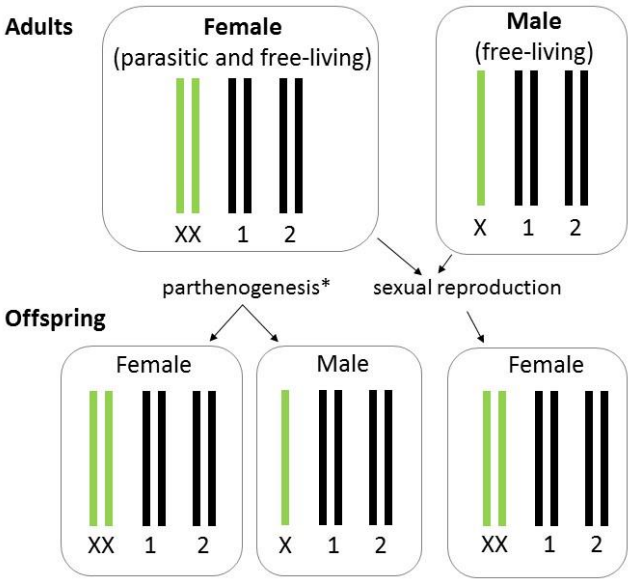


Figure 6. *Strongyloides stercoralis* and *S. ratti* chromosome arrangement where two autosomes are accompanied by two X chromosomes in females and a single X in males. Parasitic females produce both male and female offspring by parthenogenesis. Free-living

male and females reproduce sexually but all surviving offspring are female. *Parthenogenesis is not documented in free-living *S. stercoralis* or *S. ratti* but has been shown to occur in *S. planiceps*[38].

1.2.2 Strongyloidiasis

Clinical presentation

Uncomplicated strongyloidiasis is often quoted as being asymptomatic, possibly because symptoms are non-specific and because it is under-diagnosed in endemic regions. However, in a study by Becker *et al.* (2011)[21], frequent stomach ache was associated with *S. stercoralis* infection and a majority of infected respondents had wheezing, abdominal pain and general malaise. These findings concur with an earlier study of 192 patients presenting to the Hospital for Tropical Diseases in London, UK, where 71-78% of infected people had a strongyloidiasis-related symptom including abdominal or skin symptoms[39].

Hyperinfection occurs when the nematode replicates uncontrolled inside the host. As a result of the within-host migration, larvae can cause damage to lungs and gut as they pass through in large numbers[40, 41]. Dissemination of larvae may also occur into organs distant from the understood migration path including spinal fluid, kidneys and liver, among others[7, 42]. In addition, the compromising of internal membranes by larvae may cause bacterial infections, such as meningitis, which add additional complexity to the diagnosis and treatment of these urgent cases[42].

Epidemiology

Epidemiological data on strongyloidiasis is lacking, largely due to the fact that *Strongyloides* is not detected by the diagnostic methods routinely used for other soil transmitted helminth (STH) surveys. Estimates of global strongyloidiasis prevalence range from 30-100 million[22], to 370 million[43], which vary due to the lack of epidemiological studies using reliable and *Strongylodies*-specific diagnostic methods. This has led to the call for high quality epidemiological data to inform disease control and prevent potentially fatal hyperinfection[43, 44].

Regions with endemic transmission are those with the particular conditions of soil humidity and temperature, as well as unsanitary conditions, to support the nematode's life cycle. However, with the movement of people as refugees and migrants, there is evidence of

Strongyloides-infected people moving into non-endemic regions[45-48]. *S. stercoralis* has also been reported in a small number of travellers[49, 50]. As well as the humid tropics, the infection also occurs in more temperate areas and has been detected at high altitude in Bolivia[23], regions of Europe[24, 51] and the USA[52].

Risk factors for acquiring *S. stercoralis* infection are predominantly linked to poverty[53]. These include not wearing shoes[54], and open defecation. Latrine use and improved sanitation could reduce cases sustainably[55-57]. In addition, alcoholism and HIV are risk factors for strongyloidiasis. Once infected with *S. stercoralis*, risk factors for developing potentially fatal hyperinfection and dissemination syndrome are: immunosuppression, often by corticosteroid treatment of an unrelated, or presumed unrelated, condition[19]; solid organ transplant[58, 59]; co-infection with human T-cell lymphotropic virus (HTLV-1)[60]; alcoholism[61]; certain cancers. Hyperinfection and disseminated strongyloidiasis are fatal in up to 70% of cases due to the advanced state of illness by the time of correct diagnosis, inappropriate treatment and complications caused by underling health problems or bacterial co-infection[7, 42, 62-64].

Immune response

S. stercoralis infection, as other parasitic helminths, induces a T_H2 immune response. Various mechanisms are involved which together act to directly kill migrating larvae and the embedded parasitic adult worms, prevent attachment of the nematode to the gut by increasing mucus secretion, and promote smooth muscle contractility to expel the parasite[19]. However, the response is characterised by a regulatory predominance which enables worms to survive in the host and leads to interaction with other infections.

Eosinophilia (>500 eosinophils/µl of blood) is frequently a clinical finding that leads to diagnosis of asymptomatic strongyloidiasis[65-68] but which commonly resolves soon after successful treatment[69, 70]. Eosinophilia is mostly absent in hyperinfection and disseminated infection[7, 42]. These cells defend against pathogens that are too large to be phagocytosed by binding to the surface and targeting release of cytotoxic proteins[71]. Conversely, eosinophils may promote helminth survival by reducing the production of nitric oxide by macrophages and neutrophils[71, 72]. They may also initiate the anti-helminth antibody response which is

particularly IgG4 and IgE[73]. IgG4 dampens the immune response to frequently encountered immunogens, facilitating worm survival[74].

In the host, parasitic helminths produce excretory and secretory (E/S) products; a complex mixture of proteins, glycans, small molecules and vesicles[75-78]. Together, this material has direct interaction with the host immune system, inhibiting dendritic cells, promoting regulatory T-cells (T_{reg}) and stimulating the humoral response[79].

Along with T_H2 upregulation, the T_H1 response is suppressed during helminth infection. Studies on *S. stercoralis* co-infections indicate that this suppression can exacerbate the co-infection. Two of the most prevalent infectious causes of mortality worldwide, tuberculosis (TB) and malaria, share endemic regions and populations with *S. stercoralis*[80]. Both pathogens primarily stimulate a T_H1 immune response but the interaction between them and helminth co-infections is less clear-cut.

Malaria, caused by the intracellular parasite *Plasmodium* spp., has a complex immunological interaction with helminth co-infections and contradictory outcomes have been reported, including susceptibility to, and protection from, severe malaria[81, 82]. Immune interaction with malaria appears to vary by species of helminth. Specifically, fewer malaria episodes occurred in those with *Ascaris lumbricoides* infection[83]. However, many other factors may influence the interaction[82].

TB appears to have a more direct interaction with helminth infections as the two diseases exacerbate each other, including in the case of *S. stercoralis*[84]. In addition, helminth infections are known to reduce the efficacy of the Bacille Calmett-Guérin (BCG) vaccine against TB, an effect that may be reversed upon successful anthelminthic treatment[85].

As well as malaria and TB, *Trypanosoma cruzi* promotes a T_H1 immune response which is suppressed by *Strongyloides* co-infection. This was found in one study where *T. cruzi* was more likely to be detected in the blood of individuals who were also positive for *S. stercoralis* than those without the helminth[86]. Similarly, *Leishmania braziliensis* symptoms were more severe and more difficult to treat in nematode-infected individuals[87]. In addition to infectious interactions, helminth infection interacts with human fertility due to the T_H2 bias induced by both conditions, and different nematode species have been found to either enhance or decrease host fertility[88]. Following effective treatment, immunity to *S. stercoralis* is not complete and re-infection is commonplace[54].

Immunology of hyperinfection

In hyperinfection, the immune mechanisms to control the autoinfective cycle of the worm are weakened. This may be triggered by various factors, primarily Human T cell lymphotropic virus type 1 (HTLV-1) co-infection, corticosteroid medication and alcoholism.

HTLV-1 co-infection is a major risk factor for developing severe strongyloidiasis. The retrovirus infects T cells and induces their proliferation and secretion of T_H1 cytokines, particularly interferon gamma (IFN-γ). In fewer than 5%, HTLV-1 causes potentially fatal lymphoma. However, the potent T_H1 predominance in turn reduces T_H2 cytokine production, leading to loss of control of *S. stercoralis* in the co-infected host[60]. Despite their similarities, HIV co-infection does not promote hyperinfection as HTLV-1 does, although it is associated with increased *S. stercoralis* infection prevalence[22]. This is likely to be due to the T_H1 suppressive effect of HIV which preserves and even stimulates the T_H2 response, in contrast to the T_H1 enhancing/T_H2 inhibiting effect of HTLV-1[89].

Corticosteroid medication is increasingly used to control inflammatory disease in conditions including allergy and autoimmunity, organ transplant, and even symptoms caused by undiagnosed strongyloidiasis itself[42]. Corticosteroids act within cell nuclei, inhibiting gene expression, particularly of T_H2 cytokines[90, 91]. In addition, the drugs prevent recruitment and maturation of multiple immune cell types, having a broad inhibitory effect on both T_H1 and T_H2 cell populations[90].

Alcoholism predisposes to *Strongylodies* hyperinfection by its multiple interconnected impacts on an affected person. Alcohol intoxication activates endogenous corticosteroid production by the adrenal glands, and thus has a similar immunosuppressive effect to corticosteroid drugs. Additional effects of alcoholism are malnutrition, poor hygiene and decreased intestinal integrity, all of which facilitate the *S. stercoralis* autoinfective cycle[61].

Treatment

Ivermectin, an oral drug given in one or two doses of 200 mg/kg, is the first line treatment for strongyloidiasis. To treat hyperinfection or disseminated strongyloidiasis, ivermectin is given daily until the patient has been free of larvae in the stool for at least 2 weeks[92]. This ensures that all internally-migrating larvae are also killed.

Ivermectin, in the avermectin chemical group, is derived from the soil bacterium *Streptomyces avermictilis*. It binds to the glutamate-gated sodium channels of invertebrate nerve cells which leads to hyperpolarisation, paralysis and death[93] (Figure 7). Ivermectin is used on an individual patient basis and in research, against *S. stercoralis* but has been used more widely for mass drug administration (MDA) for the control of the nematode diseases, lymphatic filariasis and onchocerciasis[94]. In these areas, the MDA has also reduced the prevalence of strongyloidiasis[95, 96]. There is a plea to initiate MDA with ivermectin specifically for strongyloidiasis in high prevalence areas due to the potential for fatal hyperinfection[43], however, the drug remains unlicensed for human use in many countries, although it is used for livestock[44].

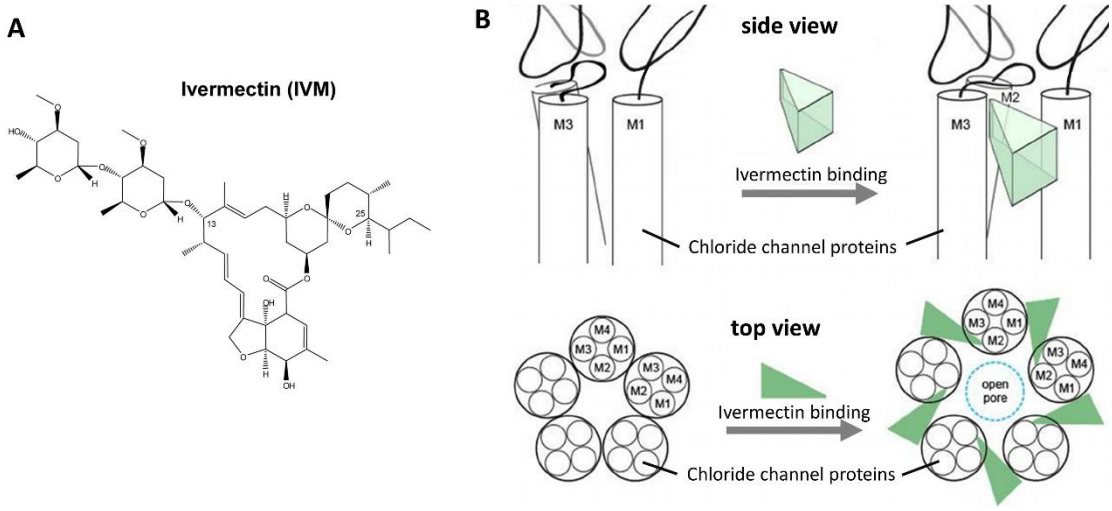


Figure 7. Ivermectin. A) molecular structure and B) binding in the invertebrate glutamate-gated chloride channel which irreversibly opens the channel. Image A from Ménez *et al.* (2012)[97], reproduced under a Creative Commons license, image B from Wolstenholme (2012)[98], copyright the authors, reproduced with permission.

While there are limited accounts of drug resistance of *Strongyloides* to ivermectin *in vivo*[99], it is known that the standard dosing regimen is not always sufficient to achieve complete clearance, which may indicate emerging resistance[70]. However, the poor sensitivity of stool-based diagnostics and lack of a timely test-of-cure preclude the identification of drug resistance as distinct from re-infection. The intensive use of ivermectin and other anthelminthics in agriculture has driven resistance in various livestock helminths, including *Strongylodies* species, suggesting that this may also occur in humans[100, 101]. Conversely, this may simply be a result of insufficient dosing and incomplete killing of unhatched and internally-migrating larvae.

To avoid the development of drug resistance to the avermectins, the combination of ivermectin and albendazole has been tested, which would also be beneficial for controlling the other STH[102, 103]. However, albendazole has lower efficacy than ivermectin against *S. stercoralis* therefore this combination would still rely on ivermectin to treat strongyloidiasis[104].

Other anthelminthics are two other 'azoles'; mebendazole and thiabendazole. The latter has equivalent efficacy to ivermectin against *S. stercoralis* but is not recommended due to frequent adverse reactions[104, 105]. The choice of drug must be made considering risk of *Loa loa* co-infection, where ivermectin could lead to fatal encephalopathy, and patients with neurocysticersosis, cystic *Taenia solium* infection, for whom albendazole may increase risk of complications.

Moxidectin, in the milbemycin class is chemically related to the avermectins and also binds to glutamate-gated chloride channels. Moxidectin has shown similar performance to ivermectin in a non-inferiority trial[106]. Therefore, this drug, commonly used for veterinary helminth treatment may be an alternative to ivermectin for human strongyloidiasis, pending further validation. In addition to existing anthelminthics, screening of neuromodulatory drugs has identified three promising candidates for repurposing as anthelminthics[107].

Diagnosis

There is no single gold standard or single reference test for strongyloidiasis but instead a combination is used. The methods that are in use are rarely standardised between laboratories, countries and research studies[108]. Definitive diagnosis is based on visualising larvae in a stool sample, however the commonly-used methods of stool microscopy for other STH, such as Kato Katz, are unsuitable for detecting *S. stercoralis* because larvae are scanty and do not survive well in solutions used to concentrate eggs. Therefore *S. stercoralis* specific methods are required.

Microscopy

Specific methods for microscopically visualising *S. stercoralis* exploit the natural mobility of the larvae and the free-living life stage. Samples are usually stool but in hyperinfection and

dissemination, larvae may be seen in sputum[109]. Stool microscopy techniques include the Koga agar plate culture, in which several grams of stool are incubated in the centre of an agar plate for up to 5 days, during which time tracks are made by migrating larvae and a 10% formalin solution is used to collect the larvae for detailed microscopic identification[110]. The second of the most widely used microscopic methods is the Baermann funnel; several grams of stool are suspended in a sieve in warm water and larvae that migrate into the water sink to a collection tube and are examined microscopically (Figure 8).

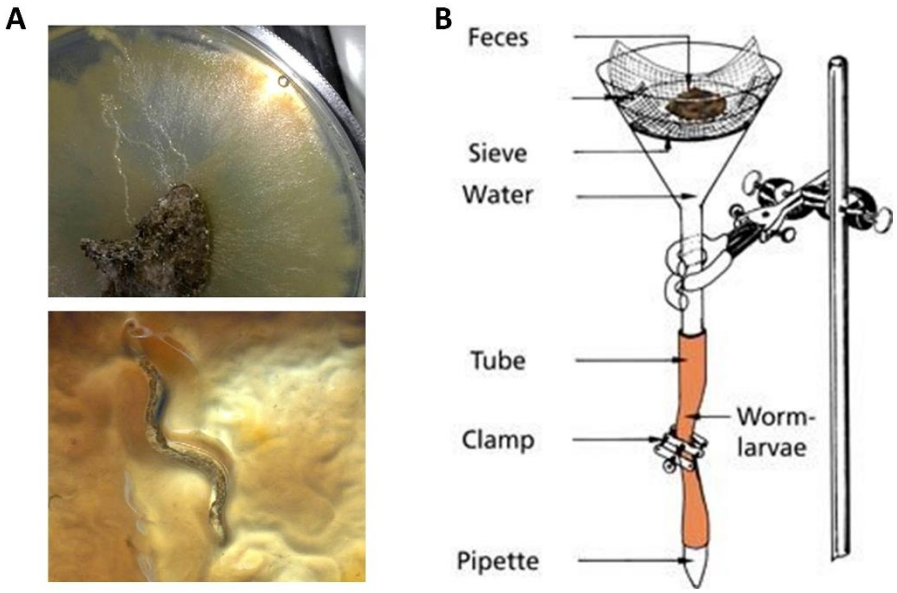


Figure 8. Stool-based definitive diagnostic methods for isolating *S. stercoralis* larvae. A) Koga agar plate, B) Baermann funnel. Images from Strkolcova *et al.* (2017)[111] and Encyclopedia of Parasitology[112], both with permission.

Sensitivity of microscopic techniques is limited due to low and irregular larval excretion but is increased when more than one stool sample is examined on consecutive days[113], or by taking multiple samples from the same stool[114]. Schar *et al.* (2014)[115] found that two consecutive samples increased sensitivity from 75% to 95% in low intensity infections (≤1 larva per gram of stool) (Table 1).

PCR and qPCR

Real-time polymerase chain reaction (qPCR) is used in high-income countries on DNA extracted from stool. The most widely used assay, developed by Verweij *et al.* (2009)[116] detects a region of the 18S rRNA gene. This assay has been widely used on DNA extracted from fresh, frozen or ethanol preserved stool samples with sensitivity higher than microscopy but highly dependent on the diagnostic reference method and whether this

includes serology (Table 1). Methods of stool preservation and DNA extraction vary between labs and studies and include DNA extracted by various commercial kits or in-house methods[117]. The DNA extraction method has a significant impact on the apparent sensitivity of qPCR[118]. Regarding specificity, molecular methods are superior to microscopy because hookworm larvae may occasionally be mistaken for *S. stercoralis*; one cause of microscopy positive, qPCR negative results[119]. As an alternative and more readily-available sample type, urine has been assayed for *S. stercoralis* DNA with sensitivity of 74.7%, making it worthy of additional study as a diagnostic option[120] (Table 1).

Serology

Serology for detection of anti *Strongyloides* IgG has the highest sensitivity for active infection of the existing diagnostic methods. The antigen used is not standardised and various *Strongyloides* species are used to produce in-house seroantigen, including *S. ratti* or *S. venezuelensis* from laboratory rat infections, or *S. stercoralis* from previous human or dog infections[121, 122]. A comparison of *S. stercoralis* and *S. ratti* antigen revealed equivalent sensitivity for human strongyloidiasis[123]. Therefore these assays are likely to be useful for increasing diagnostic sensitivity when used in conjunction with other tests. However, serological assays based on crude lysate antigens have been found to cross-react occasionally, with infections with trematodes *Opisthorchis* and *Fasciola*[123], and filarial nematodes such as *Mansonella*, *Loa* and *Onchocerca* but also *T. cruzi*[124]. Apparent false positives could also be due to otherwise undetected strongyloidiasis if other methods are not used to confirm the diagnosis[125]. Detection of IgG in urine has proven somewhat successful in two studies in Thailand where urine ELISAs using various antigen preparations were about 80% sensitive and 40%-55% specific, slightly lower than serum tested by the same assay but of potential use in initial population screening[126, 127].

In addition to high sensitivity, serology has the highest negative predictive value so that if a person is negative by serology, they are almost certainly not infected. However, there are instances where this is not the case and an infected person may present with negative serology and atypical symptoms/clinical findings[128], particularly newly-infected travellers, as opposed to chronically exposed individuals[39].

A *Strongylodies* specific seroantigen termed NIE was identified by Ravi *et al.* (2002)[129] using immuno-screening of a genomic library of *S. stercoralis* L3 larvae. The 31 kDa protein was later found to be related to insect venom allergens[130] but remains uncharacterised in *Strongyloides*. In LIPS assay format, NIE had 100% specificity and 97% sensitivity. An

additional recombinant protein antigen '*S. stercoralis* immunoreactive antigen' (SsIR) was incorporated into a LIPS assay along with NIE, increasing both sensitivity and specificity to 100%[131] (Table 1). The LIPS assay involves a synthetic fusion protein comprising the antigen of choice and the enzyme luciferase, expressed from a mammalian cell expression system[1]. The fusion protein is incubated with the serum sample, followed by incubation with protein A or G coated beads to capture IgG from the serum, beads are washed to remove all unbound material, a luciferase substrate is added and light output is measured. LIPS produces low background and a higher dynamic range than ELISA, giving clearer separation between positive and negative samples[131].

IgG4, an IgG antibody sub-class particularly stimulated by parasitic infections, showed improved specificity over total IgG in NIE assays[125], as did IgG1, IgG4 and IgE with *S. ratti* somatic antigen[122]. However, IgG4 had lower sensitivity than IgG in another NIE LIPS study[131]. Anti *Strongyloides* IgA has been detected in saliva, although this was not sufficiently sensitive to replace conventional IgG serology[132]. Another potential seroantigen, termed 'rSs1a', was recently identified using a similar genomic library screening strategy to NIE. It had 96% sensitivity and 93% specificity in initial testing[133]. Therefore, serology for strongyloidiasis is a growing field and is becoming widely recognised for its advantages.

Challenges with serology include defining seropositivity cut-offs, particularly in light of inadequate reference tests and non-standardised antigens[134]. Secondly, *S. stercoralis* serology has limited utility for monitoring treatment outcome and although a declining antibody titre is seen, this is measurable only several months after treatment[135]. This makes most serology unsuitable for monitoring treatment outcome in the very short term, prior to immunosuppression or organ transplant for example, but useful in epidemiological studies where a subsequent round of sampling can be performed[136]. However, this may not differentiate treatment failure from re-infection. In HIV co-infected individuals, serology is less accurate in the severely immunocompromised thus empiric treatment with ivermectin has been recommended where travel history and eosinophilia suggest infection[137].

Antigen detection

Assays detecting *Strongyloides* antigens are as yet limited to early development in research. Several assays have been developed against rat-infective *Strongyloides* species, the first of which, by Nageswaran *et al.* (1994)[138], used somatic adult and larval worm antigens, with the anti-adult worm assay tracking the progression of infection in rats.

However, as the authors concluded, a better defined antigen or monoclonal antibody would have increased sensitivity.

Goncalves *et al.* (2010)[139] used L3 larvae of *S. venezuelensis* to raise a polyclonal antibody which was used successfully to capture and detect antigen in infected rat faeces. This was possible from early in the infection and even when faecal egg output fell as the infection progressed and eggs were not detected microscopically (unlike *S. stercoralis*, *S. venezuelensis* eggs are seen in faeces).

Rabbit anti *S. ratti* E/S material was also successful when used by Sykes and McCarthy (2011)[140] on a *S. stercoralis* infected human stool sample and was negative with E/S material of certain other helminths. This highlights the advantage of using a polyclonal antibody, but the down side was cross-reactivity with rat gut components which had been carried over into the antigen preparation used to raise the antibody.

In addition to stool antigens, immune complexes of *Strongyloides* antigen with attached host antibody, have been detected in blood[141] and saliva[142] with limited sensitivity. However, the pilot coproantigen assays are proof of principle and show promising results but emphasise the need for known antigen targets and more specific antibodies.

Diagnostic sensitivities

Table 1 presents examples of the sensitivities and specificities of the various methods for diagnosing strongyloidiasis. In the absence of a single reference test, different methods are used as reference standard, which have a large effect on the apparent sensitivity. Specificity is not always calculated due to the difficulty in guaranteeing absence of infection. The recommended practice for diagnosing strongyloidiasis is a combination of serology and stool microscopy, confirmed with qPCR, dependent on the circumstances[117]. For epidemiological screening and diagnostic test evaluation, a composite reference standard or latent class analysis should be used to measure the sensitivity of the individual techniques and the true prevalence in a population[117, 143].

Table 1. Sensitivity of different techniques for diagnosing strongyloidiasis, illustrating the many variables and the absence of a single reference method. P, parasitological (includes any microscopic method); S, serology; CRS, composite reference standard (including various combinations of methods); w/o, without; LCA, latent class analysis; CrAg, crude antigen; IFAT, immunofluorescence antibody test.

Diagnostic method	**Sensitivity (%) (95% Confidence interval (CI))**		**Specificity (%)**	**Reference method**	**Reference**
Baermann funnel	86.3 (76.7-92.9) 2 samples		Not stated	CRS (w/o S)	Barda *et al.* (2018)[118]
	83.6 (71.2-92.2)		Presumed 100	CRS (w/o S)	Knopp *et al.* (2014)[144]
	43.6 (25.7-70.4)		97.9 (96.5-98.9)	LCA	Krolewiecki *et al.* (2018)[120]
	37.5 (24.9-51.5)		Not stated	CRS (w/o S)	Becker *et al.* (2015)[145]
Koga agar plate	21.4 (11.6-34.4)		Not stated	CRS (w/o S)	Becker *et al.* (2015)[145]
qPCR	77.5 (66.8-86.1)		Not stated	CRS (w/o S)	Barda *et al.* (2018)[118]
	76.8 (63.6-87.0)		Not stated	CRS (w/o S)	Becker *et al.* (2015)[145]
	66.2		Not stated	CRS (w/o S)	Kristanti *et al.* (2018)[146]
	64.4 (46.2-77.7)		93.9 (90.3-96.3)	P	Buonfrate *et al.* (2018)[117]
	56.5 (39.2-72.4)		95.4 (91.7-97.5)	P + S	Buonfrate *et al.* (2018)[117]
	31.0 (21.3-42.6)		Not stated	CRS (w/o S)	Barda *et al.* (2018)[118]
	30.9		96.2	CRS (w/o S)	Knopp *et al.* (2014)[144]
PCR	Urine	74.7 (53.8-91.8)	77.1 (71.7-83.7)	LCA	Krolewiecki *et al.* (2018)[120]
	Stool	71.8 (52.2-85.5)	93.5 (90.4-95.6)	P	Buonfrate *et al.* (2018)[117]

		61.8 (42.0-78.4)	95.3 (92.0-97.2)	P + S	Buonfrate *et al.* (2018)[117]

Table 1 continued.

Diagnostic method		**Sensitivity (%) (95% CI)**	**Specificity (%)**	**Reference method**	**Reference**
Serology	IFAT	94.6 (90.7-98.5)	87.4 (83.4-91.3)	CRS	Bisoffi *et al.* (2014)[124]
	NIE LIPS	97.8	100	P	Krolewiecki *et al.* (2010)[134]
		97	100	P	Ramanathan *et al.* (2008)[131]
		83.9 (77.5-90.1)	99.6 (98.9-100)	CRS	Bisoffi *et al.* (2014)[124]
	SsIR LIPS	97	97	P	Ramanathan *et al.* (2008)[131]
	NIE/SsIR LIPS	100	100	P	Ramanathan *et al.* (2008)[131]
		91.2	100	P	Krolewiecki *et al.* (2010)[134]
	NIE Luminex	93 (88-96)	95 (93-97)	P	Rascoe *et al.* (2015)[125]
	NIE IgG4 ELISA	95 (92-97)	93 (90-96)	P	Rascoe *et al.* (2015)[125]
	NIE ELISA	97	95	P	Ramanathan *et al.* (2008)[131]
		84	100	P	Krolewiecki *et al.* (2010)[134]
		76.7 (67.1-85.1)	71.6 (65.7-77.4)	LCA	Krolewiecki *et al.* (2018)[120]
		72.3 (61.8-80.8)	93.6 (88.2-96.6)	CRS	Fradejas *et al.* (2018)[147]
		70.8 (63.0-78.6)	91.1 (87.7-94.5)	CRS	Bisoffi *et al.* (2014)[124]
	CrAg ELISA (various)	37.5-100	73.4-100	various	Levenhagen and Costa-Cruz (2014)[148] (review)
		77.1% (serum) 80.0% (urine)	61.4% (serum) 55.1% (urine)	CRS	Eamudomkarn *et al.* (2018)[126]
		89.2 (80.7-94.2)	89.3 (83.1-93.4)	CRS	Fradejas *et al.* (2018)[147]

1.2.3 *Leishmania donovani*

The Leishmania donovani complex

There are approximately 20 species of *Leishmania* that infect humans and cause a wide range of clinical manfestations including cutaneous lesions, mucosal destruction and visceral disease, the last is addressed by this thesis and is caused by members of the *Leishmania donovani* complex (Figure 9). This includes two species: *L. donovani* and *L. infantum*. However, they have different phlebotomine sand fly vector species and their geographic distribution differs in that *L. donovani* is present in Africa and Asia while *L. infantum* occurs in the Middle East, Europe and South America.

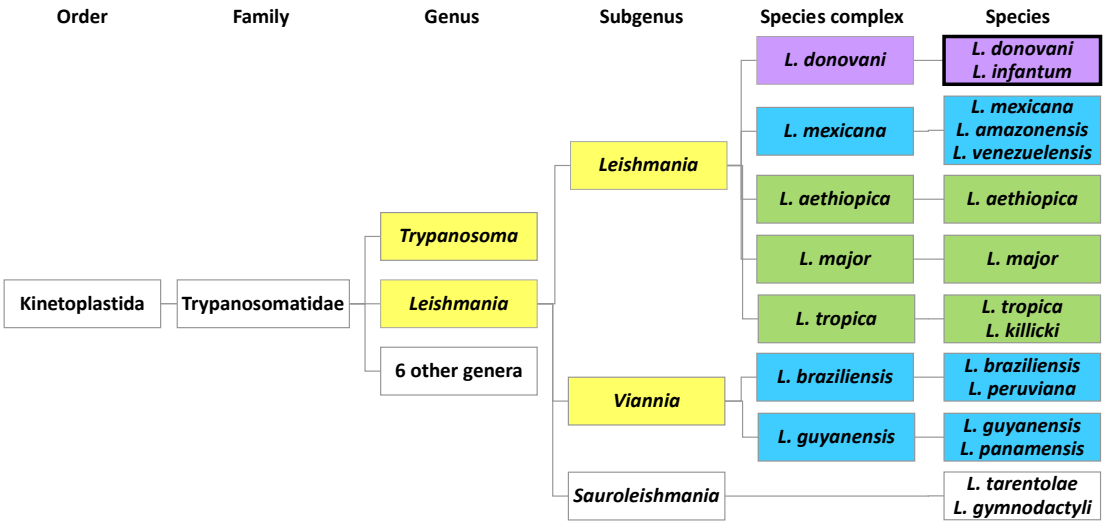


Figure 9. Taxonomy of *Leishmania*. Yellow, genera and subgenera which contain human-infective species; purple, species causing visceral leishmaniasis; green, species causing cutaneous leishmaniasis outside of the Americas; blue, species causing cutaneous or mucocutaneous leishmaniasis in the Americas. Box with thick outline indicates the species addressed in this report. Figure made by the author, based on information from Real *et al.* (2013)[149].

Leishmania was identified as a parasite and the causative agent of 'Dum Dum fever' by Major William Boog Leishman[150] and separately, Captain Charles Donovan[151] in 1903, in Indian patients and British miltary personnel formerly stationed in Dum Dum (near Calcutta), India. The following year, the same parasite was confirmed in humans in Sudan[152]. Genetic analysis of *Leishmania* has indicated separation between the different populations of *L. donovani* in Africa and Asia[153]. However, molecular analysis proved that *L. infantum*, formerly known as *L. chagasi*, in South America orignated from Europe in recent

history, possibly making the sea voyage several times in infected dogs on board the ships of European settlers[154].

Life cycle

Infection is initiated by a bite from an infectious female phlebotomine sand fly, during which the highly motile promastigote form of the parasite is injected into the host skin where it enters a macrophage (Figure 10).

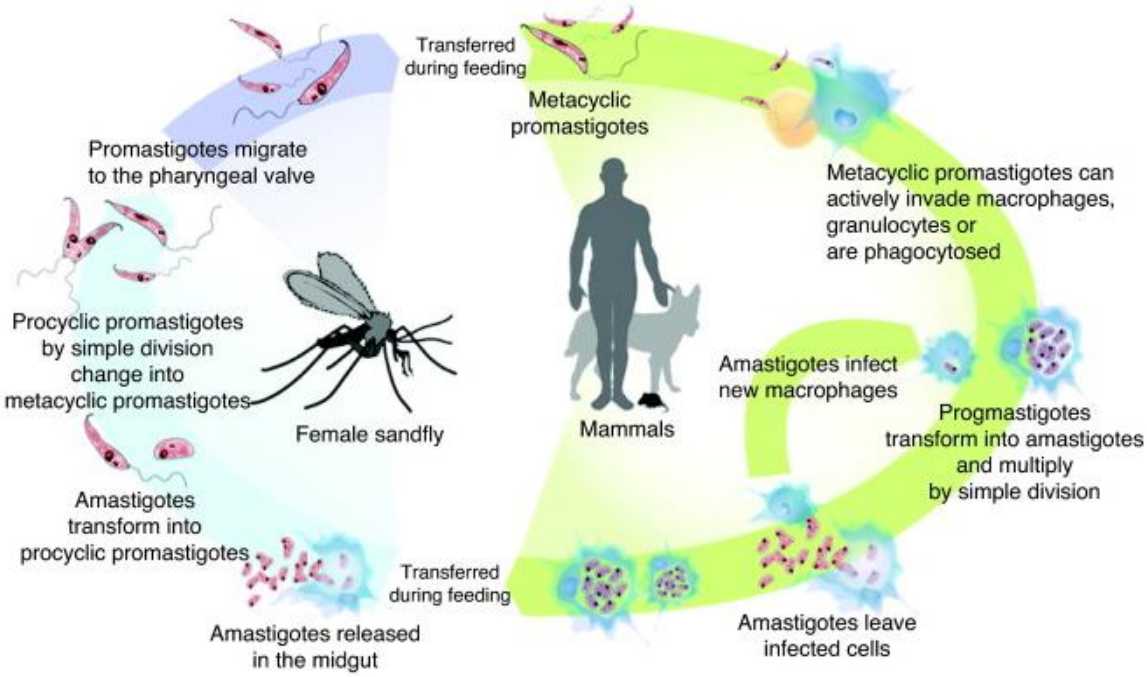


Figure 10. Life cycle of *Leishmania* spp. as motile promastigotes in the sand fly vector and intracellular amastigotes in the mammalian host. Image from Harhay *et al.* (2011)[155], reproduced with permission.

The parasite evades destruction in the macrophage by inhibiting cell signalling pathways which would otherwise trigger its destruction, turning a phagosome into a parasitophorous vacuole. In this parasitophorous vacuole, *Leishmania* transforms into the non-flagellated amastigote which replicates by binary fission[156] (Figure 11). Among other mechanisms, *Leishmania* evades macrophage killing by release of its surface-bound metalloprotease, GP63[156].

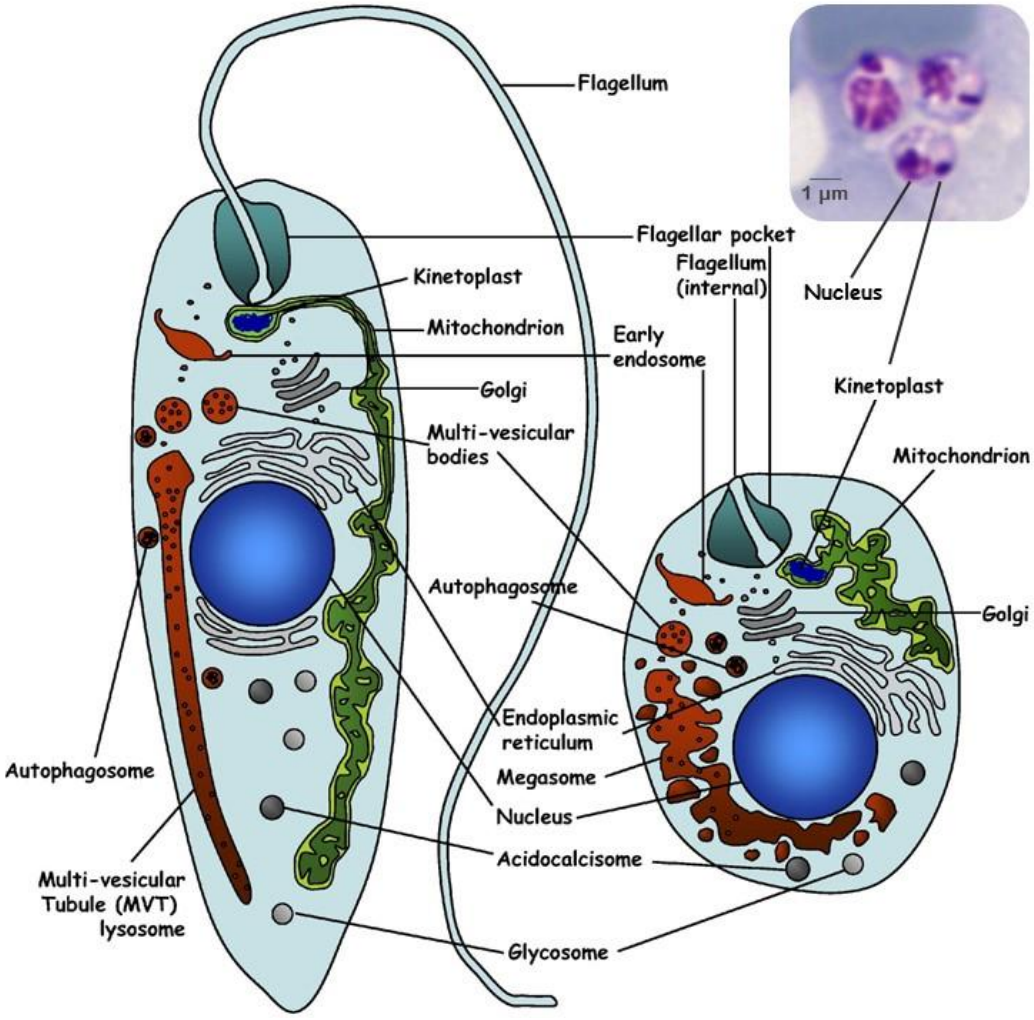


Figure 11. Diagram of a *Leishmania* cell in promastigote form, left, and amastigote form, right. When appropriately stained, the nucleus and kinetoplast DNA (dark blue) form characteristic diagnostic features of a large diffuse spot and a small, dense bar, inset. Diagram from Besteiro *et al.* (2007)[157], reproduced under a Creative Commons license. Inset photo by the author from a Giemsa stained impression smear kindly provided by Claire Rogers.

1.2.4 Visceral leishmaniasis

Pathology and immunology

The majority of *L. donovani* and *L. infantum* infections are asymptomatic. Often defined as seropositivity, asymptomatic individuals outnumber VL cases by 4-18 to 1[158-160]. However, circulating parasites and parasite DNA have also been recovered from asymptomatic individuals, including from skin microbiopsies, indicating the potential of asymptomatic individuals to transmit to sand flies[161-165]. Of those that develop disease, the clinical

features are fever of long duration (>2 weeks; often weeks or months if untreated), enlargement of the spleen and/or liver (spenomegaly/ hepatomegaly), wasting and weakness, and in advanced cases, skin hyperpigmentation and immune compromise leading to bacterial co-infections[166].

As an intracellular pathogen, an effective reponse to *Leishmania* is the initiation of a T_H1 inflammatory response which destroys the parasite via nitric oxide (Figure 12)[167]. Such an immune response leads to immune memory so that the majority of cured individuals will not be re-infected, but inadequate parasite clearance and weaker immunity can lead to relapse of VL. In the case of a dominant T_H2 humoral response this leads to hypergammaglobulimemia, a common characteristic of VL, which can result in immune complex damage to the kidneys and is ineffective at overcoming the infection[168]. The negative impact of a strong serological response is supported by the observation that recent, strong seroconversion is associated with disease[169].

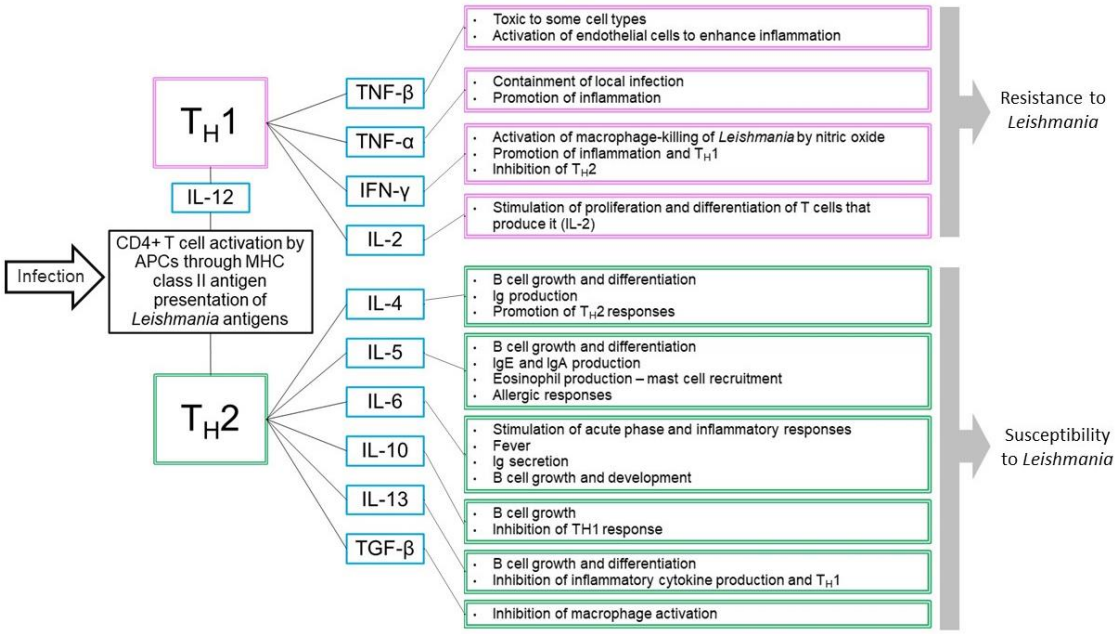


Figure 12. The contribution of T_H1 and T_H2 polarised immune responses to susceptibility or resistance to *Leishmania* infection. Figure adapted from Dunning (2009)[167] under a Creative Commons license. CD, cluster of differentiation; MHC, major histocompatibility complex; T_H, T helper cell; IL, interleukin; TNF-α/β, tissue necrosis factor alpha/beta; IFN-γ, interferon gamma, TGF-β, transforming growth factor beta; Ig, immunoglobulin.

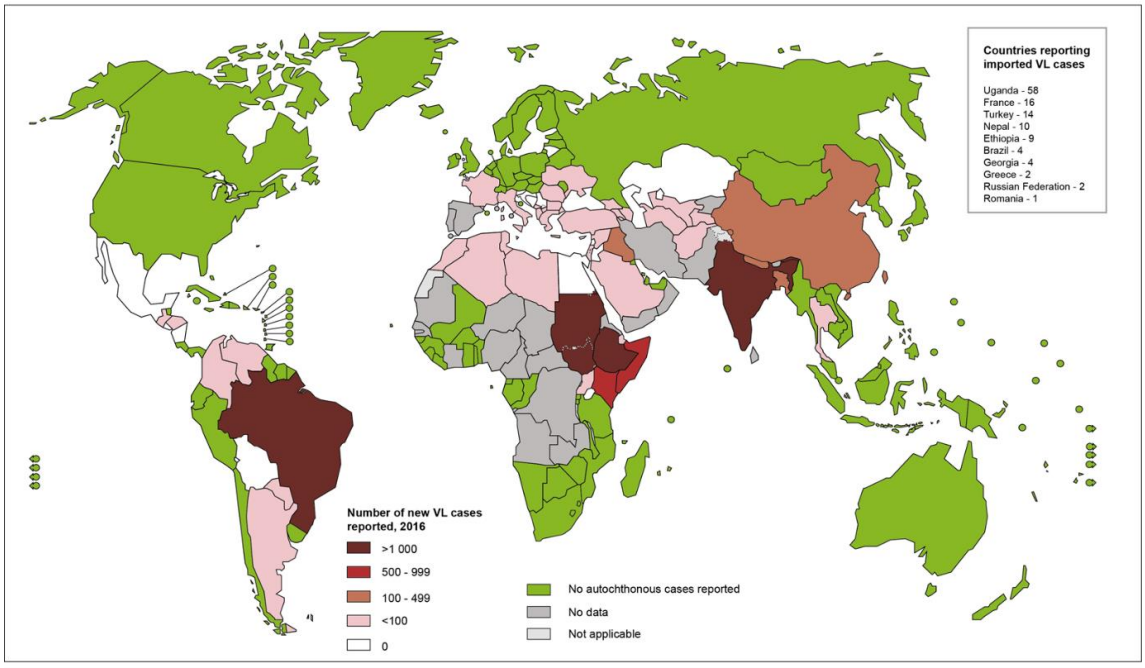
HIV-VL co-infection leads to more severe VL disease which does not achieve complete cure, even following multiple treatement rounds[18, 170]. In turn, VL exacerbates HIV progression[171].

This synergistic effect is in part due to the destruction of a beneficial T_H1 response by HIV which could otherwise confer resistance to *Leishmania* (Figure 12).

In addition to infectious disease interaction, VL is worse in malnourished individuals. In particular low meat protein consumption, and retinol (vitamin A) and zinc deficiencies have also been found in VL patients[158, 172]. Zinc is a co-factor in over 300 enzymes and its deficiency is known to have detrimental effects on many types of immune cells[173]. Providing zinc to *Leishmania*-infected mice was both therapeutic and prophylactic[174]. The contribution of vitamin A is less well-defined than that of zinc but it may play a role in immune regulation[175]. However, very little evidence is available on the effect of providing nutritional supplements with VL treatment in humans, and a Cochrane Review on the subject in 2018 failed to identify any eligible studies, therefore more research is needed[176]. Occurence of PKDL, which tends to appear on sun-exposed parts of the body, may be related to endogenous production of vitamin D_3 which has potent immune interaction[177].

Epidemiology and control

Leishmaniasis is endemic in about 98 countries, however in 2014, 90% of reported VL cases were from just six countries: Brazil, India, Bangladesh, Ethiopia, South Sudan and Sudan[178, 179] (Figure 13).



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement. © WHO 2018. All rights reserved

Data Source: World Health Organization
Map Production: Control of Neglected Tropical Diseases (NTD)
World Health Organization

World Health Organization

Figure 13. Global visceral leishmaniasis endemicity in 2016, as reported to the World Health Organization. From World Health Organization (2018)[180], with permission.

Risk factors for VL include poverty and malnutrition, with a tendency for younger children to be affected more than older people, males more than females, and a possible role of the proximity of cattle to the household[12, 181]. Cases of VL and asymptomatic infection, cluster at village and household level[158].

VL has historically caused large epidemics as it emerges in a susceptible population. A very severe epidemic occured in South Sudan as a result of the movement of infected people into a previously non-endemic region due to armed conflict. VL caused approximately 100,000 deaths, one third of the population, in a ten year period from 1984[9]. An epidemic cycle of about 8-10 years has now developed in parts of Sudan[182]. In India and neighbouring countries, epidemics have occurred on a cyclic basis for over a century[183]. This cycle, recently of about 15 years, is likely to be related to herd immunity with cases increasing as the younger, previously unexposed, population reach sufficient numbers. This is evidenced by incidence peaks of both VL and PKDL in children and young adults aged from 5-29 years, and increasing cellular immunity (i.e., disease resistance) with age in adults, as determined by the leishmanin skin test[159, 184-186]. In addition, or alternatively, to herd immunity, the cyclic incidence may be a consequence of withdrawal of insecticide spraying campaigns and other factors[10, 183]. According to the most recent data, India is continuing to see a reduction in the number of VL cases and Bangladesh is on track to meet its elimination goal (Figure 14)[187]. However, to what extent this decline is due to control efforts, or the inter-epidemic period, is not yet clear.

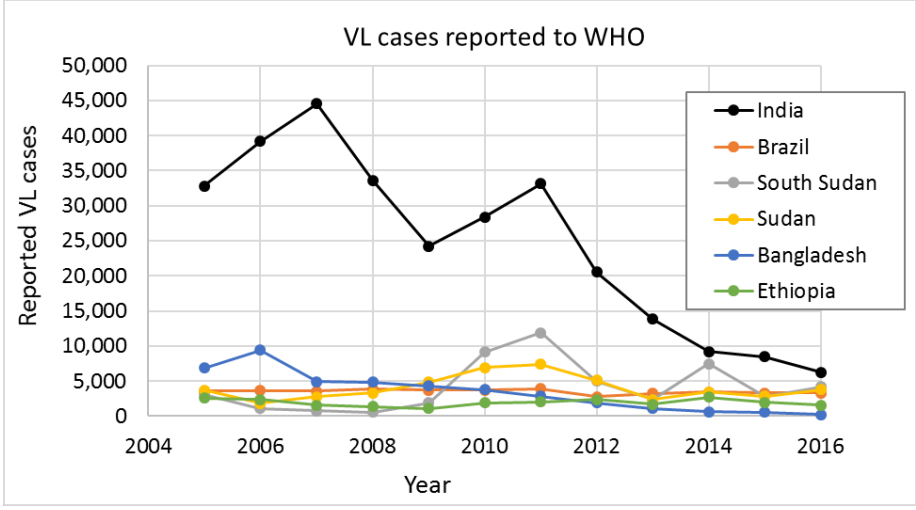


Figure 14. Number of VL cases reported to the World Health Organization (WHO) by the 6 countries accounting for 90% of global cases. Reporting methods changed in 2013 to separate autochthonous from imported cases, only the former is shown here from 2013. Data from the World Health Organization (2017)[188].

Relapse

Relapse of VL after apparently successful treatment i.e. resolution of clinical symtoms, is reported in up to 10% of cases in the Indian subcontinent and fewer than 4% in regions of eastern Africa[189]. These patients undergo repeated biopsy to confirm the disease and then repeated treatment, which may require hospitalisation. Relapse is more common in cases of HIV co-infection, even when the patient is taking anti-retrovirals[18, 190]. In immunocompetent VL patients in India, the majority of relapses occur 6 to more than 12 months post-treatment with AmBisome[16] but within 6 months after miltefosine treatment[191]. In Sudan, relapse rates increased when a short course of SSG and PM was used for VL treatment, instead of longer duration SSG alone[192]. Therefore incomplete treatment is likely a factor contributing to relapse, which may not be due to drug resistance in the parasite[193].

PKDL

PKDL is a non-painful but potentially temporarily disfiguring sequel of VL, occuring mostly after treatment but also during, or without, a prior VL episode[194]. It is characterised by skin pathoglogy (nodules, macules and/or papules), typically beginning on sun-exposed areas such as the face[184], and proven to be a source of infection for sand flies[195, 196]. Therefore, PKDL is one mechanism by which the parasite is maintained in a human population between epidemics[181]. Presentation of PKDL differs between Africa and Asia, occuring months to years after VL treatment in the Indian subcontinent[194] but weeks to months

after, or during, treatment in Sudan[197]. In both regions, it may occur without prior VL, 10% in Sudan and up to 23% of cases in India[182, 198]. PKDL is much less reported in *L. infantum* endemic areas[199]. Treatment of PKDL in the Indian subcontinent is lengthy, requiring weeks or months of regular injections, however, it may self-heal over months to years if untreated[184]. In Sudan, PKDL is largely self-healing and usually not treated[197, 200].

Asymptomatic VL and progression

Asymptomatic individuals may be another reservoir of parasites as 2%-3.5% of these will progress to VL disease[201, 202], a number that can rise to 12% among those with a high anti-*Leishmania* antibody titre[169]. Asymptomatic infection is most often characterised as seropositivity in the absence of previous VL, a state which outnumbers VL cases at between 4-10 to 1[169] and may cluster around households with VL cases, although evidence is conflicting[198, 203]. However, seropositivity alone does not necessarily indicate presence of circulating parasites in the blood and over 80% of seropositive individuals revert to seronegative within a year[201]. Given that recent seroconversion and high titre seropositivity have been identified as risk factors for progression to VL[169], and that asymptomatics are not treated, these individuals need to be identified and followed so that when disease manifests, they can be treated without delay to avoid morbidity, mortality and onward disease transmission.

Zoonotic VL

The epidemiology of VL varies greatly across its geographic range. This is because *L. infantum* is a zoonotic parasite with a main reservoir in domestic dogs, whereas *L. donovani* in the Indian subcontinent is considered anthroponotic, but may have less-well characterised animal reservoirs[204] and *L. donovani* transmission in Sudan has moderate influence of sylvatic animal reservoirs[182], as well as dogs[205].

Control

VL control measures differ across the different geographies. In Brazil, where VL is zoonotic, animal reservoir control, for example, culling of seropositive dogs has been attempted, but proved ineffective as well as higly unpopular[206]. A vaccine for dogs is available to pet owners with about 70% efficacy[207]. Reducing the contact between sand flies and dogs by the use of insecitide-releasing collars is also practiced. However, Brazil in particular experienced a rise in VL morbidity and mortality between 1990 and 2016[208]. In some regions this was particularly associated with HIV/AIDS co-incidence, although improved nutrition and overall vaccination coverage (of standard childhood vaccines) in children has reduced incidence in younger age groups[209]. Since these figures were released, the Pan

American Health Organization set goals to reduce mortality from VL by strengthening surveillance, diagnosis and treatment systems[210]. However, weak political will, the low perception of VL as a problem, and the need for extensive research on transmission dynamics and vaccine development, could continue to hinder control programs in Brazil[211].

In the Mediterranean region where *L. infantum* infects dogs, sylvatic mammals and humans, control is based on case detection and treatment, both of infected dogs and humans[206]. Injecting drug users, immunosuppressed, and HIV positive individuals are at increased risk of VL[212, 213]. An unusual outbreak of (human) VL in Madrid in 2009-15 was traced to wild hares in an urban park[214], thus the culling of hares and rabbits in the park, use of insecticide-impregnated dog collars and complete treatment of infected individuals brought the outbreak under control[215].

VL is targeted for elimination as a public health problem in India, Bagladesh and Nepal, a region where there is presumed to be no animal reservoir. Priorities of the elimination strategy are: early VL case detection with complete treatment, for which the drug AmBisome is donated by the manufacturer Gilead; and vector control using indoor residual spraying (IRS) of pyrethroid insecticides[14]. Additionally, health system strengthening and community engagement have been important in Bangladesh[187]. Historically, VL in India was reduced to negligible levels during extensive DDT insecticide spraying to control malaria in the 1950s and 60s, which had a knock-on effect on sand flies, demonstrating that intensive vector control is effective. However, by the 1970's, when DDT spraying had been ceased, VL cases returned. Likewise, long-term success of the current control program must focus on PKDL which is likely to be a reservoir of parasites between outbreaks of VL, is underreported due to being asymptomatic and involves lengthy treatment that may not be tollerated by patients[182]. In addition to IRS for vector and transmission control, bed nets have shown benefit in protecting against VL when used consistently during summer months, even if they were not insecticide treated[158].

VL control in eastern Africa relies primarily on diagnosis and treatment to prevent morbiditity and mortality. Many treatment centres across the region are provided by Médecins Sans Frontières (MSF) and face challenges of high HIV-coinfection rates, and in some areas, disruption due to armed conflict[216]. A study of the most effective treatments in African VL identified the now first-line combination therapy which reduced treatment time[217]. In addition, a large collaborative and internationally-funded project 'KalaCORE' is

supporting country efforts to increase capacity for VL control in East Africa addressing diagnosis, treatment, vector control and supply chains, among other requirements[218].

Treatment

There are many drugs available against *Leishmania*; antimonials, pentamidine, paromomycin, amphotericin B and miltefosine (Figure 15). The first class of drugs, used for VL treatment from 1912, was based on the element antimony (Sb, latin: *stibium*) which was since developed into sodium stibogluconate (SSG)[219]. This pentavalent antimonial continues to be used for VL in regions where the parasite remains susceptible, although it has potentially fatal cardio-toxicity which must be monitored during the month-long treatment[219].

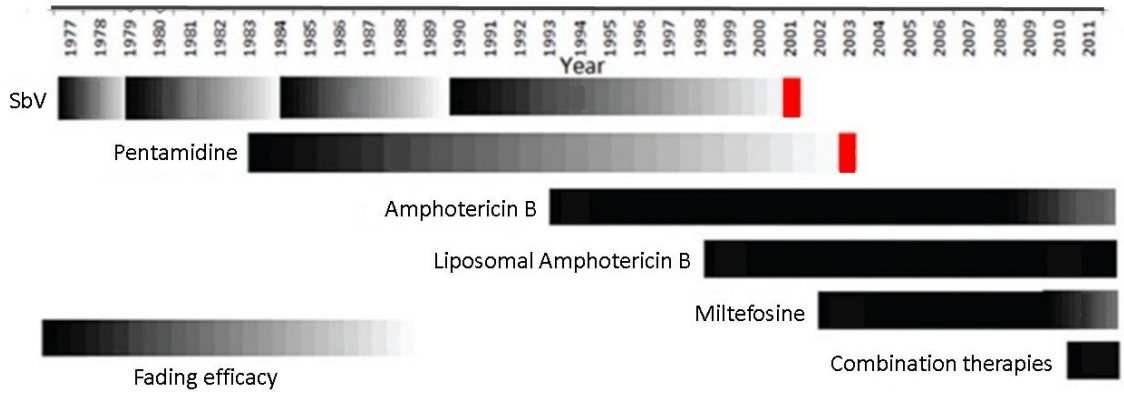


Figure 15. Timeline of drugs in use against visceral leishmanianisis in India, 1977 until 2011. SbV: pentavalent antimony. Red blocks indicate the cessation of use of these drugs as monotherapies in India. Figure adapted from Muniaraj (2014)[10] under a Creative Commons license.

Paromomycin (PM) is used in combination with SSG as first-line treatment in South Sudan given as intramuscular injections; the combination shortens treatment from 30 to 17 days but has been associated with higher relapse rates[192, 220]. Amphotericin B, formulated into liposomes as the brand AmBisome, is the safest drug for VL and is donated by the manufacturer Gilead for the VL control programmes in India, Bangladesh, Nepal, South Sudan and Sudan[13]. In India, single dose intravenous liposomal amphotericin B is a first line treatment[221]. Amphotericin B was first discovered in the 1950s from the soil bacterium *Streptomyces nodosus*[222]. Also functional as an antifungal agent, amphotericin B acts on

Leishmania membrane sterols and increases permeability so that the cells leak[223] (Figure 16).

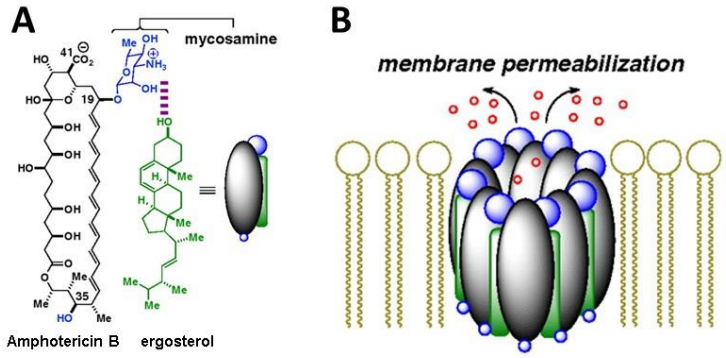


Figure 16. Amphotericin B. (A) Molecular structure and binding to ergosterol. (B) Mode of action by creating pores in the cell membrane. Images from Gray *et al.* (2012)[224] with permission.

Miltefosine is the only oral drug available for VL treatment. However it has teratogenic effects which limit its use in women of childbearing age and drug resistance is emerging in the parasite[225]. To avoid additional drug resistance, combinations of existing drugs are increasingly used, including AmBisome and miltefosine in the Indian subcontient, among other combinations[226, 227].

Diagnosis

Microscopy

Diagnosis of VL is made primarily by serological testing, either by RDT, ELISA, direct agglutination test (DAT) or immunofluorescence antibody test (IFAT), the choice of which differs between countries and regions. Positive serology is followed by parasitological confirmation to visualise amastigotes in aspirated fluid from the spleen, bone marrow or lymph nodes (Figure 17)[228].

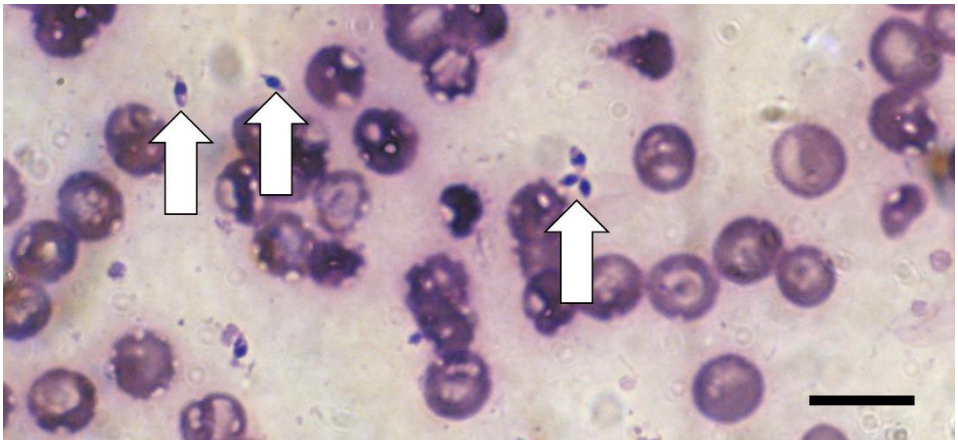


Figure 17. Amastigotes of *L. donovani* (arrowed) alongside red blood cells in a Giemsa stained Indian splenic aspirate. Scale bar is approximately 10 um. Photograph by the author.

Serology

The discovery of K39, a highly sero-reactive antigen with high specificity to VL was made by screening a genomic library of *L. infantum* with the serum of an infected patient[229]. This protein, a 298 amino acid (aa) section of a kinesin, consists of 6.5 repeats of a 39 aa sequence (Figure 18). The discovery and production of recombinant K39 (rK39) revolutionised VL diagnosis due to its high sensitivity and specificity, and field applicability in RDT format[230].

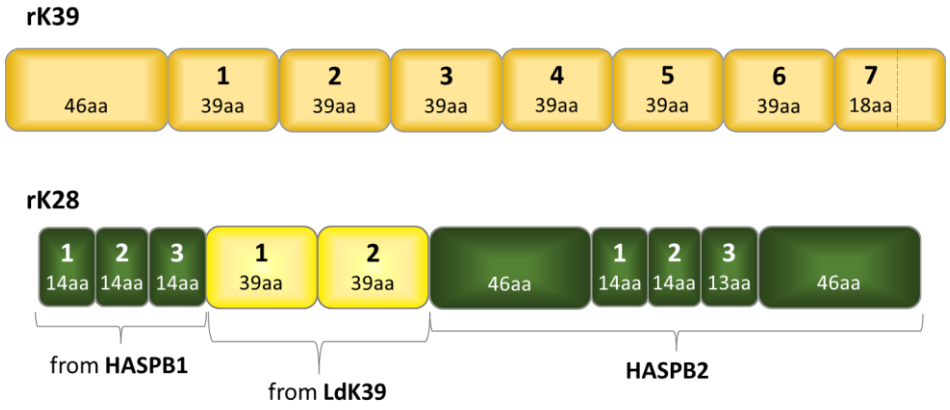


Figure 18. Representation of the rK39 and rK28 serological antigens of *Leishmania donovani*, used to diagnose VL. Yellow shades indicate sequences derived from kinesin proteins and green indicates HASPB derived sequences.

Highly effective in the Indian subcontinent, rK39-based serological assays have sensitivity and specificity in the range of 93-100%[228]. However, this antigen has lower accuracy in

eastern Africa, ranging between senstivities of 75-85% and specificities of 70-92% in Ethiopia, Kenya and Sudan[231]. The poorer performance of rK39 for African VL is due to a combination of differences in human serological responses and parasite antigen sequence differences between Africa and Asia[231-233]. To address the need for a reliable seroantigen in eastern Africa, a 253 aa fusion protein termed rK28 was created from a synthetic gene that combines repeats of an African *L. donovani* rK39 orthologue with sections of two antigenic HASPB proteins[234, 235] (Figure 18). An rK28-based RDT tested in Sudan had specificity of 100% and 97.6% with whole blood and serum respectively. Corresponding sensitivitites were 92.5% and 94.5%[236]. Beyond eastern Africa, rK28 performed equally to rK39 when tested in India[237] and Bangladesh[234]. In Brazil, rK28 serology has proven useful for identifying seropositive dogs, a key component of zoonotic VL control in that region[238]. rK39 serology has also shown high sensitivity and specificity in Brazil, including 100% negativity in patients with cutaneous leishmaniasis[239, 240]. However, both rK39 and rK28-based diagnostic tests had much reduced sensitivity (61-67%) in HIV co-infected individuals in Brazil[241].

The detection of antibodies has also been evaluated using urine samples, against recombinant antigens: rKRP42 with 94% sensitivity and 99.6% specificity[242]; rK39 with 96.4% sensitivity but insufficient specificity of 62% in one study[243] but similar sensitivity and 100% specificity in another study[244]. rK28 had 95.4% sensitivity and specificity of at least 93% with urine[245]. The reason for the presence of anti-*Leishmania* antibodies in urine is not well defined and given the variable results in various studies, the use of urine with rapid diagnostic tests or ELISA would need additional validation prior to incorporation into control programs or for individual diagnosis.

PCR and qPCR

Molecular diagnostics for VL include PCR and qPCR of DNA extracted from periferal blood, bone marrow, or buffy coat (a white cell and platelet layer formed when blood is separated by centrifugation). However, molecular tests are mostly limited to highly resourced laboratories due to the need for advanced equipment, and methods are not standardised. Commonly used PCR targets are the 18S rRNA gene and the kinetoplast, among others[246, 247]. The most sensitive PCR amplifies the minicircles, DNA circles within the kinetoplast, which occur at about 10,000 copies per parasite[247]. In immunocompetent VL patients, PCR of blood and bone marrow have almost equal sensitvity[246]. In India, PCR of DNA from blood yielded sensitivity and specificity of 96.4% and 98.5% respectively, and in Bangladesh, amplification of buffy coat DNA had sensitivity and specificity of 90.7% & 100%

respectively[247]. qPCR, has been found to distinguish between asymptomatic and active infection based on the quantity of circulating parasite DNA[248]. VL urine has also been tested by qPCR and although it had lower sensitivity than serological and parasitological methods, it became negative soon after treatment[249].

Despite these promising results in research trials, PCR and qPCR are expensive, therefore, they are generally not used in clinical practice in endemic areas where they do not add additional value above existing tests[228]. However, PCR may be a solution for diagnosis of VL in HIV positive individuals in whom serology is less reliable, and as regions move towards elimination, when a measure of active infection will be needed. According to a meta analysis, PCR of buffy coat or bone marrow in HIV-VL co-infection, both have high sensitivity and specificty of around 93% and 96% respectively[246].

LAMP

Alternatives to PCR and qPCR are methods that rely on isothermal (single-temperature) DNA amplification. Loop-mediated isothermal amplification (LAMP) is one such technique that uses 4 or 6 primers giving high specificity, and is carried out in a single tube at about 60°C, after which results can be observed visually. For VL, a LAMP assay had the ability to detect less than 1 parasite's worth of *L. donovani* DNA and had very high sensitivity and specificity with DNA from human blood and bone marrow samples[250]. The assay was also positive with a small number of post-treatment samples from patients who went on to relapse. In Sudan, a commercial LAMP kit was evaluated for VL diagnosis and had very high sensitivity using DNA extracted from blood or buffy coat by either a commercial kit, or a low-cost method using common lab reagents and equipment, indicating its potential application to confirm infection in lower-resourced settings[251]. Although LAMP removes the need for a thermocycler and fluorescence detection equipment, which theoretically would make it more field-applicable, it still requires the DNA to be purified from the sample, is susceptible to enzyme inhibitors carried through from samples, and is complex to develop due to the large number of primers.

RPA

As well as LAMP, the recombinase polymerase amplification (RPA) technique is another isothermal nucleic acid amplification method with the advantage of working at lower temperatures, including ambient temperature. RPA involves not only primers and polymerase, but also a recombinase enzyme that first binds to the primers and catalyses their attachment to the template. This removes the need for temperature cycles. While the

recombinase technology is provided by the manufacturing company, TwistDx, RPA assays (primers) have been developed for many infectious agents, including *Leishmania*. Castellanos-Gonzalez *et al.* (2015)[252] detected *L. infantum* in dog blood and mucosal samples using their RPA and demonstrated high specificity, and sensitivity similar to qPCR. An additional advantage of RPA is the possibility to provide a visual result readout on a lateral flow strip, thus giving the technique the potential to be used closer to point of care. To make the RPA truly field-applicable, Mondal *et al.* (2016)[253] incorporated all the components needed for an RPA into a 'suitcase laboratory' for detection of *L. donovani* in blood, buffy coat or PKDL skin biopsies. DNA extraction was performed in 20 minutes from one suitcase and RPA from another, taking another 20 minutes and powered with electricity from a battery and solar panel. As with the *L. infantum* RPA, the target was minicircle DNA, and 100% sensitivity and specificity was achieved on a small sample set, equalling that of qPCR. The same suitcase RPA was applied to samples from Sri Lanka where *L. donovani* causes cutaneous lesions[254]. Specificity remained at 100% but although sensitivity was lewer at 65.5%, the method was more sensitive than microscopy and was 3 times cheaper than PCR.

Antigen detection

Antigen detection is an alternative to serology and molecular diagnostics. Detecting parasite antigens in urine has the advantage of being non-invasive, and theoretically of indicating cure in a timely way, being effective in HIV-VL co-infected patients, and being cheaper and more field-applicable than PCR. The KAtex test is the only urine antigen test that has been assessed widely for VL diagnosis. It is a latex agglutination test comprising antibody coated latex particles which when mixed with an infected urine sample on a card, agglutinate to indicate the presence of a *Leishmania* antigen. Uptake of the KAtex in VL control programs has been hampered by the variable and often low sensitivity of 35.8% to 94%, overall assessed as 63.6% (95% CI: 40.9-85.6%) by a meta-analysis[255], despite high specificity of close to 100%[231, 256]. In addition, there is a need to boil urine samples before processing to remove false positive reactions. However, the carbohydrate antigen has been well characterised and a monoclonal antibody against it has been developed into ELISA format, which removed the need for boiling samples and had 94.1% sensitivity and 100% specificity[257].

Studies reported in the literature have identified *Leishmania infantum* antigens in VL urine, including nuclear transport factor 2, tryparedoxin and iron superoxide dismutase[258]. Abeijon *et al.* (2012)[258] developed ELISAs to capture these antigens but with limited

sensitivity. This was improved by combining the three analytes[259] and successfully detected a drop in antigen levels immediately at the end of treatment at day 30 compared with pre-treatment, in a small pilot study in India[260]. The assay has since been developed further by the use of more robust camelid antibodies[261].

In an alternative approach to urine antigen detection, Vallur *et al.* (2015)[262] used an antigen-affinity purified polyclonal antibody against crude *Leishmania* lysate which was developed into a standardised ELISA kit for commercialisation. This kit had 88-100% sensitivity which varied by region, and 100% specificity, for VL from Africa, South America and Asia[262], indicating very good potential for wider validation and use, but potentially limited by the need for polyclonal antibodies.

1.3 Aims and Objectives

1.3.1 *Strongyloides*

Aim

To identify protein antigens of *S. stercoralis* that can contribute to development of a coproantigen rapid diagnostic test.

Objectives

a) Gather appropriate open access 'omics' data sets and analyse these for candidate antigens, based on: likely presence in stool; specificity to *Strongyloides* or *S. stercoralis*; antigenicity.

b) Investigate relative gene expression, sequence homology, predicted epitopes and glycosylation, 3D structure and other features of the candidate antigens.

c) Compile a shortlist of candidate coproantigens meeting the above criteria, and synthesise peptide coproantigens to be used to raise antibodies.

d) Investigate gene and protein sequence diversity/conservation of candidate coproantigens from diverse geographical regions, by designing primers and amplifying and sequencing regions of interest from *S. stercoralis*-infected stool DNA from regions of Africa and Asia.

e) Collect stool samples from a community in Guinea Bissau, storing samples in ethanol for DNA analysis, and in formalin for antigen analysis for development of assays to detect coproantigen and to test prototypes.

1.3.2 Visceral leishmaniasis

Aim

Contribute to development of diagnostic assays for visceral leishmaniasis through assessment of the capacity of IgG1 anti-*L. donovani* responses to discriminate between different VL disease states, and to identify *L. donovani* protein antigens in human urine from VL cases, which may be the targets of a urine antigen capture assay.

Objectives

f) Assess the ability of different diagnostic assays for IgG1 anti-*L. donovani*, to distinguish between active or progressive VL and asymptomatic carriers; post-

chemotherapy relapse versus cure, and to diagnose PKDL, with sera from Sudan and India, including results of laboratory research performed at Banares Hindu University, Varanasi, India.

g) Use western blot and mass spectrometry to investigate the different antigens involved in individual anti-*L. donovani* IgG1 responses of active, aymptomatic, cured, relapsed and PKDL individuals.

h) Immunocapture *L. donovani* antigens from VL patient urine using anti *L. donovani* antibodies and Indian and Sudanese urine. Identify proteins in the captured material from urine using mass spectrometry, and analyse this for potential *Leishmania*-specific candidate antigens and epitope regions.

CHAPTER 2: Comparative 'omics' discovery of *Strongyloides stercoralis* coproantigens for diagnostic development

Marlais T, Talavera-López C, Le H, Chowdhury F, Miles MA. Comparative 'omics' discovery of diagnostic coproantigens for *Strongyloides stercoralis* infection.

Key points, novel results and implications

- A novel analysis of open access genomic, transcriptomic and proteomic data, literature on existing seroantigens and published genomic analysis, using online computational tools for the discovery of coproantigens of *Strongyloides stercoralis*.
- The result is a rationally selected collection of *S. stercoralis* candidate coproantigens that are being synthesised as peptides, and others that can be expressed as recombinant proteins, and used to raise antibodies for the development of a coproantigen assay.
- These results can assist in the development of a coproantigen capture assay for strongyloidiasis, which would be a valuable basis for a rapid and robust diagnostic assay.
- In addition, this study provides a model for the application of this method to numerous other helminths of medical and veterinary importance, especially those that are difficult to culture.

Candidate's contribution

The candidate co-wrote the project proposal for the funding application then co-developed the detailed methodology. The candidate performed the literature review followed by bioinformatic and computational processes and associated analyses including, but not limited to: selecting and downloading all relevant datasets from online databases; constructing phylogenetic trees; sequence alignments; epitope prediction; 3D protein modelling; custom BLAST database creation and searching. The candidate compiled all results and figures, made the final selection of candidate antigens and prepared the manuscript. The candidate was assisted with the initial transcriptomic data analysis and received advice on which tools to use throughout the project but worked largely independently.

Contributions of co-authors

Carlos Talavera-López, Wellcome Sanger Institute, designed the bioinformatics approach and provided the candidate with the basic knowledge to carry out this work. Carlos performed the differential gene expression analysis of the transcriptomic data which formed the starting point of further analysis by the candidate. He also recommended particular bioinformatics tools used for other analyses and advised the candidate on their use.

Hai Le, LSHTM, designed primers and carried out PCR and sequencing, and qPCR of *Strongyloides* genes.

Fatima Chowdhury, LSHTM, began the initial process of identifying *S. ratti* E/S orthologues in *S. stercoralis* before the improved Hunt et al. (2016) genomic dataset was available.

Michael Miles, LSHTM, co-wrote the funding application and supervised the project.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Tegwen Marlais
Principal Supervisor	Prof. Michael A. Miles
Thesis Title	Diagnostics development for the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Frontiers in Cellular and Infection Microbiology
Please list the paper's authors in the intended authorship order:	Tegwen Marlais, Carlos Talavera-López, Hai Le, Fatima Chowdhury, Michael A. Miles
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I co-developed the detailed methodology, performed the large majority of the bioinformatic and computational analysis, compiled all results, figures, the selection of candidate antigens and wrote the manuscript.

Student Signature: [redacted] **Date:** 16/2/19

Supervisor Signature: [redacted] Date: 15/2/19

Comparative 'omics' discovery of *Strongyloides stercoralis* coproantigens for diagnostic development

Tegwen Marlais[1], Carlos Talavera-López[2], Hai Le[1], Fatima Chowdhury[1], Michael A. Miles[1]

1 Pathogen Molecular Biology Department, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK

2 Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

Abstract

Background: Infection with the gut nematode *Strongyloides stercoralis* persists unless effectively treated and is potentially fatal in immunosuppression. In addition, detailed epidemiological data are lacking due to inadequate diagnostic options. A rapid antigen detection test is needed to screen populations, to validate cure after treatment and prior to immunosuppressive therapy.

Method: We analysed open access 'omics' data sets and existing seroantigens, and used online predictors to identify *S. stercoralis* proteins that are likely to be present in infected human stool, *Strongyloides*-specific, and antigenic.

Results: Transcriptomic data from gut- and non-gut-dwelling life stages of *S. stercoralis* revealed 328 proteins differentially expressed. Excreted and secreted (E/S) proteome data of *S. ratti* were 'converted' to *S. stercoralis*, giving 1,057 orthologues. Two effective seroantigens and seven parasitism-associated protein families were compared phylogenetically between *S. stercoralis* and outgroups. Proteins with least homology to outgroups were selected. Proteins that overlapped the datasets were analysed by multiple sequence alignment, epitope prediction and 3D structure modelling to reveal peptides/protein sequences that form a selection of candidate coproantigens for *S. stercoralis*. A subset of short peptide antigen candidates are being synthesised for assay development.

Conclusions: We detail multiple candidate antigens for *S. stercoralis* coproantigen detection assay development, identified using open access data and freely-available protein analysis

tools. This powerful approach has potential to be applied to multiple neglected parasitic infections with available 'omic' data to speed up the development of rapid, specific, diagnostic assays for laboratory and field use.

Introduction

Biology

The nematode *Strongyloides stercoralis* is a soil transmitted helminth (STH) occurring in faecally-contaminated humid soils in tropical and sub-tropical regions. Strongyloidiasis is estimated to affect up to 40% of people in endemic regions, where it is particularly a disease of poverty [1, 2]. Infection occurs when infective third stage (iL3) larvae penetrate the skin and migrate within the host via the bloodstream and airways, before being swallowed into the digestive tract. The parasitic adult female burrows into the epithelium of the duodenum where it feeds on host tissue. Clinical signs and symptoms of strongyloidiasis may be eosinophilia, digestive pain or discomfort, urticaria and respiratory symptoms [3-5].

S. stercoralis is unusual among human parasitic nematodes as it can complete its life cycle within the host and thus sustain infection for many years or decades if untreated [6-8]. Autoinfection occurs when larvae reach iL3 stage before leaving the host and penetrate the gut or perianal skin to complete development. However, during immunosuppression, particularly due to corticosteroid treatment of co-morbidities, or HTLV-1 co-infection, very large numbers of larvae may undergo this autoinfective cycle, causing hyperinfection or dissemination [9, 10]. These conditions are frequently diagnosed late due to their non-specific presentations, are difficult to manage, and have a fatality rate of over 60% [10-12]. Diagnosis of strongyloidiasis and validation of cure after treatment are therefore imperative.

Treatment with the first-line drug ivermectin, is well tolerated and has a reported efficacy of between 57% and 100%, with two doses preferred over a single dose. However, accurate determination of cure depends on follow-up and the diagnostic method used[8, 12, 13]. Albendazole and mebendazole, used to treat infection with other STH (*Ascaris lumbricoides*, hookworm and *Trichuris trichiura*), are less effective, or ineffective, against *Strongyloides*[14]. However, albendazole and ivermectin might be used together for joint benefits, pending further validation[15, 16]. In addition to ivermectin, moxidectin has shown equivalent effectiveness against *S. stercoralis* in early trials[17].

Diagnosis

Diagnosis of strongyloidiasis can be made by microscopy using agar plate culture[18] or Baermann funnel, both of which isolate larvae from fresh stool. qPCR on extracted stool DNA is used in research and highly resourced laboratories and improves sensitivity over microscopy[19-22]. However, sensitivity may be reduced by polymerase inhibitors present in stool[23], low stool volume used in testing (as little as 0.1 g compared to 1 gram or more used for culture or larval concentration methods), and irregular larval excretion[13]. Serology, detecting antibodies against either whole worm or recombinant antigens NIE and SsIR, has sensitivity of 70-98% and high specificity[24] but cannot distinguish cure in a timely way as it remains positive for months to years after treatment[25-27]. Therefore, there is a need for a rapid diagnostic test (RDT) that detects parasite material and can be used for screening as well as for confirmation of cure. Pilots of such assays have shown proof of principle under research conditions using antibodies against *Strongyloides ratti* and *Strongyloides venezuelensis* somatic or excretory/secretory (E/S) antigens[28-30]. However, none of these studies has identified specific *S. stercoralis* proteins antigens that would enable production of standardised diagnostic tests at large scale.

Reverse vaccinology

The wealth of 'omic' data now available in the public domain, coupled with online protein analysis tools, enables a computational approach to antigen discovery. This concept was termed reverse vaccinology when used in vaccine candidate discovery[31]. The approach begins with genomic analysis, as opposed to biochemical or serological methods and it also has significant potential in diagnostic antigen discovery. It has the advantages of not requiring culture of the organism and of revealing antigens that may be less abundant or difficult to purify *in vitro*. Incorporation of transcriptomic data can inform candidate gene/protein selection in parasites with multiple life stages, such as *S. stercoralis*[31].

Our approach was facilitated by the publication of the genomes of *S. stercoralis* and three related species[32]. Here, we have applied a series of computational analyses to open access transcriptomic, genomic and proteomic data from *Strongyloides* species and other helminths. We have used common bioinformatic tools to identify *Strongyloides* protein antigens that may be diagnostic targets detectable in human stool in a coproantigen capture RDT. We focus on 8 protein families and suggest that the method could be used to explore other protein families of significance in parasitism.

Methods

Data sources

This study used data obtained from public databases, as detailed in Supplementary file S1. *S. stercoralis* transcript sequences identified by the prefix 'SSTP' can be obtained via UniProtKB (www.uniprot.org) or WormBase ParaSite (WBPS: www.parasite.wormbase.org).

Overview of method

Three criteria were applied for candidate antigen selection: presence in infected stool; specificity to *Strongyloides* and/or *S. stercoralis*; antigenicity, to facilitate raising sensitive antibodies. Datasets and analyses used to make this selection are shown in Figures 1 and 2, and detailed subsequently.

Figure 1. Three requirements of a *S. stercoralis* coproantigen candidate (filled boxes, top row), and the analyses and open-access data sets used to assess against each criterion (un-filled boxes).

Figure 2. Starting data and analyses applied towards selection of *Strongyloides stercoralis* candidate coproantigens.

Parasitism-associated proteins

Protein families reportedly associated with parasitism, 'priority protein families', formed our focus for identifying coproantigens. Analysis of the *Strongyloides* genomes by Hunt *et al.* (2016)[33] revealed seven protein families associated with parasitic as opposed to free-living environmental nematodes. Of the seven, four protein families contained more genes in *S. stercoralis* than in *S. ratti*. Additional evidence for particular protein families containing potential coproantigens was found in published literature on the E/S proteomes of various parasitic helminth species.

Transcriptomics

Stoltzfus *et al.* (2012)[34] analysed the transcriptome of 7 life stages of *S. stercoralis* (Figure 3). For the present study, these RNA data[34] were downloaded from the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) and grouped by potential location of the life stage, either in the host gut or not, in order to represent parasite life stages which could be excreting or secreting antigens into human stool (Table 1). Each life stage consisted of triplicate RNA reads.

Figure 3. *S. stercoralis* life stages. Asterisks indicate life stages for which transcriptomic data were obtained by Stoltzfus *et al.* (2012)[34]. For the present study, transcriptomic data were grouped into two groups: those potentially present in the host gut, indicated by boxed asterisks, and those outside the host gut. Figure modified from Stoltzfus *et al.* (2012)[34], under a CC BY license. P Females, parasitic females; PP, post parasitic; FL, Free-living; PFL, post free-living; L1, stage 1 larva; L3, stage 3 larva; iL3, infectious third stage larva; L3+, tissue-migrating larva; L3a, autoinfective L3.

Table 1. Grouping of transcriptomic data from Stoltzfus *et al*. (2012) into gut and non-gut life stages. Accession numbers are given for NCBI SRA for the triplicate reads.

Gut		Non-gut	
Life stage	Accession numbers	Life stage	Accession numbers
Parasitic female	ERR146959 ERR146960 ERR146961	Free-living female	ERR146941 ERR146942 ERR146943
Post-parasitic L1 larvae	ERR146953 ERR146954 ERR146955	Post free-living L1 larvae	ERR146950 ERR146951 ERR146952
Post-parasitic L3 larvae	ERR146956 ERR146957 ERR146958	iL3 (environmental infectious larvae)	ERR146947 ERR146948 ERR146949
		L3+ (tissue-migrating larvae)	ERR146944 ERR146945 ERR146946

We calculated relative abundance of transcripts using RSEM and bowtie2[35] and subsequently grouped RNA data from the 4 non-gut-dwelling stages and 3 gut-dwelling stages separately to give two groups. We then analysed differential gene expression between these two groups using ebseq in RSEM[35]. We selected only the genes with

normalized read count differences that we could be 100% confident were differentially expressed (DE) between the two groups of life stages.

DE protein family identification

ClustalW[36] alignment of the DE genes was used to perform a multiple alignment and produce a phylogenetic tree which was annotated with iTOL[37]. The tree, labelled only with *S. stercoralis* gene accession numbers, facilitated the grouping of these DE genes into protein/gene families but protein identities remained unknown at this stage.

DE proteins were identified by two methods. In the first, amino acid sequences of selected clusters on the tree were submitted to three domain-finding tools: Delta BLAST[38], InterPro[39] and ExPASy Prosite[40, 41] and a consensus of all three was used to obtain a probable protein identity. The second method was to submit all DE protein sequences to BlastKOALA[42]. BlastKOALA protein identities were considered alongside the previously obtained consensuses or were used if there was no existing identity from the first three tools.

E/S proteomics

Source of proteomic data

Soblik *et al.* (2011)[43] submitted excretory/secretory (E/S) material of *S. ratti* parasitic females to mass spectrometry and identified the constituent proteins. However, this was in the absence of a detailed *Strongyloides* genome against which to search the mass spectra. Therefore, in their 2016 publication of the *Strongyloides* genomes, Hunt *et al.* (2016)[32] re-analysed these spectral data and obtained protein identities from the corresponding genomic data of *S. ratti*. We acquired the list of parasitic female E/S proteins, with *S. ratti* genome accession numbers and protein identities, from Supplementary Table 19 of Hunt *et al.* (2016)[32] and subsequently obtained the amino acid sequences for these accession numbers from the *S. ratti* protein file (WBPS v8) using samtools[44].

S. stercoralis orthologues to the *S. ratti* E/S proteome

At the time of conducting this study, there were no E/S proteomic data available for *S. stercoralis*, therefore we obtained *S. stercoralis* orthologues of the *S. ratti* E/S proteins. To do this, we searched the *S. ratti* E/S proteins against a custom blast+ database consisting of the *S. stercoralis* protein file (WBPS v8). We used blastp with word size 2 and e-value -50. *S. stercoralis* hits, in the form of accession numbers, were extracted from the resulting table and duplicated hits were removed. Corresponding *S. stercoralis* amino acid sequences were extracted from the *S. stercoralis* protein file using samtools. VENNY 2.1[45] was used to reveal

177 the *S. stercoralis* accession numbers that occurred in both the DE proteins and the E/S
178 orthologues. All the E/S orthologues were submitted to BlastKOALA as before, to obtain
179 protein family identities, as well as matching them with the protein identities reported for
180 the original *S. ratti* E/S proteins by Hunt *et al.* (2016)[32]. Separately, differential gene
181 expression data from analysis of the Stoltzfus *et al.* (2012)[34] transcriptomic dataset was
182 extracted for the E/S orthologues that occurred in both datasets.

183 SignalP

184 Signal peptide prediction, which provides evidence that a protein is secreted, was carried
185 out on the *S. stercoralis* DE proteins and E/S orthologues using SignalP 4.1[46].

186 Phylogenetic diversity

187 We used phylogenetic comparison to indicate *S. stercoralis* proteins with least homology to
188 those of other relevant species, followed by sequence alignment to identify exact regions
189 of specificity.

190 Selecting outgroups

191 A custom blast+ database was created from the genome-derived proteomes of selected
192 outgroup species (Table 2). The outgroups were selected to represent parasitic and non-
193 parasitic nematodes, as well as trematodes, cestodes and human.
194

Table 2. Outgroup species used for phylogenetic comparison to *S. stercoralis*.

Species		Rationale for selection
1	*Ancylostoma duodenale*	Hookworm. Prevalent nematode of humans
2	*Ascaris lumbricoides*	Prevalent nematode of humans
3	*Caenorhabditis elegans*	Free-living nematode and model organism
4	*Clonorchis sinensis*	Trematode of humans
5	*Enterobius vermicularis*	Prevalent nematode of humans
6	*Homo sapiens**	Human, present in all samples
7	*Necator americanus*	Hookworm. Prevalent nematode of humans
8	*Onchocerca volvulus*	Non-gut nematode of humans
9	*Parastrongyloides trichosuri*	Relative of *S. stercoralis*. Facultative parasite of possums
10	*Strongyloides ratti*	Close relative of *S. stercoralis*. Serological cross-reactivity with *S. stercoralis*[47].
11	*Strongyloides venezuelensis*	Close relative of *S. stercoralis*
12	*Strongyloides papillosus*	Close relative of *S. stercoralis*
13	*Syphacia muris*	Nematode of mice and rats. Serological cross-reactivity with *Strongyloides* in experimental animal infections[48].
14	*Taenia solium*	Cestode of humans
15	*Trichuris trichiura*	Prevalent nematode of humans

*Human protein/coding sequence (CDS) data were provided by the Human Genome Project at the Wellcome Trust Sanger Institute and can be obtained from Ensembl via ftp://ftp.ensembl.org/pub/release-87/fasta/homo_sapiens/pep/Homo_sapiens.GRCh38.pep.all.fa.gz. Accession numbers for all data are available in Supplementary file S1.

Specificity of DE proteins to *S. stercoralis*

S. stercoralis DE proteins were searched as separate families, against the custom outgroup database with blast+ criteria: word size 2 and e-value of -5 or -10, as appropriate, to obtain about 100 to 1,000 hits (S2 file). In cases where there were very few DE proteins in a particular family, the DE protein(s) were also searched against a custom database consisting of only *S. stercoralis* genome-derived proteins (S2 file). This was intended to increase the number of *S. stercoralis* proteins to enable species-specific clusters to be revealed on phylogenetic trees.

Protein hits from each outgroup, and the additional *S. stercoralis* hits where appropriate, were aligned with their respective original blast queries by multiple sequence alignment (MSA) using ClustalW, and phylogenetic trees were constructed. Trees were annotated with iTOL[37] to show proteins from each species in a different colour. *S. stercoralis* proteins which formed a distinct cluster, or clusters, on each phylogenetic tree were viewed in MSA along with the most and least similar proteins from each of the outgroups on that tree. These ClustalW alignments were analysed by eye for *Strongyloides* and *S. stercoralis*-specific regions which were submitted to BLASTP search against the NCBI non-redundant (nr) database to validate their specificity.

Antigenic potential

Existing seroantigens

Published literature was searched for uses of specific antigens for the serodiagnosis of strongyloidiasis, which therefore have proven antigenicity. Relevant proteins were submitted to analysis alongside other proteins, as described here.

Epitope prediction

BepiPred 1.0[49] and Bcepred[50] were used to predict epitopes within the DE proteins and E/S proteome orthologues. Both tools give a score to each amino acid residue signifying its likelihood of being part of an epitope. A BepiPred threshold of 1.3 (range -4 to 4) was selected for maximum specificity of 96%, with corresponding 13% sensitivity[51], in order to minimise the chance of false positive epitope predictions. Minimum length was 9 amino acids with no maximum. In the DE proteins, longer sequences with an overall very high epitope score were allowed to contain small regions scoring below 1.3.

Bcepred criteria were based on the reported highest accuracy of 58.7%, which was achieved using a threshold of 2.38 for the average score of four amino acid properties: hydrophobicity, flexibility, polarity and exposed surface [52]. In addition to BepiPred 1.0 and

Bcepred, we also used BepiPred version 2.0 for certain candidate antigens. This version became available only after the majority of the analysis, and offered improved prediction of conformational epitopes. BepiPred 2.0 was used with the same epitope length criteria and an epitope score threshold of 0.55 (range 0 to 1) which provided specificity of 81.7% and sensitivity of 29.2% on epitope predictions.

Outputs from the two prediction tools were compared, initially for proteins present in both outputs. The predicted epitope regions of these proteins were then examined for sequence overlap. Prior to selection as candidate antigens, predicted epitopes were assessed for their specificity to *Strongyloides.* Sequences were searched using BLASTP against the NCBI nr database. The "expect threshold" in BLASTP was increased if no results were obtained with default parameters. BLAST output was examined by eye for the sequence identity and biological relevance, i.e. likelihood of presence in a human stool sample, of non-*Strongyloides* results in the list.

3D modelling

Selected proteins of interest i.e. containing predicted epitopes, *Strongyloides*-specific regions and in a priority protein family or an existing seroantigen, were submitted to Phyre2[53] for 3D structure modelling against known crystal structures, using the intensive mode. UCSF Chimera[54] was used to visualise and annotate 3D models to highlight specific sequences of interest on the model. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

Glycosylation prediction

N-linked glycosylation was predicted with NetNGlyc[55] to account for the potential of a glycan to obscure protein antigen regions, or conversely to contribute to antigenicity. The prediction tool identified asparagine (N) residues with a high probability of being glycosylated via their amide nitrogen. Prediction was based on the motif N-X-S/T, where X is any residue except proline (P), and along with the presence of a signal peptide or trans-membrane domain on that protein, indicates that potential glycosylation sites are likely to be glycosylated. Intracellular, intramembrane regions, or signal peptides of a protein are unlikely to be glycosylated. If present, a glycosylation site close to a candidate antigen region on the 3D protein could indicate that the protein is less likely to be accessible to antibodies in a capture assay and therefore a lower priority candidate, pending *in vitro*

screening. Alternatively, glycosylated proteins may be investigated as antigens, however, this was not a focus of this study.

Results

Priority protein families

Seven protein families were found to be associated with *Strongyloides* parasitism by Hunt *et al.* (2016)[33]: sperm-coating-proteins/Tpx-1/Ag5/PR-1/Sc7 (SCP/TAPS), astacin, transthyretin-like (TTL), acetylcholinesterase (AChE), prolyl oligopeptidase (POP), trypsin-like inhibitors and aspartic peptidases. The first four of these have more proteins per family in *S. stercoralis* than *S. ratti*[33]. We included all seven protein families in initial analysis, although two, astacin and trypsin-like inhibitors, were not used to generate candidate coproantigens. In total, we identified 46 candidate coproantigens from the other 5 priority protein families and existing seroantigens. These candidate antigens match some or all of our criteria: antigenicity; specificity to *Strongylodies*; presence in stool.

Differential gene expression in gut life stages

Of a total of 13,098 *S. stercoralis* genes identified in RNA-seq data, we found 328 which we could be 100% confident were DE between gut-dwelling and non-gut-dwelling life stages (Figure 4). Of these, 203 were expressed more in gut-dwelling life stages than non-gut (S3 file). As well as the 7 parasitism-associated priority protein families, we also included collagen in further analysis due to its occurrence in the DE dataset and high number of predicted epitopes (Figure 4). Along with these 8, another 20 protein families were identified among the DE proteins, accounting for 193 (58.8%) of the proteins, with the remainder either not identified (22%), or given a disorder prediction (19.2%) indicating that they do not have a fixed conformation and are difficult to assign to a particular function or family.

Figure 4. *S. stercoralis* proteins differentially expressed with 100% certainty between gut-dwelling compared to non-gut-dwelling life stages, excretory/secretory (E/S) orthologues and predicted epitopes. Colours refer to different protein families or features, those containing several proteins are labelled. Protein families of particular interest as sources of coproantigens are in bold, underlined font. Dark grey markers on the three outermost circles of the tree indicate (from inside to outside): proteins orthologous to *S. ratti* E/S

298 proteins; proteins containing predicted epitopes according to BepiPred (middle circle) and
299 Bcepred (outermost circle) respectively.

300 Excreted/Secreted proteome

301 We identified 1,057 *S. stercoralis* proteins that had high homology to the 584 proteins in
302 the published E/S proteome of *S. ratti* [32]. Original *S. ratti* E/S proteins were given as 582
303 accession numbers, however two were found to have alternative isoforms, which we also
304 included here. Multiple sequence alignment indicated that 550 (94.2%) of the *S. ratti* E/S
305 proteins had a *S. stercoralis* orthologue at the selected similarity level (e-value 1E-50) and
306 that 284 (51.6%) of these had multiple homologues in *S. stercoralis* (S4 file).

307 In order to identify possible *S. stercoralis* E/S proteins among the DE gut stage proteins
308 already listed, we compared the 1,057 *S. stercoralis* orthologues with the 328 identified as
309 DE in the transcriptomic data. Seventy seven proteins were shared between both data sets,
310 of which 58 were shared between the 203 gut-stage DE proteins and the E/S orthologues.

311 To investigate gene expression of the E/S orthologues, we extracted the differential
312 expression data for these 1,057 proteins from the DE analysis which included all 13,098
313 active *S. stercoralis* genes. Protein families were assigned by a combination of BlastKOALA,
314 which identified 537 (50.8%) of the E/S orthologues, and the original Hunt *et al.* (2016)[32]
315 protein identities (S5 file). Proteins in the seven 'priority protein families' were identified
316 among the protein datasets.

317 Predicted epitopes

318 Within the 328 DE proteins, BepiPred and Bcepred jointly predicted epitopes in 104
319 proteins; 78 and 62 proteins respectively, with 36 proteins containing epitopes predicted
320 by both tools (Figure 4 and S3 file). Within the 78 proteins, BepiPred predicted 125
321 epitopes, and within 62 proteins Bcepred predicted 108 epitopes (Table 3). Fifty six
322 epitopes contained overlap or identity between the two prediction tools (S6 file). Predicted
323 epitopes ranged from the selected minimum of 9 residues to entire proteins of up to 651
324 aa with BepiPred and 8 to 66 aa with Bcepred (S7 file). These regions were given greater
325 scrutiny in the context of species specificity and antibody accessibility.

326 *S. stercoralis* orthologues (n = 1,057) of *S. ratti* E/S proteome were also submitted to
327 BepiPred and Bcepred, which predicted 747 and 62 epitope regions respectively, in a total

of 324 of the proteins. These ranged from 9 to 99 residues and contained 49 epitope sequences that overlapped, originating from 40 proteins (S6 and S7 files).

Table 3. Number of *S. stercoralis* proteins in differentially expressed (DE) and E/S orthologous protein datasets containing predicted epitopes according to two different prediction tools and number of predicted epitopes within those proteins.

Dataset	Number of	Epitopes predicted by			Total
		BepiPred only	Bcepred only	Both tools	
DE proteins (n = 328)	Proteins containing epitopes	42	26	36	104
	Epitopes	98 (125 total)	81 (108 total)	56[a]	233 (125 + 108)
E/S orthologues (n = 1,057)	Proteins containing epitopes	280	4	40	324
	Epitopes	728 unique (747 total)	59 unique (62 total)	49[a]	787 (728 + 59)

[a] Overlapping sequences but not necessarily identical lengths.

S. stercoralis-specific candidates identified by phylogenetic comparisons

S. stercoralis protein families linked to parasitism and represented in the DE proteins, and collagen, were analysed for *S. stercoralis* genus or species specificity. Separate BLAST searches of each protein family against 15 outgroups revealed which *S. stercoralis* proteins were most likely to contain specific regions (Figure 5).

Figure 5. *Strongyloides stercoralis* proteins (dark green) cluster to indicate possible species specificity (arrowed). Each circular tree is a different *S. stercoralis* protein family. Each colour represents a different outgroup species. *S. stercoralis* proteins were differentially expressed between gut and non-gut dwelling life stages and were aligned with BLAST hits from outgroups. Colours reaching the outer edge represent *Strongyloides* species whereas shorter colour bands are non-*Strongyloides* outgroups. A) SCP/TAPS, B) transthyretin-like (TTL), C) acetylcholinesterase (AChE), D) aspartic peptidase, E) prolyl oligopeptidase (POP), F) astacin-like metallopeptidase, G) trypsin-like inhibitor, H) collagen.

S. stercoralis proteins from the clusters identified in the phylogenetic trees were examined in alignment with outgroup-representative homologues with the most and least similarity. From the alignments, the *S. stercoralis* regions with least homology to outgroups were selected as candidate antigens. In some cases this included the whole protein. Results of this analysis are detailed below.

SCP/TAPS coproantigen candidates

Seven of the 19 SCP/TAPS proteins in the DE dataset formed a cluster (Figure 5A, arrowed). Multiple sequence alignment and BLAST searching showed that SCP/TAPS proteins contained several regions of possible *S. stercoralis* species specificity (S8 file) where most regions were within the NIE-homologous region as described below. It also revealed that the sequence HGVPPLTY was conserved across many species and should therefore be excluded from any candidate antigen.

Of the 7 clustered SCP/TAPS proteins, four were predicted to contain epitopes and among these were several with high similarity to the seroantigen 'NIE' (GenBank accession AAD46493)[56], indicating high potential antigenicity. These included SSTP_0001008900, SSTP_0000511800 and SSTP_0001008500 (Figure 6). The region from amino acid (aa) 1-126 of protein 1008900 was species-specific according to sequence alignment with outgroups. BepiPred 1.0-predicted epitopes in this protein were in the region aa 280-376 which does not include the NIE homologous region (not shown), whereas Bcepred did not predict any epitopes in these proteins. BepiPred 2.0 was used with protein SSTP_0001008900 and did predict an epitope within the NIE region from residues 64-101, as well as at aa 240-457.

Figure 6. Amino acid sequence similarity between regions of three *S. stercoralis* SCP/TAPS proteins identified from transcriptomic data and the existing *S. stercoralis* serological antigen NIE (AAD46493).

S. stercoralis SCP/TAPS protein SSTP_0000511800 was modelled to the 3D protein model of a pathogenesis-related protein 1-like (PR-1-like) protein (Figure 7). The Phyre2-generated model, based on two templates, covered 236 residues (91%; residues 17-252) of the sequence with >90% confidence (high probability that sequence and model are homologous). Regions outside the template were modelled *ab initio* which is highly unreliable. Residues 1-22 are a predicted signal peptide sequence on this protein and in SSTP_0001008900.

Figure 7. An *S. stercoralis* SCP/TAPS protein with high homology to seroantigen NIE. Views of the modelled protein (accession number SSTP_0000511800) to illustrate BepiPred 2.0 predicted epitopes spanning the following residues: yellow, 23-33; orange, 41-61; blue, 64-110; red, 156-166; green, 193-204; purple, 242-251. A signal peptide sequence from residues 1-22 that would not be present on the secreted protein is coloured dark grey.

SCP/TAPS protein and NIE homologue, SSTP_0000511800, contained only one potential N-linked glycosylation site at position 2, however this was within the signal peptide (residues 1-22) (Figure 7) and would therefore not be present on the secreted molecule. Eight of the SCP/TAPS, including NIE homologues, also appeared as *S. stercoralis* orthologues of *S. ratti* E/S proteins where they were labelled as CAP domain-containing proteins (S5 file). SCP/TAPS protein sequences, including NIE-homologues, that form candidate coproantigens are detailed in Supplementary files S8 and S9.

TTL coproantigen candidates

Of ten TTL proteins (Fig 5B, arrowed) viewed in alignment, five (SSTP_0000700800, 700900, 701300, 701400 and SSTP_0001222100) appeared to have no sequence similarity to any of the outgroup homologues. Other proteins (SSTP_0001133200, 1222000, 1222200, 1222300 and 0485800) contained one or more regions of species or genus specificity. A total of 11 potential species-specific sequences/peptides were selected (S8 file). BLAST search of these

peptides against the NCBI nr database revealed candidates for *S. stercoralis* specific sequences (Table 4 and S8 file). In addition, it revealed that the amino acid sequence VTCDGKPL in protein SSTP_0000485800 is conserved across many nematode genera and should therefore be avoided in a candidate antigen.

TTL protein SSTP_0000700800 was 3D modelled to a template generated from TTR-52 protein of *C. elegans* (PDB accession: 3UAF) (Figure 8). This model aligned to 89 residues (aa 2-90; 50% of the sequence) with 99.9% confidence. This protein contained a predicted glycosylation site at position N165 and although it was not predicted to have a signal peptide, it was indicated as an 'extracellular or secreted' protein by UniProt (uniprot.org) and is a known E/S protein family, therefore it is more likely to be glycosylated.

Figure 8. A transthyretin-like (TTL) protein of *S. stercoralis* showing surface-exposed antigen regions and N-linked glycosylation sites. Accession number SSTP_0000700800. Blue, BepiPred predicted epitope aa's 124-143; pink, Bcepred predicted epitope aa's 99-123; green, predicted N-linked glycosylation site at N165 in the motif NVS. The Bcepred epitope (pink) extends to residue 141 but has been shortened to show the BepiPred epitope (blue) which overlaps with it.

AChE coproantigen candidates

DE proteins in the AChE family formed two distinct clusters when aligned with BLAST hits from outgroups, one of 4 *S. stercoralis* proteins, the other of 10 (arrowed in Figure 5C). The cluster of 4 consisted of SSTP_0000274700, SSTP_0000671000, SSTP_0000638700 and SSTP_0000670800, all of which contained several regions of potential *S. stercoralis* specificity (examples in S8 file). Equally, the 10 proteins within the other cluster contained multiple regions of possible specificity.

AChE protein SSTP_0000509400 was modelled with 100% confidence to 6 templates which jointly covered aas 17-551 (95%). Peptides with potential *S. stercoralis* species specific sequences were annotated on the model to view their surface exposure (Figure 9). BepiPred 1.0 and Bcepred both failed to identify epitopes. However, BepiPred 2.0 did identify epitope regions, with moderate specificity and surface exposure (Figure 9). AChE protein SSTP_0000509400 contained multiple potential glycosylation sites, three of which,

at positions 38, 89 and 319, were predicted with high confidence on this known glycoprotein (Figure 9).

Figure 9. *S. stercoralis* acetylcholinesterase showing surface-exposed epitope regions and N-linked glycosylation sites. Accession number SSTP_0000509400. Predicted features indicated in colour: blue, BepiPred 2.0 predicted epitope at residues 85-103; and yellow at 390-411; green, predicted N-linked glycosylation sites at N38 (NVT), N89 (NFS) and N319 (NLT). The site at N89 is within the blue epitope sequence.

Aspartic peptidase coproantigen candidates

Only 2 aspartic peptidases appeared in the DE proteins therefore these were analysed for *S. stercoralis* specificity alongside homologous proteins from *S. stercoralis* itself and the outgroups (Figure 5D, arrowed). These two DE proteins (SSTP_0000164500 and 164700) had little sequence similarity with each other and did not cluster with other *S. stercoralis* BLAST hits. However, BLAST of possible species-specific regions from the alignment against NCBI nr revealed higher similarity with other gut nematodes and microflora than we found for other proteins analysed here. One peptide was identified from SSTP_0000164500 that was specific to *S. stercoralis* (Table 4 and S8 file). This protein was modelled with less reliability as only 63% (526 residues) was modelled with confidence, therefore a large part was modelled *ab initio* and was not deemed reliable for selecting epitope regions. No epitopes were predicted in either of the DE aspartic peptidases.

Astacin-like metallopeptidase coproantigen candidates

The phylogenetic tree for the DE astacin-like metallopeptideases did not reveal any significant clustering of *S. stercoralis* proteins (Figure 5F) and no candidate coproantigens or individual proteins were analysed from the astacin-like protein family. However, all *S. stercoralis* proteins were on branches with proteins from other *Strongyloides* species and the closely-related *Parastrongyloides*, separate from other outgroups, and therefore have potential for genus specificity. In addition, this family of proteins was highly represented in the *S. stercoralis* orthologues of *S. ratti* E/S proteome, accounting for 197 (18.6%) of the 1,057 and is one of the protein families with more members in *S. stercoralis* than *S. ratti*.

Prolyl oligopeptidase coproantigen candidates

In the phylogenetic tree, all 5 *S. stercoralis* DE POP proteins clustered with other *Strongyloides* proteins and away from non-*Strongyloides* outgroups, but did not form a single cluster (Figure 5E). Four of these proteins (SSTP_0000289100, 0001108800, 0001108500, 0001019400) were indicated to have more species-specificity, clustering with other *S. stercoralis* BLAST hits (as arrowed in Figure 5E). Of these 4 POP proteins, SSTP_0000289100 contained two Bepipred epitopes and was very highly expressed in parasitic females, as well as being in the E/S orthologues (S3 file). However, the epitope regions had moderate sequence identity to other nematodes, a *Staphylococcus* species and to *Plasmodium vivax*, suggesting widespread conservation. Therefore, although listed in the candidate coproantigen table (S8 file) for completeness, these peptides may not be sufficiently specific to *Strongylodies*.

The other four DE POP family proteins (other than 0000289100), were highly expressed in parasitic females and among the E/S orthologues, despite only two containing a signal peptide. However, they did not contain predicted epitopes (S3 file). One of these (SSTP_0001108800) was analysed for species-specific regions by MSA and subsequently BLASTP against NCBI nr and nematoda (taxid: 6231) databases, which revealed several regions of high specificity for *S. stercoralis*. These are therefore possible coproantigens (Table 4 and S8 file). Amino acid motifs conserved across genera were removed from the sequences originally selected from the MSA, these included: DKLEN, KTDSK, RNAH and DIFAFI.

Other predicted epitopes and existing serological antigens

Among the DE gut-stage proteins identified here, SSTP_0001226800 contained two BepiPred-predicted epitope peptides, one of these, pep120, had high homology to existing seroantigens, including *S. stercoralis* immunoreactive antigen 'SsIR'. It was therefore considered a potential coproantigen for its high antigenicity potential and specificity to *Strongyloides*. This protein was not initially assigned to a protein family but instead given a disorder prediction by BlastKOALA. However 77% of residues were modelled at >90% confidence to 'collagen I alpha 1' protein by Phyre[2] with 86% of the sequence predicted to be disordered (Figure 10). Homology, as analysed by a BLAST search through the WBPS website revealed similarity to a protein described as acetylcholinesterase collagenic tail peptide, also known as ColQ, further supporting the identity of the model template.

Figure 10. *S. stercoralis* candidate coproantigen, a collagenic tail peptide of acetylcholinesterase- ColQ (accession number SSTP_0001226800) with homology to seroantigen 'SsIR'. [A] BepiPred predicted epitopes 'pep120' (blue) and 'pep119' (yellow); green, N-linked glycosylation site at N-terminus. [B] Residues coloured to indicate repeat regions; pale blue, proline; purple, glutamic acid.

The *S. stercoralis* protein, SSTP_0001226800, identified here with high similarity to SsIR as a collagenic tail of AChE, contained a trans-membrane domain or signal peptide and a glycosylation site at residue 53, which would be extracellular and therefore likely to be correctly predicted (Figure 10).

Summary of candidate coproantigens

Table 4 presents selected candidate coproantigen protein regions which satisfy the criteria of being present in stool, specific to *Strongyloides* or *S. stercoralis* and being antigenic. A complete list of candidates resulting from this analysis is given in Supplementary file S8 and amino acid sequences in fasta format in file S9.

Table 4. Selected candidate coproantigens of *S. stercoralis* based on importance and upregulation in parasitism, amino acid sequence specificity to *Strongyloides* or *S. stercoralis* and one or both of the following: containing a predicted epitope, or homology to existing seroantigens. Gene accession numbers are given for WormBase ParaSite or UniProtKB. Additional candidate coproantigens are given in file S8 and full sequences for all are in file S9.

Protein family	Gene	Residues	Length	Evidence/Rationale for selection as potential coproantigen
SCP/TAPS	SSTP_0001008900 SSTP_0000511800	1-256 1-255	256 255	Known highly antigenic, species specific (NIE antigen homologues), E/S orthologues
TTL	SSTP_0000700800	1-177	177	Contains predicted epitopes, surface exposed
AChE	SSTP_0000274700 SSTP_0000509400	187-207 390-411	21 21	DE in gut stages, high *Strongyloides* specificity, E/S orthologue. Predicted epitope, surface exposed
POP	SSTP_0001108800	221-266	46	E/S orthologue, high expression in parasitic female, high *Strongyloides* specificity
Aspartic peptidase	SSTP_0000164500	107-140	34	E/S orthologue, high *Strongyloides* specificity
ColQ protein	SSTP_0001226800	174-266	93	SsIR seroantigen homologue, Predicted epitope
Collagen	SSTP_0001040300	1-123	123	Predicted epitope, likely antigenic coil structure

Short peptide sequences from within the candidate antigen list have been synthesised and are listed in Table 5. Longer candidate coproantigens will be expressed as recombinant proteins.

532 **Table 5.** Peptides being synthesised commercially for testing as coproantigens. Accession
533 numbers are given for WBPS ('SSTP') or GenBank ('AA').

	Protein family	Parent protein	Residues	Length	Peptide
1	SCP/TAPS	SSTP_00005118 00	41-61	21	NGNDYDTKEKLEDAIQKDYPD
2			156-166	11	DLEHDPNNEIE
3			193-204	12	LYDFSKQGHSAE
4	TTL	SSTP_00011332 00	109-123	15	LPFGKITQKPGKDLI
5	POP	SSTP_00002891 00	101-127	27	KPPYCKPPPCKPIPPPTCEPVPPPTCE
6			143-175	33	LKPSKPSKPPKPSTPQKPSTPQKPKTTP KGTTT
7	ColQ / 'SsIR'	AAB97359	1-27	27	NSARVENQDQKDQLENQDQKDQLEN QD
8			28-54	27	QKNQLKNQSENQDQKNQLKNQSENQ DQ
9			55-79	25	KKPIKKPIKKPGPKPIRPIVKPKPK
10		SSTP_00012268 00	183-212	30	PEEPEGPEEPEGPEEPEGPEGPEEPEGP EE
11			240-252	13	PEEPEGPAGPEEP
12			253-266	14	RDDDDGVDEEDERD
13	SCP/TAPS 'NIE'	AAD46493	18-32	15	DYDTKEAMEDAIQRD
14			39-53	15	TFGGDNNNGKKRKID
15			58-73	16	KGNNTFSNKIFDEIWE

534

535

Discussion

Diagnostic needs

The paucity of data on *S. stercoralis* infection prevalence and its low profile compared with the other STH species are largely due to inadequate diagnostics and lack of a single gold standard method. While serology has the highest sensitivity for active disease, it is unsuitable for monitoring treatment outcome or defining cure in a timely manner, remaining positive for months to years[57]. Inadequate treatment and reinfection post-treatment may occur[8, 58, 59]. Therefore, this study aimed to identify specific diagnostic targets from the nematode itself that could be captured by a rapid antigen detection test for use on stool samples. Such coproantigen assays are commercially available for *Giardia* and *Cryptosporidium* and have been developed for a wide range of human and animal parasites including *Fasciola*[60], *Echinococcus*[61], *Strongyloides ratti*[28], *S. venezuelensis*[30], *Opisthorcis*[62] *Toxocara*[63] and *Entamoeba histolytica*[64], among others. These assays have been developed using either somatic, E/S material, or known antigens as targets and many have shown high sensitivity for active infection, as well as high specificity using both polyclonal and monoclonal antibodies.

Protein families and E/S proteomes

We used open access data sources, published literature and freely-accessible online protein analysis tools to shortlist candidate antigens based on three requirements: presence in infected stool; *Strongyloides* or *S. stercoralis* specificity; antigenicity. We focused on proteins that were differentially expressed between gut-dwelling and non-gut-dwelling life stages of *S. stercoralis* in RNA-seq data produced by Stoltzfus *et al.* (2012)[34] and that were from particular protein families. Seven of these protein families were identified by other studies as expanded in the genomes of parasitic nematodes, and upregulated in parasitic life stages[33]. In addition, a study of the hookworm, *Ancylostoma*, E/S proteome identified some of the same protein families, namely: SCP/TAPS, various proteases and TTL protein family members, adding evidence that they are likely to be detectable in stool[65]. We also obtained *S. stercoralis* orthologues of the *S. ratti* E/S proteome as published by Soblik *et al.* (2011)[43] and Hunt *et al.* (2016)[32].

We found limited overlap of 77 proteins (5.9% of the total) between E/S orthologues and DE proteins. This could be due to post-transcriptional gene regulation[32] which means that gene expression does not equal presence of protein, therefore the use of DE proteins would have low specificity for selecting candidate coproantigens. Conversely, our possible oversimplification of the transcriptomic dataset by grouping together life stages with very different gene expression profiles could have underestimated the number of candidates if high expression was not shared by all gut life stages. This would lead to low sensitivity of the method for detecting candidate coproantigens.

E/S proteomics may be a more reliable starting point for coproantigen discovery. However, when we investigated this by viewing gene expression data for all the E/S orthologues, there was no single life stage that accounted for all the parasitic female E/S proteome orthologues and some genes showed very low expression across all life stages. This indicated that although the E/S orthologues were broadly representative of the *S. ratti* E/S proteome dataset, there may be limitations due to species-specific differences. This was evident in the large number of orthologues in *S. stercoralis* (1,057) compared to the original 550 from *S. ratti* which had BLAST hits.

Alternative approaches, in the absence of the E/S proteome of the exact species of interest, would be to exclude orthologues with very low gene expression or no signal peptide, thus focusing on the E/S orthologues that also have transcriptomic evidence and are more likely to be secreted. In addition, applying a more stringent e-value to the BLAST search would ensure only orthologues with the highest identity to E/S proteins are selected. If the E/S proteomic data contains the semi-quantitative relative abundance information, as the data from Hunt *et al.* (2016)[32] does, the highly abundant proteins could be prioritised for identifying orthologues. This was not accounted for in our analysis but there does not appear to be correlation between protein abundance in *S. ratti* E/S proteome and transcription level in *S. stercoralis* orthologues. Therefore differential gene expression between species is also likely to be a complicating factor and ultimately, the E/S proteome of the species of interest would be most desirable for coproantigen discovery. However, shared epitopes between these closely related species still warrant the use of all available data[47].

Priority protein families and species specificity

598 Phylogenetic trees of *S. stercoralis* DE proteins, and their homologues in a selection of
599 outgroups enabled focused analysis of proteins within the priority protein families that may
600 contain genus or species-specific regions.

601 SCP/TAPS

602 The SCP/TAPS protein family is among the CAP domain-containing proteins and is proposed
603 to have a role in modulating the host immune response[33]. In this study we identified 21
604 SCP/TAPS proteins and 6 additional CAP-domain-containing proteins in the DE proteins.
605 Several of these were homologues or potential parent proteins of the serological antigen
606 NIE in the *S. stercoralis* transcriptome. NIE was originally identified by genome library
607 panning[56] and subsequently shown to have homology with insect venom allergens[66]. These
608 allergens in turn are in the CAP domain-containing proteins[67, 68]. We investigated NIE as a
609 potential coproantigen due to its high immunogenicity and specificity to *Strongyloides*.

610 We used the transcriptomic dataset to determine the likelihood of NIE being present in
611 infected stool. The closest sequence homology was with proteins expressed at their highest
612 in infectious larvae from the environment and the subsequent tissue-migrating stage,
613 suggesting that this family of proteins is highly important during skin penetration by the
614 nematode and generates a strong antibody response during the passage of the larvae via
615 host blood and tissue. This may also be the case in autoinfective larvae which would be
616 present and detectable in stool, although these were not represented in the transcriptomic
617 data. However, their presence in E/S orthologous proteins makes NIE and other SCP/TAPS
618 proteins good candidates that can meet all three of our requirements for a coproantigen.

619 The closest homologue of NIE was modelled to an *Ancylostoma caninum* SCP/TAPS 3D
620 structure originally generated by Osman *et al.* (2012)[69]. However, despite being the best
621 match available, the structure model did not include the NIE region at the start of this
622 protein. This could be because the *A. caninum* protein is only a partial protein, or that this
623 region is not present in this hookworm species, thus contributing to its high *S. stercoralis*
624 specificity in serology[70, 71]. In addition, it was unexpected that neither epitope predictor
625 identified the NIE region (in SSTP_0001008900) as an epitope. This is supported by the fact
626 that conformation of epitopes is of utmost importance for B-cell antigens, and suggests
627 that linear sequence should not be relied on alone to guarantee species-specificity.
628 However, computational tools are constantly being improved upon and an updated version
629 of BepiPred (2.0), released after the majority of the present analysis, takes better account
630 of conformational epitopes[72] and did predict an epitope region within the NIE sequence.

A comparison was made by Ramanathan *et al.* (2008)[57] of the serological performance of two NIE-based assays: ELISA and luciferase immunoprecipitation system (LIPS). The full-length protein (229 aa) was used in ELISA and a luciferase fusion protein in the LIPS, comprising 88 residues of NIE. The comparison revealed that better specificity was achieved with the LIPS which may suggest that the full-length antigen generated false positive reactions. Therefore, as a coproantigen, the shorter sequence may also be a more specific target. In addition, co-infecting nematodes which may cause serological cross reactivity e.g. filarial worms, are less likely to produce coproantigens due to their location outside the host gut.

Acetylcholinesterase

Secreted AChE has a possible role in enabling certain parasitic helminths to evade host expulsion mechanisms from mucosal surfaces[33, 73]. The transcriptomic and E/S proteomic data strongly supported this, with AChE family proteins in the E/S orthologues being expressed almost exclusively in the parasitic female life stage (S5 file). Secreted AChE differs from the neuromuscular protein in structure, gene family and substrate, being less specific to acetylcholine[73]. We found moderate homology between the predicted epitope regions of a *S. stercoralis* AChE and other nematode species, however, nematode secreted AChE commonly contains an amino acid sequence insertion (when compared with the well-studied AChE of *Torpedo californica*, the Pacific electric ray) which is surface exposed and differs between species[73]. In the *S. stercoralis* protein analysed here by 3D model (SSTP_0000509400), this region may be the 13 residues at 351-363 (SNLHDYIYNCKLD) which is on the surface, close to the predicted epitope region and glycosylation site. However, this 'insertion' was not sufficiently unique to *Strongyloides* to merit its use as an antigen.

Collagen as a coproantigen

Collagens identified in the DE proteins were also highly represented in the predicted epitopes and contain a similar proline-repeat-induced coil structure and flexibility to ColQ protein.

Sequence similarity of SSTP_0001226800 was also found with four immunoreactive peptides identified by Ramachandran *et al.* (1998)[74] as well as to 'SsIR'[57] (Figure 11). Other homologous DE proteins were SSTP_0000270800 and 0000811600. Ramachandran *et al.* (1998)[74] identified several sequences from a *S. stercoralis* cDNA library using sera from

infected patients originating from Vietnam, Cambodia and Laos[75]. These antigens were named 5a, 8a, 12a and 19a with original accession numbers AAB65139, AAB65140, AAB65141 and AAB65142 respectively, and varied between 90 and 152 residues in length. Subsequently Ramanathan *et al.* (2008)[57] reported the use of SsIR (AAB97359) of 156 aa, for serodiagnosis and as a potential vaccine. We found that SsIR is most similar to peptides 5a and 8a (Figure 11). Through a literature survey we found that Peptide 5a was identified among peptides digested from the surface of L3 larvae in a proteomic study[76].

Figure 11. Epitope peptide 120 (pep120) from *S. stercoralis* ColQ protein SSTP_0001226800, aligned with seroreactive antigens 5a, 8a, 12a, 19a (Ramachandran et al., 1998) and SsIR (Ramanathan et al., 2008).

Collagen has not been found as a secreted protein but is instead a main component of the nematode cuticle[77]. Analysis of the DE transcriptome revealed that collagen was most highly expressed in iL3 and L3+ life stages (S3 file), possibly linked with moults between larval stages[77] but oddly not upregulated in all larval stages. As a coproantigen, collagen could enable detection of components of whole or parts of excreted worms. We have suggested a collagen sequence as a coproantigen (S8 and S9 files). However the nematode cuticle is also glycan rich[78], therefore cuticle collagen protein antigens may not be freely-accessible to antibodies, whereas the glycans may form a better target coproantigen in this case.

Epitope prediction

We performed epitope prediction on the DE proteins and E/S proteome homologues using two open access online tools that yielded many predicted epitope peptides. An alternative to this would be to scan the entire genome for epitopes, a method applied to vaccine candidate discovery[79]. The challenge faced by this approach is the complexity of conformational epitopes compared with linear peptide epitopes. Antibodies frequently bind to conformational epitopes formed by the 3D structure of the antigen which therefore cannot easily be detected by sequence analysis alone[72]. The use of information on existing seroantigens is beneficial in informing antigenicity but does not confirm their presence in stool, indeed the SsIR and NIE homologues were expressed at very high levels in infectious L3 and tissue migrating larvae and much less in gut dwelling life stages, therefore the

availability of E/S proteome data to validate life stage protein secretion is crucially important to the search for coproantigens.

The availability of 3D protein models can assist with selecting conformational epitopes by modelling a sequence onto the structure of a homologous protein and revealing adjacent amino acids on the surface of the protein[80]. Models do not necessarily have high sequence identity to the query sequence but this does not decrease the confidence in the model. Confidence in 3D models of >90% indicates that the protein adopts the overall folds of the model but may differ from the native protein in surface loops[53], thus this method is not guaranteed, but provides a good indication for selecting candidate antigens. *Ab initio*-modelled regions, where the sequence was not covered by the model, have very poor accuracy and should therefore be interpreted with caution and not used as the sole basis for selecting conformational epitopes.

Differences in predicted epitopes were seen based on a computational versus 'by eye' approach to selecting epitope regions. The DE protein dataset contained longer predicted epitopes due to the decision, where relevant, to extend predicted epitopes across a short region of lower epitope score whereas the computational selection worked only on the exact score threshold and would not join two adjacent high scoring regions.

Glycoprotein antigens were not considered in this study, apart from potential N-linked glycosylation sites on candidate protein antigens. Glycans form existing species-specific, highly antigenic diagnostic antigens, including CCA and CAA of *Schistosoma mansoni* and *Schistosoma* genus trematodes respectively[81], LAM of *Mycobacterium tuberculosis*[82] and have been implicated in lysate seroantigen of *S. stercoralis*[83]. In helminths, glycan structures may not only be species-specific, but also life-stage specific[84]. Glycans may obscure some of the protein epitopes predicted here, particularly in the secreted AChE, which is highly glycosylated. This is to be expected as glycans form many of the host-parasite interactions[84]. In addition, secreted candidate antigens may also contain O-linked glycans, via oxygen atoms of serine or threonine, which are not easily predicted. Although we have excluded potential glycan epitopes, they could be accounted for to some extent by expressing antigens of interest, ideally in a closely-related system, potentially *C. elegans* or even *Strongyloides* itself[85]. As an alternative, glycans could be excluded altogether by synthesising peptides or expressing recombinant proteins in bacteria, thus focusing the antigen search purely on proteins, as we have done here.

Disorder prediction and antigenicity

Disordered proteins are proteins with a mobile structure that may alter conformation depending on interactions with other molecules[86]. For this reason, their structures and epitopes may be less reliably predicted[87]. We identified a disordered protein containing a signal peptide and predicted epitope region. This protein did not occur in the E/S orthologues but had very high gene expression in iL3 larvae. Further evidence for its high antigenicity came from sequence similarity to the serological antigen 'SsIR'. SsIR and the identified matches/parent proteins contained repeats of the proline-rich sequence PEEPEG, suggesting a twisting or helical structure due to the distortion provided by proline. This was confirmed by the 3D model template, an AChE collagenic tail peptide. These peptides, also known as ColQ, bind to AChE in the neuromuscular junction and link it to an anchoring protein, MuSK, in the cell membrane[88]. However, immunoreactive peptide '5a' (NCBI accession 2290388/ AAB65139), a homologue of SsIR, was identified in a trypsin digest of undamaged whole larvae[76], and antiserum against SsIR was seen to bind to the outside of the nematode cuticle[89], both suggesting that ColQ is exposed on the worms surface and would thus be an ideal target as a coproantigen.

Epitope regions were predicted in the SsIR homologous sequences, probably due to its twisting structure and repeated sequence, both of which are known to create strong epitopes and are a feature of other serological antigens including rK39 for visceral leishmaniasis diagnosis, and immunoreactive *Trypanosoma cruzi* proteins[90, 91]. ColQ in neuromuscular junctions is a coil of three molecules therefore, conformational epitopes may be formed by the proximity of these, rather than, or as well as, those on a single protein. The mammalian (rat) ColQ protein consists of a proline-rich sequence, however, the intervening amino acids differ from those in *S. stercoralis*, therefore the conformational epitopes of this triple helix are also likely to differ[92]. This ColQ was predicted with high reliability to be glycosylated via residue N53, which may create additional epitopes but seems unlikely to obscure predicted protein epitopes due to the physical separation.

Future work

The next stage in coproantigen assay development using the protein and peptide candidates detailed here in Table 4 and Supplementary file S8 is the synthesis of those that are short peptides, up to 30 residues, or, if longer to express and purify them as recombinant proteins in a bacterial expression system. To target correctly glycosylated

antigens, the coproantigens could potentially be expressed and purified from nematode systems but we have not assessed the potential specificity of glycans in the present study.

The antigens could then be screened for immuno-reactivity, as a proxy for antigenicity, using known positive human serum or serum from a laboratory infected animal in a simple direct ELISA format. Highly reactive candidates would then be synthesised at sufficient quantity and high purity, and inoculated into mice or rabbits to generate antiserum. Antibodies should be purified from the serum and can be affinity purified against the original antigen for better specificity, prior to being developed into an antigen capture pair. As the target format of this assay is a rapid dipstick test, the successful antibody pair(s) would be optimised in a lateral-flow format on nitrocellulose and gold or latex nanoparticles. In addition, the assay could be optimised in ELISA format. Antigen targets may also be combined to increase sensitivity and specificity of the coproantigen assay.

The approach presented here to identify candidate diagnostic antigens using open access 'omic' data could feasibly be applied to other nematode and indeed non-nematode helminth infections, including animal parasites. Genomic data availability is improving and increasing at a rapid pace. Alongside this are proteomic studies of E/S material from multiple helminths. This data, and other genome analysis, have revealed protein families common to the E/S proteome of many parasitic species: SCP/TAPS, astacins and TTL, among others.

Other potential proteins of interest for coproantigen detection assays include those identified in E/S material of other helminths. For example: enolase, common to *Schistosoma japonicum*[93], *Echinostoma* and *Fasciola*[94], *Onchocerca*[95] and *Trichuris*[96], as well as in the *S. stercoralis* E/S orthologous proteins reported here and with constitutive high expression across all life stages (S5 file); protein 14-3-3 from E/S and somatic extract of *Strongyloides, Ascaris, Schistosoma* and *Ancylostoma*[65, 76, 93, 97, 98]. A summary of E/S proteomic studies is presented by Ditgen *et al.* (2014)[99] which collectively may reveal other common E/S protein family candidates. Other proteins have been tested as vaccine candidates against *S. stercoralis*: sodium potassium ATPase (SsEAT), tropomyosin, and a galectin (LEC-5)[89, 100] suggesting that these may also generate effective antibodies for antigen capture, providing they are detectable in stool. In addition, as-yet uncharacterised genes[32] may encode ideal targets because of their potential species specificity due to low homology with known proteins.

Arifin *et al.* (2018)[101] identified a previously unreported seroantigen with high sensitivity and specificity to IgG4 of *S. stercoralis* infected individuals. As with the discovery of NIE and SsIR, they screened a cDNA library, in this case identifying the reactive protein as similar to *S. ratti* immunoglobulin binding protein 1 (BiP) in the TAP42-like family, which they designated 'rSs1a'. This is a cytoplasmic phosphatase-associated protein, involved in translational control in *C. elegans*[102] and unless it has multiple purposes in parasitic nematodes, seems unlikely to be secreted. We found that the *S. ratti* homologue of this protein, as given by Arifin *et al.* (2018)[101] (CEF66010.1), contained multiple alpha helices when modelled to a protein phosphatase subunit template, which was confirmed by structural studies of the *Saccharomyces* yeast TAP42[103]. We identified a single orthologue of the *S. ratti* protein, encoded in the *S. stercorlais* genome (SSTP_0000328000). The highly coiled structure suggests that one or more of these regions could form possible epitopes for a coproantigen assay that may detect somatic worm antigen in faeces.

Stool-based rapid immunoassays could potentially be adapted to detect parasite antigens in more readily available samples such as urine, which has been found to contain antigens of other helminth infections[104] and *Strongyloides* DNA[105]. Whatever the sample type, there is a need for well-characterised sample collections in order to pursue diagnostics development, particularly in this case, stool samples preserved in 10% formalin stored at -20°C or below for preservation of protein antigens[28].

In addition to coproantigen assay development, which can be informed by existing seroantigens, the analysis conducted here of antigenicity, species specificity and life stage expression of proteins can indicate potential seroantigens. The development of a coproantigen assay will include an antigenicity screening step where candidates are screened in ELISA format against well-characterised seropositive and seronegative sera, prior to raising antibodies. This will reveal potentially strong seroreactive antigens. Other DE proteins that were not present in the parasitic female E/S orthologues and have highest expression in non-gut dwelling life stages could also be pursued as potential seroantigens.

Antigenic variation should be taken into account when an antigen has been selected, due to geographic differences between nematode strains[106]. This has impacted on diagnosis and vaccination for other parasitic infections[107, 108] and can be investigated by amplifying and sequencing the genes of interest from a wide geographic range of samples. For *S. stercoralis*, this is especially relevant as the reference genome strain PV001 (also

sometimes denoted PV0001 or PV1) originates from a dog infection with UPD (University of Pennsylvania dog) strain[109], which may differ from human-infective strains[110].

Summary

This work presents detailed analysis of *S. stercoralis* proteins leading to the selection of diagnostic coproantigens. We have identified multiple *S. stercoralis* candidate protein antigen sequences with evidence for their specificity to *Strongyloides* or *S. stercoralis* from phylogenetic and sequence comparison with relevant other species, as well as serological specificity reported in the literature. Evidence supporting their presence in infected stool was assessed by belonging to parasitism-associated protein families, upregulation in gut-dwelling life stages, presence in E/S material of other helminths and in *S. stercoralis* orthologues of the *S. ratti* E/S proteome. Antigenicity was predicted using epitope prediction tools, information from existing seroantigens and 3D structure modelling. Short peptide candidate antigens analysed and presented here are currently being synthesised for screening as coproantigens. Longer peptides and proteins form promising candidates for expressing as recombinant proteins, raising antibodies against and capturing *S. stercoralis* antigen in stool.

Funding statement

TM was funded by the Sir Halley Stewart Trust (http://www.sirhalleystewart.org.uk/). The views expressed within this report are those of the authors and not necessarily those of the Trust. TM was additionally supported by the John Henry Memorial Fund (http://www.johnhenrymf.org/).

Supplementary files

These files are available in the accompanying CD.

S1 table: Sources and details of open access data used in this study.

S2 table: BLAST conditions and number of sequences included in phylogenetic trees of each protein family.

S3 spreadsheet: *S. stercoralis* genes differentially expressed in gut-dwelling versus non-gut-dwelling life stages of the nematode (n = 328) and those containing predicted epitopes. Transcriptomic data from Stoltzfus et al. (2012). Gene accession numbers (starting SSTP) are given for WormBase Parasite or UniProtKB and life stage transcriptomic data accession numbers (starting ERR) for NCBI SRA. Expression levels are in normalized read counts.

S4 spreadsheet: BLAST output- *S. stercoralis* orthologues of *S. ratti* E/S proteins at 1e-50.

S5 spreadsheet: *S. stercoralis* orthologues of *S. ratti* E/S proteins – gene expression, protein families, epitope predictions and signal peptide predictions.

S6 spreadsheet: Predicted epitope sequences from E/S orthologues and DE genes with length and position in parent proteins, and overlaps between BepiPred and Bcepred.

S7 text file: *S. stercoralis* predicted epitope sequences from DE proteins and E/S orthologues using two prediction tools, in fasta format.

S8 table: Details of *S. stercoralis* candidate coproantigens selected for capture assay development.

S9 text file: *S. stercoralis* candidate coproantigens for capture assay development, in fasta format.

References

1. Schar, F., et al., *Strongyloides stercoralis*: global distribution and risk factors. PLoS Negl Trop Dis, 2013. **7**(7): p. e2288.
2. Forrer, A., et al., *Strongyloides stercoralis* and hookworm co-infection: spatial distribution and determinants in Preah Vihear Province, Cambodia. Parasit Vectors, 2018. **11**(1): p. 33.
3. Buonfrate, D., et al., Epidemiology of *Strongyloides stercoralis* in northern Italy: results of a multicentre case-control study, February 2013 to July 2014. Euro Surveill, 2016. **21**(31).
4. Becker, S.L., et al., Diagnosis, clinical features, and self-reported morbidity of *Strongyloides stercoralis* and hookworm infection in a co-endemic setting. PLoS Negl Trop Dis, 2011. **5**(8): p. e1292.
5. Minciullo, P.L., Cascio, A., and Gangemi, S., Association between urticaria and nematode infections. Allergy Asthma Proc, 2018. **39**(2): p. 86-95.
6. Atkins, N.S., et al., Humoral responses in human strongyloidiasis: correlations with infection chronicity. Trans R Soc Trop Med Hyg, 1997. **91**(5): p. 609-13.
7. Abrescia, F.F., et al., Reemergence of strongyloidiasis, northern Italy. Emerg Infect Dis, 2009. **15**(9): p. 1531-3.
8. Repetto, S.A., et al., Strongyloidiasis outside endemic areas: long-term parasitological and clinical follow-up after ivermectin treatment. Clin Infect Dis, 2018: p. cix1069-cix1069.
9. Kinjo, T., et al., Acute respiratory distress syndrome due to *Strongyloides stercoralis* infection in a patient with cervical cancer. Intern Med, 2015. **54**(1): p. 83-7.
10. Tefé-Silva, C., et al., Hyperinfection syndrome in strongyloidiasis, in Curr Top Trop Med. 2012.
11. Buonfrate, D., et al., Severe strongyloidiasis: a systematic review of case reports. BMC Infect Dis, 2013. **13**: p. 78.
12. Luvira, V., Watthanakulpanich, D., and Pittisuttithum, P., Management of *Strongyloides stercoralis*: a puzzling parasite. Int Health, 2014. **6**(4): p. 273-81.
13. Schar, F., et al., *Strongyloides stercoralis* larvae excretion patterns before and after treatment. Parasitology, 2014. **141**(7): p. 892-7.

14. Henriquez-Camacho, C., et al., Ivermectin versus albendazole or thiabendazole for *Strongyloides stercoralis* infection. Cochrane Database Syst Rev, 2016(1): p. Cd007745.
15. Palmeirim, M.S., et al., Efficacy and safety of co-administered ivermectin plus albendazole for treating soil-transmitted helminths: a systematic review, meta-analysis and individual patient data analysis. PLoS Negl Trop Dis, 2018. **12**(4): p. e0006458.
16. Marti, H., et al., A comparative trial of a single-dose ivermectin versus three days of albendazole for treatment of *Strongyloides stercoralis* and other soil-transmitted helminth infections in children. Am J Trop Med Hyg, 1996. **55**(5): p. 477-81.
17. Barda, B., et al., Efficacy of moxidectin versus ivermectin against *Strongyloides stercoralis* infections: a randomized controlled non-inferiority trial. Clin Infect Dis, 2017.
18. Koga, K., et al., A modified agar plate method for detection of *Strongyloides stercoralis*. Am J Trop Med Hyg, 1991. **45**(4): p. 518-21.
19. Meurs, L., et al., Diagnosing polyparasitism in a high-prevalence setting in Beira, Mozambique: detection of intestinal parasites in fecal samples by microscopy and real-time PCR. PLoS Negl Trop Dis, 2017. **11**(1): p. e0005310.
20. Repetto, S.A., et al., Comparison between PCR and larvae visualization methods for diagnosis of *Strongyloides stercoralis* out of endemic area: a proposed algorithm. Acta Trop, 2016. **157**: p. 169-77.
21. Incani, R.N., et al., Diagnosis of intestinal parasites in a rural community of Venezuela: advantages and disadvantages of using microscopy or RT-PCR. Acta Trop, 2016.
22. Becker, S.L., et al., Real-time PCR for detection of *Strongyloides stercoralis* in human stool samples from Cote d'Ivoire: diagnostic accuracy, inter-laboratory comparison and patterns of hookworm co-infection. Acta Trop, 2015. **150**: p. 210-217.
23. Repetto, S.A., et al., An improved DNA isolation technique for PCR detection of *Strongyloides stercoralis* in stool samples. Acta Trop, 2013. **126**(2): p. 110-4.
24. Buonfrate, D., et al., Novel approaches to the diagnosis of *Strongyloides stercoralis* infection. Clin Microbiol Infect, 2015.
25. Buonfrate, D., et al., Accuracy of five serologic tests for the follow up of *Strongyloides stercoralis* infection. PLoS Negl Trop Dis, 2015. **9**(2): p. e0003491.
26. Vargas, P., et al., Serologic monitoring of public health interventions against *Strongyloides stercoralis*. Am J Trop Med Hyg, 2017. **97**(1): p. 166-172.
27. Lindo, J.F., et al., Parasite-specific serum IgG following successful treatment of endemic strongyloidiasis using ivermectin. Trans R Soc Trop Med Hyg, 1996. **90**(6): p. 702-3.
28. Sykes, A.M. and McCarthy, J.S., A coproantigen diagnostic test for *Strongyloides* infection. PLoS Negl Trop Dis, 2011. **5**(2): p. e955.
29. Nageswaran, C., Craig, P.S., and Devaney, E., Coproantigen detection in rats experimentally infected with *Strongyloides ratti*. Parasitology, 1994. **108 (Pt 3)**: p. 335-42.
30. Goncalves, A.L., et al., A new faecal antigen detection system for *Strongyloides venezuelensis* diagnosis in immunosuppressed rats. Exp Parasitol, 2010. **125**(4): p. 338-41.

31. Rappuoli, R., Reverse vaccinology. Curr Opin Microbiol, 2000. **3**(5): p. 445-50.
32. Hunt, V.L., et al., The genomic basis of parasitism in the *Strongyloides* clade of nematodes. Nat Genet, 2016.
33. Hunt, V.L., et al., The genome of *Strongyloides* spp. gives insights into protein families with a putative role in nematode parasitism. Parasitology, 2016: p. 1-16.
34. Stoltzfus, J.D., et al., RNAseq analysis of the parasitic nematode *Strongyloides stercoralis* reveals divergent regulation of canonical dauer pathways. PLoS Negl Trop Dis, 2012. **6**(10): p. e1854.
35. Li, B. and Dewey, C.N., RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 2011. **12**: p. 323.
36. Thompson, J.D., Higgins, D.G., and Gibson, T.J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 1994. **22**(22): p. 4673-80.
37. Letunic, I. and Bork, P., Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res, 2016. **44**(W1): p. W242-5.
38. Boratyn, G.M., et al., Domain enhanced lookup time accelerated BLAST. Biol Direct, 2012. **7**: p. 12.
39. Hunter, S., et al., InterPro: the integrative protein signature database. Nucleic Acids Res, 2009. **37**(Database issue): p. D211-5.
40. Sigrist, C.J., et al., PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res, 2010. **38**(Database issue): p. D161-6.
41. Sigrist, C.J., et al., New and continuing developments at PROSITE. Nucleic Acids Res, 2013. **41**(Database issue): p. D344-7.
42. Kanehisa, M., et al., KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res, 2016. **44**(D1): p. D457-62.
43. Soblik, H., et al., Life cycle stage-resolved proteomic analysis of the excretome/secretome from *Strongyloides ratti* - identification of stage-specific proteases. Mol Cell Proteomics, 2011. **10**(12): p. M111.010157.
44. Li, H., et al., The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009. **25**(16): p. 2078-9.
45. Oliveros, J.C. Venny. An interactive tool for comparing lists with Venn's diagrams. 2007-2015 [cited 12/7/18]; Available from: http://bioinfogp.cnb.csic.es/tools/venny/index.html.
46. Petersen, T.N., et al., SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods, 2011. **8**(10): p. 785-6.
47. Eamudomkarn, C., et al., Comparative evaluation of *Strongyloides ratti* and *S. stercoralis* larval antigen for diagnosis of strongyloidiasis in an endemic area of opisthorchiasis. Parasitol Res, 2015.
48. Sousa, J.E., et al., Serological cross-reactivity between *Strongyloides venezuelensis* and *Syphacia muris* in Wistar rats (*Rattus norvegicus*). Parasitol Int, 2016. **65**(2): p. 137-45.
49. Larsen, J.E., Lund, O., and Nielsen, M., Improved method for predicting linear B-cell epitopes. Immunome Res, 2006. **2**: p. 2.

50. Saha, S. and Raghava, G.P.S. BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. in Artificial Immune Systems: Third International Conference, ICARIS 2004. 2004. Catania, Sicily, Italy: Springer Berlin Heidelberg.
51. BepiPred website. Output format. [cited 2018; Available from: http://www.cbs.dtu.dk/services/BepiPred-1.0/output.php.
52. Bcepred website. Bcepred: prediction of linear B-cell epitopes, using physico-chemical properties. [cited 2018; Available from: http://crdd.osdd.net/raghava/bcepred/.
53. Kelley, L.A., et al., The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protocols, 2015. **10**(6): p. 845-858.
54. Pettersen, E.F., et al., UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem, 2004. **25**(13): p. 1605-12.
55. Gupta, R. and Brunak, S., Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput, 2002: p. 310-22.
56. Ravi, V., et al., Characterization of a recombinant immunodiagnostic antigen (NIE) from *Strongyloides stercoralis* L3-stage larvae. Mol Biochem Parasitol, 2002. **125**(1-2): p. 73-81.
57. Ramanathan, R., et al., A luciferase immunoprecipitation systems assay enhances the sensitivity and specificity of diagnosis of *Strongyloides stercoralis* infection. J Infect Dis, 2008. **198**(3): p. 444-51.
58. Khieu, V., et al., *Strongyloides stercoralis* infection and re-infection in a cohort of children in Cambodia. Parasitol Int, 2014. **63**(5): p. 708-12.
59. Forrer, A., et al., Ivermectin treatment and sanitation effectively reduce *Strongyloides stercoralis* infection risk in rural communities in Cambodia. PLoS Negl Trop Dis, 2016. **10**(8): p. e0004909.
60. Sarkari, B. and Khabisi, S.A., Immunodiagnosis of human fascioliasis: an update of concepts and performances of the serological assays. J Clin Diagn Res, 2017. **11**(6): p. Oe05-oe10.
61. Morel, N., et al., A monoclonal antibody-based copro-ELISA kit for canine echinococcosis to support the PAHO effort for hydatid disease control in South America. PLoS Negl Trop Dis, 2013. **7**(1): p. e1967.
62. Teimoori, S., et al., Chicken IgY-based coproantigen capture ELISA for diagnosis of human opisthorchiasis. Parasitol Int, 2016.
63. Elsemore, D.A., et al., Enzyme-linked immunosorbent assays for coproantigen detection of *Ancylostoma caninum* and *Toxocara canis* in dogs and *Toxocara cati* in cats. J Vet Diagn Invest, 2017. **29**(5): p. 645-653.
64. Haque, R., et al., Diagnosis of pathogenic *Entamoeba histolytica* infection using a stool ELISA based on monoclonal antibodies to the galactose-specific adhesin. J Infect Dis, 1993. **167**(1): p. 247-9.
65. Mulvenna, J., et al., Proteomics analysis of the excretory/secretory component of the blood-feeding stage of the hookworm, *Ancylostoma caninum*. Mol Cell Proteomics, 2009. **8**(1): p. 109-21.
66. Ravi, V., et al., *Strongyloides stercoralis* recombinant NIE antigen shares epitope with recombinant Ves v 5 and Pol a 5 allergens of insects. Am J Trop Med Hyg, 2005. **72**(5): p. 549-53.

67. Yamazaki, Y. and Morita, T., Structure and function of snake venom cysteine-rich secretory proteins. Toxicon, 2004. **44**(3): p. 227-31.
68. Cantacessi, C., et al., A portrait of the "SCP/TAPS" proteins of eukaryotes — Developing a framework for fundamental research and biotechnological outcomes. Biotechnol Adv, 2009. **27**(4): p. 376-388.
69. Osman, A., et al., Hookworm SCP/TAPS protein structure—A key to understanding host–parasite interactions and developing new interventions. Biotechnol Adv, 2012. **30**(3): p. 652-657.
70. Bisoffi, Z., et al., Diagnostic accuracy of five serologic tests for *Strongyloides stercoralis* infection. PLoS Negl Trop Dis, 2014. **8**(1): p. e2640.
71. Krolewiecki, A.J., et al., Improved diagnosis of *Strongyloides stercoralis* using recombinant antigen-based serologies in a community-wide study in northern Argentina. Clin Vaccine Immunol, 2010. **17**(10): p. 1624-30.
72. Jespersen, M.C., et al., BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res, 2017. **45**(W1): p. W24-W29.
73. Selkirk, M.E., Lazari, O., and Matthews, J.B., Functional genomics of nematode acetylcholinesterases. Parasitology, 2005. **131**(S1): p. S3-S18.
74. Ramachandran, S., et al., Recombinant cDNA clones for immunodiagnosis of strongyloidiasis. J Infect Dis, 1998. **177**(1): p. 196-203.
75. Nutman, T.B., et al., Eosinophilia in Southeast Asian refugees: evaluation at a referral center. J Infect Dis, 1987. **155**(2): p. 309-13.
76. Marcilla, A., et al., Proteomic analysis of *Strongyloides stercoralis* L3 larvae. Parasitology, 2010. **137**(10): p. 1577-83.
77. Kingston, I.B., Nematode collagen genes. Parasitol Today, 1991. **7**(1): p. 11-5.
78. Page, A.P. and Johnstone, I.L., The cuticle, in WormBook, J.M. Kramer and D.G. Moerman, Editors. 2007.
79. Goodswen, S.J., Kennedy, P.J., and Ellis, J.T., Vacceed: a high-throughput *in silico* vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. Bioinformatics, 2014. **30**(16): p. 2381-3.
80. Russi, R.C., et al., *In silico* prediction of T- and B-cell epitopes in PmpD: first step towards to the design of a *Chlamydia trachomatis* vaccine. Biomed J, 2018. **41**(2): p. 109-117.
81. Corstjens, P.L., et al., Tools for diagnosis, monitoring and screening of *Schistosoma* infections utilizing lateral-flow based assays and upconverting phosphor labels. Parasitology, 2014. **141**(14): p. 1841-55.
82. Lawn, S.D., Point-of-care detection of lipoarabinomannan (LAM) in urine for diagnosis of HIV-associated tuberculosis: a state of the art review. BMC Infect Dis, 2012. **12**: p. 103.
83. Ines, E.J., et al., The role of glycosylated epitopes in the serodiagnosis of *Strongyloides stercoralis* infection. Diagn Microbiol Infect Dis, 2013. **76**(1): p. 31-5.
84. Hokke, C.H. and van Diepen, A., Helminth glycomics - glycan repertoires and host-parasite interactions. Mol Biochem Parasitol, 2017. **215**: p. 47-57.
85. Lok, J.B., et al., Transgenesis in *Strongyloides* and related parasitic nematodes: historical perspectives, current functional genomic applications and progress towards gene disruption and editing. Parasitology, 2016: p. 1-16.

86. Oldfield, C.J. and Dunker, A.K., Intrinsically disordered proteins and intrinsically disordered protein regions. Annu Rev Biochem, 2014. **83**: p. 553-84.
87. MacRaild, C.A., et al., Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. PLoS One, 2015. **10**(3): p. e0119899.
88. Campanari, M.-L., et al., Neuromuscular junction impairment in amyotrophic lateral sclerosis: reassessing the role of acetylcholinesterase. Front Mol Neurosci, 2016. **9**(160).
89. Abraham, D., et al., Immunization with the recombinant antigen Ss-IR induces protective immunity to infection with *Strongyloides stercoralis* in mice. Vaccine, 2011. **29**(45): p. 8134-40.
90. Burns, J.M., et al., Molecular characterization of a kinesin-related antigen of *Leishmania chagasi* that detects specific antibody in African and American visceral leishmaniasis. Proc Natl Acad Sci U S A, 1993. **90**(2): p. 775-779.
91. Goto, Y., Carter, D., and Reed, S.G., Immunological dominance of *Trypanosoma cruzi* tandem repeat proteins. Infect Immun, 2008. **76**(9): p. 3967-74.
92. Krejci, E., et al., The mammalian gene of acetylcholinesterase-associated collagen. Vol. 272. 1997. 22840-7.
93. Gao, H., et al., Assessment of the diagnostic efficacy of enolase as an indication of active infection of *Schistosoma japonicum*. Parasitol Res, 2016. **115**(1): p. 151-164.
94. Marcilla, A., et al., Extracellular vesicles from parasitic helminths contain specific excretory/secretory proteins and are internalized in intestinal host cells. PLoS One, 2012. **7**(9): p. e45974.
95. Eberle, R., et al., Isolation, identification and functional profile of excretory-secretory peptides from *Onchocerca ochengi*. Acta Trop, 2015. **142**: p. 156-66.
96. Liu, Z., et al., Identification, characterization, and structure of Tm16 from *Trichuris muris*. J Parasitol Res, 2017. **2017**: p. 4342789.
97. Wang, T., et al., Proteomic analysis of the excretory-secretory products from larval stages of *Ascaris suum* reveals high abundance of glycosyl hydrolases. PLoS Negl Trop Dis, 2013. **7**(10): p. e2467.
98. Rodpai, R., et al., Identification of antigenic proteins in *Strongyloides stercoralis* by proteomic analysis. Parasitol Res, 2017.
99. Ditgen, D., et al., Harnessing the helminth secretome for therapeutic immunomodulators. Vol. 2014. 2014. 964350.
100. Kerepesi, L.A., et al., DNA immunization with Na+-K+ ATPase (Sseat-6) induces protective immunity to larval *Strongyloides stercoralis* in mice. Infect Immun, 2005. **73**(4): p. 2298-305.
101. Arifin, N., et al., Identification and preliminary evaluation of a novel recombinant protein for serodiagnosis of strongyloidiasis. Am J Trop Med Hyg, 2018.
102. Rhoads, R.E., Dinkova, T.D., and Korneeva, N.L., Mechanisms and regulation of translation in *C. elegans*, in WormBook, The *C. elegans* Research Community, Editor. 2006.
103. Yang, J., et al., The structure of Tap42/alpha4 reveals a tetratricopeptide repeat-like fold and provides insights into PP2A regulation. Biochemistry, 2007. **46**(30): p. 8807-15.
104. Worasith, C., et al., Advances in the diagnosis of human opisthorchiasis: development of *Opisthorchis viverrini* antigen detection in urine. PLoS Negl Trop Dis, 2015. **9**(10): p. e0004157.

105. Lodh, N., et al., Diagnosis of Strongyloides stercoralis: detection of parasite-derived DNA in urine. Acta Trop, 2016.
106. Qiang, S., et al., Variation between ASP-1 molecules from *Ancylostoma caninum* in China and the United States. J Parasitol, 2000. **86**(1): p. 181-5.
107. Neafsey, D.E., et al., Genetic diversity and protective efficacy of the RTS,S/AS01 malaria vaccine. N Engl J Med, 2015. **373**(21): p. 2025-2037.
108. Bhattacharyya, T., Boelaert, M., and Miles, M.A., Comparison of visceral leishmaniasis diagnostic antigens in African and Asian *Leishmania donovani* reveals extensive diversity and region-specific polymorphisms. PLoS Negl Trop Dis, 2013. **7**(2): p. e2057.
109. Stoltzfus, J.D., et al., *Strongyloides stercoralis age-1*: a potential regulator of infective larval development in a parasitic nematode. PLoS One, 2012. **7**(6): p. e38587.
110. Jaleta, T.G., et al., Different but overlapping populations of *Strongyloides stercoralis* in dogs and humans - dogs as a possible source for zoonotic strongyloidiasis. PLoS Negl Trop Dis, 2017. **11**(8): p. e0005752.

Figures

Figure 1.

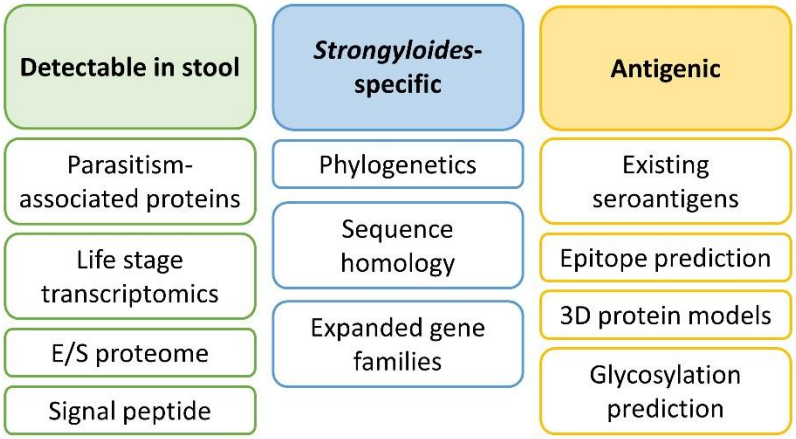


Figure 2.

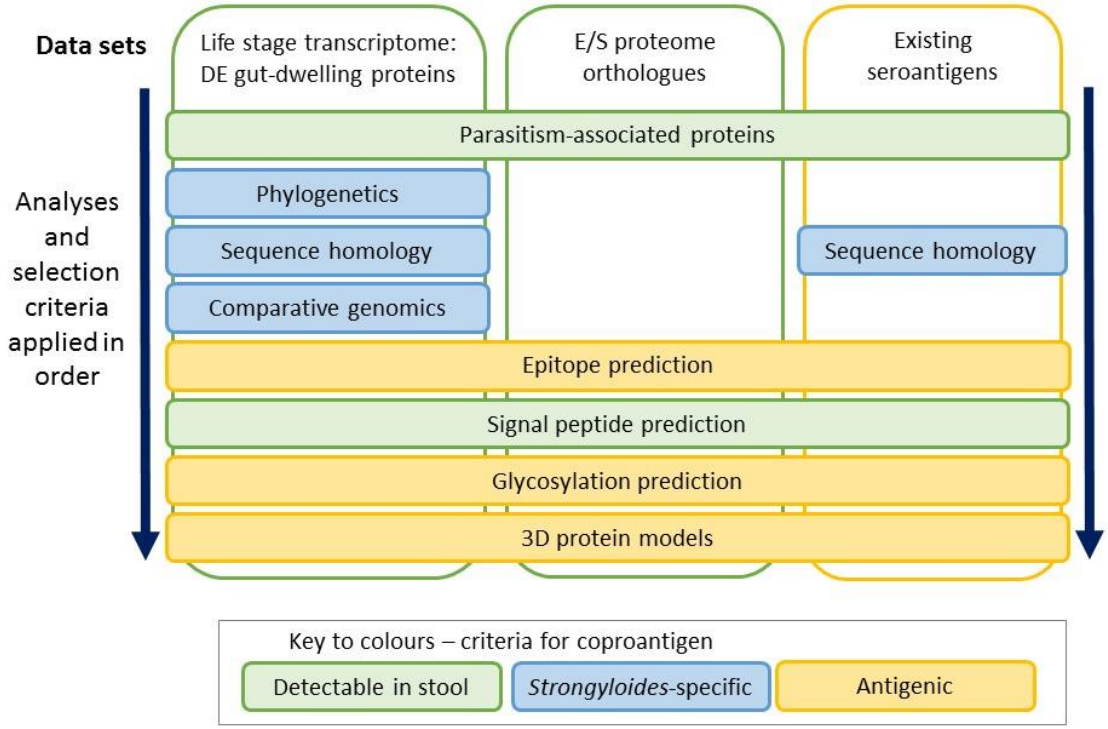


Figure 3.

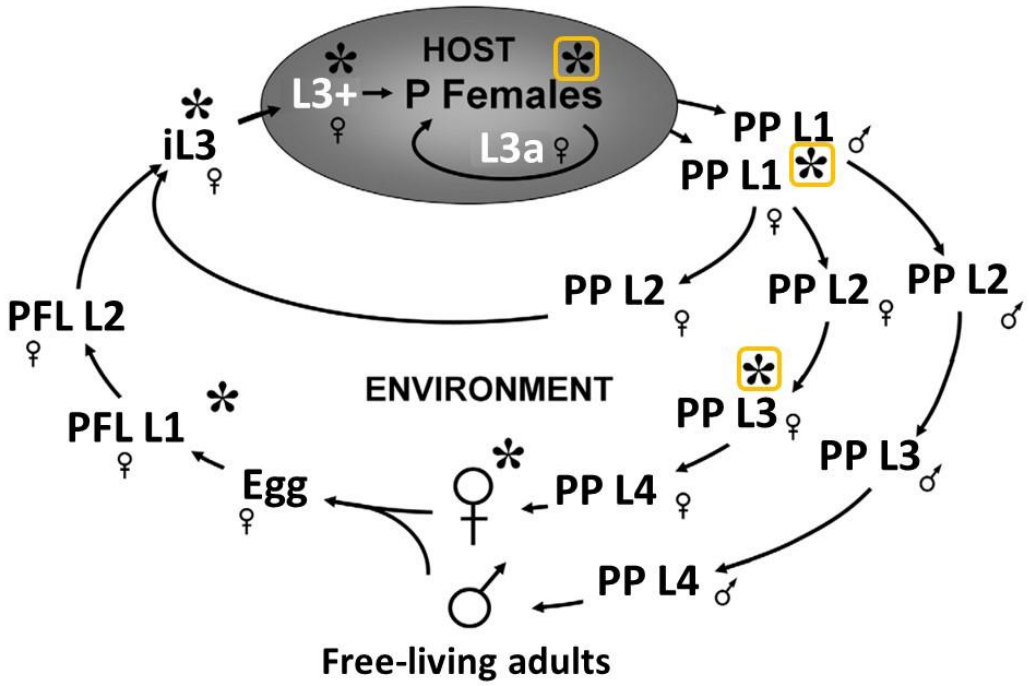


Figure 4.

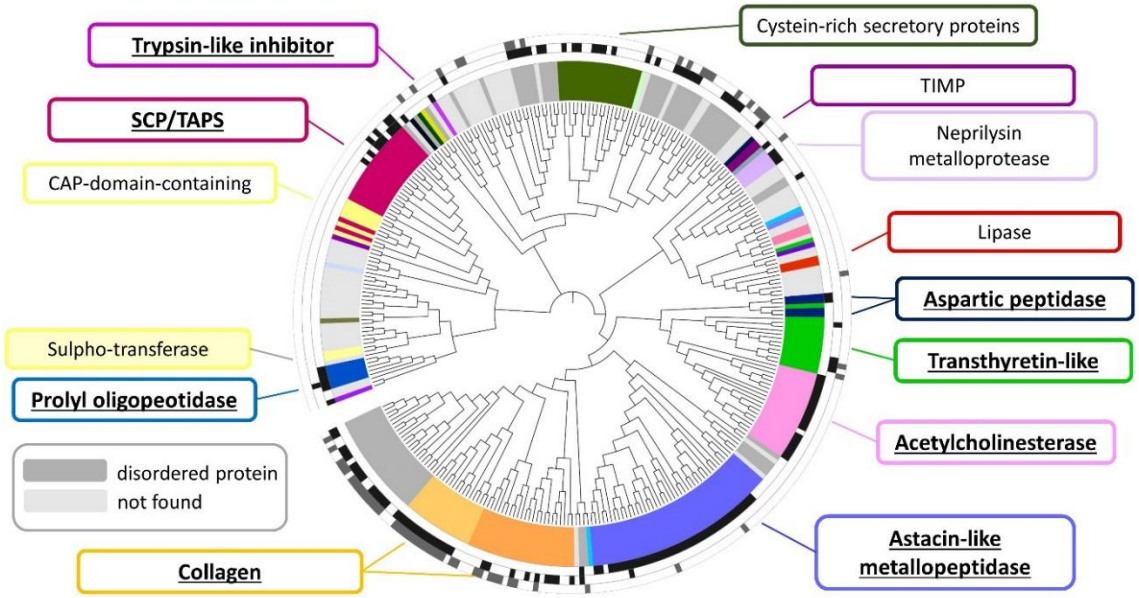


Figure 5.

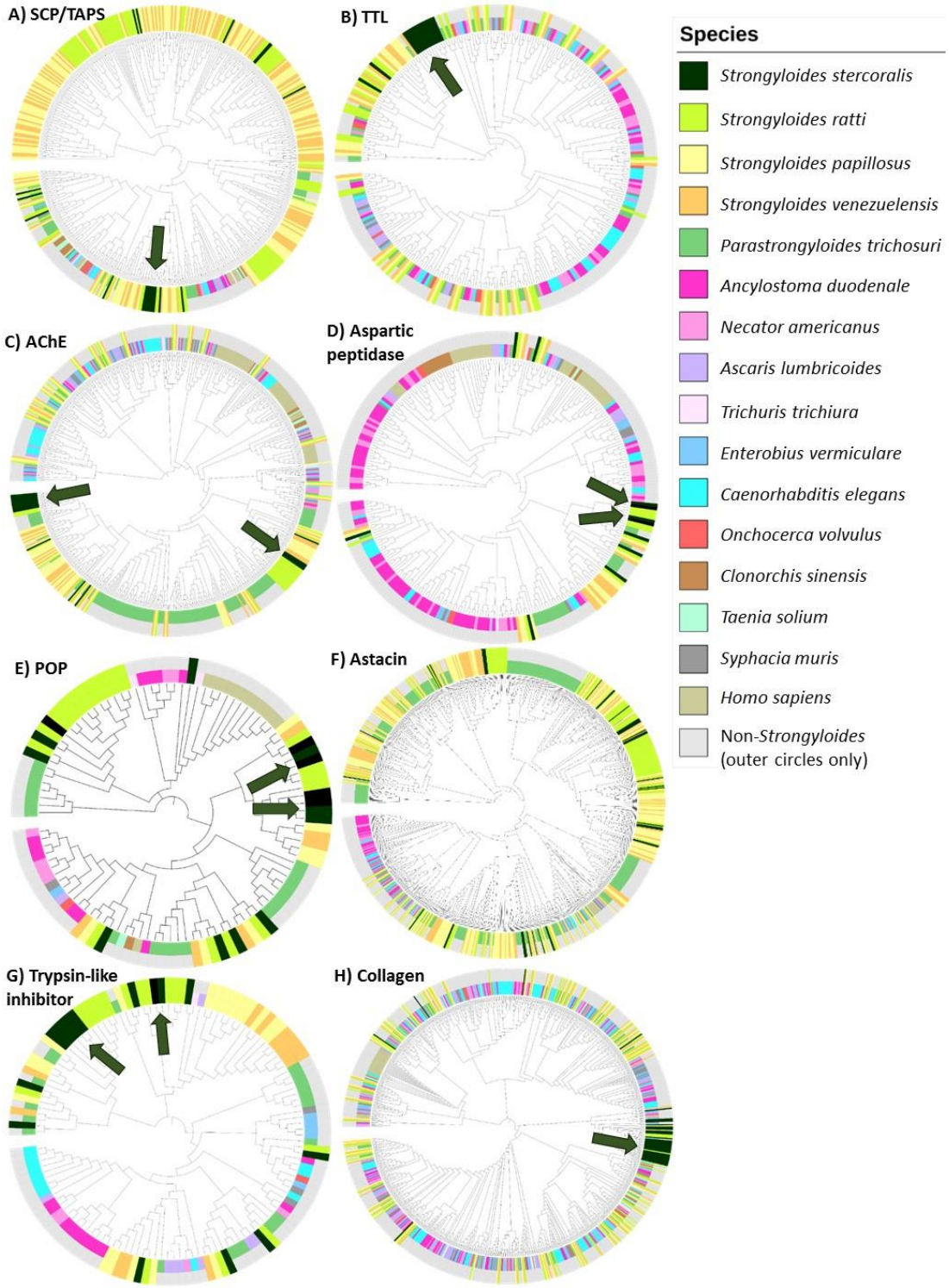

A) SCP/TAPS
B) TTL
C) AChE
D) Aspartic peptidase
E) POP
F) Astacin
G) Trypsin-like inhibitor
H) Collagen
Species
Strongyloides stercoralis
Strongyloides ratti
Strongyloides papillosus
Strongyloides venezuelensis
Parastrongyloides trichosuri
Ancylostoma duodenale
Necator americanus
Ascaris lumbricoides
Trichuris trichiura
Enterobius vermiculare
Caenorhabditis elegans
Onchocerca volvulus
Clonorchis sinensis
Taenia solium
Syphacia muris
Homo sapiens
Non-Strongyloides (outer circles only)

Figure 6.

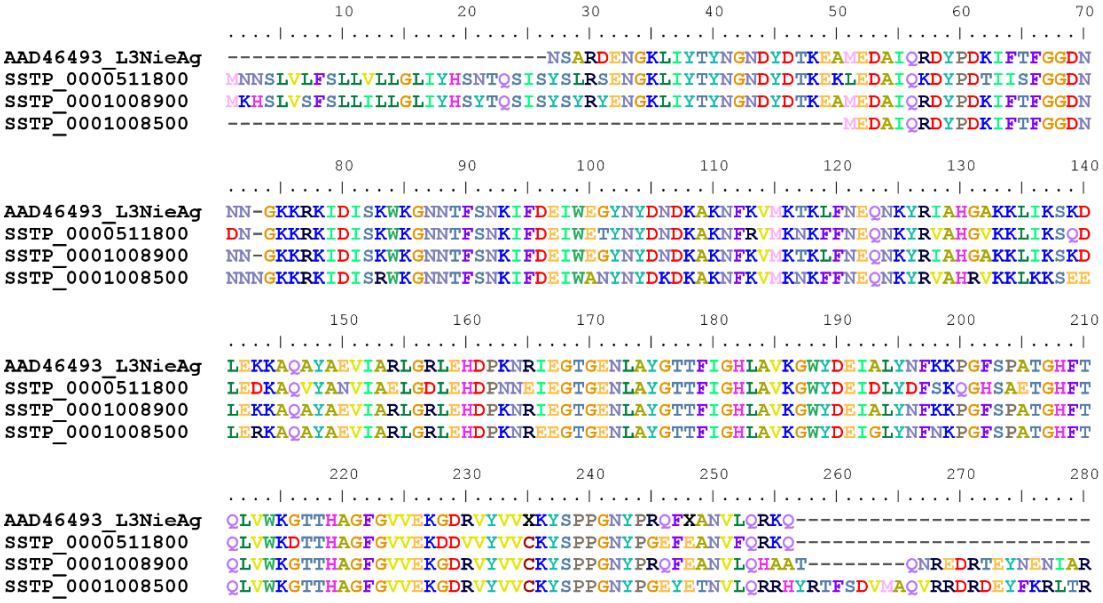


Figure 7.

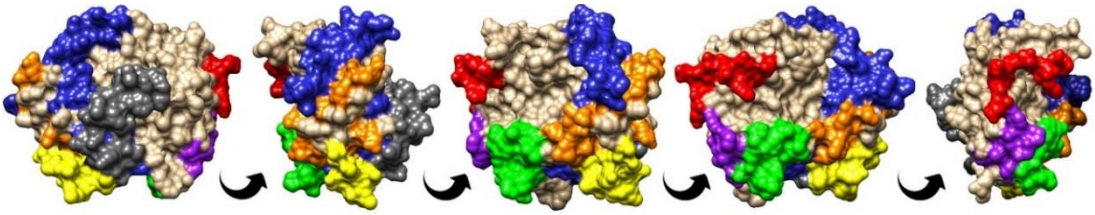


Figure 8.

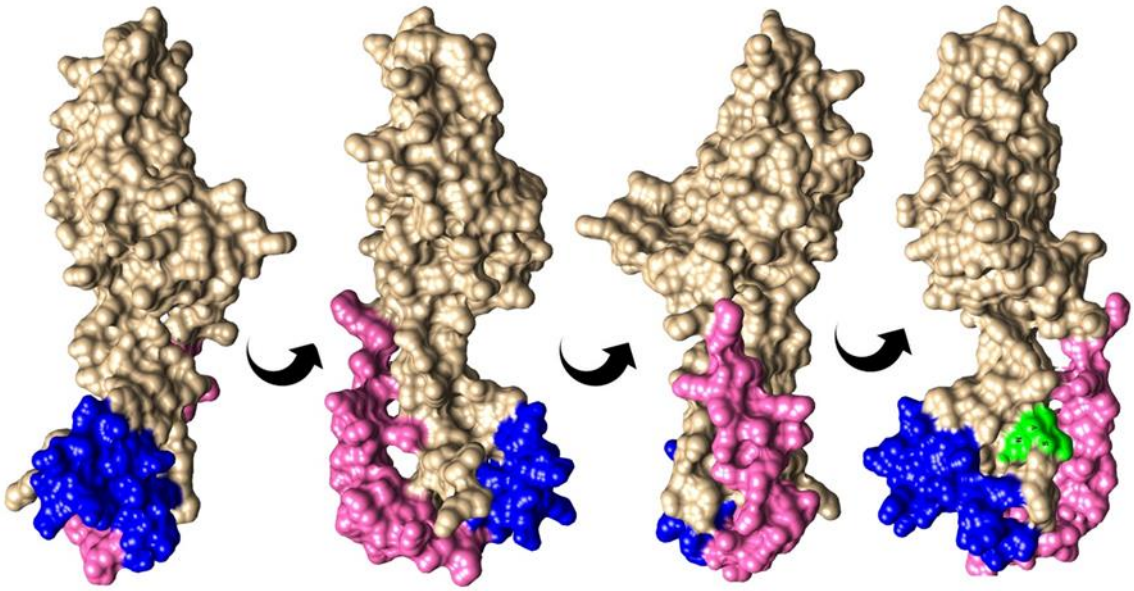


Figure 9.

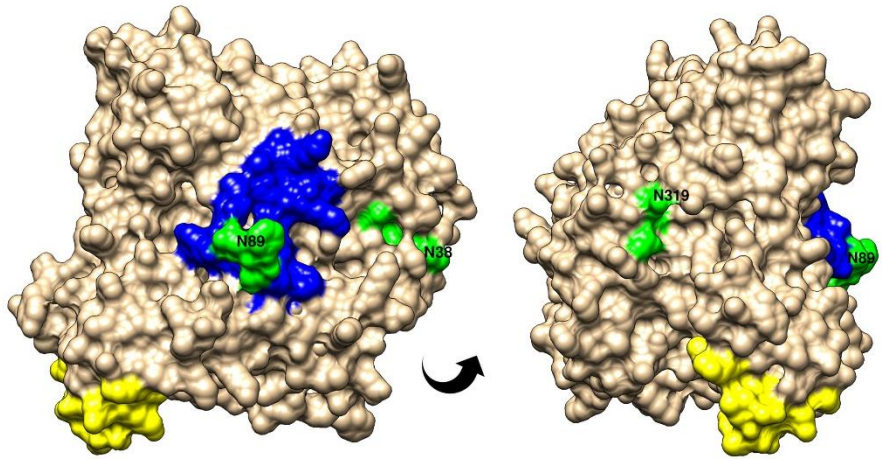


Figure 10.

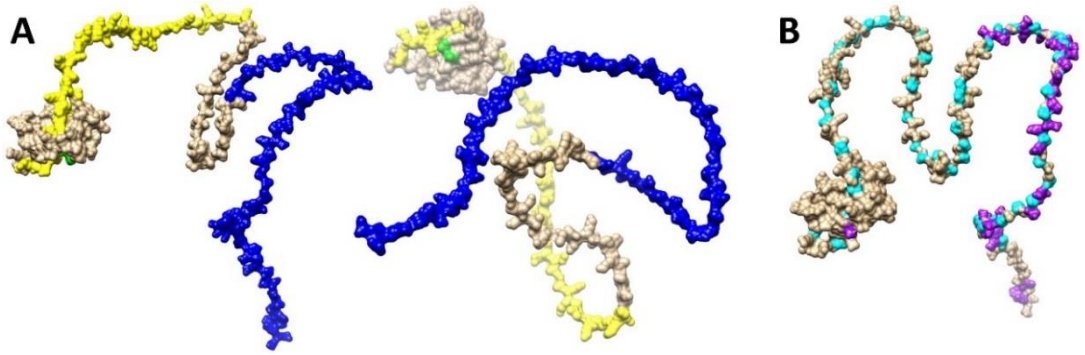


Figure 11.

```
                                10        20        30        40        50        60        70
                                |....|....|....|....|....|....|....|....|....|....|....|....|
SSTP_0001226800 pep120          ------------------------------------------------------------
AAB65139 (5a)                   NSARD----------QKNQLENQDQKNQLKNQSENQDQKKPIKKPIRKPGPKKPIKKPIKKPGPKPI
AAB65140 (8a)                   NSARG------------------------------------------------------------
AAB65141 (12a)                  NSA--------------------------------------------------------------
AAB65142 (19a)                  NSARE----------------------------------------------------PIKKPIKKPGPKPI
AAB97359 (SsIR)                 NSARVENQDQKDQLENQDQKDQLENQDQKNQLKNQSENQDQKKPIKKPIKKPGPKPI

                                        80        90        100       110       120       130       140
                                |....|....|....|....|....|....|....|....|....|....|....|....|....|
SSTP_0001226800 pep120          -----PKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGPEEPEGPEEPEGPAGPEEPEGPAGPEEPE
AAB65139 (5a)                   RPIVKPKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGP---AGPEEPEGPAGPEEPEGPAGPEEPE
AAB65140 (8a)                   -PIVKPKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGP---AGPEEPEGPAGPEEPEG--------
AAB65141 (12a)                  RAIVKPKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGP---AGPEEPEGPAGPEEPEGPAGPEEPE
AAB65142 (19a)                  RPIVKPKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGP---AGPEEPEGPAGPEEPEGPAGPEEPE
AAB97359 (SsIR)                 RPIVKPKPKTTTQAPEEPEGPEEPEGPEEPEGPEGPEEPEGP---AGPEEPEGPAGPEEPEG--------

                                        150       160       170       180       190
                                |....|....|....|....|....|....|....|....|....|....|...
SSTP_0001226800 pep120          GPEEPEGPAGPEEP------------------------------RDDDDGVDEEDERD
AAB65139 (5a)                   GPEEPEGPAGPEEP------------------------------RDDDDGVDEEDERD
AAB65140 (8a)                   -PEEPEGPAGPEEP------------------------------RDDDDGVDEEDERD
AAB65141 (12a)                  GPEEPEGPAGPEEPEGPEEPEGPEEPEEPEGPEEPEEPEGPEEPRDNDDGIDEEDERD
AAB65142 (19a)                  GPEEPEGPAGPEEPEGPEEPEGPEEPEEPEGPEEPEEPEGPEEPRDNDDGIDEEDERD
AAB97359 (SsIR)                 -PEEPEGPAGPEEP------------------------------RDDDDGVDEEDERD
```

CHAPTER 3: Visceral leishmaniasis IgG1 rapid monitoring of cure vs. relapse, and potential for diagnosis of post kala-azar dermal leishmaniasis

Marlais T, Bhattacharyya T, Singh OP, Mertens P, Gilleman Q, Thunissen C, Hinckel BC, Pearson C, Gardner BL, Airs S, de la Roche M, Hayes K, Hafezi H, Falconar AK, Eisa O, Saad A, Khanal B, Bhattarai NR, Rijal S, Boelaert M, El-Safi S, Sundar S, Miles MA. Visceral leishmaniasis IgG1 rapid monitoring of cure vs. relapse, and potential for diagnosis of post kala-azar dermal leishmaniasis. (2018). Frontiers in Cellular and Infection Microbiology, 8:427. doi: 10.3389/fcimb.2018.00427

Key points, novel results and implications

- This manuscript further investigates the potential of IgG1 to be a marker of cure, relapse and PKDL after treatment of VL, and for the first time, applies the assay to paired sera taken pre- and post-treatment. Further validation of the assay in rapid diagnostic test (RDT) format is presented after testing with Indian and Sudanese VL sera.
- 78% of 104 Indian VL patients were positive by the IgG1 RDT before treatment. 84% of these had reduced or negative IgG1 levels at 6 months when deemed cured. 85% of 33 Indian relapse sera were IgG1 RDT positive, a significant difference from those deemed cured. 78% of 63 PKDL patients were positive by IgG1 RDT.
- The RDT was more likely than IgG1 ELISA to be positive pre-treatment and negative at 6 months post-treatment.
- This IgG1 RDT could contribute to VL disease control by indicating those at greater risk of relapse after treatment and assisting with the diagnosis of relapse and PKDL. These are major diagnostic gaps in the current tools available to control programmes.

Candidate's contribution

The candidate co-supervised the laboratory work, collated and analysed the data, and wrote the original manuscript and subsequent revisions.

Co-author's contributions

Tapan Bhattacharyya, LSHTM, co-ordinated and supervised the laboratory work, provided guidance on data analysis and provided feedback on the manuscript.

Om Prakash Singh, Banares Hindu University, provided samples and hosted lab work in India.

Pascal Mertens, Quentin Gilleman and Caroline Thunissen, Coris BioConcept, manufactured the RDTs.

Bruno Hinckel, Callum Pearson, Bathsheba Gardner, Stephanie Airs, Marianne de la Roche, Kiera Hayes and Hannah Hafezi, LSHTM, conducted the lab work.

Andrew Falconar, Universidad del Norte, advised on the development of the immunoassays.

Osama Eisa and Alfarazdeg Saad, University of Khartoum, collected samples in Sudan.

Basudha Khanal, Narayan Bhattarai and Suman Rijal, BPKIHS, provided samples and hosted lab work in Nepal.

Marleen Boelaert, Sayda El-Safi, Shyam Sundar, Michael Miles, obtained funding, hosted the project in their respective labs and provided critical review of the manuscript.

Evidence of copyright retention for publication:

Marlais *et al*. 2018. Visceral Leishmaniasis IgG1 Rapid Monitoring of Cure vs. Relapse, and Potential for Diagnosis of Post Kala-Azar Dermal Leishmaniasis. Frontiers in Cellular and Infection Microbiology, 8:427.

- **What is the Frontiers Copyright Policy?**

Under the Frontiers Terms and Conditions, authors retain the copyright to their work. All Frontiers articles are Open Access and currently distributed under the terms of the Creative Commons Attribution License, which permits the use, distribution and reproduction of material from published articles, provided the original authors and source are credited, and subject to any copyright notices concerning any third-party content. For earlier articles, please check the online page for license details.

Updated : March 20, 2017 11:43

From: https://zendesk.frontiersin.org/hc/en-us/articles/201904552-What-is-the-Frontiers-Copyright-Policy-

- **Do I need permission to reprint my article or parts of my article published with Frontiers?**

As long as you cite the original publication with Frontiers and no third-party licenses apply within the article you are free to reprint your article.

Frontiers does not provide any formal permissions for reuse.

Updated: March 13, 2017 15:34

From: https://zendesk.frontiersin.org/hc/en-us/articles/201904562--Do-I-need-permission-to-reprint-my-article-or-parts-of-my-article-published-with-Frontiers-

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Tegwen Marlais
Principal Supervisor	Prof. Michael A. Miles
Thesis Title	Diagnostic development for the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Frontiers in Cellular and Infection Microbiology		
When was the work published?	December 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	n/a		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I co-supervised the laboratory work, collated and analysed the data and wrote the original manuscript and subsequent revisions.

Student Signature: [redacted] **Date:** 22/7/19

Supervisor Signature: [redacted] **Date:** 22/7/19

Visceral leishmaniasis IgG1 rapid monitoring of cure versus relapse, and potential for diagnosis of post kala-azar dermal leishmaniasis

Authors

Tegwen Marlais[1]*‡, Tapan Bhattacharyya[1]‡, Om Prakash Singh[2], Pascal Mertens[3], Quentin Gilleman[3], Caroline Thunissen[3], Bruno C. Bremer Hinckel[3,4], Callum Pearson[1], Bathsheba L. Gardner[1], Stephanie Airs[1], Marianne de la Roche[1], Kiera Hayes[1], Hannah Hafezi[1], Andrew K. Falconar[5], Osama Eisa[6], Alfarazdeg Saad[6], Basudha Khanal[7], Narayan Raj Bhattarai[7], Suman Rijal[8], Marleen Boelaert[9], Sayda El-Safi[6], Shyam Sundar[2], Michael A. Miles[1]

1 Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom,

2 Department of Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India,

3 Coris BioConcept, Gembloux, Belgium,

4 Department of Biomedical Sciences, University of Antwerp, Belgium,

5 Departamento de Medicina, Universidad del Norte, Barranquilla, Colombia,

6 Faculty of Medicine, University of Khartoum, Khartoum, Sudan,

7 Department of Microbiology, B.P. Koirala Institute of Health Sciences, Dharan, Nepal,

8 Department of Internal Medicine, B.P. Koirala Institute of Health Sciences, Dharan, Nepal,

9 Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium

‡ These authors contributed equally to this work.

Correspondence:

Tegwen Marlais

tegwen.marlais@lshtm.ac.uk

Abstract

Background: There is a recognised need for an improved diagnostic test to assess post-chemotherapeutic treatment outcome in visceral leishmaniasis (VL) and to diagnose post

kala-azar dermal leishmaniasis (PKDL). We previously demonstrated by ELISA and a prototype novel rapid diagnostic test (RDT), that high anti-*Leishmania* IgG1 is associated with post-treatment relapse versus cure in VL.

Methodology: Here, we further evaluate this novel, low-cost RDT, named VL Sero K-SeT, and ELISA for monitoring IgG1 levels in VL patients after treatment. IgG1 levels against *L. donovani* lysate were determined. We applied these assays to Indian sera from cured VL at 6 months post treatment as well as to relapse and PKDL patients. Sudanese sera from pre- and post-treatment and relapse were also tested.

Results: Of 104 paired Indian sera taken before and after treatment for VL, when deemed clinically cured, 81 (77.9%) were positive by VL Sero K-SeT before treatment; by 6 months, 68 of these 81 (84.0%) had a negative or reduced RDT test line intensity. ELISAs differed in positivity rate between pre- and post-treatment ($p = 0.0162$). Twenty eight of 33 (84.8%) Indian samples taken at diagnosis of relapse were RDT positive. A comparison of Indian VL Sero K-SeT data from patients deemed cured and relapsed confirmed that there was a significant difference ($p < 0.0001$) in positivity rate for the two groups using this RDT. Ten of 17 (58.8%) Sudanese sera went from positive to negative or decreased VL Sero K-SeT at the end of 11-30 days of treatment. Forty nine of 63 (77.8%) PKDL samples from India were positive by VL Sero K-SeT.

Conclusion: We have further shown the relevance of IgG1 in determining clinical status in VL patients. A positive VL Sero K-SeT may also be helpful in supporting diagnosis of PKDL. With further refinement, such as the use of specific antigens, the VL Sero K-SeT and/or IgG1 ELISA may be adjuncts to current VL control programmes.

Words: approx. 4,800 Figures: 2 Tables: 2

Introduction

Visceral leishmaniasis (VL; kala-azar), is caused by the protozoan parasites *Leishmania donovani* in Asia, Africa and the Middle East and *Leishmania infantum* in Europe and South America. These parasites are transmitted by blood-feeding female phlebotomine sand flies. Symptomatic VL is usually fatal if untreated. Symptoms include prolonged fever >14 days, wasting, splenomegaly, hepatomegaly and anaemia (Sundar and Rai, 2002). While VL is present in about 75 countries, the majority (90%) of cases in 2015 occurred in India, Sudan, South Sudan, Ethiopia, Somalia, Kenya and Brazil (World Health Organization., 2017), where it

is closely linked to poverty, both as cause and effect (Boelaert et al., 2009; Sarnoff et al., 2010).

Following clinical suspicion of VL, serology is used for diagnosis. Techniques vary by region and include the immunofluorescence antibody test (IFAT), direct agglutination test (DAT), enzyme linked immunosorbent assay (ELISA), and detection of IgG against recombinant antigens rK39 or rK28 (Singh and Sundar, 2015). In India the DAT and rK39 serology are used, with a positive result in either test indicative of exposure to infection with *L. donovani*. For confirmatory parasitological diagnosis, seropositive individuals undergo spleen, bone marrow or lymph node biopsy to search for the intracellular amastigote stage in films of Giemsa-stained aspirates. These are invasive, costly and potentially hazardous techniques with low and variable sensitivities ranging from 53-99% (Singh and Sundar, 2015).

VL is treated with antimonials, miltefosine, paromomycin, amphotericin B, liposomal amphotericin (AmBisome) or drug combinations (World Health Organization, 2010). Currently, post-treatment outcome is determined by assessment of clinical signs and symptoms, initially on the last day of drug treatment and, in India, again six months after administration of the last dose (World Health Organization, 2010). Possible outcomes are: cure; relapse; death (by VL or not); post kala-azar dermal leishmaniasis (PKDL); loss to follow up. However, recent studies from India and Nepal have reported relapse rates of between 1.4% and 20%, including up to and beyond 12 months following the end of treatment (Burza et al., 2013; Rijal et al., 2013; Burza et al., 2014). In Sudan, relapse rates around 6% have been reported (Gorski et al., 2010; Atia et al., 2015). Patients who relapse face a further biopsy procedure to confirm the presence of parasites.

PKDL is a non-painful sequela of VL occurring in over 50% of cases in Sudan (Zijlstra et al., 2003) but is far less prevalent in South Asia (Zijlstra et al., 2003; Uranw et al., 2011). PKDL is less reported in *L. infantum* endemic regions where cases have mostly been associated with HIV/AIDS (Ridolfo et al., 2000; Bittencourt et al., 2003; Celesia et al., 2014), other co-infections (Trindade et al., 2015) or immune suppression (Roustan et al., 1998). PKDL manifests between 0.5 months to one or more years after apparently successful VL treatment (Musa et al., 2002; Uranw et al., 2011; Singh et al., 2012; Moulik et al., 2017) and may occasionally occur without a prior episode of VL (el Hassan et al., 1992; Zijlstra et al., 2003; Das et al., 2012; Das et al., 2016). PKDL is suspected based on dermal manifestations that are non-specific and diagnosis is made on previous VL treatment history and confirmed parasitologically by microscopy of slit skin smear or biopsy or PCR (Zijlstra et al., 2017).

Conventional serology is likely to remain positive from the earlier VL and there is no test in use to predict PKDL (Gidwani et al., 2011). The high parasite density in PKDL skin provides a source of infection to sand flies and thus sustains long term transmission and endemicity (Molina et al., 2017; Mondal et al., 2018).

An unresolved crucial question is how to identify asymptomatic infected individuals simply and reliably (as defined by seropositivity, lack of clinical symptoms and no prior history of VL) who will progress to active VL. High DAT and/or rK39 ELISA titres have been associated with increased risk of progression in the Indian subcontinent but as yet there is no single rapid test in use for this purpose (Hasker et al., 2014; Chapman et al., 2015). To improve outcome monitoring of VL and disease control, the World Health Organization has identified the vital need for a marker of post-chemotherapeutic cure, and the high priority incorporation of such a biomarker into a point-of-care rapid diagnostic test (RDT) (World Health Organization, 2012). Such a test should meet the "ASSURED" criteria of being: affordable; sensitive (few false negatives); specific (few false positives); user-friendly, requiring minimal training; rapid; robust, not requiring cold-storage; equipment-free, and deliverable to those who need it (Peeling et al., 2006).

We have previously shown that high anti-*Leishmania* IgG1 ELISA titres are associated with treatment failure, whereas, in cases deemed to be cured following chemotherapy, IgG1 levels diminish significantly by six months post-treatment and only IgG1 gave this level of discrimination (Bhattacharyya et al., 2014a). We demonstrated this by ELISA against *L. donovani* whole cell lysate, and then adapted the assay to a prototype lateral flow immunochromatographic RDT. Here, we present further evaluation of this RDT, called VL Sero K-SeT, to indicate cure after VL treatment in a larger, paired, sample set and to confirm relapse. We also performed western blot on the same sample set. Additionally we show the potential utility of VL Sero K-SeT and other IgG1 assays to confirm PKDL.

Methods

Ethics statement

In India, the collection of samples was approved by the Ethics Committee of Banaras Hindu University, Varanasi. In Sudan approval was by the Ethical Research Committee, Faculty of Medicine, University of Khartoum and the National Health Research Ethics Committee, Federal Ministry of Health, Sudan. Written informed consent was obtained from adult subjects included in the study or from the parents or guardians of individuals less than 18 years of age. In Nepal, informed consent was obtained from all the participants and the ethical committee of the B.P. Koirala Institute of Health Sciences (BPKIHS) approved the

study. This research was also approved, as part of the EC NIDIAG project, by the London School of Hygiene and Tropical Medicine Ethics Committee.

Sources of sera/plasma

We retrospectively selected sera or plasma from an archive of different VL disease states. Samples had been collected in VL endemic regions, namely Muzaffarpur in Bihar, India after 2007 and in 2013 in Gedaref, Sudan. Sample sizes used during this evaluation were dependent on availability of appropriate samples and reagents.

In India, cases of VL had been diagnosed by positive rK39 serology and/or parasitologically by microscopy of splenic aspirates. In Sudan active cases of VL had been diagnosed by microscopy of bone marrow or lymph node aspirates in conjunction with serological assays. These diagnoses were made according to their respective national procedures, prior to the present study. Sera/plasma were stored at -80°C until use. All patients were HIV negative. We have previously observed that serum and plasma derived from the same sample show no difference in IgG titre in ELISA against *L. donovani* lysate (unpublished observations), although we have not specifically assessed IgG1 with both sample types.

India

Indian sample types are described in Table 1. We have previously found that in Indian VL, IgG1 titre up to day 30 post-treatment initiation is not statistically significantly different from pre-treatment (Bhattacharyya et al., 2014a) and therefore we consider these as 'pre-treatment' in paired samples for the purposes of this study. Treatment of VL was with single-dose AmBisome alone or with 10 days of miltefosine. PKDL was treated with miltefosine for 84 days. DAT and rK39 ELISA were conducted prior to the present study as part of standard diagnostic procedures in India.

Sudan

Sudanese paired serum samples (n = 17 pairs) were taken on day of diagnosis of VL and at the end of treatment at 11 days (AmBisome), 17 days (sodium stibogluconate (SSG) + paromomycin) or 30 days (SSG only). These samples were previously tested for IgG1 by ELISA (Bhattacharyya et al., 2014a). Additional Sudanese serum samples used in the present study were unpaired treated individuals (n = 2) taken an unknown time after treatment, and relapse (n = 1). Sudanese EHC samples had previously been tested by the IgG1 ELISA using the same antigen and were negative (Bhattacharyya et al., 2014a) but were not retested here.

Antigen production

Whole cell lysate of *L. donovani* strain MHOM/IN/80/DD8 isolated from India, and MHOM/SD/97/LEM3458 isolated from Sudan, was prepared as described previously (Bhattacharyya et al., 2014a). Lysate antigen was used for VL Sero K-SeT development (strain LEM3458), ELISA and western blot (strain DD8). Antigen preparation for western blot strips contained 50 µl of protease inhibitor cocktail (P8340, Sigma, UK) per 1 ml of *L. donovani* cells; centrifugation after sonication was 16,160 x *g* for 45 min at 4°C.

ELISA for IgG1 anti *L. donovani*

Duplicate ELISA plates were coated overnight at 4°C with *L. donovani* DD8 strain antigen prepared as above, at 2 µg/ml in coating buffer (35 mM $NaHCO_3$, 15 mM $NaCO_3$, pH 9.6), 100 µl/well. Plates were washed 3 times with phosphate buffered saline (PBS) + 0.05% Tween 20 (PBST) prior to blocking with 200 µl/well PBS + 2% w/v non-fat milk powder (Premier International Foods, UK) (PBSM) for 2 hours at 37°C, followed by three PBST washes. Sera/plasma were diluted 1/100 in PBST+ 2% w/v non-fat milk powder (PBSTM) and applied at 100 µl/well, incubated for 1 hour at 37°C then washed 6 times with PBST. Mouse anti human IgG1-horse radish peroxidase (HRP) (ab99774, Abcam, UK) was diluted 1:5,000 in PBSTM and incubated at 100 µl/well, 37°C for 1 hour. Plates were washed 6 times with PBST before the addition of 100 µl/well of freshly prepared substrate solution (50 mM citric acid, 50 mM Na_2HPO_4, 2 mM *o*-phenylenediamine HCl, 0.009% H_2O_2). Plates with substrate were incubated in the dark at room temperature for 10-15 minutes when the reaction was stopped with 100 µl/well of 1 M H_2SO_4 and absorbance read at 490 nm. Each plate contained an EHC sample as a negative serological control for determining the positivity cut-off and a known seropositive VL sample as positive control. All ELISA results reported are the mean A_{490} of duplicate plates.

RDT production and use

Whole cell lysate was prepared as described previously (Bhattacharyya et al., 2014a) from *L. donovani* strain MHOM/SD/97/LEM3458. The VL Sero K-SeT lateral flow immunochromatographic tests were developed at Coris BioConcept and consisted of a cassette with a nitrocellulose membrane, a sample pad, a conjugate pad and an absorbent pad, backed with a plastic strip. The nitrocellulose membranes were sensitized with the *L. donovani* lysate antigen and anti-human IgG1-specific antibody labelled with colloidal gold was dried on the conjugate pad. This strip was housed in a plastic cassette with two windows: the smaller buffer well and the long central test window.

To perform the test, 3.5 µl of serum/plasma was applied to the test window at the point indicated by a dot (•) on the cassette, followed immediately by 120 µl of supplied running

195 buffer to the buffer well (Figure 1). Devices were incubated flat, at ambient temperature for
196 15 minutes before being assessed visually. Any test line at position T was considered a
197 positive result if a control line was also present at position C. Positive test line intensity was
198 assessed visually for samples from pre- and post-treatment VL (Figure 1). A subset of samples
199 was tested on different batches of the VL Sero K-SeT. Readers of the RDTs were blinded to all
200 the corresponding ELISA results.

201 ### Western blotting

202 Western blots were performed to visualise antigen recognition in patients from the different
203 clinical groups, as described in Supplementary Material S1. Briefly, tricine SDS-PAGE gels were
204 made as per Schägger (2006). *L. donovani* DD8 lysate was used as antigen with sera/plasma
205 diluted 1 in 300 (Sudan) or 1 in 400 (India) and detection was by mouse anti human IgG1-
206 HRP.

207 ### Statistical analysis

208 We performed a two-tailed Fisher's exact test on Indian VL Sero K-SeT and IgG1 ELISA data to
209 calculate p-values between samples from pre- and matched 6 months post-treatment
210 (deemed cured), separately between post-treatment and relapse, and between post-
211 treatment and PKDL. Cut-off for the IgG1 ELISA was calculated as the mean absorbance of the
212 EHC samples plus 3 standard deviations.

213

214 ## Results

215 ### IgG1 diminishes by 6 months in cured VL patients

216 Samples taken from Indian patients before or at the outset of therapy, were compared by VL
217 Sero K-SeT and IgG1 ELISA with paired samples taken 6 months later when the individuals
218 were deemed cured. Both IgG1 assay methods showed a statistically significant difference in
219 positivity rate between pre- and post-treatment samples ($p = 0.0162$ and $p < 0.0001$ for ELISA
220 and RDT respectively) (Figure 2). A consistent and strongly significant difference was also
221 observed between cured versus relapsed samples ($p < 0.0001$), again with both IgG1 assay
222 methods (Figure 2).

223 A subset of pre- and post-treatment of the cured pairs samples (n = 87) was tested on
224 different batches of the VL Sero K-SeT, with agreement between individual RDTs of 92.0%
225 (80/87).

226 ### Changes in IgG1 levels by ELISA and VL Sero K-SeT

ELISA absorbance and corresponding VL Sero K-SeT results for individual samples are given in Supplementary Material S2. Of the 80 Indian paired samples tested for anti *L. donovani* IgG1 by ELISA, 54 (67.5%) were positive before treatment. Of these, 51/54 declined in titre: 21/51 (representing 26.3% of the total 80) went from positive to negative and 30/51 (representing 37.5% of the total 80) had reduced IgG1 at 6 months when deemed cured (Figure 2 and Table 2). Twenty one (26.3%) paired cured sera were negative by IgG1 ELISA before treatment and remained so at 6 months.

Overall, including those negative at the start, at 6 months after treatment 79/104 (76.0%) were negative by VL Sero K-SeT (Table 2). VL Sero K-SeT results were additionally assessed according to whether the Indian 6 month post-treatment (cure) sample had a decreased or not decreased test line intensity compared to the paired pre-treatment sample. Of the 104 paired samples tested by VL Sero K-SeT from deemed cured Indian VL patients, 81 (77.9%) were positive at start of treatment (Table 2); of these, 68/81 (84.0%) had either become negative or had a visibly reduced test line intensity at 6 months when deemed cured. Thirteen (12.5%) initially RDT positive individuals showed no visible decrease in RDT band intensity at 6 months, despite being deemed cured, and none became positive from negative.

Ninety four percent of samples positive by ELISA at pre-treatment, decreased in seropositivity; for VL Sero K-SeT, this proportion was 84%. However, at 6 months post-treatment, the ELISA was more likely to remain positive than the RDT, using the cut-off value established for the IgG1 ELISA.

Seventy nine Indian samples were tested by both VL Sero K-SeT and ELISA. Of these samples, the RDT was more likely than the ELISA to be positive with the Day 0 samples (78.5% versus 67.1%) and negative with the 6 month samples (78.5% versus 53.2%). Of samples which remained positive at 6 months by both methods (n = 14), the change in intensity of RDT test line generally mirrored the change in ELISA absorbance value for the same sample. Three of the Indian samples increased markedly in IgG1 titre by ELISA at 6 months (Figure 2). Two of these accorded with a corresponding rise in VL Sero K-SeT test line intensity; for the third sample, both pre-treatment and 6 month RDTs were negative (Figure 2A).

Sudanese paired samples taken before and immediately after treatment (11 to 30 days later) were similarly assessed (Table 2). For Sudanese paired samples prior to treatment, 13/17 (76.5%) were positive by VL Sero K-SeT and at completion of treatment, 10/13 (76.9%) had a negative or reduced test line intensity (Table 2). If taken as a single time point at the end of treatment, 10/17 (58.8%) Sudanese VL patients had negative VL Sero K-SeT result. Four (23.5%) of the Sudanese treated individuals were negative pre-treatment, similar to the

proportion of Indian samples (22.1%). Two additional un-paired treated Sudanese samples were negative by RDT (not shown).

IgG1 western blot confirmed negative/declined RDT in cure

For a subset of 25 of the paired Indian samples, western blots mirrored the VL Sero K-SeT RDT findings, in that IgG1 declined dramatically in all but one VL patient at 6 months follow up after treatment (Supplementary Material S3). As with the RDT, the blots showed that samples that were positive and detecting many antigens before treatment had become negative or reduced in intensity by 6 months. Corresponding RDT images are shown in Supplementary Material S4.

Elevated IgG1 in VL relapse

For 33 Indian patients for whom we had unpaired samples at the time of relapse, the VL Sero K-SeT was 84.8% (28/33) positive and ELISA 91.3% (21/23) positive, confirming relapse. Of the 23 samples tested by ELISA that were also tested by RDT, 19 gave the same result by both assays (Supplementary Material S2). The single available Sudanese relapse sample was IgG1 positive (Supplementary Material S5). Twenty five of the Indian samples and the single Sudanese sample were also tested on western blot for IgG1 against *L. donovani* lysate antigen and showed concordance between the RDT and blots (Supplementary Material S2 and S5). For two of the 33 Indian relapse samples, a paired pre-treatment sample was available. Both individuals were VL Sero K-SeT positive at both time points.

All samples from other diseases, namely malaria, tuberculosis, dengue fever, rheumatoid arthritis and multiple myeloma were negative by VL Sero K-SeT, as were all samples from endemic healthy controls.

VL Sero K-SeT can provide evidence for PKDL but not for its cure

Of the 63 PKDL samples tested, 49 (77.8%) were positive by VL Sero K-SeT and of the subset of 45 tested by IgG1 ELISA, 43 (95.6%) were positive (Supplementary Material S2). A subset of 10 VL Sero K-SeT-positive PKDL samples were tested by western blot, of which 9 showed discernible bands. Images of the blots and their corresponding VL Sero K-SeT RDTs are shown in Supplementary Material S6. There was a highly statistically significant difference between post-treatment cured samples at 6 months and PKDL by both VL Sero K-SeT and IgG1 ELISA (Fisher's exact $p < 0.0001$ for both assays).

Seventeen of the 63 individuals with PKDL provided between 1 and 5 additional sequential follow-up samples over intervals ranging from 15 to 365 days post-treatment. These PKDL post-treatment sequential samples retained the initial RDT result in 12/17 (70.6%) cases,

decreased in 3/17 (17.6%), increased slightly in one case (5.9%) and varied between positive and negative over time in one case (5.9%).

IgG1 can indicate progression from asymptomatic status

When samples from asymptomatic seropositive individuals who later progressed to symptomatic disease (progressors, n = 4) were tested on the VL Sero K-SeT, all gave a positive test result (Supplementary Material S7). In contrast, 4 individuals who were seropositive but did not develop symptomatic VL were negative by VL Sero K-SeT. Thus, in our limited sample size, elevated IgG1 levels, as detected by VL Sero K-SeT, were associated with progression to symptomatic disease. This result was corroborated by ELISA and western blot (Supplementary Material S7).

Discussion

Conventional serology for VL diagnosis relies on detecting the overall IgG response. This has been reported to remain elevated, often for years, after treatment (Bhattarai et al., 2009; Gidwani et al., 2011; Srivastava et al., 2013). This makes current serology unsuitable for timely monitoring of treatment outcome. We have previously found using ELISA that a decreased or negative anti *Leishmania* IgG1 titre at 6 months post-treatment can be indicative of VL cure, whereas elevated IgG1 levels are associated with post-chemotherapeutic relapse (Bhattacharyya et al., 2014a).

Monitoring of post-treatment outcomes

Here we used a larger panel of paired samples to assess the IgG1 response as detected by the rapid test, VL Sero K-SeT, where 77.9% of Indian samples were positive before treatment and of these 69.1% had become negative 6 months later when deemed cured (Table 2). In total, 76% of 6 month samples were negative, a significant difference from pre-treatment ($p < 0.0001$). Of those still positive at 6 months using this RDT, we found that a diminished test line intensity was also consistent with cure. This decline was corroborated by ELISAs, and despite slight differences in the antigen preparations. We have found no difference in performance of the VL Sero K-SeT when DD8 strain antigen is used instead of LEM3458 (unpublished observations). Thus, the VL Sero K-SeT is a promising innovation, although there is a need to improve further its discriminative capacity.

Sudanese samples declined from positive to negative or decreased VL Sero K-SeT test line intensity in 76.9% of patients immediately after treatment, no more than 30 days after the first sample. This apparently rapid drop in IgG1 was not seen in Indian samples and could be

due to the overall lower IgG titre observed in Sudanese samples (Bhattacharyya et al., 2014b; Abass et al., 2015). Thus, a small drop in IgG1 titre could have taken these samples below the detection limit of the VL Sero K-SeT. This may suggest that the VL Sero K-SeT can be used before 6 months to indicate cure or relapse in eastern Africa. The unexpectedly low sensitivity of the VL Sero K-SeT at the start of treatment for both Indian (77.9%) and Sudanese (76.5%) samples does not hinder the subsequent assessment of cure at 6 months, because a negative IgG1 result at 6 months can indicate cure. In addition, we do not propose to use IgG1 assays as a diagnostic for active VL but rather to assist with confirming cure, relapse and PKDL, all of which currently lack an appropriate diagnostic test. With Indian samples, where there was discrepancy between VL Sero K-SeT and ELISA, the RDT was generally more accurate, being positive with pre-treatment and negative with 6 month samples (Supplementary Material S2). As for the Indian sera, the strength of RDT test line intensity broadly corresponded with ELISA signal for an individual sample.

Elevated levels of IgG1 were associated with VL relapse in both assays here for Indian samples. Likewise, the single Sudanese relapse patient was positive by VL Sero K-SeT, whilst 2 treated individuals were negative. We do not know the length of time between treatment and relapse for relapsed individuals (India and Sudan), or the outcome of treated Sudanese individuals. Burza et al. (2014) advised that patient follow-up should be extended from 6 to 12 months as 50-85% of relapses have been found to occur 6 to ≥ 12 months post-treatment (Rijal et al., 2013; Burza et al., 2014). Our evaluations of a limited number Nepalese relapse samples eluted from filter paper indicated that, although encouraging, elution volumes and conditions need further optimisation before they can be more extensively used with VL Sero K-SeT (data not shown).

We found that in Indian cases who relapsed, the RDT positivity rate was significantly different from 6 month samples from patients deemed cured ($p < 0.0001$). Thus the VL Sero K-SeT, with Indian samples, can contribute to distinguishing patients deemed cured from those who have relapsed. Of the 13 Indian patients deemed cured at 6 months but who had no clear decrease in VL Sero K-SeT test line intensity (Table 2), none is known to have relapsed with VL. However, the quantitative ELISA did detect an IgG1 decrease in these samples, consistent with cure. Apparent relapses might however, occasionally include re-infections given the highly endemic locations (Morales et al., 2002). Although beyond the scope of the present study, the inclusion of parasite genotyping in a future study would be an advantage.

Cases co-infected with HIV and VL were not included in the present study. Serological diagnosis is less reliable in HIV/VL co-infection (Cota et al., 2012; Abass et al., 2015) and the

dynamics of IgG1 response in HIV/VL co-infections need to be determined. Other techniques such as a loop mediated isothermal amplification (LAMP) or qPCR detecting parasite DNA might have the potential to discriminate cure from relapse in HIV/VL patients but are currently less accessible than immunological tests (Mukhtar et al., 2018).

PKDL

Indian individuals with PKDL tested here were defined as being with or without a previous history of VL, presenting with a dermal macular, papular or nodular rash often starting on the face with further spread to other parts of the body without loss of sensation. VL Sero K-SeT and IgG1 ELISA results suggest that these assays might contribute to PKDL case detection, as found by a study by Saha et al. (2005), whereas conventional serology may be of limited utility (Gidwani et al., 2011). Our data did not assess the predictive value of IgG1 for development of PKDL.

Where the information was available with our sample set, we did not observe an association between elevated IgG1 and macular versus polymorphic PKDL presentation, this is in contrast to the report of Mukhopadhyay et al. (2012). For a subset of these PKDL samples, we also tested sequential samples taken up to one year after the initial sample. We did not observe a consistent decrease in IgG1 after PKDL treatment.

Progression from asymptomatic to active VL

Asymptomatic, seropositive cases outnumber active VL cases (Bern et al., 2007; Ostyn et al., 2011; Hasker et al., 2013; Hirve et al., 2016; Saha et al., 2017) but a proportion are at elevated risk of progressing to active VL (Gidwani et al., 2009; Topno et al., 2010; Ostyn et al., 2011). Asymptomatics have been reported to occasionally have detectable parasites by PCR or culture of blood (le Fichoux et al., 1999; Costa et al., 2002; Bhattarai et al., 2009; Srivastava et al., 2013). Therefore, neither standard seropositivity nor parasitaemia are indicators of progression to clinical disease. Gidwani et al. (2009) found that this progression to VL occurred up to 2 years after serological positivity.

Our limited sample size of seropositive asymptomatic individuals were identified during a community serological screening study, before the present study. Those who later progressed to clinical VL were positive by VL Sero K-SeT and ELISA, whilst those who did not progress were negative by both assays. High titres in both DAT and rK39 ELISA have been indicative of progression in larger studies (Ostyn et al., 2011; Hasker et al., 2014). However, this combination of tests requires laboratory facilities, therefore it would be desirable to have an RDT that could predict progression.

Additional validation of the VL Sero K-SeT should compare larger cohorts who do and do not progress to VL.

Potential clinical application of IgG1 tests

On the basis of the IgG1 responses reported here by VL Sero K-SeT and ELISA, we propose that IgG1 levels may contribute to monitoring the therapeutic outcome of VL, irrespective of whether there is a pre-treatment sample or result. With further development and validation, IgG1 assays, including the VL Sero K-SeT, which can be produced in large-scale at a cost of a few Euros per test, can be used as an adjunct to the clinical assessment of VL status following treatment. A negative, or defined decrease in IgG1 result at 6 months post treatment in India could be supportive of the clinical assessment of cure. Conversely, an un-paired positive or non-decreased paired positive result at 6 months could indicate the need for additional follow-up. In Sudan, the test may be applicable for defining cure before 6 months. A positive IgG1 result in suspected PKDL or relapse could support the presence of leishmaniasis compared to differential diagnoses. Although western blots were supportive of the use of IgG1, we did not specifically assess banding patterns, and do not propose their use in VL diagnosis. However, we are investigating the discriminative diagnostic potential of antigens separated on acrylamide gels.

Recommendations for further validation of IgG1 assays

We propose that a prospective study, with extended follow-up of a larger cohort of treated VL patients, should be used to validate the use of IgG1 ELISA and the VL Sero K-SeT for confirming cure in all endemic areas and defining the optimal time for testing, which may differ between regions. This longer follow-up would also indicate the potential of elevated IgG1 to predict relapse and PKDL and in turn, link these with different treatment regimens. A more extensive study of PKDL is required to determine the potential role of IgG1 in identifying PKDL as distinct from leprosy and fungal skin diseases (Saha et al., 2005; Mondal and Khan, 2011). In addition, use of the IgG1 assays on a much larger panel of seropositive asymptomatic individuals would help to define its role in predicting progression to VL. In all cases, comparison with existing diagnostics, including definitive parasitological methods, would directly assess the advantage of IgG1 assays.

Technical refinement of the VL Sero K-SeT should consider the use of electronic RDT readers to give an objective assessment of test band intensity. In addition, the identification of specific antigens suitable to replace the use of parasite lysate would obviate issues regarding batch-to-batch variation. These developments could improve precision of IgG1 readings and reproducibility. A comparison of whole blood and serum/plasma is also required for point-of-

care use, although a study in Bangladesh on various VL RDTs did find high agreement between the two sample types (Ghosh et al., 2015).

Conclusion

IgG1 assays, particularly in the VL Sero K-SeT RDT format, may be a useful adjunct in the assessment of VL treatment outcome and diagnosis of PKDL, which have been identified as research priorities for VL (World Health Organization, 2012). With additional refinement and validation, the VL Sero K-SeT and IgG1 ELISA could contribute to life-saving follow-up of treated patients and to control programme monitoring, surveillance, and targeting of strategies for long-term control of VL.

Acknowledgements

We thank Abhishek Kumar Singh and Keshav Rai for their kind assistance during the laboratory work.

Author contributions

Conceived and designed the study: MM. Wrote the manuscript: TM, TB, MM. Performed the experiments/Collected data: TM, TB, CP, BG, SA, MR, KH, HH. Analysed data: TM, TB, CP, SA, MR, KH, HH. Provided materials: OS, PM, QG, CT, BH, AF, OE, AS, BK, NB, SR, SE, SS. Provided feedback on final draft: AF, MM, OS, PM. Supervised the project: MM, TB, TM. Obtained funding: MB, MM.

Conflict of interest statement

PM, QG and CT are employed by Coris BioConcept which developed the VL Sero K-SeT.

Keywords

Visceral leishmaniasis, diagnosis, serology, treatment, relapse, cure, IgG1, RDT, PKDL.

Supplementary Material

These files are available online in the published version of this chapter by Marlais *et al*. in Frontiers in Cellular and Infection Microbiology.

S1: Detailed SDS-PAGE and western blotting methods.

S2: Spreadsheet with ELISA, RDT & blot results for Indian cured paired, PKDL and relapse samples.

S3: Images of western blots for Indian paired cured samples.

S4: Images of VL Sero K-SeT for Indian paired cured samples, corresponding to those in S3.

S5: Images of VL Sero K-SeT and western blots for Indian and Sudanese relapsed samples.

S6: Images of VL Sero K-SeT and western blots for Indian PKDL samples.

S7: Images of VL Sero K-SeT and western blots for Indian asymptomatic progressors and non-progressors.

Funding

This work was part of the NIDIAG network research partnership supported by the European Commission under the Health Cooperation Work Programme of the 7th Framework Programme (Grant agreement no. 260260, http://cordis.europa.eu/fp7/home_en.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

TM is funded by the Sir Halley Stewart Trust (http://www.sirhalleystewart.org.uk/). The views expressed within this report are those of the authors and not necessarily those of the Trust. TM was additionally supported by the John Henry Memorial Fund (http://www.johnhenrymf.org/).

BH is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642609. (https://ec.europa.eu/programmes/horizon2020/)

References

Abass, E., Kang, C., Martinkovic, F., Semiao-Santos, S.J., Sundar, S., Walden, P., et al. (2015). Heterogeneity of Leishmania donovani parasites complicates diagnosis of visceral leishmaniasis: comparison of different serological tests in three endemic regions. *PLoS One* 10(3), e0116408. doi: 10.1371/journal.pone.0116408.

Atia, A.M., Mumina, A., Tayler-Smith, K., Boulle, P., Alcoba, G., Elhag, M.S., et al. (2015). Sodium stibogluconate and paromomycin for treating visceral leishmaniasis under routine conditions in eastern Sudan. *Trop Med Int Health* 20(12), 1674-1684. doi: 10.1111/tmi.12603.

Bern, C., Haque, R., Chowdhury, R., Ali, M., Kurkjian, K.M., Vaz, L., et al. (2007). The epidemiology of visceral leishmaniasis and asymptomatic leishmanial infection in a highly endemic Bangladeshi village. *Am J Trop Med Hyg* 76(5), 909-914.

Bhattacharyya, T., Ayandeh, A., Falconar, A.K., Sundar, S., El-Safi, S., Gripenberg, M.A., et al. (2014a). IgG1 as a potential biomarker of post-chemotherapeutic relapse in visceral

494 leishmaniasis, and adaptation to a rapid diagnostic test. *PLoS Negl Trop Dis* 8(10)**,**
495 e3273. doi: 10.1371/journal.pntd.0003273.
496 Bhattacharyya, T., Bowes, D.E., El-Safi, S., Sundar, S., Falconar, A.K., Singh, O.P., et al. (2014b).
497 Significantly lower anti-Leishmania IgG responses in Sudanese versus Indian visceral
498 leishmaniasis. *PLoS Negl Trop Dis* 8(2)**,** e2675. doi: 10.1371/journal.pntd.0002675.
499 Bhattarai, N.R., Van der Auwera, G., Khanal, B., De Doncker, S., Rijal, S., Das, M.L., et al.
500 (2009). PCR and direct agglutination as Leishmania infection markers among healthy
501 Nepalese subjects living in areas endemic for kala-azar. *Trop Med Int Health* 14(4)**,**
502 404-411. doi: 10.1111/j.1365-3156.2009.02242.x.
503 Bittencourt, A., Silva, N., Straatmann, A., Nunes, V.L., Follador, I., and Badaro, R. (2003). Post-
504 kala-azar dermal leishmaniasis associated with AIDS. *Braz J Infect Dis* 7(3)**,** 229-233.
505 Boelaert, M., Meheus, F., Sanchez, A., Singh, S.P., Vanlerberghe, V., Picado, A., et al. (2009).
506 The poorest of the poor: a poverty appraisal of households affected by visceral
507 leishmaniasis in Bihar, India. *Trop Med Int Health* 14(6)**,** 639-644. doi: 10.1111/j.1365-
508 3156.2009.02279.x.
509 Burza, S., Nabi, E., Mahajan, R., Mitra, G., and Lima, M.A. (2013). One-year follow-up of
510 immunocompetent male patients treated with miltefosine for primary visceral
511 leishmaniasis in Bihar, India. *Clin Infect Dis* 57(9)**,** 1363-1364. doi: 10.1093/cid/cit508.
512 Burza, S., Sinha, P.K., Mahajan, R., Lima, M.A., Mitra, G., Verma, N., et al. (2014). Risk factors
513 for visceral leishmaniasis relapse in immunocompetent patients following treatment
514 with 20 mg/kg liposomal amphotericin B (Ambisome) in Bihar, India. *PLoS Negl Trop*
515 *Dis* 8(1)**,** e2536. doi: 10.1371/journal.pntd.0002536.
516 Celesia, B.M., Cacopardo, B., Massimino, D., Gussio, M., Tosto, S., Nunnari, G., et al. (2014).
517 Atypical presentation of PKDL due to Leishmania infantum in an HIV-infected patient
518 with relapsing visceral leishmaniasis. *Case Rep Infect Dis* 2014**,** 370286. doi:
519 10.1155/2014/370286.
520 Chapman, L.A.C., Dyson, L., Courtenay, O., Chowdhury, R., Bern, C., Medley, G.F., et al.
521 (2015). Quantification of the natural history of visceral leishmaniasis and
522 consequences for control. *Parasit Vectors* 8**,** 521. doi: 10.1186/s13071-015-1136-3.
523 Costa, C.H., Stewart, J.M., Gomes, R.B., Garcez, L.M., Ramos, P.K., Bozza, M., et al. (2002).
524 Asymptomatic human carriers of Leishmania chagasi. *Am J Trop Med Hyg* 66(4)**,** 334-
525 337.
526 Cota, G.F., de Sousa, M.R., Demarqui, F.N., and Rabello, A. (2012). The diagnostic accuracy of
527 serologic and molecular methods for detecting visceral leishmaniasis in HIV infected

patients: meta-analysis. *PLOS Negl Trop Dis* 6(5)**,** e1665. doi: 10.1371/journal.pntd.0001665.

Das, V.N., Ranjan, A., Pandey, K., Singh, D., Verma, N., Das, S., et al. (2012). Clinical epidemiologic profile of a cohort of post-kala-azar dermal leishmaniasis patients in Bihar, India. *Am J Trop Med Hyg* 86(6)**,** 959-961. doi: 10.4269/ajtmh.2012.11-0467.

Das, V.N.R., Pandey, R.N., Siddiqui, N.A., Chapman, L.A.C., Kumar, V., Pandey, K., et al. (2016). Longitudinal study of transmission in households with visceral leishmaniasis, asymptomatic infections and PKDL in highly endemic villages in Bihar, India. *PLOS Negl Trop Dis* 10(12)**,** e0005196. doi: 10.1371/journal.pntd.0005196.

el Hassan, A.M., Ghalib, H.W., Zijlstra, E.E., Eltoum, I.A., Satti, M., Ali, M.S., et al. (1992). Post kala-azar dermal leishmaniasis in the Sudan: clinical features, pathology and treatment. *Trans R Soc Trop Med Hyg* 86(3)**,** 245-248.

Ghosh, P., Hasnain, M.G., Ghosh, D., Hossain, F., Baker, J., Boelaert, M., et al. (2015). A comparative evaluation of the performance of commercially available rapid immunochromatographic tests for the diagnosis of visceral leishmaniasis in Bangladesh. *Parasit Vectors* 8**,** 331. doi: 10.1186/s13071-015-0935-x.

Gidwani, K., Kumar, R., Rai, M., and Sundar, S. (2009). Longitudinal seroepidemiologic study of visceral leishmaniasis in hyperendemic regions of Bihar, India. *Am J Trop Med Hyg* 80(3)**,** 345-346. doi: 10.4269/ajtmh.2009.80.345.

Gidwani, K., Picado, A., Ostyn, B., Singh, S.P., Kumar, R., Khanal, B., et al. (2011). Persistence of Leishmania donovani antibodies in past visceral leishmaniasis cases in India. *Clin Vaccine Immunol* 18(2)**,** 346-348. doi: 10.1128/cvi.00473-10.

Gorski, S., Collin, S.M., Ritmeijer, K., Keus, K., Gatluak, F., Mueller, M., et al. (2010). Visceral leishmaniasis relapse in Southern Sudan (1999-2007): a retrospective study of risk factors and trends. *PLoS Negl Trop Dis* 4(6)**,** e705. doi: 10.1371/journal.pntd.0000705.

Hasker, E., Kansal, S., Malaviya, P., Gidwani, K., Picado, A., Singh, R.P., et al. (2013). Latent infection with Leishmania donovani in highly endemic villages in Bihar, India. *PLoS Negl Trop Dis* 7(2)**,** e2053. doi: 10.1371/journal.pntd.0002053.

Hasker, E., Malaviya, P., Gidwani, K., Picado, A., Ostyn, B., Kansal, S., et al. (2014). Strong association between serological status and probability of progression to clinical visceral leishmaniasis in prospective cohort studies in India and Nepal. *PLoS Negl Trop Dis* 8(1)**,** e2657. doi: 10.1371/journal.pntd.0002657.

Hirve, S., Boelaert, M., Matlashewski, G., Mondal, D., Arana, B., Kroeger, A., et al. (2016). Transmission dynamics of visceral leishmaniasis in the Indian subcontinent - a

systematic literature review. *PLoS Negl Trop Dis* 10(8)**,** e0004896. doi: 10.1371/journal.pntd.0004896.

le Fichoux, Y., Quaranta, J.F., Aufeuvre, J.P., Lelievre, A., Marty, P., Suffia, I., et al. (1999). Occurrence of Leishmania infantum parasitemia in asymptomatic blood donors living in an area of endemicity in southern France. *J Clin Microbiol* 37(6)**,** 1953-1957.

Molina, R., Ghosh, D., Carrillo, E., Monnerat, S., Bern, C., Mondal, D., et al. (2017). Infectivity of post-kala-azar dermal leishmaniasis patients to sand flies: revisiting a proof of concept in the context of the kala-azar elimination program in the Indian subcontinent. *Clin Infect Dis* 65(1)**,** 150-153. doi: 10.1093/cid/cix245.

Mondal, D., Bern, C., Ghosh, D., Rashid, M., Molina, R., Chowdhury, R., el al. (2018). Quantifying the infectiousness of post-kala-azar dermal leishmaniasis towards sandflies. *Clin Infect Dis*, ciy891. doi: 10.1093/cid/ciy891.

Mondal, D., and Khan, M.G.M. (2011). Recent advances in post-kala-azar dermal leishmaniasis. *Curr Opin Infect Dis* 24(5)**,** 418-422. doi: 10.1097/QCO.0b013e32834a8ba1.

Morales, M.A., Cruz, I., Rubio, J.M., Chicharro, C., Cañavate, C., Laguna, F., et al. (2002). Relapses versus reinfections in patients coinfected with Leishmania infantum and Human Immunodeficiency Virus Type 1. *J Infect Dis* 185(10)**,** 1533-1537. doi: 10.1086/340219.

Moulik, S., Chaudhuri, S.J., Sardar, B., Ghosh, M., Saha, B., Das, N.K., et al. (2017). Monitoring of parasite kinetics in Indian post kala azar dermal leishmaniasis. *Clin Infect Dis*. doi: 10.1093/cid/cix808.

Mukhopadhyay, D., Das, N.K., De Sarkar, S., Manna, A., Ganguly, D.N., Barbhuiya, J.N., et al. (2012). Evaluation of serological markers to monitor the disease status of Indian post kala-azar dermal leishmaniasis. *Trans R Soc Trop Med Hyg* 106(11)**,** 668-676. doi: 10.1016/j.trstmh.2012.07.005.

Mukhtar, M., Ali, S.S., Boshara, S.A., Albertini, A., Monnerat, S., Bessell, P., et al. (2018). Sensitive and less invasive confirmatory diagnosis of visceral leishmaniasis in Sudan using loop-mediated isothermal amplification (LAMP). *PLoS Negl Trop Dis* 12(2)**,** e0006264. doi: 10.1371/journal.pntd.0006264.

Musa, A.M., Khalil, E.A., Raheem, M.A., Zijlstra, E.E., Ibrahim, M.E., Elhassan, I.M., et al. (2002). The natural history of Sudanese post-kala-azar dermal leishmaniasis: clinical, immunological and prognostic features. *Ann Trop Med Parasitol* 96(8)**,** 765-772. doi: 10.1179/000349802125002211.

596 Ostyn, B., Gidwani, K., Khanal, B., Picado, A., Chappuis, F., Singh, S.P., et al. (2011). Incidence
597 of symptomatic and asymptomatic Leishmania donovani infections in high-endemic
598 foci in India and Nepal: a prospective study. *PLoS Negl Trop Dis* 5(10)**,** e1284. doi:
599 10.1371/journal.pntd.0001284.
600 Peeling, R.W., Holmes, K.K., Mabey, D., and Ronald, A. (2006). Rapid tests for sexually
601 transmitted infections (STIs): the way forward. *Sex Transm Infect* 82(suppl 5)**,** v1.
602 Ridolfo, A.L., Gervasoni, C., Antinori, S., Pizzuto, M., Santambrogio, S., Trabattoni, D., et al.
603 (2000). Post-kala-azar dermal leishmaniasis during highly active antiretroviral therapy
604 in an AIDS patient infected with Leishmania infantum. *J Infect* 40(2)**,** 199-202.
605 Rijal, S., Ostyn, B., Uranw, S., Rai, K., Bhattarai, N.R., Dorlo, T.P.C., et al. (2013). Increasing
606 failure of miltefosine in the treatment of kala-azar in Nepal and the potential role of
607 parasite drug resistance, reinfection, or noncompliance. *Clin Infect Dis* 56(11)**,** 1530-
608 1538. doi: 10.1093/cid/cit102.
609 Roustan, G., Jimenez, J.A., Gutierrez-Solar, B., Gallego, J.L., Alvar, J., and Patron, M. (1998).
610 Post-kala-azar dermal leishmaniasis with mucosal involvement in a kidney transplant
611 recipient: treatment with liposomal amphotericin B. *Br J Dermatol* 138(3)**,** 526-528.
612 Saha, P., Ganguly, S., Chatterjee, M., Das, S.B., Kundu, P.K., Guha, S.K., et al. (2017).
613 Asymptomatic leishmaniasis in kala-azar endemic areas of Malda district, West
614 Bengal, India. *PLoS Negl Trop Dis* 11(2)**,** e0005391. doi:
615 10.1371/journal.pntd.0005391.
616 Saha, S., Mazumdar, T., Anam, K., Ravindran, R., Bairagi, B., Saha, B., et al. (2005). Leishmania
617 promastigote membrane antigen-based enzyme-linked immunosorbent assay and
618 immunoblotting for differential diagnosis of Indian post-kala-azar dermal
619 leishmaniasis. *J Clin Microbiol* 43(3)**,** 1269-1277. doi: 10.1128/jcm.43.3.1269-
620 1277.2005.
621 Sarnoff, R., Desai, J., Desjeux, P., Mittal, A., Topno, R., Siddiqui, N.A., et al. (2010). The
622 economic impact of visceral leishmaniasis on rural households in one endemic district
623 of Bihar, India. *Trop Med Int Health* 15 Suppl 2**,** 42-49. doi: 10.1111/j.1365-
624 3156.2010.02516.x.
625 Schägger, H. (2006). Tricine-SDS-PAGE. *Nat Protoc* 1(1)**,** 16-22. doi: 10.1038/nprot.2006.4.
626 Singh, O.P., and Sundar, S. (2015). Developments in diagnosis of visceral leishmaniasis in the
627 elimination era. *J Parasitol Res* 2015**,** 239469. doi: 10.1155/2015/239469.
628 Singh, R.P., Picado, A., Alam, S., Hasker, E., Singh, S.P., Ostyn, B., et al. (2012). Post-kala-azar
629 dermal leishmaniasis in visceral leishmaniasis-endemic communities in Bihar, India.
630 *Trop Med Int Health* 17(11)**,** 1345-1348. doi: 10.1111/j.1365-3156.2012.03067.x.

631 Srivastava, P., Gidwani, K., Picado, A., Van der Auwera, G., Tiwary, P., Ostyn, B., et al. (2013).
632 Molecular and serological markers of Leishmania donovani infection in healthy
633 individuals from endemic areas of Bihar, India. *Trop Med Int Health* 18(5)**,** 548-554.
634 doi: 10.1111/tmi.12085.
635 Sundar, S., and Rai, M. (2002). Laboratory diagnosis of visceral leishmaniasis. *Clin Diagn Lab*
636 *Immunol* 9(5)**,** 951-958.
637 Topno, R.K., Das, V.N., Ranjan, A., Pandey, K., Singh, D., Kumar, N., et al. (2010).
638 Asymptomatic infection with visceral leishmaniasis in a disease-endemic area in
639 Bihar, India. *Am J Trop Med Hyg* 83(3)**,** 502-506. doi: 10.4269/ajtmh.2010.09-0345.
640 Trindade, M.A., Silva, L.L., Braz, L.M., Amato, V.S., Naafs, B., and Sotto, M.N. (2015). Post-
641 kala-azar dermal leishmaniasis and leprosy: case report and literature review. *BMC*
642 *Infect Dis* 15**,** 543. doi: 10.1186/s12879-015-1260-x.
643 Uranw, S., Ostyn, B., Rijal, A., Devkota, S., Khanal, B., Menten, J., et al. (2011). Post-kala-azar
644 dermal leishmaniasis in Nepal: a retrospective cohort study (2000-2010). *PLoS Negl*
645 *Trop Dis* 5(12)**,** e1433. doi: 10.1371/journal.pntd.0001433.
646 World Health Organization (2010). "Technical Report Series. 949. Control of the
647 leishmaniases", in: *WHO Technical Report Series.* (Geneva, Switzerland: World Health
648 Organization).
649 World Health Organization (2012). "Technical Report Series. 975. Research priorities for
650 Chagas disease, human African trypanosomiasis and leishmaniasis", in: *WHO*
651 *Technical Report Series.* (Geneva, Switzerland: World Health Organization).
652 World Health Organization. (2017). *Leishmaniasis fact sheet* [Online]. Available:
653 http://www.who.int/mediacentre/factsheets/fs375/en/ [Accessed 10/1/2018].
654 Zijlstra, E.E., Alves, F., Rijal, S., Arana, B., and Alvar, J. (2017). Post-kala-azar dermal
655 leishmaniasis in the Indian subcontinent: a threat to the South-East Asia region kala-
656 azar elimination programme. *PLoS Negl Trop Dis* 11(11)**,** e0005877. doi:
657 10.1371/journal.pntd.0005877.
658 Zijlstra, E.E., Musa, A.M., Khalil, E.A., el-Hassan, I.M., and el-Hassan, A.M. (2003). Post-kala-
659 azar dermal leishmaniasis. *Lancet Infect Dis* 3(2)**,** 87-98.

660

Tables and Figures

Table 1. Indian sample types and total numbers tested by IgG1 assays.

Sample type	Definitions	n
Pre- and post-treatment pairs, deemed cured	Treated for VL, with improvement in clinical symptoms and no evidence of relapse at any time 6 months after treatment. Samples were taken at or around the start of treatment and at 6 months.	105 pairs
Relapse	VL treated and subsequently relapsed to active disease. Sampled at the time of relapse diagnosis.	33
PKDL	Samples taken at or up to 30 days after diagnosis of PKDL. Parasite infection was confirmed by PCR or a slit-skin smear or biopsy.	63
Asymptomatic	Asymptomatic seropositive, on the basis of DAT and/or rK39 ELISA, without symptoms or history of VL. Progressors (n = 4) developed VL after sampling. Non-progressors (n= 4) did not develop VL during follow-up of at least 3 years.	8
Other diseases	Malaria (n = 5); tuberculosis (n = 5), rheumatoid arthritis (n = 1); dengue (n = 1); multiple myeloma (n = 1).	13
Endemic healthy control	Resident in VL endemic area, seronegative by DAT and rK39 ELISA, no history of VL, healthy.	30

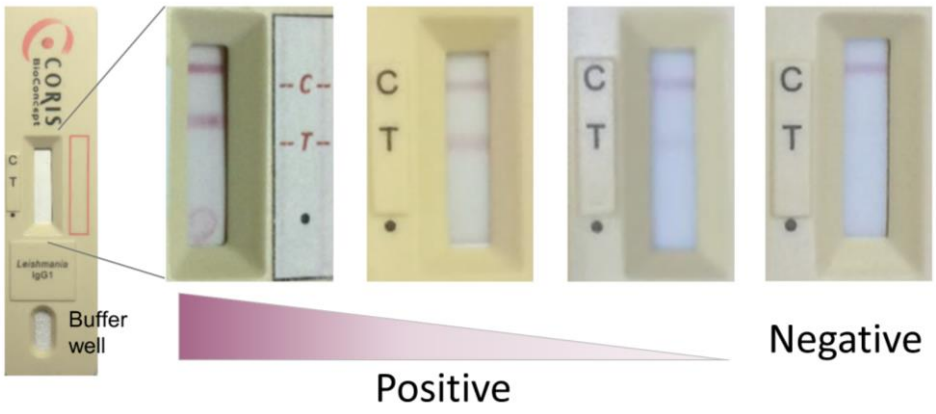


Figure 1. Representative examples of VL Sero K-SeT test line intensity. C: control line, T: test line, dot (•) indicates where sample is applied. Test strip manufacture was identical despite being housed in different cassettes. Image is best viewed in digital, colour format.

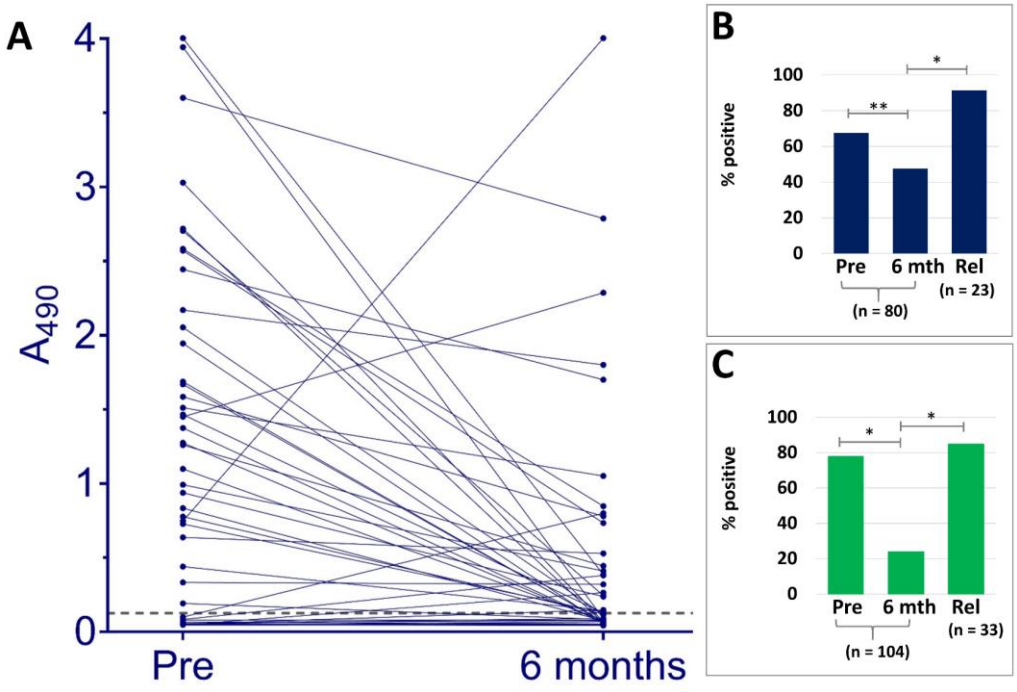


Figure 2. IgG1 anti *L. donovani* assays with Indian VL samples detect differences according to treatment outcome. (A) ELISA A_{490} change between paired pre-treatment samples and at 6 months post-treatment when deemed cured. Dashed line indicates cut-off (A_{490} = 0.128). Positivity rates with paired pre-treatment and cured samples at 6 months (6 mth), and non-paired relapse (Rel) for (B) ELISA, (C) VL Sero K-SeT. * $p < 0.0001$, ** $p = 0.0162$.

Table 2. Change in IgG1 response of pre- and post-treatment paired samples from India and Sudan.

Change in IgG1 response	**ELISA A_{490}**	**VL Sero K-SeT test line intensity**	
	India n (%)	**India n (%)**	**Sudan n (%)**
Positive to negative	21 (26.3%)	56 (53.8%)	7 (41.2%)
Positive clear decrease	30 (37.5%)	12 (11.5%)	3 (17.6%)
Positive no clear decrease	3 (3.8%)	13 (12.5%)	3 (17.6%)
Negative no change	21 (26.3%)	23 (22.1%)	3 (17.6%)
Negative to positive	5 (6.3%)	0 (0%)	1 (5.9%)

Total	80 (100%)	104 (100%)	17 (100%)

677

CHAPTER 4: Capture and identification of *Leishmania donovani* protein antigens in human urine during visceral leishmaniasis

Marlais T, Bhattacharyya T, Pearson C, Gardner BL, Marhoon S, Airs S, Singh OP, Reed SG, El-Safi S, Sundar S, Miles MA. Capture and identification of *Leishmania donovani* protein antigens in human urine during visceral leishmaniasis. Submitted.

Key points, novel results and implications

- Immuno-capture methods were used to concentrate *L. donovani* antigens from urine of VL patients prior to identification of proteins by mass spectrometry and subsequent *Leishmania*-specific epitope identification within these proteins using online tools.
- Seven *L. donovani* proteins were identified which have not previously been identified in VL urine.
- This manuscript presents a rationally-selected list of potential protein antigens specific to *Leishmania* that can be synthesised and used to raise antibodies for development of a urine antigen capture assay.
- This work furthers the available information on VL urine antigens and can contribute to the development or improvement of a non-invasive diagnostic test for VL which could be used to diagnose and confirm cure. This has implications to improve long-term VL disease control among the poorest and most neglected communities.

Candidate's contribution

The candidate performed all laboratory components on Sudanese VL urine. This included: rabbit antibody purification from serum; analysing the response of the rabbit antibody by ELISA; conjugating rabbit antibody to an affinity column; urine preparation and immunocapture; SDS PAGE gels and western blot; conjugating rabbit antibodies to the enzyme horseradish peroxidase. The candidate wrote the method and co-supervised the laboratory work with Indian urine. Preparation and submission of all samples to Oxford University for mass spectrometry was by the candidate, as was the large majority of mass spectrometry data analysis, all computational protein analysis, and manuscript writing. Sudanese urine processing was carried out prior to the candidate's PhD registration. Indian

urine processing, mass spectrometry data analysis, computational protein analysis and manuscript writing were carried out during the candidate's PhD.

Contribution of others

Tapan Bhattacharyya, supervised lab work in India and contributed to revision of the manuscript.

Callum Pearson, Bathsheba Gardner and Stephanie Airs, carried out lab work in India.

Safiyyah Marhoon, assisted with interpretation and analysis of mass spec data.

Om Prakash Sing, Steven Reed, Sayda El Safi, Shyam Sundar provided materials including human urine samples and antibodies.

Michael Miles co-ordinated the study.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Tegwen Marlais
Principal Supervisor	Prof. Michael A. Miles
Thesis Title	Diagnostics development for the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Journal of Clinical Microbiology
Please list the paper's authors in the intended authorship order:	Tegwen Marlais, Tapan Bhattacharyya, Callum Pearson, Bathsheba L. Gardner, Safiyyah Marhoon, Stephanie Airs, Kiera Hayes, Andrew K. Falconar, Om Prakash Singh, Steven G. Reed, Sayda El-Safi, Shyam Sundar, Michael A. Miles
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I performed all laboratory components on Sudanese VL urine. This included: rabbit antibody purification from serum; analysing the response of the rabbit antibody by ELISA; conjugating rabbit antibody to an affinity column; urine preparation and

	immunocapture; SDS PAGE gels and western blot; conjugating rabbit antibodies to the enzyme horseradish peroxidase. I devised the method and supervised the laboratory work with Indian urine. I prepared and submitted all samples to Oxford University for mass spectrometry, performed the large majority of mass spectrometry data analysis, all computational protein analysis, and manuscript writing.

Student Signature: [redacted] **Date:** 16/2/19

Supervisor Signature [redacted] **Date:** 15/2/19

Isolation and characterisation of *Leishmania donovani* protein antigens from human urine during visceral leishmaniasis

Running title: *Leishmania* antigens in visceral leishmaniasis urine

Tegwen Marlais[a#], Tapan Bhattacharyya[a], Callum Pearson[a], Bathsheba L. Gardner[a], Safiyyah Marhoon[a], Stephanie Airs[a], Kiera Hayes[a], Andrew K. Falconar[b], Om Prakash Singh[c], Steven G. Reed[d], Sayda El-Safi[e], Shyam Sundar[c], Michael A. Miles[a]

a Pathogen Molecular Biology Department, Faculty of Infectious and Tropical Disease, London School of Hygiene and Tropical Medicine, London, UK

b Departamento de Medicina, Universidad del Norte, Barranquilla, Colombia

c Department of Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India

d Infectious Disease Research Institute, Seattle, Washington, United States of America

e Faculty of Medicine, University of Khartoum, Khartoum, Sudan

Corresponding author tegwen.marlais@lshtm.ac.uk

Abstract

Background and Aim: Diagnosis of visceral leishmaniasis (VL) relies on invasive and risky aspirate procedures, and confirmation of cure after treatment is unreliable. Detection of *Leishmania donovani* antigens in urine has the potential to provide both a non-invasive diagnostic and a test of cure. We searched for *L. donovani* antigens in urine of VL patients from India and Sudan to contribute to the development of urine antigen capture immunoassays.

Method: VL urine samples from India and Sudan were incubated with immobilised anti-*L. donovani* polyclonal antibodies and captured material was eluted. Sudanese eluted material and concentrated VL urine were analysed by western blot. Immunocaptured and immunoreactive material from Indian and Sudanese urine was submitted to mass spectrometry for protein identification.

Results: We identified seven *L. donovani* proteins from VL urine. Named proteins were 40S ribosomal protein S9, kinases, and others were hypothetical. Thirty five epitope regions were predicted with high specificity in the 7 proteins. Of these, 22 were highly specific to *Leishmania* spp.

Conclusion: We identified *L. donovani* protein antigens in VL urine from Indian and Sudanese patients. Predicted epitope regions of these proteins are highly suitable for raising antibodies for the subsequent development of an antigen capture assay. We provide the amino acid sequences of these epitopes. Combining these antigens for antibody generation, or as assay targets, could improve sensitivity and specificity of the prospective assay.

Introduction

Visceral leishmanisis (VL) is most commonly caused by *Leishmania donovani* in the Indian subcontinent and East Africa, whereas *L. infantum* is the agent in the Mediterranean, Middle East and South America. Both species are transmitted by female phlebotomine sand flies and symptomatic infection is considered fatal if untreated, therefore accurate diagnosis is crucial to patient outcome. India, Bangladesh and Nepal are aiming to eliminate VL as a public health problem and this relies on rapid case detection and confirmation of cure after treatment (1).

Diagnosis of VL is based on serology, commonly using the recombinant rK39 or rK28 antigens, followed by microscopic visualisation of the parasite in spleen, bone marrow or lymph node aspirate as confirmation. Conventional serology, which detects anti-*Leishmania* IgG antibodies, has several drawbacks: it is ineffective at confirming cure or relapse because it can remain positive for many years after successful treatment (2-6); it is also less reliable in HIV co-infected cases where a negative result does not rule out leishmaniasis (7).

An ideal diagnostic for both primary VL cases and validating cure is the detection of parasite material in non-invasive samples such as urine or saliva, or a serological test that is specific for active infection (8). As well, there is the need for low-cost, rapid and equipment-free diagnostics that can be used in low-resource settings at point-of-care with minimal training. Such assays may detect parasite DNA, for example by loop-mediated isothermal amplification (LAMP) (9, 10) or by recombinase polymerase amplification (RPA) (11), or may detect parasite antigens.

Several urine antigen capture immunoassays have been developed with the best established being the KAtex, a latex particle agglutination test that detects a carbohydrate antigen (12, 13). The KAtex has a specificity of 84-100%, but poorer sensitivity of 47-87% (14-17) with the drawback that urine samples must be boiled

before testing. However, the test is rapid, giving a result in less than 10 minutes and becoming negative for most patients 30 days post-treatment (17). In addition, this urine antigen assay has shown utility in HIV/VL co-infection (18, 19). Monoclonal and polyclonal antibodies against the antigen in the KAtex test were later adapted to ELISA format, which removed the need for boiling urine (20).

Other assays have been reported that detect particular protein antigens of *L. infantum* in urine (21). This approach required first identifying *Leishmania* proteins in VL urine by mass spectrometry, expressing them as recombinant antigens and raising antibodies that could be produced as highly specific and sensitive polyclonal or monoclonal antibodies (21).

An alternative approach is to raise antibodies to lysed whole parasite cells, containing a wide diversity of antigens and to use these to capture a range of undefined antigens from VL patient urine. A. C. Vallur et al. (22) reported the development of an ELISA using an affinity purified polyclonal rabbit antibody against *L. donovani* whole cell lysate. The assay was optimised by those authors and developed into an ELISA kit that performed well in detecting urine antigen in VL patients from both *L. infantum* and *L. donovani* endemic regions. We undertook a study using this and other polyclonal anti-*Leishmania* antibodies to detect and capture antigens from Indian and Sudanese VL urine. We then identified these parasite proteins using mass spectrometry.

Methods

Ethics statement

Ethical permission was granted by the LSHTM Ethical Review Committee with approval number 11478, and as part of the EC-funded NIDIAG project. In India, the collection of samples was approved by the Ethics Committee of Banaras Hindu University, Varanasi. In Sudan, approval was by the Ethical Research Committee, Faculty of Medicine, University of Khartoum and the National Health Research Ethics Committee, Federal Ministry of Health, Sudan. Written informed consent was obtained from adult subjects included in the study or from the parents or guardians of individuals less than 18 years of age.

Two rabbits were used here to raise antiserum. The inoculation and subsequent collection of serum from the rabbits was performed at the Royal Veterinary College, London, UK, in accordance with animal welfare standards required by law.

Processing and analysis of samples

Figure 1 depicts the overall workflow of the urine samples in this study.

Figure 1. Sample processing and analysis workflow used in this study. Asterisks indicate immunocapture methods. Anti-1S2D, -DD8 and -LV9 are rabbit antibodies against these respective strains of *Leishmania donovani*. VLu, concentrated proteins from VL urine; EHCu, concentrated proteins from urine of Endemic Healthy Control participants; uAg, material captured from VL urine by anti-*L. donovani* antibodies; LC-MS/MS, liquid chromatography tandem mass spectrometry.

Urine samples

Urine samples were from VL patients and healthy controls in two distinct VL endemic regions, namely the Indian state of Bihar and the Sudanese state of Gedaref. India: urine samples were collected from VL patients attending the Kala-Azar Medical Research Centre (KAMRC) clinic in Muzzarfarpur, Bihar. Sudan: VL and endemic healthy control (EHC) urine samples were collected from field sites in Gedaref.

Generation of rabbit anti-*L. donovani* antibodies

A. C. Vallur et al. (22) described the generation of a rabbit anti-*L. donovani* polyclonal antibody against whole promastigote lysate antigen, which was affinity purified against soluble promastigote lysate antigen of *L. donovani* strain MHOM/SD/00/1S-2D. This antibody is hereafter referred to as anti-1S2D.

Here, we prepared soluble promastigote lysate antigens of *L. donovani* strains MHOM/IN/80/DD8 and MHOM/ET/67/LV9 (also known as HU3), as described (23). Log phase promastigotes were washed 3 times in PBS, pelleted by centrifugation, flash frozen in liquid nitrogen and thawed 3 times, sonicated (Soniprep 150, MSE, UK) at 12 microns intensity for 3 x 30 seconds at 2-minute intervals on ice and finally centrifuged at 12,000 *x g* for 1 minute at 4°C. The supernatants containing soluble antigens were retained and used subsequently.

Soluble antigen (DD8 and LV9) was used to raise antiserum in two rabbits that were immunised percutaneously with two doses of the respective lysate. The first inoculation was with 200 µg - 500 µg of lysate antigen (originating from roughly 10^7 parasites) with Freund's complete adjuvant (F5881, Sigma-Aldrich). A second inoculation, four weeks later, was with the same amount of antigen, plus Freund's incomplete adjuvant (F5506, Sigma-Aldrich). Five millilitres of blood was taken before inoculations and about 50 ml five months after the second inoculation.

The resulting polyclonal IgG was purified from both DD8 and LV9 antisera on protein A agarose (P3476, Sigma, UK), eluted with 0.1 M glycine, pH 2.7 and immediately neutralised with 60 µl of 1 M Tris pH 9.0. The purified IgG elutions are hereafter

referred to as anti-DD8 and anti-LV9. An ELISA was performed to confirm the reactivity of the two antibodies with *L. donovani* antigen and that the pre-immune sera were non-reactive.

Precipitation and panning of Indian VL urine proteins

Total protein was precipitated from Indian VL patient urine: approximately 130 ml of urine (83 ml fresh and 52 ml previously frozen) was combined from 31 Indian VL patients and protease inhibitor cocktail (P8340, Sigma-Aldrich) was added at a dilution of 1:500. The urine was centrifuged at 2,500 x g for 3 minutes to pellet cellular debris, which was discarded. To the supernatant, we added 85 g solid ammonium sulphate to achieve 90% saturation then incubated the solution for one hour at ambient temperature (about 28°C) with mixing to allow proteins to precipitate. The urine/ammonium sulphate solution was centrifuged in aliquots at 3,900 *x g* at 20°C for 40 minutes and the supernatant was discarded. Protein pellets were then resuspended as two aliquots in a total of 4.25 ml PBS then desalted using PD10 columns (170851-01, GE Healthcare) by the centrifugation protocol and each replicate made up to 3 ml with PBS.

To capture *L. donovani* antigens from the precipitated Indian VL urine protein, we performed immuno-panning (Figure 1, upper left). Anti-1S2D antibody was coated onto 5 cm diameter plastic dishes at 8 µg/ml (Dish 8) and 35 µg/ml (Dish 35) in 1.5 ml coating buffer (0.1 M $NaHCO_3$, pH 8.6) by incubation at 4°C overnight. The unbound antibody was removed, replaced with blocking buffer (coating buffer containing 5 mg/ml bovine serum albumin) and the dishes were incubated at 4°C for 1 hour. After blocking, both dishes were washed 5 times using PBS containing 0.05% v/v Tween 20 (PBST), with gentle agitation.

The first aliquot of urine protein solution was incubated in Dish 8 for 1 hour at room temperature with agitation before the unbound material was moved into Dish 35. Dish 8 was washed 3 times with PBST then another 3 times with PBS before bound material was eluted with a 15 minute incubation with 1 ml of 0.1 M glycine pH 2.7. Eluate was pipetted into a tube containing 32 µl neutralising buffer (1 M Tris pH 9.0) and sodium azide to a final concentration of 0.02% (w/v) then stored at -80°C.

Unbound material from Dish 35 was transferred back to Dish 8 for repetition of the whole process. Four elutions were made from each aliquot of urine protein solution, 2 from each plate, to maximise the chance of capturing *Leishmania* antigens. After each elution, the dishes were washed 10 times with PBS prior to re-use. The second aliquot of urine protein solution was put through the same process yielding another 4 eluates (thus n = 8 in total).

Eluates containing captured proteins were pooled into two volumes: one from Dish 8, the other from Dish 35. These were buffer exchanged into PBS and spin concentrated at 5°C in 3 kDa molecular weight cut-off (MWCO) Amicon Ultra filters (UFC500324, Millipore) at 14,000 *x g* in several 20 minute spins until the total volume of each was reduced to 100 µl.

Concentration and affinity chromatography of Sudanese VL urine

Sudanese VL patients' urine from seven individuals was pooled, 0.4 ml from each to give 2.8 ml, and spin concentrated to 190 µl in Amicon Ultra 3 kDa MWCO tubes, with a final protein concentration of 12.5 mg/ml (hereafter 'VLu' for VL urine). Pooled EHC urine, 0.5 ml each from 14 individuals, was spin concentrated as above, down to about 100 µl (EHCu). Both urine concentrates were then washed to remove salts by making up to 0.5 ml with PBS and spinning down five times. These two concentrates, VLu and EHCu, were run in SDS-PAGE as described further below (Figure 1, upper right).

Separately, another pool of Sudanese VL urine comprising 0.5 ml each of 57 individual urines with 1/100 protease inhibitor (P8340, Sigma-Aldrich) was concentrated with a ProteoSpin Urine Protein Concentration Micro Kit (17400, Norgen Biotek Corp) as per manufacturer's instructions and in batches to avoid overloading the capacity of the kit to a final volume of 3.7 ml. To capture *L. donovani* antigens in urine concentrated in this manner, rabbit anti-DD8 was coupled to a cyanogen bromide-activated Sepharose matrix (C9142, Sigma-Aldrich) following manufacturer's instructions and the 1.75 ml of gel was loaded into a disposable column. The column was equilibrated using ten column volumes of PBS prior to use. Urine protein concentrate was incubated on the anti-DD8 column for 10 minutes, drained, washed 3 times with 10 ml PBS and bound material was eluted in 8 x 1 ml volumes of 0.1 M glycine pH 2.7, neutralised with 32 µl neutralising buffer per 1 ml eluate. Eluted fractions indicated to contain antigen by a dot blot were spin concentrated to give 'VL urine antigen (uAg)'. Briefly, the dot blot consisted of 2 µl of each fraction dried onto nitrocellulose, blocked with PBS + 3% w/v non-fat milk powder (Premier International Foods, Spalding, UK) (PBSM), probed with anti-DD8 and detected with anti-rabbit IgG-horseradish peroxidase (HRP) (711-035-152, Jackson Immunoresearch) and DAB/CN substrate, with details as described below.

SDS-PAGE and western blotting

Both of the Sudanese VL urine products, namely the urine concentrate (VLu) and immune-selected urine antigen (uAg), were subjected to non-reducing SDS-PAGE in duplicate lanes of a 10% acrylamide Tricine gel as described by H. Schägger (24).

Half of the gel was stained firstly for carbohydrate using Pro Q Emerald 300 (P21857, ThermoFisher Scientific) visualised at 280 nm (Gel Doc, ThermoFisher Scientific) then the same gel was stained for protein using Sypro Ruby (S12000, ThermoFisher Scientific) and visualised with 488 nm excitation and 610 nm emission wavelengths (Typhoon Trio phosphorimager, Amersham Biosciences) (Figure 1).

The other identical gel half was blotted onto 0.2 µm pore size nitrocellulose (10600015, GE Healthcare) at 160 mA for 120 minutes and air-dried. The membrane was then blocked using PBSM overnight at 4°C then washed in PBST, three times for 10 min each. Rabbit anti-DD8 and anti-LV9 were conjugated to HRP using a Lightning Link kit (701-0000, Innova Bioscience) then diluted together at 1:400 and 1:500 respectively in PBST + 3% milk powder (PBSTM) and incubated on the blot for 2 hours with gentle agitation. The membrane was washed with PBST six times for 5 minutes each before addition of DAB/CN substrate solution (30 µg 4-chloro-1-naphthol dissolved in 5 ml methanol then added to 40 ml PBS, followed by the addition of 10 µg of 3,3'-diaminobenzidine dissolved in 5 ml methanol and finally 15 µl of 30% H_2O_2). The blot was incubated for 15 mins with gentle agitation in the dark before stopping the reaction by washing in deionised water several times.

Due to poor signal gained from these rabbit polyclonals conjugated directly to HRP, the blot was stripped using stripping buffer (glycine 15 g/L, SDS 1 g/L, Tween 20 10 ml/L, pH 2.2) for 10 min with agitation, followed by 2 x 10 min washes in PBS and 2 x 5 min washes in Tris buffered saline with Tween (20 mM Tris, 150 mM NaCl, 0.05% Tween 20, pH 7.6) (TBST). After stripping away the previous reaction, the blot was re-blocked and probed with the same two rabbit anti-*L. donovani* antibodies without HRP conjugation, followed by anti-rabbit IgG-HRP at 1:1000 and DAB/CN substrate as before. Bands in the carbohydrate/protein-stained gel half, corresponding to immunoreactive blot bands in the two VL urine sample types, were excised for mass spectrometry.

Mass spectrometry

The excised gel bands from Sudanese VLu and uAg, as well as the two Indian VL urine solutions eluted from the panning dishes, were submitted to trypsin digest and liquid chromatography tandem mass spectrometry (LC-MS/MS) at the Advanced Proteomics Facility, Oxford University, UK.

For all samples, mass spectra were assigned to *L. donovani* peptides by a simultaneous search using Mascot (Matrix Science, UK) against both the *L. donovani* BPK282A1 reference and UniProt *Homo sapiens* protein databases, which contained all the proteins of these organisms, as deduced from their genomes (25). The

simultaneous search was used to ensure that peptides that continued to be identified as *Leishmania* were almost certainly not derived from human proteins. Criteria for identifying *Leishmania* peptides were being rank 1 (the peptide with the best match for a particular mass spectrum) and having a ppm of -10 to 10 (error on experimental peptide mass values, fraction expressed as parts per million). Proteins were identified when they contained 2 or more different peptides matching the above criteria and were taken forward to further analysis (Figure 1, lower panel). Once identified, peptides and proteins from VLu and uAg from Sudanese samples were considered together for subsequent analysis of their properties, as detailed below.

Epitope prediction

The complete amino acid sequences of proteins identified by 2 or more peptides were obtained through UniProtKB (25) and submitted to linear B-cell epitope prediction by BepiPred 2.0 (26). The epitope score threshold was set at 0.65 (on a scale of 0 to 1), which gave a specificity of 99% and sensitivity of 2%, with an alternative threshold of 0.55 for higher sensitivity (specificity 81.6%, sensitivity 29%). Minimum length was 8 residues, with no maximum.

Specificity of proteins and epitopes to *L. donovani*

Proteins and predicted epitope peptides from these proteins were assessed for sequence similarity to other species using a BLASTP search against the NCBI non-redundant (nr) database with no species restriction and using the default settings. For epitope sequences, the BLASTP search was optimised for short sequences. BLAST output was assessed for the sequence identity between the query and biologically-relevant species, i.e., human pathogens and commensals. Sequence similarity to other *Leishmania* species was tolerated if *L. donovani* was a complete match.

Signal peptide prediction

Proteins were submitted to SignalP 4.1 (27) to predict the presence of a signal peptide, which provided evidence that the protein may be secreted from the cell by this method. SignalP 4.1 is optimised to distinguish between transmembrane domains and signal peptides (27).

Glycosylation prediction

N-linked glycosylation of proteins was predicted using NetNGlyc (28), which identified potential glycosylation sites via a nitrogen atom of asparagine in the Asn-X-Ser/Thr motif, where X could be any amino acid residue except Pro.

Protein properties

The possible identities of hypothetical proteins identified by mass spectrometry were sought using online tools that identify domains and protein features based on sequence similarity with known proteins: NCBI domain enhanced lookup time (DELTA) BLAST tool (29); InterPro (30); BlastKOALA (KEGG Orthology and Links Annotation) via KEGG (31). Additional information on protein features was from Expasy Prosite (32) and published literature.

Results

Using mass spectrometry, we identified five *L. donovani* proteins in Indian VL urine and two in Sudanese VL urine. Epitope prediction returned 35 linear B-cell epitopes of 8 or more amino acids, within the complete protein sequences of the 7 proteins. Of the 35, 22 had complete identity only with *Leishmania* spp. in a BLAST search of the NCBI nr database of all organisms. Details of peptides and proteins from each urine source are described further below.

Indian VL urine antigen capture

In total from Indian VL urine, 20 different peptides were identified as those of *L. donovani* after a search of the mass spectra against *L. donovani* and human protein databases simultaneously; 12 from each panning dish, with four present in both dish eluates (Table S1).

Five proteins, one of which occurred on both panning dishes, were confidently identified by more than one peptide (Table 1). In addition to proteins, 9 solo peptides (i.e., the only representatives of their parent protein) were identified with a highly reliable ID by mass spectrometry and high specificity to *L. donovani* (Table S2).

Table 1. *Leishmania donovani* proteins identified by more than one peptide in urine of Indian and Sudanese visceral leishmaniasis patients.

VL urine origin	Origin of peptide	Peptide	Parent protein	UniProtKB / GenBank accession
India	Dish 8	EYEELR	Hypothetical protein, conserved	LdBPK_191140 / XP_003860289.1
		ALAEGQER		
		AKAEAEAAR		
	Dish 8 and 35	LSRSMEVR	40S ribosomal protein S9, putative	LdBPK_070760 / XP_003865205.1
		LSSVQAGEVR		
	Dish 8	VAVVLEGR	Hypothetical protein, conserved	LdBPK_363030 / XP_003865418.1
		RVAVVLEGR		
	Dish 35	FLDKLR	Hypothetical protein, conserved	LdBPK_323250 / XP_003863736.1
		RSSQSSTSATYR		
	Dish 35	ALISPSVLR	Protein kinase, putative	LdBPK_262110 / XP_003861796.1
		LSDAPRVCR		
Sudan	VLu	EFVVSGAALR	Hypothetical protein, conserved	LdBPK_160110 / XP_003859699.1
		ITDMQREIR		
	VLu	ITSDEVLR	Protein kinase, putative	LdBPK_351070 / XP_003864692.1
		TVNEDLSR		

Sudanese VL urine antigen capture

L. donovani peptides and proteins were identified in mass spectrometry data from excised gel bands corresponding to immunocaptured urine antigen (uAg) and immunoreactive VL urine components (VLu) (Figure 2). Protein and carbohydrate

staining of the SDS PAGE gel revealed that very few of these antigens were glycoproteins (Figure 2, B and C).

Figure 2. Sudanese VL patients' urine (VLu) and immunocaptured urine antigen (uAg) detected by rabbit anti-*L. donovani* DD8 and -LV9 by western blot. Regions submitted to mass spectrometry are broadly indicated by boxes. A) western blot, B) corresponding gel stained for proteins, C) the same gel as B, stained for carbohydrates. Molecular mass marker (M) sizes in kilodaltons. Lanes are: lysate antigen of *L. donovani* strain DD8; uAg, VL urine material eluted from an anti-*L. donovani* DD8 immunochromatography column; VLu, concentrated VL urine; EHCu, concentrated urine from endemic healthy controls.

Mass spectrometry of the Sudanese VL urine material (uAg and VLu) excised from the SDS PAGE gel confidently (>1 peptide) identified two proteins and nine additional solo peptides (Table 1 and Table S2). The peptides with a confident mass spectrometry ID had complete identity only to proteins from *Leishmania* genus according to a BLAST search, but part of some sequences occurred in other human pathogens or commensals. Specificity to *Leishmania* was improved when two peptides from the same protein were considered together.

Specificity of proteins to *L. donovani*

In total, we identified seven proteins with confidence (>1 peptide) from Indian (5 proteins) and Sudanese (2 proteins) VL urine that showed high specificity to *Leishmania* spp., although not necessarily to *L. donovani*. From Indian VL urine, protein kinase (LdBPK_262110) and hypothetical proteins (LdBPK_191140, LdBPK_363030, LdBPK_323250) showed very high sequence similarity to several *Leishmania* species including *L. infantum* and species causing cutaneous leishmaniasis, and moderate identity to *Trypanosoma cruzi* proteins. The 40S ribosomal protein S9 (LdBPK_070760) also had high sequence identity to that of other *Leishmania* species, but little to any other genera.

The two *L. donovani* proteins identified in Sudanese VL urine, hypothetical protein (LdBPK_160110) and protein kinase (LdBPK_351070), also had very high homology to *L. infantum* with minor sequence differences between these and other *Leishmania* species. Moderate homology was found between the hypothetical protein and those of *T. cruzi*, whereas the protein kinase had high homology with *T. cruzi* proteins.

In addition to the proteins, 18 solo peptides were identified in total by all methods, 9 each from Indian and Sudanese VL urine (Table S2). Six of these were from named proteins and all others were from hypothetical proteins. Interestingly, there was no overlap between *L. donovani* peptides from Indian and Sudanese VL urine.

Epitope prediction

All 7 proteins identified with confidence (>1 peptide) from the Indian and Sudanese VL urine were submitted to epitope prediction using BepiPred 2.0 and together contained 35 peptides at a score threshold of 0.65 which optimised specificity and 147 epitopes at a threshold score of 0.55, which provided higher sensitivity (Table S3). None of the 7 proteins contained signal peptides and 4 contained potential N-linked glycosylation sites (Table S3).

Potential VL urine antigens

The 35 epitope peptides predicted with high specificity, and one with lower specificity, were assessed for specificity to *L. donovani*. Twenty two of the 36 epitope sequences were selected as they had complete sequence identity to *L. donovani* and little or no identity to sequences from other relevant species *i.e.*, human pathogens or commensals. Specific peptides or epitope regions indicated for production of antibodies for use in antigen capture assays are detailed in Table 2.

Table 2. Epitope sequences within *L. donovani* proteins identified in Indian and Sudanese VL patient urine. Epitopes were predicted by BepiPred 2.0 with score ≥0.65 and selected based on high homology to *Leishmania* and low homology to other relevant genera.

Sample of origin	Parent protein (UniProtKB accession)	Epitope sequences with high specificity to *Leishmania*	Length (aa)
Indian VL urine	Hypothetical protein (LdBPK_191140)	VDDRTHREA	9
		QRQRQHAHA	9
		RRQRHTSP	8
		RNRPESSH	8
	40S ribosomal protein S9 (LdBPK_070760)	SSRRASTTKPGPPPRAS	17
		GMQLVGELNDSLD	13
		LDQQPSVGTTT	11
	Hypothetical protein	GSRHEKGQRGGRRS	14
		QKSQQQQQSTSS	12

	(LdBPK_363030)		
	Hypothetical protein (LdBPK_323250)	SDNGASPGSRSPRSSRRSSQSSTS	24
		SPAHQRSRAGASRSASRQG	19
		STKRPRQSAVYG	12
	Protein kinase (LdBPK_262110)	NSSSYSGSLGSPAS	14
		VSPVRRNSSSTAL	13
		ANGGNSSSNSYT	12
		QQQQQSSNRPS	11
		AGTARLGSSS	10
		RSTPRAGMP	9
Sudanese VL urine	Hypothetical protein (LdBPK_160110)	VRFRPNASLADGDAKSSAHGTVTQYGSPA‡	29
	Protein kinase (LdBPK_351070)	ANDDSESATRVEGLQVMSDINSIPL	25
		DGQQIKVSSSGGGSSSKGSSNSTGS	25
		KEERQRMHA	9

‡ This peptide was identified by the lower epitope score threshold of 0.55 because the protein did not contain epitopes >8 residues at a threshold of 0.65.

Discussion

Validated tests for urine antigen in infectious diseases include those for schistosomiasis (33), tuberculosis (34), *Legionella pneumophila* and *Streptococcus pneumoniae* (35). Here, we identified by mass spectrometry seven *L. donovani* proteins from urine of VL patients from India and Sudan. These proteins were predicted to contain highly species-specific epitopes that therefore make good candidates for targets of a non-invasive urine antigen capture immunoassay for VL. We additionally identified 18 single peptides with very high identity to *L. donovani*, which provide evidence of additional parasite proteins in VL urine.

Studies using similar methods to identify pathogen antigens in urine include that of C. Abeijon et al. (21) who identified the *L. infantum* proteins tryparedoxin peroxidase, superoxide dismutase and nuclear transport factor 2, by mass spectrometry of concentrated Brazilian VL urine excised from gel bands. We did not identify any of these proteins, and in addition, none of the peptides and proteins that we found occurred in both Indian and Sudanese samples. This suggests that there could be many proteins or protein fragments of parasite origin in the urine of VL patients.

The proteins identified here are likely to be intracellular, based on their identities and features. However, *Leishmania* has various secretion pathways, therefore the proteins may be secreted by non-classical mechanisms(36).

Therefore, although they did contain potential glycosylation sites, these are unlikely to be glycosylated *in vivo* as this is a feature more common to surface proteins involved directly in host-parasite interactions (37, 38). The carbohydrate and protein staining of the SDS-PAGE gel also indicated that very few proteins were glycosylated, and those that were, were not detected by antibodies on the corresponding blot. This may be expected as the rabbit antibodies were raised against a soluble lysate antigen from a preparation method that favours intracellular contents rather than cell membranes. By comparison, the KAtex assay that detects a carbohydrate antigen uses an antibody raised against whole intact parasites (12). Either antigen, lysate or whole cell, has the potential to generate an effective antigen-capturing assay.

Four of the 7 proteins identified in VL urine were hypothetical proteins, defined by the presence of start and stop codons in their genomic sequences but without experimental evidence for the protein itself. These proteins were submitted to several protein domain identification tools to reveal possible similarities to well characterised proteins and reveal their features and possible functions in *Leishmania* (Table 3).

Table 3. Functions of *Leishmania donovani* proteins identified in urine of Indian and Sudanese patients with visceral leishmaniasis.

VL urine origin	Protein accession number	Protein features	Functions
India	LdBPK_191140	Hypothetical protein with predicted zinc finger RING-type domain; predicted coil regions, a feature of predicted and existing antigens (39, 40).	Ubiquitination pathway and other intracellular protein processing pathways.
India	LdBPK_363030	Hypothetical protein with similarity to cilia- and flagella-associated protein 65, which in humans is an	In humans, may be linked to cellular differentiation and RNA binding (43) but role in *Leishmania* not studied.

		intracellular or membrane-anchored protein expressed in the lung and reproductive tract (41, 42).	
India	LdBPK_323250	Hypothetical protein with very weak homology to a yeast protein, DUF1630-domain-containing protein. May contain DENN domain. Protein largely specific to kinetoplastids by sequence analysis.	DENN domains are involved in GTP/GDP exchange and occur in other proteins that regulate membrane traffic in eukaryotes (44).
Sudan	LdBPK_160110	MORN repeat motif. Possible 2-isopropylmalate synthase.	Precise function of MORN motif unknown, but is present in the proteins of many eukaryotes, including trypanosomes (45). Possibly an acyltransferase involved in metabolism.
India	LdBPK_070760	40S ribosomal S9 protein	Protein subunit of the 40S ribosomal subunit. In *Leishmania*, a homologue has also been identified as a subunit of the unusual 45S small subunit in the kinetoplast, which may indicate greater species specificity (46, 47). *Leishmania* 40S ribosomal protein S12 was successfully assayed in VL blood (48).

India	LdBPK_262110	Protein kinase	Add phosphate groups to other proteins in cell signalling pathways.
Sudan	LdBPK_351070	Protein kinase	Add phosphate groups to other proteins in cell signalling pathways.

The great difference in western blot and SDS-PAGE gel profiles between the concentrated Sudanese VL urine (VLu) and urine antigens (uAg) purified from the same sample type was unexpected. However, the sensitivity of antibodies may detect proteins that are not readily visible on the gel. In addition, the process of stripping and re-probing the blot could have somewhat altered the composition of the blot, particularly if the stripping process was not complete and rabbit antibodies remained. However, the EHCu remained negative, indicating very little non-specific reaction. Heating the urine, which we did not do, could have led to more specific reactions (49), however, this favours carbohydrate antigens as it denatures proteins and we sought to retain as many of the conformational protein epitopes as possible by running non-reduced samples on gels.

Coiled protein regions are often made by repeats of a few of the same amino acids, as identified here in hypothetical protein LdBPK_191140, and this repetition can lead to high antigenicity, making them good targets for immunodiagnostics (39, 40, 50). The geographic overlap of VL with Chagas disease, caused by *T. cruzi*, particularly in Brazil, makes it imperative to identify a *Leishmania* antigen that will not cross-react with this other trypanosomatid. Although the proteins identified here had very high sequence identity with *L. donovani* or *L. infantum*, some had moderate and one had high homology with *T. cruzi* proteins. Therefore, a polyclonal antibody against a complete *L. donovani* protein could also react with *T. cruzi* proteins. For this reason, as done here, selecting shorter and more species-specific peptide sequences is preferable because these contain fewer epitopes and generate a more specific polyclonal antibody.

Based on the species specificity of the epitope regions, we suggest that combining several of these, either for raising a mixed polyclonal antiserum, or later combining individual antibodies in the assay, may improve specificity and sensitivity of the prospective assay. This was found by C. Abeijon and A. Campos-Neto (51) where the combined assay had improved performance over individual analyte assays.

Strengths

Our criteria for identifying parasite proteins in VL urine took into account the large number and amount of human protein in the sample (52). Initial steps to enrich for *Leishmania* antigens using specific antibodies assisted in achieving peptide identification by mass spectrometry. Further, by searching mass data simultaneously against both human and *L. donovani* protein databases, we were able to exclude mass spectra that had a better match to human peptide sequences, thus excluding potential false positive matches to *L. donovani*. The dual searching did not exclude possible matches to bacterial peptides, therefore we searched both the peptides and the proteins against the NCBI nr database to exclude this possibility.

Limitations

While we have presented *L. donovani* protein identities here, there is no certainty that these proteins exist as whole proteins in VL urine as we can only identify short peptides by mass spectrometry. However, the presence of the peptides suggests that at least fragments of the proteins do occur in urine, along with many host-derived proteins (52, 53). Progression from identifying VL urinary antigens to having a prototype RDT relies on follow-up to synthesise antigens and raise antibodies. An alternative 'shortcut' method was used by U. S. Markakpo et al. (54) to develop a urine antigen capture RDT for malaria by generating monoclonal antibodies directly to precipitated urine proteins from malaria patients and subsequently screening them for specificity and sensitivity.

Conclusions

We used various methods to capture *L. donovani* antigens from VL urine from India and Sudan, including panning of urine proteins with immobilised anti-*L. donovani* antibodies and visualisation of immunoreactive bands in VL urine on a western blot. All methods yielded *L. donovani* proteins by mass spectrometry of captured or immunoreactive material and 6 of 7 proteins were highly specific to *Leishmania*. In addition, epitope prediction revealed 22 species-specific B-cell epitopes that make ideal candidates for synthesis and to generate antiserum for antigen capture assay development.

Supplemental Material

These files are available in the Appendix.

S1 Table: Number of *L. donovani* peptides and proteins identified by mass spectrometry of antigens captured from Indian VL urine with anti-1S2D antibody.

S2 Table: Solo peptides of *Leishmania donovani* identified in urine of Indian and Sudanese patients with visceral leishmaniasis.

S3 Table: Number of epitope peptides of ≥8 amino acids in *Leishmania donovani* proteins identified from VL patients' urine.

Acknowledgements

The authors thank Svenja Hester and Monika Stegmann at the Advanced Proteomic Facility at Oxford University for conducting the mass spectrometry and advice on data analysis. We thank Marleen Boelaert of the Institute of Tropical Medicine, Antwerp, Belgium, for securing funding for the NIDIAG project, of which this work was part.

References

1. World Health Organization. 2012. Regional Strategic Framework for elimination of kala-azar from the South East Asia Region (2011–2015).
2. Reiter-Owona I, Rehkaemper-Schaefer C, Arriens S, Rosenstock P, Pfarr K, Hoerauf A. 2015. Specific K39 antibody response and its persistence after treatment in patients with imported leishmaniasis. Parasitol Res doi:10.1007/s00436-015-4801-8:1-9.
3. Gidwani K, Picado A, Ostyn B, Singh SP, Kumar R, Khanal B, Lejon V, Chappuis F, Boelaert M, Sundar S. 2011. Persistence of *Leishmania donovani* antibodies in past visceral leishmaniasis cases in India. Clin Vaccine Immunol 18:346-8.
4. De Almeida Silva L, Romero HD, Prata A, Costa RT, Nascimento E, Carvalho SF, Rodrigues V. 2006. Immunologic tests in patients after clinical cure of visceral leishmaniasis. Am J Trop Med Hyg 75:739-43.
5. Singh DP, Sundar S, Mohapatra TM. 2009. The rK39 strip test is non-predictor of clinical status for kala-azar. BMC Res Notes 2:187.
6. Zijlstra EE, Nur Y, Desjeux P, Khalil EA, El-Hassan AM, Groen J. 2001. Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. Trop Med Int Health 6:108-13.
7. Cota GF, de Sousa MR, Demarqui FN, Rabello A. 2012. The diagnostic accuracy of serologic and molecular methods for detecting visceral leishmaniasis in HIV infected patients: meta-analysis. PLoS Negl Trop Dis 6:e1665.
8. World Health Organization. 2012. Technical Report Series. 975. Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. World Health Organization, Organization WH, Geneva, Switzerland.

9. Verma S, Singh R, Sharma V, Bumb RA, Negi NS, Ramesh V, Salotra P. 2017. Development of a rapid loop-mediated isothermal amplification assay for diagnosis and assessment of cure of *Leishmania* infection. BMC Infect Dis 17:223.
10. Mukhtar M, Ali SS, Boshara SA, Albertini A, Monnerat S, Bessell P, Mori Y, Kubota Y, Ndung'u JM, Cruz I. 2018. Sensitive and less invasive confirmatory diagnosis of visceral leishmaniasis in Sudan using loop-mediated isothermal amplification (LAMP). PLoS Negl Trop Dis 12:e0006264.
11. Mondal D, Ghosh P, Khan MA, Hossain F, Bohlken-Fascher S, Matlashewski G, Kroeger A, Olliaro P, Abd El Wahed A. 2016. Mobile suitcase laboratory for rapid detection of *Leishmania donovani* using recombinase polymerase amplification assay. Parasit Vectors 9:281.
12. Attar ZJ, Chance ML, El-Safi S, Carney J, Azazy A, El-Hadi M, Dourado C, Hommel M. 2001. Latex agglutination test for the detection of urinary antigens in visceral leishmaniasis. Acta Trop 78:11-6.
13. Sarkari B, Chance M, Hommel M. 2002. Antigenuria in visceral leishmaniasis: detection and partial characterisation of a carbohydrate antigen. Acta Trop 82:339-48.
14. Sundar S, Agrawal S, Pai K, Chance M, Hommel M. 2005. Detection of leishmanial antigen in the urine of patients with visceral leishmaniasis by a latex agglutination test. Am J Trop Med Hyg 73:269-71.
15. Diro E, Techane Y, Tefera T, Assefa Y, Kebede T, Genetu A, Kebede Y, Tesfaye A, Ergicho B, Gebre-Yohannes A, Mengistu G, Engers H, Aseffa A, Desjeux P, Boelaert M, Hailu A. 2007. Field evaluation of FD-DAT, rK39 dipstick and KATEX (urine latex agglutination) for diagnosis of visceral leishmaniasis in northwest Ethiopia. Trans R Soc Trop Med Hyg 101:908-14.
16. Rijal S, Boelaert M, Regmi S, Karki BMS, Jacquet D, Singh R, Chance ML, Chappuis F, Hommel M, Desjeux P, Van der Stuyft P, Le Ray D, Koirala S. 2004. Evaluation of a urinary antigen-based latex agglutination test in the diagnosis of kala-azar in eastern Nepal. Trop Med Int Health 9:724-729.
17. Salam MA, Khan MG, Mondal D. 2011. Urine antigen detection by latex agglutination test for diagnosis and assessment of initial cure of visceral leishmaniasis. Trans R Soc Trop Med Hyg 105:269-72.
18. Riera C, Fisa R, Lopez P, Ribera E, Carrio J, Falco V, Molina I, Gallego M, Portus M. 2004. Evaluation of a latex agglutination test (KAtex) for detection of *Leishmania*

antigen in urine of patients with HIV-*Leishmania* coinfection: value in diagnosis and post-treatment follow-up. Eur J Clin Microbiol Infect Dis 23:899-904.

19. Vogt F, Mengesha B, Asmamaw H, Mekonnen T, Fikre H, Takele Y, Adem E, Mohammed R, Ritmeijer K, Adriaensen W, Melsew Y, van Griensven J, Diro E. 2018. Antigen detection in urine for noninvasive diagnosis and treatment monitoring of visceral leishmaniasis in human immunodeficiency virus coinfected patients: an exploratory analysis from Ethiopia. Am J Trop Med Hyg doi:10.4269/ajtmh.18-0042.

20. Sarkari B, Chance M, Hommel M. 2005. A capture ELISA for the diagnosis of visceral leishmaniasis using a monoclonal antibody against a leishmanial urinary antigen. Iran Biomed J 9:117-122.

21. Abeijon C, Kashino SS, Silva FO, Costa DL, Fujiwara RT, Costa CH, Campos-Neto A. 2012. Identification and diagnostic utility of *Leishmania infantum* proteins found in urine samples from patients with visceral leishmaniasis. Clin Vaccine Immunol 19:935-43.

22. Vallur AC, Tutterrow YL, Mohamath R, Pattabhi S, Hailu A, Abdoun AO, Ahmed AE, Mukhtar M, Salam MA, Almeida ML, Almeida RP, Mondal D, Albertini A, Ghalib H, Duthie MS, Reed SG. 2015. Development and comparative evaluation of two antigen detection tests for visceral leishmaniasis. BMC Infect Dis 15:384.

23. Bhattacharyya T, Ayandeh A, Falconar AK, Sundar S, El-Safi S, Gripenberg MA, Bowes DE, Thunissen C, Singh OP, Kumar R, Ahmed O, Eisa O, Saad A, Silva Pereira S, Boelaert M, Mertens P, Miles MA. 2014. IgG1 as a potential biomarker of post-chemotherapeutic relapse in visceral leishmaniasis, and adaptation to a rapid diagnostic test. PLoS Negl Trop Dis 8:e3273.

24. Schägger H. 2006. Tricine-SDS-PAGE. Nat Protoc 1:16-22.

25. The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158-D169.

26. Jespersen MC, Peters B, Nielsen M, Marcatili P. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 45:W24-W29.

27. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785-6.

28. Gupta R, Brunak S. 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput:310-22.

29. Boratyn GM, Schaffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. Biol Direct 7:12.

30. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY, Mitchell AL. 2017. InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 45:D190-d199.

31. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44:D457-62.

32. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L, Xenarios I. 2013. New and continuing developments at PROSITE. Nucleic Acids Res 41:D344-7.

33. Corstjens PL, De Dood CJ, Kornelis D, Fat EM, Wilson RA, Kariuki TM, Nyakundi RK, Loverde PT, Abrams WR, Tanke HJ, Van Lieshout L, Deelder AM, Van Dam GJ. 2014. Tools for diagnosis, monitoring and screening of *Schistosoma* infections utilizing lateral-flow based assays and upconverting phosphor labels. Parasitology 141:1841-55.

34. Suwanpimolkul G, Kawkitinarong K, Manosuthi W, Sophonphan J, Gatechompol S, Ohata PJ, Ubolyam S, Iampornsin T, Katerattanakul P, Avihingsanon A, Ruxrungtham K. 2017. Utility of urine lipoarabinomannan (LAM) in diagnosing tuberculosis and predicting mortality with and without HIV: prospective TB cohort from the Thailand Big City TB Research Network. Int J Infect Dis 59:96-102.

35. Athlin S, Iversen A, Ozenci V. 2017. Comparison of the ImmuView and the BinaxNOW antigen tests in detection of *Streptococcus pneumoniae* and *Legionella pneumophila* in urine. Eur J Clin Microbiol Infect Dis 36:1933-1938.

36. Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ, Reiner NE. 2008. Proteomic analysis of the secretome of Leishmania donovani. Genome Biol 9:R35.

37. Anish C, Martin CE, Wahlbrink A, Bogdan C, Ntais P, Antoniou M, Seeberger PH. 2013. Immunogenicity and diagnostic potential of synthetic antigenic cell surface glycans of *Leishmania*. ACS Chem Biol 8:2412-22.

38. Rodrigues JA, Acosta-Serrano A, Aebi M, Ferguson MA, Routier FH, Schiller I, Soares S, Spencer D, Titz A, Wilson IB, Izquierdo L. 2015. Parasite glycobiology: a bittersweet symphony. PLoS Pathog 11:e1005169.

39. Goto Y, Coler RN, Reed SG. 2007. Bioinformatic identification of tandem repeat antigens of the *Leishmania donovani* complex. Infect Immun 75:846-51.
40. Burns JM, Shreffler WG, Benson DR, Ghalib HW, Badaro R, Reed SG. 1993. Molecular characterization of a kinesin-related antigen of *Leishmania chagasi* that detects specific antibody in African and American visceral leishmaniasis. Proc Natl Acad Sci U S A 90:775-779.
41. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM, Nilsson P, Oksvold P, Rockberg J, Schutten R, Schwenk JM, Sivertsson Å, Sjöstedt E, Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M, Mardinoglu A, Pontén F, von Feilitzen K, Lilley KS, Uhlén M, Lundberg E. 2017. A subcellular map of the human proteome. Science 356.
42. The Human Protein Atlas. Available from www.proteinatlas.org. Accessed
43. Lee WR, Na H, Lee SW, Lim WJ, Kim N, Lee JE, Kang C. 2017. Transcriptomic analysis of mitochondrial TFAM depletion changing cell morphology and proliferation. Sci Rep 7:17841.
44. Marat AL, Dokainish H, McPherson PS. 2011. DENN domain proteins: regulators of Rab GTPases. J Biol Chem 286:13791-800.
45. Morriswood B, Schmidt K. 2015. A MORN repeat protein facilitates protein entry into the flagellar pocket of *Trypanosoma brucei*. Eukaryot Cell 14:1081-93.
46. Maslov DA, Spremulli LL, Sharma MR, Bhargava K, Grasso D, Falick AM, Agrawal RK, Parker CE, Simpson L. 2007. Proteomics and electron microscopic characterization of the unusual mitochondrial ribosome-related 45S complex in *Leishmania tarentolae*. Mol Biochem Parasitol 152:203-212.
47. Shalev-Benami M, Zhang Y, Matzov D, Halfon Y, Zackay A, Rozenberg H, Zimmerman E, Bashan A, Jaffe CL, Yonath A, Skiniotis G. 2016. 2.8-Å cryo-EM structure of the large ribosomal subunit from the eukaryotic parasite *Leishmania*. Cell Reports 16:288-294.
48. Zhang W-W, Ghosh AK, Mohamath R, Whittle J, Picone A, Lypaczewski P, Ndao M, Howard RF, Das P, Reed SG, Matlashewski G. 2018. Development of a sandwich ELISA to detect *Leishmania* 40S ribosomal protein S12 antigen from blood samples of visceral leishmaniasis patients. BMC Infect Dis 18:500.
49. Brito-Santos F, Ferreira MdF, Trilles L, Muniz MdM, Veloso dos Santos VG, Carvalho-Costa FA, Meyer W, Wanke B, Lazéra MdS. 2017. Preheating of urine improves the

specificity of urinary cryptococcal antigen testing using the lateral flow assay. PLoS Negl Trop Dis 11:e0005304.

50. Goto Y, Carter D, Reed SG. 2008. Immunological dominance of *Trypanosoma cruzi* tandem repeat proteins. Infect Immun 76:3967-74.

51. Abeijon C, Campos-Neto A. 2013. Potential non-invasive urine-based antigen (protein) detection assay to diagnose active visceral leishmaniasis. PLoS Negl Trop Dis 7:e2161.

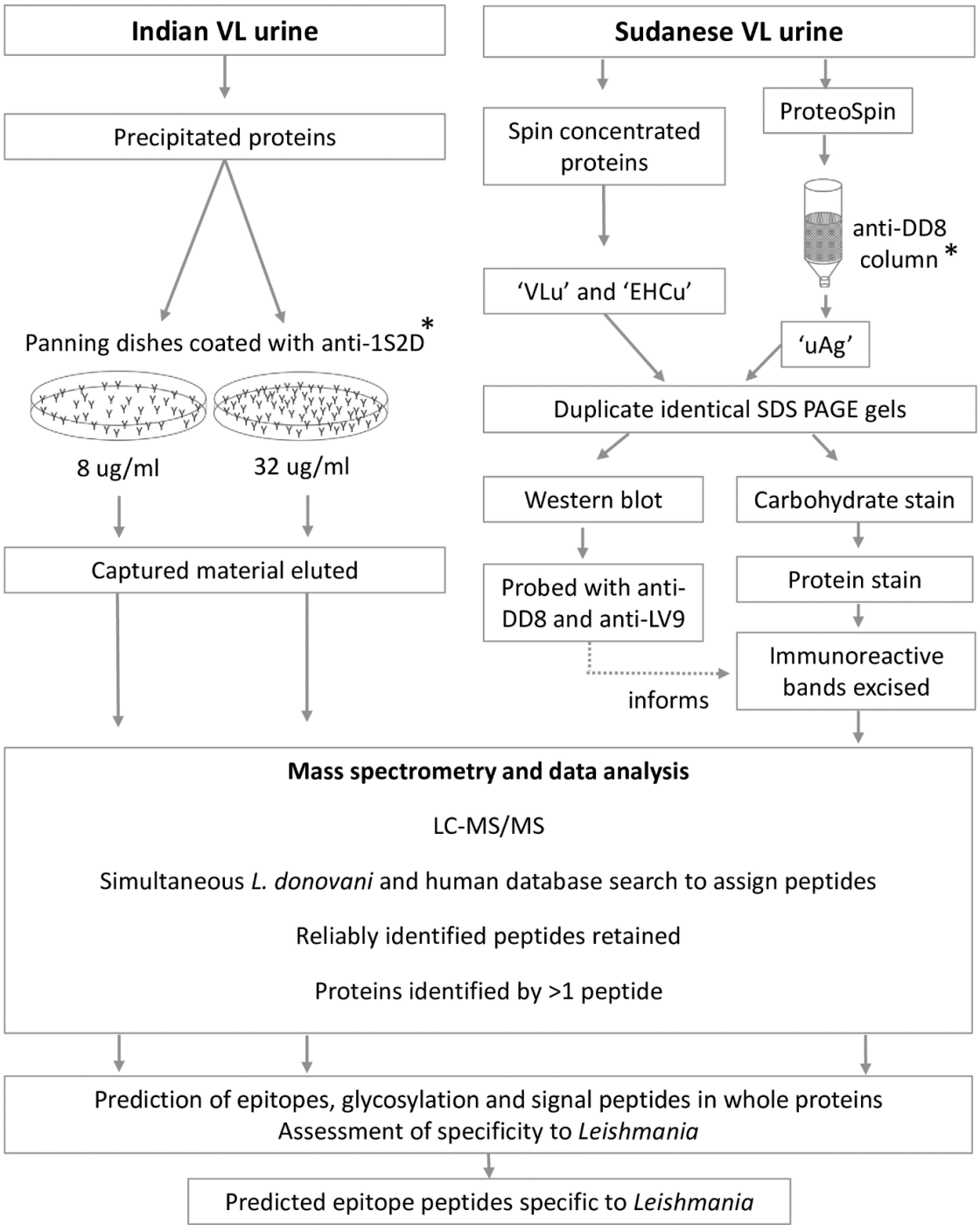
52. Santucci L, Candiano G, Petretto A, Bruschi M, Lavarello C, Inglese E, Righetti PG, Ghiggeri GM. 2014. From hundreds to thousands: widening the normal human urinome. J Proteom doi:http://dx.doi.org/10.1016/j.jprot.2014.07.021.

53. Adachi J, Kumar C, Zhang Y, Olsen J, Mann M. 2006. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol 7:R80.

54. Markakpo US, Bosompem KM, Dzodzomenyo M, Danso-Appiah A, Essuman EE, Anyan WK, Suzuki M, Stephens JK, Anim-Baidoo I, Asmah RH, Ofori MF, Madjitey P, Danquah JB, Frempong NA, Kwofie KD, Amoa-Bosompem M, Sullivan D, Jr., Fobil JN, Quakyi IA. 2016. Minimising invasiveness in diagnostics: developing a rapid urine-based monoclonal antibody dipstick test for malaria. Trop Med Int Health 21:1263-1271.

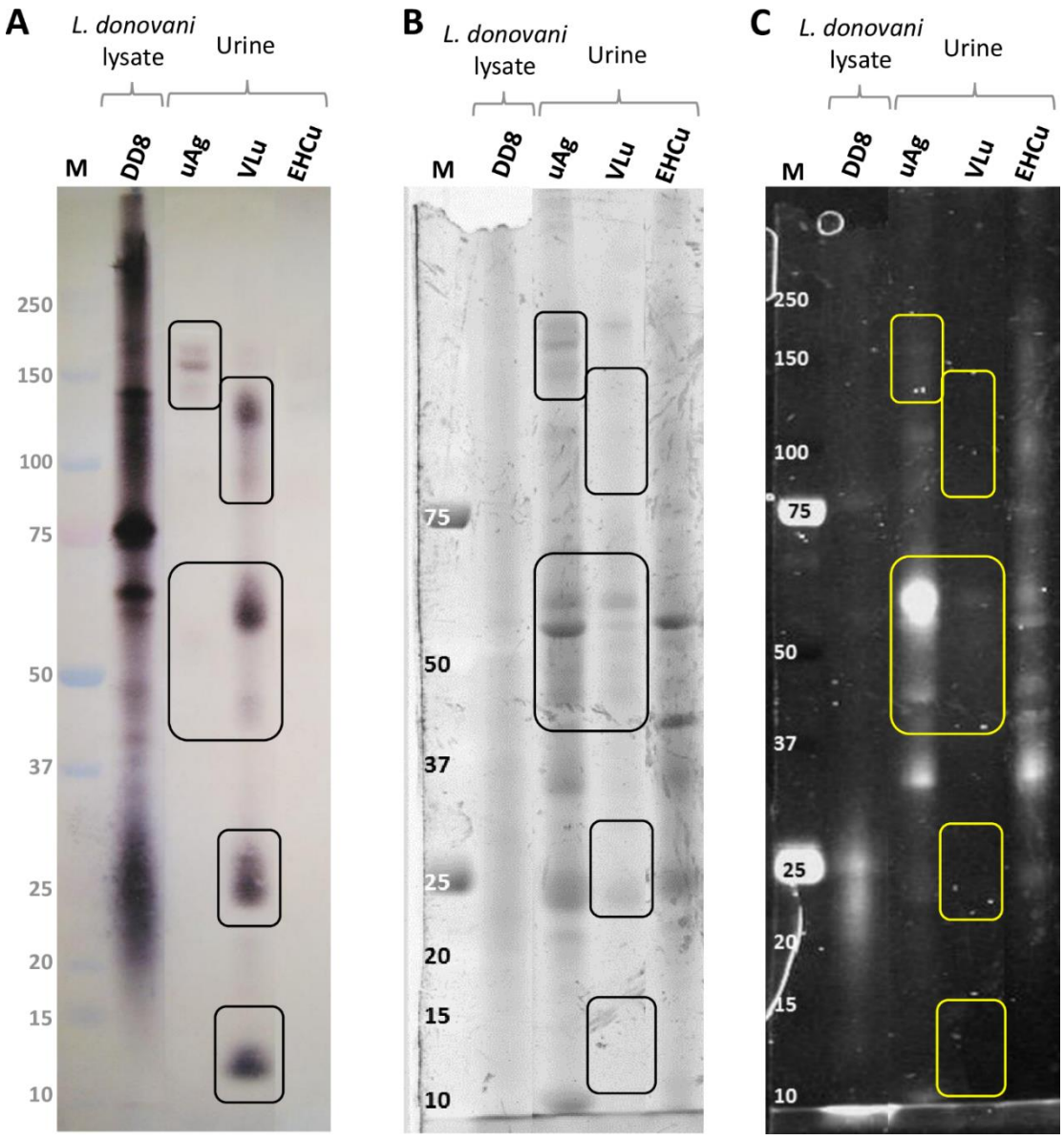
Figures

Figure 1.



662 Figure 2.

663



664

665

666

CHAPTER 5: Refining wet lab experiments with *in silico* searches: a rational quest for biomarkers of post-chemotherapeutic relapse in visceral leishmaniasis.

Hinckel BC, **Marlais T**, Airs S, Bhattacharyya T, Imamura H, Dujardin JC, El-Safi S, Singh OP, Sundar S, Falconar AK, Andersson B, Litvinov S, Miles MA, Mertens P. Refining wet lab experiments with *in silico* searches: a rational quest for biomarkers of post-chemotherapeutic relapse in visceral leishmaniasis. (2019). PLoS NTD, 13(5). doi: 10.1371/journal.pntd.0007353

Key points, novel results and implications

- IgG1 serology has potential to distinguish cure from relapse after treatment for the potentially fatal disease visceral leishmaniasis (VL) but the specific antigens involved are not known.
- This work used western blotting and mass spectrometry to identify individual proteins recognised by IgG1 in active and relapsed VL, and applied computational filters to the protein data in order to select antigenic peptides.
- Synthetic peptides were produced based on *in silico* filtering and analysis, and tested for reactivity with serum from Sudanese VL cases and healthy controls.
- A single IgG1-reactive peptide from a *Leishmania* protein was identified which reacted with IgG1 from VL patients in a rapid diagnostic test.
- This work has implications for using this *in silico* method and high-throughput peptide screening to define seroantigens for VL and other diseases in need of improved serology.

Candidate's contribution

The candidate performed SDS PAGE gels, prepared western blot strips and optimised the immunoassay of the strips. The candidate planned and supervised the assaying of Indian sera with the western blot strips and analysed the blot results. The candidate jointly carried out repeat SDS PAGE gels and prepared samples from these for mass spectrometry. The candidate jointly carried out initial data quality assessment of the mass spectrometry data, all of which form the starting point for this manuscript. The candidate also provided feedback on the draft manuscript.

Co-author's contributions

Bruno Hinckel conducted the *in-silico* analysis pipeline to analyse mass spec data, performed lab work with the synthetic peptides, analysed the data from this and wrote the manuscript.

Other authors provided materials and/or advice and supervised the study.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Tegwen Marlais
Principal Supervisor	Prof. Michael A. Miles
Thesis Title	Diagnostics development for the neglected parasitic diseases strongyloidiasis and visceral leishmaniasis

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	PLoS Neglected Tropical Diseases		
When was the work published?	May 2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	n/a		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I performed SDS PAGE gels, prepared western blot strips and optimised the immunoassay of the strips. I planned and supervised the assaying of Indian sera with the western blot strips and analysed the blot results. I jointly carried out repeat SDS PAGE gels and prepared samples from these for mass spectrometry. I jointly carried out

	initial data quality assessment of the mass spectrometry data, all of which form the starting point for this manuscript. I also provided feedback on the draft manuscript.

Student Signature: [redacted] **Date:** 22/7/19

Supervisor Signature: [redacted] **Date:** 22/7/19

Refining wet lab experiments with *in silico* searches: a rational quest for biomarkers of post-chemotherapeutic relapse in visceral leishmaniasis

Bruno Bremer Hinckel[1,2*], Tegwen Marlais[3], Stephanie Airs[3], Tapan Bhattacharyya[3], Hideo Imamura[2], Jean-Claude Dujardin[2], Sayda El-Safi[4], Om Prakash Singh[5], Shyam Sundar[5], Andrew Keith Falconar[6], Bjorn Andersson[7], Sergey Litvinov[8], Michael A. Miles[3] and Pascal Mertens[1]

1 Coris BioConcept, Gembloux, Belgium

2 Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

3 Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

4 Faculty of Medicine, University of Khartoum, Khartoum, Sudan

5 Department of Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India

6 Departamento de Medicina, Universidad del Norte, Barranquilla, Colombia

7 Department of Cell- and Molecular Biology, Karolinska Institutet, Stockholm, Sweden

8 Aptum Biologics Ltd, Southampton, United Kingdom

* bruno.hinckel@corisbio.com

Abstract

Background

The search for biomarkers in diagnostic research has been profiting from a growing number of high quality fully sequenced genomes and freely available bioinformatic tools. These can be combined with wet-lab experiments towards a more rational search. Improved, point-of-care diagnostic tests for visceral leishmaniasis (VL) early case detection and surveillance are required. Previous investigations demonstrated the potential of IgG1 as a biomarker for monitoring clinical status.

Methodology

Immunoblots revealed *L. donovani* protein bands reacting with human IgG1 from VL patients. Upon identification of these antigens found with confidence by mass spectrometry (MS), we searched for evidence of constitutive protein expression in both parasite life stages and the presence of antigenic domains or high accessibility to B-cells. Selected candidates had their linear epitopes mapped with *in silico* algorithms. Multiple high-scoring predicted epitopes from the shortlisted proteins were screened with peptide arrays. The most promising candidate was tested on a lateral flow format, using sera positive for VL. The peptide's specificity was assessed with sera from European healthy subjects.

Results

Over 90% of the proteins identified from the immunoblots did not satisfy the searched criteria and were excluded from the downstream epitope mapping. The pilot screening with predicted epitopes from selected proteins identified the most reactive peptide, whose specificity was highly concentration dependent. This peptide showed a maximum sensitivity of 84% (95% CI 60 - 97%) with Sudanese serum samples positive for VL on a lateral flow test. None of the sera from healthy subjects (0/10) showed a positive result.

Conclusion

We employed comparative genomics and *in silico* searches to drastically reduce the extent of wet-lab experiments, focusing on promising candidates satisfying selected protein features. By profiling epitopes with *in silico* prediction algorithms we screened a large number of peptides, identifying the most promising one, which showed good results of sensitivity and specificity in a prototype rapid diagnostic test (RDT).

Author summary

Visceral leishmaniasis (VL) is a neglected tropical disease caused by protozoan parasites of the *Leishmania donovani* complex. Without appropriate treatment, VL is usually fatal. Although non-invasive diagnostic techniques, mainly based on the detection of anti-*Leishmania* antibodies are available, invasive procedures such as microscopy from spleen or bone marrow aspirates are still required to confirm cure after successful treatment. Recently, a world health organization (WHO) report included the development of new diagnostics to confirm cure as a research priority for VL. Previous investigations have outlined the potential of IgG1 as a biomarker of post-chemotherapeutic relapse for VL. Here we employed comparative genomics and bioinformatic tools to search for desired protein features in a large number of *L. donovani* antigens reacting with human IgG1 in western blots. We then used *in silico* tools to profile epitopes from the shortlisted proteins. By doing so, we screened

a large number of peptides using arrays with low reagent consumption. The most reactive peptide was adapted to a lateral flow RDT and showed promising results of both sensitivity and specificity. Thus we believe that *in silico* tools can be used to refine the extent output of wet-lab experiments towards a more rational search for diagnostic biomarkers.

Introduction

The leishmaniases comprises a group of vector-borne diseases caused by parasites of the genus *Leishmania*. The visceral form (VL, also known as kala azar - Hindi for 'black fever') affects internal organs such as liver, spleen and bone marrow, and usually leads to death if left untreated. VL is caused by parasites of the *Leishmania donovani* complex and is responsible for about 400,000 cases per year, of which between 20,000 and 40,000 are fatal [1]. Human VL is often diagnosed by a combination of clinical symptoms (prolonged fever, weight loss, hepatosplenomegaly, hypergammaglobulinemia and pancytopenia) and the detection of parasite specific immunoglobulins (IGs). The recombinant protein rK39, a fragment of the *Leishmania* kinesin-like gene, was described in 1993 [2] and remains the most widely used antigen for VL serodiagnosis. Nonetheless, novel antigens are still required to complement the use of the rK39, to improve sensitivity in eastern Africa [3, 4] and to determine cure after successful chemotherapy (versus relapse) [5-9]. The recombinant protein rK28, which is a derivative of the rK39 incorporating the first two repeats of a Sudanese kinesin flanked by HASPB1 and HASPB2, has shown improved serological sensitivity in East African VL patients [10].

There has been a dramatic reduction in genome sequencing costs, accompanied by an exponential increase in the number of available sequences in public repositories [11, 12]. The first *Leishmania* spp. genome sequencing was completed in 2005 [13]. More recently, the advent of high-throughput technologies has enabled the completion of the whole-genome sequencing of *L. donovani*, the causative agent of VL in the Indian subcontinent [14]. Moreover, the availability of next-generation sequencing made it possible to perform whole transcriptome sequencing. RNA sequencing generates data on the transcriptome at a specific stage of a pathogen life cycle or in a specific culture condition of an organism. This growing number of available high quality full genome sequences has become a central element in the area of comparative genomics, which has contributed greatly to a better understanding of multiple aspects of the leishmaniases, including determinants of disease phenotype [15], mode of action of [16] and parasite biology [17].

antibodies (ABs) are a major component of the immune system. They fight foreign pathogens by binding to their specific regions (epitopes), tagging them for clearance by the immune

system. Epitopes can be divided into linear (a contiguous stretch of amino acids (AAs)) and discontinuous (where the atoms are brought together by protein folding) and can be identified by functional and structural studies (e.g. X-ray crystallography), while *in silico* prediction algorithms are gaining popularity. The early prediction methods of linear B-cell epitopes were mainly based on propensity scales [18]. More recently, machine learning methods have been employed to improve prediction performance [19-21]. The prediction of discontinuous epitopes still depends on the availability of 3D structures. *In silico* tools can also be used to locate antigenic domains from DNA sequences or peptide sequences influencing protein localisation.

Diagnostic research can incorporate genomics, transcriptomics, proteomics as well as bioinformatic prediction algorithms for the discovery of new biomarkers. Such a systematic 'omics' approach has been applied alone for the discovery of vaccine candidates [22-24] as well as for diagnostic biomarkers [25]. These *in silico* searches can also be applied downstream of wet-lab experiments, refining their output towards a more rational search.

Mass spectrometry (MS) can be employed to reveal the identity of proteins identified from wet-lab experiments (e.g. immunoblots). Comparative genomics enables identification of species-specific genes while the expression levels of life cycle-specific proteins can be estimated using publicly available RNA-seq data. *In silico* prediction algorithms can be employed to infer protein localisation, search for antigenic domains as well as to predict linear B-cell epitopes. Peptide antigens can then be synthesised and incorporated into peptide arrays, enabling the screening of a large number of candidates in pilot serological assays. Promising candidates can then be adapted and tested as rapid diagnostic test (RDT), a format suitable for field use.

Here we employed comparative genomics as well as *in silico* algorithms to the excessive number of *L. donovani* protein candidates reacting with human IgG1 in immunoblots. By screening a large number of predicted epitopes from selected candidate antigens satisfying desired protein features, we have identified one peptide specifically recognised by human IgG1, a potential biomarker of post-chemotherapeutic relapse in VL [41]. This peptide was also tested in a prototype lateral flow RDT with serum positive for VL and from healthy subjects (non endemic healthy controls (NEHCs)), showing promising results of both sensitivity and specificity. Thus we believe the employed approach is a valid proof of concept for the discovery of diagnostic biomarkers.

Methods

Ethics statement

In India, the collection of samples was approved by the Ethics Committee of Banaras Hindu University, Varanasi while in Sudan the approval for collection and research was granted by the Ethical Research Committee, Faculty of Medicine, University of Khartoum and the National Health Research Ethics Committee, Federal Ministry of Health, Sudan. Written informed consent was obtained from all adult subjects included in the study or from the parents or guardians of individuals less than 18 years of age. This research was also approved, as part of the EC NIDIAG project, by the London School of Hygiene and Tropical Medicine Ethics Committee as well as by the Ethics Committee of the Antwerp University Hospital.

Sources of sera

India

Indian sera were selected from archived samples, collected after 2007 from active VL, relapsed and endemic healthy controls (EHCs), all from the endemic region of Muzaffarpur, Bihar state. Positive cases for VL had been diagnosed by positive rK39 serology and parasitologically by microscopy of splenic aspirates. Detailed information about the Indian samples used in this study are described in Table 1.

Sudan

Sudanese serum samples were collected in 2011 and 2012 from the endemic region of Gedaref in eastern Sudan. Active VL cases had been diagnosed by microscopy of bone marrow or lymph node aspirates in conjunction with serological assays and were all HIV negative. These diagnoses were carried out according to their respective national procedures, prior to the present study.

NEHC

Sera was obtained from whole blood collected from *Etablissement Français du Sang*. All donors are informed of the use of the blood for research purpose and that he/she gives their informed consent for the purpose of scientific research use and that all national laws and ethical principles are fulfilled. The samples are fully anonymised. All samples are certified to be tested for detection of transmissible diseases as follows: detection of antibodies for HIV-1, HIV-2, HCV, HTLV I, HTLV II, syphillis, HBc, antigen from HBs, viral genome from HIV1, HCV and HBV.

Wet-lab identification of *L. donovani* antigens reacting with IgG1

The identification of *L. donovani* promastigote antigens recognised by human IgG1 using western blots was carried out as described elsewhere (Marlais et al., submitted). Briefly, the cytosolic proteins from a whole *L. donovani* lysate (strain MHOM/IN/80/DD8) were separated by SDS-PAGE, blotted onto a nitrocellulose (NC) membrane and sliced into individual strips, to be immunoassayed with sera from individual patients from India. Upon visual identification of bands of interest on the individual strips using HRP-labelled anti-human IgG1 as secondary antibody, new gels were run and corresponding immunogenic bands were excised and analysed by liquid chromatography tandem mass spectrometry (LC-MS/MS). Protein hits were identified by matching peptide fragments against the reference *L. donovani* database (MHOM/NP/2003/BPK282).

In silico refinement of wet-lab output

Desired protein features were searched *in silico* in order to decrease the number of candidates from the MS output while shortlisting the proteins more likely to be antigenic, to have their epitopes mapped (section Epitope mapping from shortlisted proteins). Some protein features were searched in series while others in parallel, meaning that all shortlisted proteins satisfy the criteria described in either branches as shown in Fig 1. Additional information about each step of this '*in silico* filter' are detailed in the following subsections.

Figure 1. *In silico* filter applied to select desired protein features. Selected protein properties were searched *in silico* on all protein hits identified by MS from selected IgG1 sero-reactive western blot bands.

Gene expression analysis - RNA-seq

Transcriptomic data on promastigote and amastigote parasite life stages was obtained from an improved BPK282 reference sequencing as described in [27]. In this study, new gene products with their corresponding IDs - 'new IDs' - and RNAs were annotated based on *Leishmania major* annotation [13] using Companion [28] while orthologs to the 'old gene products' with their corresponding 'old IDs' [14] were identified with BLAST. The expression levels were quantified as haploid depth (HAP), as cultivated promastigotes are highly aneuploid, showing interstrain diversity and intrastrain mosaicism in ploidy number [27]. The HAP from all genes from the improved BPK282 reference sequencing were ranked for promastigotes and amastigotes separately and only those with a value greater than the first

quartile HAP>Q1) for promastigotes and amastigotes were considered as constitutively expressed and kept for downstream analysis.

Confident MS identification

Each protein hit identified by MS with the software MASCOT [29] has its own score defined by:

$$MASCOTscore = -10 \times \log_{10}(P)$$

The score translates into the probability *P* that the observed match is not a random event [30]. A score of 100 was chosen as the cut-off value to minimise the chances that any match could be found by chance.

Protein localisation

Exported *L. donovani* proteins, identified as part of exosomes by experimental approaches [31] and/ or by the presence of signal peptide were considered highly accessible to B-cells and thus selected for further analysis. The presence of a signal peptide sequence was searched for with the standalone version of the software signalP v4.0 [32], running with the option 'noTM'.

Antigenic domain selection

L. donovani proteins harbouring tandem repeats (TRs) are often B-cell antigens [33] (e.g. rK39, rK28) therefore the presence of this domain was included as a criterion correlating with antigenicity. DNA sequences were scanned for TRs with the standalone version of the program Tandem Repeats Finder [34] and searches were performed on default values.

Epitope mapping from shortlisted proteins

The B-cell epitopes of all shortlisted proteins were profiled with four algorithms separately: ABCPred [20], Bepipred1.0 [19], EpiQuest-B [35] and Ibtope [36]. For all epitopes predicted from each algorithm individually (generally a 15 or 16mer), a 'core sequence', stretching from the fifth to the eleventh AA residue was defined and scanned against the rests of the predicted antigenic peptides. For peptides sharing core sequences with other predicted peptides at any position in the AA sequence (partially overlapping peptides), only that of highest score was kept. The top 20 scoring peptides with no shared core sequences within a given algorithm were synthesised, making up a total of 80 high scoring peptides from four different prediction algorithms.

While ABCPred, lbtope and EpiQuest-B were run on default values, predicting antigenic sequences of 16, 15 and 16 AA residues, respectively, shorter sequences with high antigenic scores were also considered for EpiQuest-B. The highest scoring peptides predicted from Bepipred were obtained by calculating the area under the curve (AUC) from plots of AA residue position vs. individual residue score for all possible combinations of 16mers with one AA offset, considering all contiguous predicted antigenic sequences from all submitted proteins. Additional details about the prediction algorithms employed can be found in Table 1.

Peptide synthesis

Desalted peptides were synthesised with an N-terminal biotin molecule linked via AAG spacer so that they could bind to NC membranes (1620215, Bio-Rad, USA), which had been previously soaked with neutralite avidin (NLA) (NLA30, e-proteins, Belgium) and dried at 50°C. Lyophilised peptides were dissolved according to standard protocols [37]. Where the solvent mixture in which the peptides were dissolved exhibited stable background values, the peptide concentration was calculated based on their molar absorbances, measured either at 280 nm, for tyrosine (Y) and tryptophan (W) containing peptides [38], or at 205 nm [39], for those without W or Y. For the remaining peptides, the concentration was calculated by dividing the reported amount by the volume of the solvent mixture in which they were dissolved.

Any promising candidate identified from the pilot peptide screening was synthesised with higher purity (90%) and with a polyethylene glycol-glycine spacer in order to improve water solubility as well as to increase rotation and ensure that the full amino acid sequence could freely interact with ABs, as opposed to being adsorbed onto the solid support and therefore unavailable for AB recognition [40].

Pilot screening of desalted peptides using arrays

Array production

sciFLEXARRAYER (Scienion AG, Germany) with a PDC 70 piezo dispense capillary (type 3 coating, P-2030/ S-6051, Scienion AG, Germany) was used to spot 7 nl of the selected peptides at the required concentrations in multiple replicates onto NC membranes (1620215, Bio-Rad, USA), previously soaked with NLA and dried at 50°C. rK39 (RAG0061: Rekom Biotech, Spain) and a whole *L. donovani* lysate, obtained as described in [41], were spotted both at 0.1

mg/ml as positive controls and for orientation purposes. Arrays were incubated overnight (ON) at 50°C upon completion of the spotting, to be hybridised with serum the next day.

Hybridisation with serum/ image acquisition

Peptide arrays were blocked with phosphate-buffered saline (PBS) + 3% bovine serum albumin (BSA) (PBSB) ON at 4°C or for 2h at room temperature (RT), followed by three 5 minute washes of PBS + 0.05% Tween 20 (PBST). In order to assess the peptide recognition by sera from VL patients and NEHCs, separate arrays were hybridised with pooled Sudanese serum samples positive for VL and with commercial pooled NEHC sera (S1-100ML, EMD Millipore Corporation, USA), respectively, diluted in PBST + 3% BSA (PBSTB), for 1h. After five 5 minute washes with PBST, they were incubated with fluorescent mouse anti-Human IgG1 Fc - Alexa Fluor 488 AB (A-10631, ThermoFisher) diluted 1:1000 in PBSTB, for 1h at RT. Followed by five 5 minute washes with PBST, the arrays were incubated at 50°C until completely dry.

Images were acquired at 500ms and 20 dB gain with a digital CCD camera (ORCA-R², Hamamatsu, Japan) coupled to a fluorescence microscope (model BX53, Olympus, Germany) equipped with filter cube U-FGFP (N271350, Olympus, Germany). Spot fluorescence was quantified using the software cellSens Dimensions v.1.7 (Olympus GmbH, Germany)

Adaptation to an RDT

RDTs were composed of a NC strip sensitised with reagent and a conjugate pad, impregnated with anti-human IgG1-specific antibody conjugated to colloidal gold. The strip was either housed within a plastic cassette, with a buffer application well and a test/reading window (lateral-flow), or not (dipstick). The reagent used for NC sensitization was NLA at 3.5 mg/ml. Prior to running the RDT, equal volumes of serum and the biotinylated peptide at stock concentration were mixed and incubated at 37°C for 15 minutes. 3.5 µl of the mix was then pipetted onto the sample application zone, just above the top of the sample pad. 150 µl of buffer solution was dispensed into the buffer application well (cassette) or dipsticks were dipped into recipient vessel with the appropriate volume of running buffer. After 15 minute incubation, a test was deemed valid if a clear red control band was present in line with the ‘C’ on the cassette, and deemed positive if a second band was present in line with the ‘T’. If no band was visible at the ‘T’, then the test was deemed negative.

Statistical analysis

All statistical analyses were performed using the computing environment R [42] and relevant packages [43, 44]. The final fluorescence of each spot on the peptide arrays (spot

fluorescence) was expressed in arbitrary units and calculated by subtracting the mean background value of a given acquired image from the individual fluorescence values. Peptide specificity was expressed as 95% Fiellers confidence intervals (CIs) of the ratio between the mean spot fluorescence from the arrays hybridised with VL serum and that of arrays hybridised with NEHC serum (ratio VL/NEHC).

RDT results were compared with defined clinical status to establish sensitivity and 95% CIs with Clopper-Pearson exact method.

Results

MS analysis of selected antigenic bands identified over 1300 hits

The development of the Western blot strips immunoassayed with individual Indian serum samples positive for VL (Fig 2) and relapsed patients revealed the protein bands specifically detected by human IgG1. The strongest and most prevalent bands across active VL and relapsed patients were excised from corresponding new gels and had their composite proteins revealed by MS.

Figure 2. Western blot strips immunoassayed with sera from Indian VL patients. The arrows indicate the three antigenic bands excised from corresponding acrylamide gels and analysed by MS. Western blots strips incubated with sera from relapsed patients showed a similar band pattern while those incubated with EHCs did not show any band pattern (data not shown).

1357 putative *L. donovani* protein hits from the three selected antigenic bands, with some hits identified in multiple bands were identified. All the candidate hits were submitted to the *in silico* filter (Fig 1) to reduce the number of promising candidates while restricting to those possessing protein features positively correlating with antigenicity.

The *in silico* filter excluded out over 90% of the unique identified hits

The initial 1357 old IDs from the three selected antigenic bands (Fig 2) contained 678 unique old IDs, annotated or not. Out of these, 538 were considered to be constitutively expressed, of which 209 had a MASCOT score ≥ 100. 66 of these candidates were found to be secreted, 60 as part of exosomes and 6 via the presence of signal peptide. These 66 unique old IDs were matched against new IDs, making 81 unique new IDs. After excluding the duplicated old IDs corresponding to unique new IDs of same gene product and any unannotated gene

product from the improved sequencing, a list of 62 extracellular and constitutively expressed proteins, identified with confidence by MS was obtained.

3 candidates harbouring TR domains were found to be constitutively expressed and were identified with confidence by MS. Although none of them was found to be extracellular, they were kept for downstream analysis as such proteins are often B-cell antigens [33]. In addition, one out of the three TR proteins identified remained unannotated in the improved BPK282 reference sequencing. Results are schematically shown on Fig 3 and the detailed list with all the 65 shortlisted proteins can be found in Table in S1 Table.

Figure 3. Selected proteins for *in silico* epitope mapping. Sixty five proteins ('FINAL 65 NEW IDs') satisfied the criteria shown in Fig 1 and were shortlisted to have their epitopes mapped with *in silico* algorithms. In **bold** the features searched/added after each step of the filter.

Epitopes selected for synthesis

The complete list of the 80 shortlisted peptides from the four prediction algorithms can be found in the Table S2 Table. Purity grades of desalted peptides ranged from 30 to 82%. Although the shortlisted peptides did not share core sequences intralgorithm, two peptides (EpQ_04_MAYV and LbT_07_CELG), predicted from different algorithms shared the core sequence 'CELGPNQ'.

Six out of the 80 peptides could not be resuspended, even after initial solubilisation in DMSO. Due to presence of organic solvent, they could not be re-lyophilised and were therefore not included in the pilot peptide screening.

Pilot peptide screening

Initial hybridisation revealed the most reactive peptide

A first batch of peptide arrays was produced in order to obtain insights about the peptide reactivity with human IgG1. The visual inspection of the imaged arrays hybridised with both pooled serum groups (Sudanese positive for VL and NEHC) diluted 1:100 revealed that a high number of peptides spotted at their stock concentrations reacted with both sera, indicating poor specificity (data not shown).

The hybridisation of the same batch of arrays with pooled serum diluted 1:200 showed that a reduced number of peptides reacted with both serum groups, revealing the most reactive

peptides, as shown in Fig 4. Details about the most reactive peptide found are shown in Table 1.

Figure 4. Hybridised peptide arrays. Peptide arrays hybridised with pooled Sudanese serum positive for VL. The most reactive peptide (spotted in duplicate, indicated by the yellow arrows) was revealed by diluting the sera 1:200. rK39. Red and white arrows indicate rK39 and a whole *L. donovani* lysate, respectively.

Specificity of peptide EpQ11 is highly concentration dependent

The peptide EpQ11 was spotted at multiple concentrations in eight replicates each onto a NC membrane soaked with NLA at 1 mg/ml, dried at 50°C for 3h. The hybridisation of two separate arrays with pooled serum samples diluted 1:200 showed that the discrimination between the VL and the NEHC pooled sera was highly concentration dependent, with the highest discrimination displayed when the peptide was spotted at 1.15 mg/ml, as shown in Fig 5.

Figure 5. Peptide specificity for VL IgG1 is highly concentration dependent. Peptide EpQ11 showed the highest discrimination between VL and NEHC IgG1 when spotted at 1.15 mg/ml.

Peptide EpQ11 binds specifically to IgG1 from active VL patients on a lateral flow format

RDTs in two formats, cassettes and dipsticks, were tested with individual VL Sudanese serum samples positive for VL as well as NEHC from Europe, in order to confirm the specificity of the EpQ11 peptide for IgG1 from VL patients. None of the NEHC sera tested (n=10) developed a test line upon test completion. Sensitivity values varied between 79% (54 - 94%) and 84% (60 - 97%) when running serum samples positive for VL (n=19), depending on the format used (Table 1). The results for some of these patients are shown in Fig 6.

Figure 6. Peptide EpQ11 specifically binds to human IgG1 in cassette as well as in dipstick format. 'T' indicates the location of the NLA at which, in a positive test, there is a coloured line due to the presence of peptide-IgG1 complex. Successful migration ensured by the development of the control line - 'C'.

Discussion

Identifying a biomarker of post-chemotherapeutic relapse is a key element for VL control. Considering that most of the patients affected by VL live in poor and remote villages in developing countries, it is of utmost importance to distinguish cured patients from those still in need of future medical interventions. This is especially valid by VL given the toxicity of the current available treatments and the growing problem of drug resistance.

We have identified a peptide specifically recognised by human IgG1 from active VL patients by mapping epitopes from multiple proteins, which were selected from immunoblots by searching antigenic features using comparative genomics and *in silico* algorithms. The peptide was adapted to an RDT format and showed promising results for VL status in terms of both sensitivity and specificity. Given the free availability of various *in silico* algorithms to predict or scan multiple protein features from DNA or AA sequences and the growing number of high quality sequenced genomes, the adopted strategy can also be employed for the search of diagnostic biomarkers or vaccine candidates for infectious diseases.

The decision to make use of comparative genomics and *in silico* algorithms to refine the output from immunoblots was due to the excessive number of putative hits identified by MS and we believe that the success of our approach lies in this combination. Nonetheless, we consider that technical issues might have been related to the excessive number of hits initially identified by MS from the immunoblots. The immunogenic bands excised from the acrylamide gels that were analysed by MS were excised from wide lanes (59 mm). We now believe that excising such wide lanes from acrylamide gels contributed to the excessive number of hits: by mechanically cutting along very wide lanes, we have probably included proteins that would not have been present had we have excised narrower (e.g. 4 mm wide) lanes, although we are convinced that multiple hits would have been identified anyway.

We have identified antigens from *L. donovani* promastigotes due to ease of culturing. In the host, promastigotes are phagocytosed by macrophages or other types of mononuclear phagocytic cells, where they differentiate into amastigotes, remaining in this form until transmission to a new vector. Therefore we believe that amastigote proteins are more likely to be targeted by host IGs. We have chosen a loose HAP cut-off for both promastigotes and amastigotes (HAP>Q1) due to the weak correlation between RNA levels and protein expression in *Leishmania* (unpublished results). By doing so, we avoid missing antigens that might be constitutively expressed, despite low RNA levels. Given the excessive number of protein candidates identified by MS we could set a stricter MASCOT score cut-off, ruling out proteins that could have been identified by chance in each antigenic band. Extracellular

proteins are more accessible for binding with host ABs, thus we have included this criterion in our filter. Finally, because intracellular proteins harbouring TRs are amongst the most widely used antigens for VL diagnosis [45], we decided to select such candidates in parallel to extracellular proteins. We have employed prediction algorithms of linear B-cell epitopes (instead of discontinuous) due to the linearised structure of proteins separated by SDS-PAGE. Also, the prediction of discontinuous epitopes still rely on the scarce availability of proteins with known 3D structures.

Even though the *in silico* filter that we applied ruled out over 90% of the proteins initially identified by MS from the immunoblots, it was still unpracticable to test the antigenicity of the selected 65 proteins. By screening synthetic peptides instead of recombinant proteins we circumvented the protein expression step and could therefore test a large number of top-scoring candidates predicted from multiple algorithms with various prediction methods. Moreover, the production of peptide arrays allowed the screening of multiple peptides in various concentrations using low volumes of reagents of limited availability to us (e.g. serum positive for VL).

Our work is a proof of concept that comparative genomics and *in silico* algorithms can be employed downstream of wet-lab experiments towards a more rational search for biomarkers. Despite promising sensitivity values with an RDT prototype we have developed, the intensity of the test lines were generally weak and our procedure still requires an extra 15 minute sample pre-incubation step when compared to RDTs used in field conditions. The sensitivity of the RDT might be improved by using a larger peptide composed of multiple copies in series ('peptide trains') or in parallel (multiple antigen peptides), a strategy already used in vaccine development.

Given the evidence from the present work that the described peptide specifically binds to IgG1 from Sudanese sera and the well described decreased performance of the rK39 in East Africa, it would be relevant to express and test the antigenicity of a chimeric protein composed of the rK39 and the EpQ11 sequence.

Conclusion

Based on the interpretation of the results of our experiments we conclude that:

1. Comparative genomics as well as *in silico* algorithms are useful tools for refining large output from wet-lab experiments towards a more rational search for diagnostic biomarkers.

2. B-cell epitopes prediction algorithms represent an interesting option for epitope mapping, enabling the screening of peptides in a high-throughput manner.

3. We have identified a peptide that specifically binds to human IgG1 from active VL patients from Sudan by refining wet-lab experiments with *in silico* searches.

4. Incorporation of the peptide described in this work into either rK39 or rK28-based assays might improve the sensitivity of VL diagnosis in eastern Africa.

Supporting information

S1 Table

S2 Table

Acknowledgments

This work was carried as part of the EuroLeish.net consortium and has received funding from the European Horizon's 2020 Research and Innovation Programme under Marie-Sklodowska Currie grant agreement number 642609.

TM is funded by the Sir Halley Stewart Trust (http://www.sirhalleystewart.org.uk/). The views expressed within this report are those of the authors and not necessarily those of the Trust. TM was additionally supported by the John Henry Memorial Fund (http://www.johnhenrymf.org/)

This work was part of the NIDIAG network research partnership supported by the European Commission under the Health Cooperation Work Programme of the 7th Framework Programme (Grant agreement no. 260260, http://cordis.europa.eu/fp7/home_en.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, et al. Leishmaniasis worldwide and global estimates of its incidence. PloS one. 2012;7(5):e35671.

[2] Burns JM, Shreffler WG, Benson DR, Ghalib HW, Badaro R, Reed SG. Molecular characterization of a kinesin-related antigen of Leishmania chagasi that detects specific antibody in African and American visceral leishmaniasis. Proceedings of the National Academy of Sciences. 1993;90(2):775–779.

[3] Cunningham J, Hasker E, Das P, El Safi S, Goto H, Mondal D, et al. A global comparative evaluation of commercial immunochromatographic rapid diagnostic tests for visceral leishmaniasis. Clinical infectious diseases. 2012;55(10):1312–1319.

[4] Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, et al. Diagnostic tests for kala-azar: a multi-centre study of the freeze-dried DAT, rK39 strip test and KAtex in East Africa

and the Indian subcontinent. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2008;102(1):32–40.

[5] Gidwani K, Picado A, Ostyn B, Singh SP, Kumar R, Khanal B, et al. Persistence of Leishmania donovani antibodies in past visceral leishmaniasis cases in India. Clinical and vaccine immunology : CVI. 2011;18:346–348. doi:10.1128/CVI.00473-10.

[6] Patil RR, Muliyil JP, Nandy A, Addy A, Maji AK, Chatterjee P. Dynamics of the antibodies in cohorts of cured cases of visceral leishmaniasis: its implication on the validity of serological test, value in prognosis and in post therapeutic assessment. Human vaccines & immunotherapeutics. 2012;8:725–730. doi:10.4161/hv.19542.

[7] De Almeida Silva L, Romero HD, Prata A, Costa RT, Nascimento E, Carvalho SFG, et al. Immunologic tests in patients after clinical cure of visceral leishmaniasis. The American journal of tropical medicine and hygiene. 2006;75:739–743.

[8] Zijlstra EE, Nur Y, Desjeux P, Khalil EA, El-Hassan AM, Groen J. Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. Tropical medicine & international health : TM & IH. 2001;6:108–113.

[9] Zijlstra EE, Daifalla NS, Kager PA, Khalil EA, El-Hassan AM, Reed SG, et al. rK39 enzyme-linked immunosorbent assay for diagnosis of Leishmania donovani infection. Clinical and diagnostic laboratory immunology. 1998;5:717–720.

[10] Pattabhi S, Whittle J, Mohamath R, El-Safi S, Moulton GG, Guderian JA, et al. Design, development and evaluation of rK28-based point-of-care tests for improving rapid diagnosis of visceral leishmaniasis. PLoS neglected tropical diseases. 2010;4. doi:10.1371/journal.pntd.0000822.

[11] National Center for Biotechnology Information. GenBank and WGS Statistics; 2015. Available from: https://www.ncbi.nlm.nih.gov/genbank/statistics/.

[12] Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP);. Available from: https://www.genome.gov/sequencingcostsdata/.

[13] Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, Leishmania major. Science (New York, NY). 2005;309:436–442. doi:10.1126/science.1112680.

[14] Downing T, Imamura H, Decuypere S, Clark TG, Coombs GH, Cotton JA, et al. Whole genome sequencing of multiple Leishmania donovani clinical isolates provides insights into

509 population structure and mechanisms of drug resistance. Genome research. 2011;21:2143–
510 2156. doi:10.1101/gr.123430.111.

511 [15] McCall LI, McKerrow JH. Determinants of disease phenotype in trypanosomatid
512 parasites. Trends in parasitology. 2014;30(7):342–349.

513 [16] Kaur G, Rajput B. Comparative analysis of the omics technologies used to study
514 antimonial, amphotericin B, and pentamidine resistance in leishmania. Journal of parasitology
515 research. 2014;2014.

516 [17] Pawar H, Renuse S, Khobragade SN, Chavan S, Sathe G, Kumar P, et al. Neglected
517 tropical diseases and omics science: proteogenomics analysis of the promastigote stage of
518 Leishmania major parasite. Omics: a journal of integrative biology. 2014;18(8):499–512.

519 [18] Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic
520 determinants on protein antigens. FEBS letters. 1990;276(1-2):172–174.

521 [19] Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes.
522 Immunome research. 2006;2(1):2.

523 [20] Saha S, Raghava G. Prediction of continuous B-cell epitopes in an antigen using
524 recurrent neural network. Proteins: Structure, Function, and Bioinformatics. 2006;65(1):40–
525 48.

526 [21] Chen Z, Li J, Wei L. A multiple kernel support vector machine scheme for feature
527 selection and rule extraction from gene expression data of cancer tissue. Artificial Intelligence
528 in Medicine. 2007;41(2):161–175.

529 [22] Capecchi B, Serruto D, Adu-Bobie J, Rappuoli R, Pizza M. The genome revolution in
530 vaccine research. Current issues in molecular biology. 2004;6:17–28.

531 [23] Bai X, Borrow R. Genetic shifts of Neisseria meningitidis serogroup B antigens and the
532 quest for a broadly cross-protective vaccine. Expert review of vaccines. 2010;9(10):1203–
533 1217.

534 [24] Zhao BP, Chen L, Zhang YL, Yang JM, Jia K, Sui CY, et al. In silico prediction of binding of
535 promiscuous peptides to multiple MHC class-II molecules identifies the Th1 cell epitopes from
536 secreted and transmembrane proteins of Schistosoma japonicum in BALB/c mice. Microbes
537 and infection. 2011;13(7):709–719.

[25] Carvalho GBF, Silva-Pereira RAd, Pacífico LGG, Fonseca CT. Identification of Schistosoma mansoni candidate antigens for diagnosis of schistosomiasis. Memorias do Instituto Oswaldo Cruz. 2011;106(7):837–843.

[26] Bhattacharyya T, Falconar AK, Luquetti AO, Costales JA, Grijalva MJ, Lewis MD, et al. Development of peptide-based lineage-specific serology for chronic Chagas disease: geographical and clinical distribution of epitope recognition. PLoS neglected tropical diseases. 2014;8:e2892. doi:10.1371/journal.pntd.0002892.

[27] Dumetz F, Imamura H, Sanders M, Seblova V, Myskova J, Pescher P, et al. Modulation of aneuploidy in Leishmania donovani during adaptation to different in vitro and in vivo environments and its impact on gene expression. MBio. 2017;8(3):e00599–17.

[28] Steinbiss S, Silva-Franco F, Brunk B, Foth B, Hertz-Fowler C, Berriman M, et al. Companion: a web server for annotation and analysis of parasite genomes. Nucleic acids research. 2016;44:W29–W34. doi:10.1093/nar/gkw292.

[29] Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20:3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

[30] Matrix Science. Mascot database search: scoring;. Available from: http://www.matrixscience.com/help/scoring_help.html.

[31] Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ, et al. Proteomic analysis of the secretome of Leishmania donovani. Genome biology. 2008;9:R35. doi:10.1186/gb-2008-9-2-r35.

[32] Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods. 2011;8:785–786. doi:10.1038/nmeth.1701.

[33] Goto Y, Coler RN, Reed SG. Bioinformatic identification of tandem repeat antigens of the Leishmania donovani complex. Infection and immunity. 2007;75:846–851. doi:10.1128/IAI.01205-06.

[34] Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic acids research. 1999;27:573–580.

568 [35] Aptum Biologics LTD . EpioQuest B - find immunodominant epitopes;. Available from:
569 https://www.epiquest.co.uk/epiquest-b.

570 [36] Singh H, Ansari HR, Raghava GP. Improved method for linear B-cell epitope prediction
571 using antigen's primary sequence. PloS one. 2013;8(5):e62216.

572 [37] Genscript. Peptide solubility and stability;. Available from:
573 https://www.genscript.com/peptide_solubility_and_stablity.html.

574 [38] Layne E. Spectrophotometric and turbidimetric methods for measuring proteins. 1957;.

575 [39] Anthis NJ, Clore GM. Sequence-specific determination of protein and peptide
576 concentrations by absorbance at 205 nm. Protein science : a publication of the Protein
577 Society. 2013;22:851–858. doi:10.1002/pro.2253.

578 [40] Falconar AK. Epitope reactions can be gauged by relative antibody discriminating
579 specificity (RADS) values supported by deletion, substitution and cysteine bridge formation
580 analyses: potential uses in pathogenesis studies. BMC research notes. 2012;5(1):208.

581 [41] Bhattacharyya T, Ayandeh A, Falconar AK, Sundar S, El-Safi S, Gripenberg MA, et al. IgG1
582 as a potential biomarker of post-chemotherapeutic relapse in visceral leishmaniasis, and
583 adaptation to a rapid diagnostic test. PLoS neglected tropical diseases. 2014;8(10):e3273.

584 [42] R Core Team. R: A Language and Environment for Statistical Computing; 2018. Available
585 from: https://www.R-project.org/.

586 [43] Scherer R. PropCIs: Various Confidence Interval Methods for Proportions; 2018.
587 Available from: https://CRAN.R-project.org/package=PropCIs.

588 [44] Djira GD, Hasler M, Gerhard D, Schaarschmidt F. mratios: Inferences for ratios of
589 coefficients in the general linear model; 2012. Available from: https://CRAN.R-
590 project.org/package=mratios.

591 [45] Bhattacharyya T, Marlais T, Miles MA. Diagnostic antigens for visceral leishmaniasis:
592 clarification of nomenclatures. Parasites & vectors. 2017;10:178. doi:10.1186/s13071-017-
593 2120-x.

594

Table and Figures

Table 1. Peptides shortlisted for synthesis. The top 20 scoring peptides from each of the four prediction algorithms were shortlisted and synthesised for a pilot screening with Sudanese VL serum samples. 'C.reported' was calculated by dividing the reported amount by the volume of the solvent mixture in which they were dissolved, while 'C.280' and 'C.205' were calculated based on the molar absorbances at 280 and 205 nm, respectively.

ID	Sequence	k-mer size	MW [g/mol]	Purity	C.reported [mg/ml]	C.280 [mg/ml]	C.205 [mg/ml]
abc_01_CVTQ	CVTQEHFREAMAKTNP	16	2287.63	0.58	2	NA	3.27
abc_02_KEAP	KEAPGATEKDRAKATP	16	2095.38	0.57	2	NA	3.14
abc_03_TAYI	TAYIMRPLDHGADVTL	16	2198.57	0.69	2	2.75	NA
abc_04_MITN	MITNDDAPVRDSVLTD	16	2187.45	0.45	1.23	NA	NA
abc_05_MPTV	MPTVDERQTFMFSATF	16	2333.71	0.4	2.31	NA	NA
abc_06_TRTG	TRTGDYAFSYDKMLDM	16	2339.67	0.55	1.13	NA	NA
abc_07_KSTI	KSTISGHLLMEKGLVD	16	2153.57	0.53	2	NA	2.65
abc_08_KEAI	KEAITTFREEDPKVTD	16	2304.58	0.48	2	NA	NA
abc_09_HVYS	HVYSELGKKFGAAADP	16	2115.41	0.51	2	3.42	4.23
abc_10_LGMG	LGMGISGGEEGARKGP	16	1941.23	0.56	2	NA	2.52
abc_11_VAYQ	VAYQETPESERAELPP	16	2241.48	0.52	2	2.22	2.13
abc_12_AGTG	AGTGFPYREMMPMNAP	16	2195.61	0.72	2	3.95	NA
abc_13_TDSW	TDSWGFFGVFDGHVND	16	2225.4	0.57	1.33	NA	NA
abc_14_HKKS	HKKSTEDNDDDAFCAP	16	2218.37	0.54	1	NA	NA
abc_15_YGLA	YGLAFDPYGGTAGLYD	16	2105.33	0.54	1.33	0.74	0.76
abc_16_WEEW	WEEWGNPNEYKYYDYM	16	2612.86	0.49	1.73	1.68	1.43
abc_17_HSTI	HSTIGVHPTSAEELCS	16	2093.33	0.65	2	NA	2.68
abc_18_EKCI	EKCIELKPDFVKGYAR	16	2321.76	0.53	0.87	3.85	1.08
abc_19_PAGI	PAGINIPNYDDIRQTV	16	2211.5	0.57	0.77	1.69	0
abc_20_TGTI	TGTIDNGVVKMEKAEE	16	2146.44	0.55	2	NA	NA
BP_01_CSPP	CSPPPPSPSPHPRPPS	16	2062.38	0.7	2	NA	1.98
BP_02_GEGS	GEGSPTSPTSPKQPGS	16	1939.11	0.65	2	NA	2.08
BP_03_AEAG	AEAGAPAGSGAPPPAD	16	1760.92	0.56	2	NA	NA
BP_04_KPAP	KPAPPKPKESKEPENA	16	2172.55	0.59	2	NA	1.88
BP_05_TAKP	TAKPKQQDEDPDGAAE	16	2125.28	0.63	2	NA	NA
BP_06_GGDR	GGDRGGGTGNEDDDYE	16	2039.02	0.79	2	2.95	2.05
BP_07_MQQP	MQQPPTPQPQKQQKQQ	16	2345.71	0.53	2	NA	3.13
BP_08_EEEE	EEEEEEEEEEEPQATR	16	2417.44	0.61	1.62	NA	NA
BP_09_RSDP	RSDPSGGGGGNRDDNE	16	2015.04	0.6	2	NA	NA
BP_10_APAS	APASAPAPAPAAAPTP	16	1783.06	0.67	2	NA	2.61
BP_11_KRGG	KRGGDKDGGGESGEAA	16	1916.04	0.62	2	NA	NA
BP_12_RSQQ	RSQQGEQEPEEDEEEV	16	2343.41	0.63	2	NA	NA
BP_13_QQLS	QQLSSSPPPPRERAED	16	2219.45	0.57	2	NA	NA
BP_14_EPEV	EPEVGEEPQPEEEEDA	16	2238.31	0.72	2	NA	NA
BP_15_AAPS	AAPSGGPSGNSDEEDL	16	1928	0.68	2	NA	NA
BP_16_SDVT	SDVTGGGGGGGSGGGG	16	1560.6	0.64	1	NA	NA
BP_17_GQQQ	GQQQQQQDPPAGQQGV	16	2119.28	0.82	2	NA	NA

BP_18_AETP	AETPADDAGQPHEPEK	16	2117.26	0.71	2	NA	NA
BP_19_PPSA	PPSAGSKDGAPSDGVP	16	1864.05	0.51	2	NA	NA
BP_20_DATP	DATPAAANGEGPGKEN	16	1924.06	0.65	2	NA	NA
EpQ_01_AYAT	AYATMLKDVQWKVRKS	16	2349.82	0.58	2	6.79	6.25
EpQ_02_HEKL	HEKLVQDIWKKLEAKG	16	2347.79	0.58	2	4.28	4.01
EpQ_03_SCSV	SCSVKLGLWKNAVNNC	16	2161.55	0.57	2	6.56	6.9
EpQ_04_MAYV	MAYVCELGPNQGWK	14	2021.38	0.58	0.88	6.49	NA
EpQ_05_LKDP	LKDPKQYQSIVDAEWK	16	2373.73	0.31	2	1.88	NA
EpQ_06_ERCE	ERCEDPNAWKGPTNGG	16	2156.35	0.62	1.33	1.92	NA
EpQ_07_DNPA	DNPAGPTTWKSDEPAL	16	2124.33	0.62	1.73	1.8	1.88
EpQ_08_YGIS	YGISFPKNPMLTEWKT	16	2337.77	0.43	2	2.69	2.79
EpQ_09_GMSS	GMSSDQLLQFLLQQQQ	16	2289.63	0.48	1.86	NA	NA
EpQ_10_AAKK	AAKKKRVGCWK	11	1700.11	0.56	2	2.29	2.22
EpQ_11_NIRI	NIRIHLGDTIRIAPCK	16	2245.71	0.51	1.1	NA	2.33
EpQ_12_ERRR	ERRRVEYQQFLDVCGQ	16	2451.78	0.41	1.4	6.14	NA
EpQ_13_MDRE	MDRESLCPNWK	11	1804.11	0.41	2	2.06	NA
EpQ_14_SRQM	SRQMTMCKEERIANCK	16	2353.81	0.34	2	NA	1.85
EpQ_15_DWSI	DWSIVERGWK	10	1700.95	0.5	2.13	1.69	NA
EpQ_16_VLVQ	VLVQGAIWGINSYDQW	16	2274.6	0.39	15	NA	NA
EpQ_17_GRTI	GRTILRNHKWAGNNKV	16	2289.67	0.31	2	2.1	1.79
EpQ_18_EKVR	EKVRSGEWKGQTGKSI	16	2215.53	0.38	2	2.13	2.11
EpQ_19_PDVA	PDVAHVVQFDLPQEMD	16	2265.57	0.52	2.31	NA	NA
EpQ_20_ADVT	ADVTATLAWK	10	1500.75	0.49	1.03	1.56	NA
LbT_01_RPGG	RPGGPPGYRTPYTAK	15	2043.35	0.45	2	9.28	2.46
LbT_02_TQGD	TQGDRQKIQDAVSAA	15	2013.23	0.52	1.3	NA	NA
LbT_03_EVKS	EVKSRYNVDVSQNKR	15	2247.54	0.53	2	2.69	2.1
LbT_04_VIEM	VIEMTRAFEDDDFDK	15	2256.51	0.51	2	NA	NA
LbT_05_GSAD	GSADLTPSNLTRPAS	15	1912.12	0.47	2	NA	NA
LbT_06_VRPI	VRPIPSFDDMPLHQN	15	2191.53	0.47	2	NA	NA
LbT_07_CELG	CELGPNQGWKAVVAD	15	2012.3	0.47	1.07	1.59	NA
LbT_08_APQQ	APQQTQSGIRRVTRA	15	2094.4	0.44	2	NA	2.18
LbT_09_SAEE	SAEEKGTGKRNQITI	15	2057.33	0.42	2	NA	0.7
LbT_10_GEAE	GEAEWLEWESTVLTP	15	2172.42	0.54	0.96	1.08	0
LbT_11_EELR	EELRQRRHQGPGSPG	15	2129.36	0.41	0.87	NA	0.66
LbT_12_EWAN	EWANKPLDDLDPHPS	15	2159.38	0.55	2	1.57	1.56
LbT_13_IKEE	IKEETEMIEGEVVEV	15	2159.48	0.5	1.62	NA	NA
LbT_14_VSDF	VSDFFGGKELNKSIN	15	2080.36	0.74	2.92	NA	2.12
LbT_15_WEDV	WEDVGGLLDVKRELQ	15	2182.5	0.53	2	1.67	0
LbT_16_SGGA	SGGAGPAGGASSGPK	15	1582.72	0.53	2	NA	2.5
LbT_17_KATN	KATNGDTHLGGEDFD	15	2002.12	0.57	2	NA	NA
LbT_18_EREG	EREGKDITLIGFSRG	15	2103.4	0.53	1.56	NA	2
LbT_19_AWAS	AWASSPAPTEARTAP	15	1938.17	0.45	2	4.01	NA
LbT_20_VDRD	VDRDNKKLSSGMVCS	15	2064.4	0.51	2	NA	2.53

Figure 1

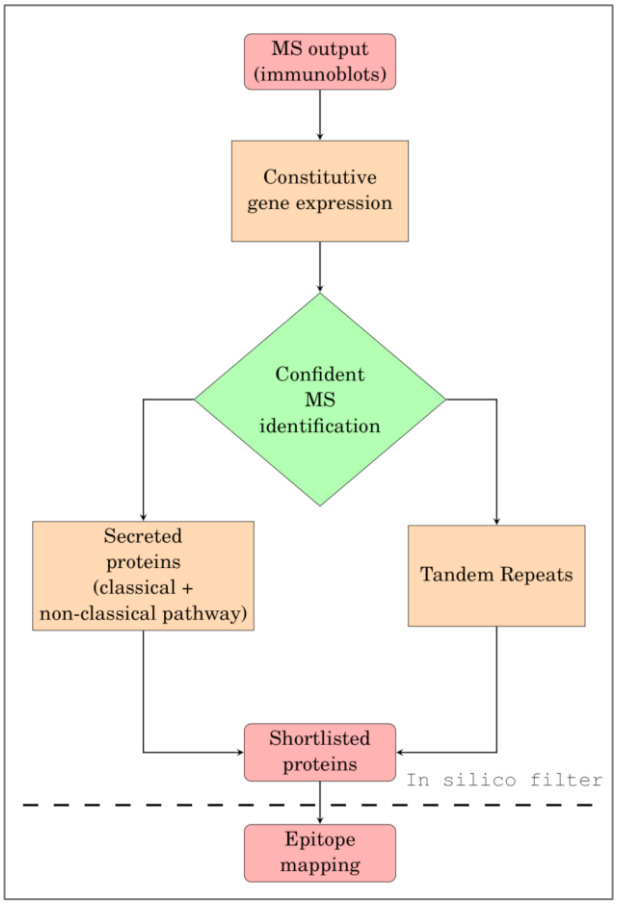


Figure 2.

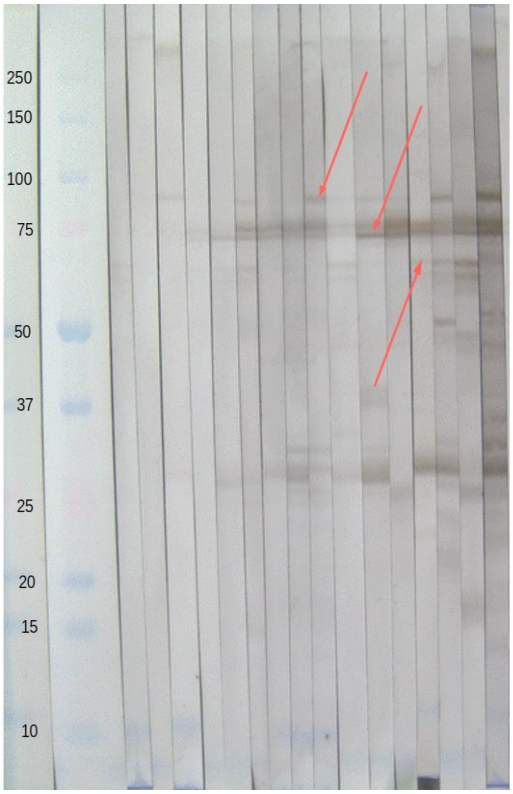


Figure 3.

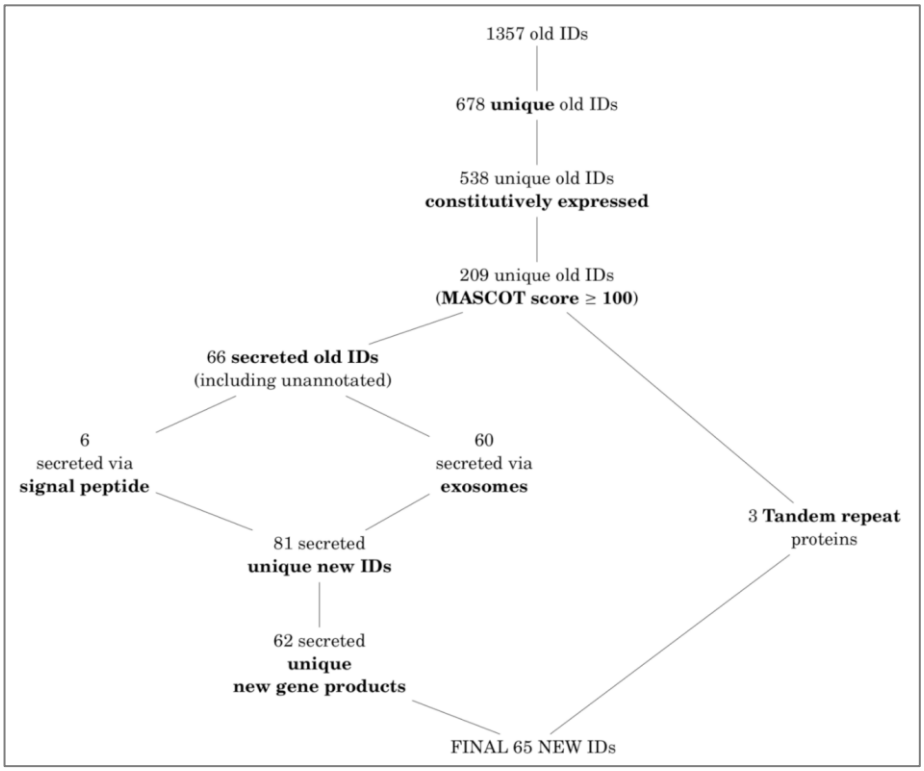


Figure 4.

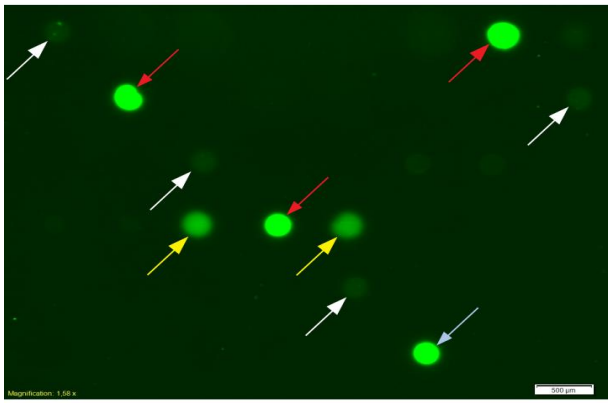
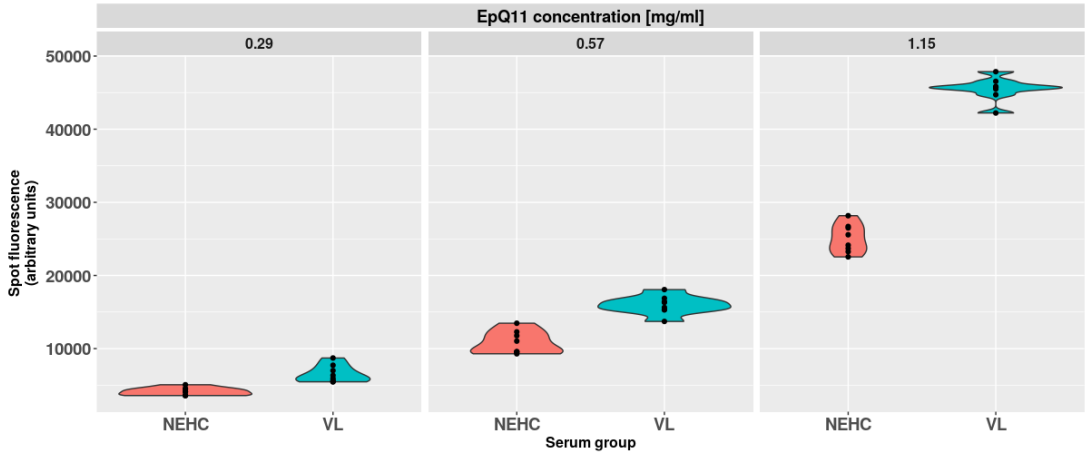
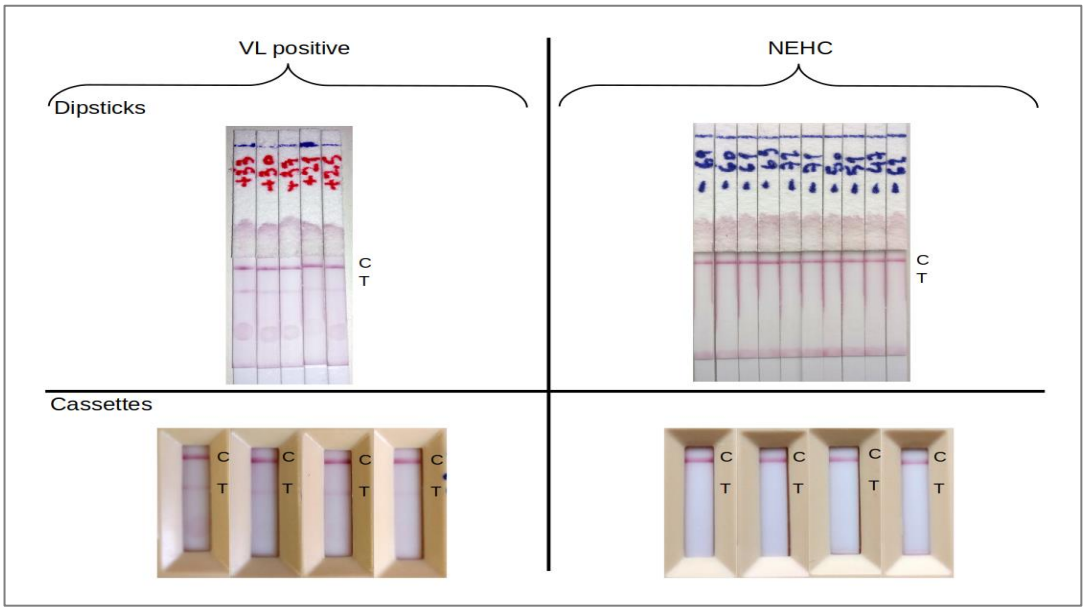


Figure 5.



EpQ11 concentration [mg/ml]	Ratio (VL/NEHC)	95% CI (VL/NEHC)
0.29	1.58	1.32 - 1.87
0.57	1.48	1.30 - 1.68
1.15	1.82	1.69 - 1.94

Figure 6.



CHAPTER 6: Results

This chapter presents results of experiments that were not part of the preceding chapters. Full methods are in the Appendix.

6.1 SCP/TAPS gene sequencing

An SCP/TAPS protein family gene of *S. stercoralis*, corresponding to existing seroantigen NIE, and under investigation as a candidate coproantigen, was the target of our PCR. The aim was to amplify and sequence this gene from geographically-diverse *Strongyloides*-infected stool samples in order to investigate antigen sequence diversity. Optimisation of the PCR to amplify the NIE homologous region and another region of gene SSTP_0001008900 from *S. stercoralis* positive stool DNA was unsuccessful, regardless of the different annealing temperatures and magnesium concentrations used (Figure 19). However, our novel primers did amplify regions of the correct size in neat larval DNA (Figure 19F). Primer set 1, covering the NIE homologous region (aa 28-252), produced an amplicon of 1308 bp and primer set 2 produced an amplicon of 1232 bp, at all annealing temperatures (55, 58 and 60°C) and at both 1.5 and 2 mM magnesium (Figure 19F). 18S rRNA primers[116], used as positive control, produced an amplicon 101 bp. Optimal annealing temperature for Set1 primers is therefore 55°C, and for Set2 primers, 60°C, both with 1.5 mM magnesium.

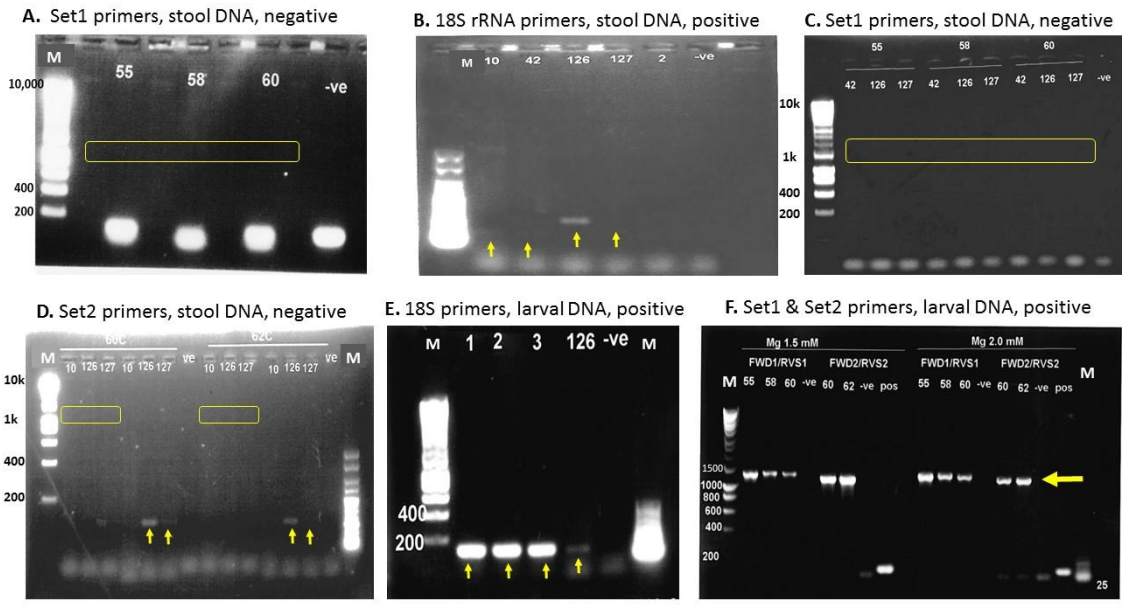


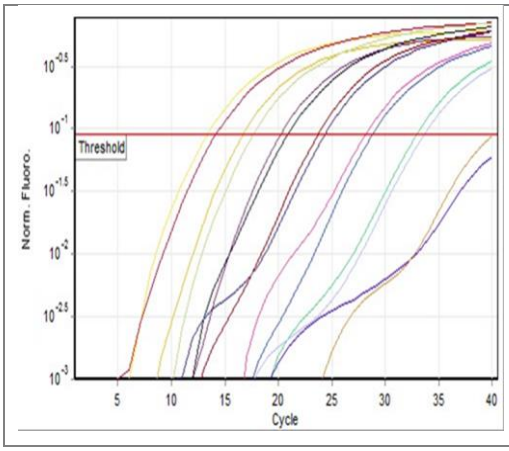
Figure 19. Amplification of *Strongyloides stercoralis* gene SSTP_0001008900, which contains a region homologous to NIE seroantigen. Amplified using novel primers and existing species-specific 18S rRNA primers (Verweij, 2009). Optimisation in preparation for sequencing. Full assay details are described in the Methods. A) Set 1 primers with positive stool DNA sample

at different annealing temperatures, all negative, box indicates expected amplicon location. B) 18S rRNA primers with four positive and one negative stool DNA samples- positive bands arrowed. C) Set 1 primers with positive stool DNA samples at different annealing temperatures, all negative, box indicates expected amplicon location. D) Two annealing temperatures (60 and 62°C) with Set 2 primers, negative, boxed, and 18S rRNA primers, positive, arrowed, with three positive stool DNA samples. E) Elutions 1, 2 and 3 of purified *S. stercoralis* larval DNA and positive stool DNA sample 126 amplified with 18S rRNA primers, positive, arrowed. F) Set 1 and Set 2 primers amplifying *S. stercoralis* larval DNA with different magnesium concentrations and annealing temperatures, all positive, arrowed, 18S rRNA primers as positive control. M: marker, sizes in base pairs (bp).

Sequences obtained from *S. stercoralis* larval DNA of the PV0001 reference strain (originating from a dog in the USA) for the two main exon regions (each containing a short intron, see figure in Appendix), did not reveal any sequence diversity. Approximately 400 bases were reliably sequenced in each direction, with a loss of 70-100 bases after each primer due to poor quality sequence trace.

6.2 *Strongyloides* qPCR trial

A trial of a widely-used qPCR assay using diluted neat *S. stercoralis* larval DNA successfully produced amplification curves (Figure 20). One operator performed dilutions of DNA and two different operators added these dilutions to the qPCR reaction tubes (Figure 20).



S. stercoralis DNA dilution	Operator 1 (Ct)	Operator 2 (Ct)	Ct difference
Neat	13.42	14.20	0.78
10	16.67	17.65	0.98
100	20.23	20.92	0.69
1000	23.69	24.30	0.61
10000	28.17	28.80	0.63
100000	32.90	33.60	0.70
Negative Control	-ve	-ve	n/a

Figure 20. *Strongyloides stercoralis* DNA amplification curves by qPCR of 18S rDNA sequence. Primers and probe as per Verweij *et al.* (2009)[116], with method as per Becker *et al.* (2015)[145]. Ct, cycle threshold.

6.3 *Strongyloides* genomic variants

Mapping of the *S. stercoralis* genome to the *S. ratti* reference revealed 13,804 variants, a rate of 1 per 3,069 bases. Of the variants, 12,737 (92.3%) were SNPs, 587 (4.3%) were insertions and 480 (3.5%) were deletions. The effects of these variants would lead to amino acid differences in 26.4% of the cases but most (73.5%) of the variants between the two species were silent, not affecting the amino acid sequence. Variants affecting protein sequences in candidate coproantigens can inform the selection of species-specific antigens.

CHAPTER 7: Discussion

7.1 *Strongyloides*

Background and diagnostic needs

Strongyloides stercoralis infection is widespread and goes hand-in-hand with poverty in tropical regions. Its capacity to cause life-long infection means that strongyloidiasis also occurs outside its endemic region in migrants and travellers. Despite being treatable, strongyloidiasis often goes undiagnosed leading to morbidity. In addition, unlike most other human helminth infections, strongyloidiasis can cause the potentially fatal conditions of hyperinfection or disseminated infection in the immunocompromised. There is an absence of information on the true global burden of strongyloidiasis, largely due to inadequate diagnostic options that fail to provide sensitive and rapid results on current infection. Therefore, there is a need for such a diagnostic. This study aimed to address this need by identifying antigens originating from the nematode that could be detected by a rapid diagnostic test performed on a stool sample.

Success of this study

This study succeeded in identifying over 40 novel potential stool antigens of *S. stercoralis*. We used an innovative method that analysed the wide range of open-access bioinformatic data now available on helminths that has not previously been applied to *Strongyloides* diagnostics development. By this combination of 'omics' data and online analysis tools, this study identified candidate proteins for a coproantigen capture immunoassay for strongyloidiasis. A similar bioinformatic approach has been applied to diagnostic and therapeutic target discovery for filarial nematode infections such as *Wuchereria bancrofti*[263], and to vaccine candidate design for the bacterial infection, brucellosis[264]. The present study used high-quality, detailed, transcriptomic data[20] as its starting point, which enabled selection of genes that are differentially expressed in gut-dwelling life stages of *S. stercoralis*, as compared with non-gut dwelling stages.

Quality of data

Open access data were carefully scrutinised before use to ensure that they were robust and relevant to the present objectives. This included selecting transcriptomic data that had multiple biological replicates and paired-end reads in high numbers, therefore giving high-quality, reliable data. In addition, the parasite culturing method used by the creators of the data was examined in detail to understand which of the analysed life stages may be found in the host gut or stool. Likewise, proteomic data from *S. ratti* was selected due to the re-

assignment of older mass spectral data to proteins by Hunt *et al.* (2016)[33] following their improvement of the *S. ratti* reference genome.

Outgroups

Genome-derived proteomes of multiple helminths and *Homo sapiens* were used to investigate the level of conservation of protein sequences between organisms. The available genomes were at various stages of assembly and annotation, which could have affected the completeness of the derived proteome data. However, the selection of these outgroups was made to include biologically relevant species that co-infect humans, as well as closely-related parasitic or free-living nematodes, covering a wide range of genera and also reflecting different levels of genome assembly.

Antigenic diversity and PCR

As part of selecting suitable diagnostic coproantigens, we aimed to investigate the diversity of these proteins by developing novel PCR assays to amplify and sequence relevant regions from geographically diverse samples of *S. stercoralis*. The selected genomic region, encoding a homologue of the NIE serological antigen, in the SCP/TAPS protein family, was successfully sequenced from neat larval DNA. The *S. stercoralis* DNA was of the same strain as the reference genome therefore little could be deduced from this with regard to geographic diversity. We were able to confirm however, that there were no heterogeneities in the sequenced region of this gene, regardless of how many copies were in the genome. We also showed that our novel primers were successful. Therefore, reasons that no amplicon was obtained from infected stool DNA could have been due to polymerase inhibition caused by stool components, or the small amount of target DNA. This SCP/TAPS PCR could be further optimised to increase sensitivity by varying the PCR assay conditions, or by using alternative stool DNA extraction procedures[118, 265]. Other gene targets may be more easily amplified from stool, particularly for candidate antigens with multi copy genes. Whenever possible, we recommend that future studies of antigenic diversity in *S. stercoralis* should obtain DNA from isolated larvae, rather than whole stool to avoid this challenge. In addition, for improved sequence coverage and reliability, sequencing primers should be designed to bind within the amplicon in order to obtain sequences of the entire gene in both directions.

Starting point

A potential limitation of this study was the use of differential gene expression as a starting point. This was as opposed to assessing standardised expression levels in transcripts per million (TPM) from the gut-dwelling life stages to reveal their highly expressed genes, regardless of the expression in non-gut life stages. The differential expression (DE) method did not capture the complexity of the RNAseq data which had triplicate measurements for 7

life stages. Not all of the seven life-stages were relevant to the antigen search, but more genes could have been used as the starting point. Looking only at parasitic females, 877 genes had normalised expression counts of >10,000. This was much higher than the 328 DE genes that I went on to analyse in phylogenetic trees and protein alignments. In addition, transcription does not necessarily correlate with protein production[33], therefore these data should ideally be used as a confirmation of genomic and proteomic data analysis rather than a foundation. However, at the time of designing and commencing this study, the E/S proteome data were not optimal and the improved and new genome assemblies of the *Strongyloides* spp. had not been released, in which case, transcriptomic data were a suitable basis for the coproantigen search.

Proteomic data conversion

An additional drawback was the lack of a proteomic dataset for *S. stercoralis*. We addressed this by obtaining E/S proteomic data for the closely-related *S. ratti* and extracting orthologous proteins from the *S. stercoralis* genome. However, there remained considerable differences between these two datasets. Therefore, the *S. stercoralis* E/S orthologues were submitted to online tools to predict the presence of signal peptides, which indicated those orthologues with most likelihood of being secreted. A similar but more extensive filtering was applied by Cuesta-Astroz *et al.* (2017)[266] to predict secreted proteomes directly from genomic data of 82 helminth species, revealing a secreted proteome of 973 proteins from *S. ratti*, similar to the experimental total of 882 found by mass spectrometry of material from both parasitic and free-living female worms[33].

Sample types

This study aimed to discover coproantigens because stool is the most likely sample to contain *Strongyloides* antigens. However, stool samples are not an ideal sample type and although they are widely collected during STH surveys, an ideal rapid test would use a urine or saliva sample that can be provided on demand and with less handling risk to both the patient and clinical staff. Evidence in favour of investigating these alternative sample types comes from a urine antigen detection assay using a monoclonal antibody against the liver fluke *Opisthorchis viverrini* which had 81% sensitivity for infection and minor cross-reactivity with *S. stercoralis*[267]. In addition, a urine-based assay for the filarial sub-cutaneous nematode infection, onchocerciasis, detected 'NATOG' a human metabolite of a nematode neurotransmitter. This test is undergoing continued development and had promising sensitivity of 85% in an initial trial[268]. *Strongyloides* DNA has been detected in urine of both human and animal infections, indicating that worm material does enter the urine[269-271]. Evidence for *Strongyloides* antigens in saliva comes from an assay detecting immune

complexes (IgG bound to antigen) and although currently-infected individuals were positive in 30% of cases, there was also 20% positivity in 'healthy' individuals, indicating weak specificity[142]. These assays suggest that urine, and potentially saliva, could be explored as sources of parasitic helminth antigens, regardless of the body site of parasitic life stages. Such analysis could be performed using antibodies raised against the candidate antigens from this study, as well as by mass spectrometry of urine and saliva.

Antigen detection application

Antigen detection as a diagnostic method has significant advantages over serology in that it can be used to monitor prevalence in areas that have been subject to mass drug administration. Assessing cure after treatment with ivermectin could be carried out as soon as 4 days after treatment as larvae cease to be excreted by this time[115]. However, a sensitive antigen detection test may require a longer period for nematode proteins to be completely cleared from the host gut. In one study, DNA was detected in stool up to 4 years after apparent cure[70]. This was presumed to be indicative of incomplete treatment, however reinfection, or contamination during the assay process are also possible causes.

Species and genus specificity

This study aimed to identify antigens of *S. stercoralis*, but given that another *Strongyloides* species, *S. fuelleborni*, is also found in humans and all other *Strongyloides* species are not, a genus-specific diagnostic would suitable. However, for certain helminth infections, there is a need for species-specific assays. Such a test has been piloted for *Taenia solium* where differentiating between this and *T. saginata* has implications for potential cysticercosis in the population[272]. Equally high specificity could be important for differentiating echinococcosis caused by *Echinococcus multilocularis* from *E. granulosus* which have different environmental reservoir hosts and treatment requirements. Such specificity is likely to be possible using monoclonal antibodies against defined antigens[273]. In other cases, genus specificity is desirable, for example to differentiate liver flukes *Opisthorcis* and *Clonorchis* that are carcinogenic, from other flukes that are not[274]. Developing such specific diagnostics could use comparative genomics to investigate variants in the genes of interest but should be supported by sequencing of those regions for confirmation. We used *S. ratti* as a reference genome for variant calling because it was the best assembled of all the *Strongyloides* genomes but by comparing other *Strongyloides* genomes to the reference, we enabled the discovery of unique variants in any of those species. Further work could expand on this aspect of the study by identifying the genome regions encoding the candidate coproantigens and identifying genus- and species-specific variants.

Future use for antigen discovery method

The computational method applied here to discover diagnostic antigen candidates has great potential to be applied to other human and animal parasitic helminth infections. Certain protein families are known to be upregulated in parasitism and in different life stages of parasitism. Analysis of these protein families from other organisms, particularly for commonly antigenic features such as repeat regions, could facilitate diagnostics development, of both serology and antigen capture, for other neglected diseases.

7.1.1. Strengths

- This study identified novel candidate coproantigens of *S. stercoralis* without the need to culture *Strongyloides* in animals or obtain material from infected humans.
- Short peptide candidate antigens have been produced commercially and will be screened for antigenicity/sero-reactivity prior to antibody generation.
- The antigen search was informed by multiple open access data types and sources, including literature on the recent detailed analyses of the *Strongyloides* genomes by Hunt *et al.* (2016)[33]; experimental proteomic and transcriptomic data; and multiple genomes.
- Candidate antigens were selected, based on comprehensive analysis and with multiple sources of evidence for their likely success as coproantigens, including specificity to *Strongyloides*, antigenicity, accessibility to antibodies and presence in stool.
- Novel PCR primers were used successfully to investigate sequence diversity of an NIE gene homologue.
- Detailed analysis of existing serological antigens of *S. stercoralis* revealed features of these antigens for the first time. This included the identity of the recently discovered rSs1a as a TAP24-like protein family member, SsIR as a collagenic tail of AChE, and potential epitope regions within NIE, as well as the possible structures of these antigens. All of this information contributes to the improvement of these antigens and to their use as potential coproantigens.

7.1.2. Limitations

- Starting with differential gene expression analysis of transcriptomic data. This may have limited the number of promising genes too soon in the analysis and may have overlooked some constitutively high-expressed genes. Future studies could use as a starting point, absolute gene expression values or comparative genomics in order to

identify protein regions and genes that are either species specific, amplified in the target organism, or contain species specific variants.

- Lack of automation of the antigen discovery pipeline was a drawback in terms of the time taken to analyse each dataset and combine the outputs of these analyses. Others have developed tools to perform similar computational tasks[275]. This approach could be applied to *Strongyloides* diagnostics development.
- Epitope prediction of conformational B-cell epitopes is often inadequate. However, shortly after this study was completed, an updated version of the online epitope prediction tool, BepiPred 2.0, was released which claimed to have better accuracy for conformational epitopes than BepiPred 1.0. It was a limitation of this study that the older version of the tool was used. However, given the pace of development of such tools, computational analyses can become outdated rapidly. Automation of the antigen discovery pipeline would assist with this problem by enabling rapid updating and repeating.

7.1.3. Further work

This study provides an example of the immense value of combining publicly available genomic, transcriptomic and proteomic data for diagnostics, vaccine development and drug target selection. A computational method can speed up and reduce the cost of selecting target molecules and such an approach is highly amenable to the neglected diseases that do not attract the financial support of other conditions. In particular, this approach could be applied to other helminth infections, most of which lack rapid, sensitive and affordable diagnostic options. This study could be followed up by the use of a bioinformatics pipeline, such as 'Vacceed', a vaccine discovery pipeline developed by Goodswen *et al.* (2014)[275]. This would facilitate the discovery of diagnostic markers for any eukaryotic pathogen with available genomic data. The WTSI is in the process of assembling the genomes of 50 helminths, all of which are open-access and include trematodes, cestodes and nematodes. In a more focused approach, protein families known to be associated with parasitism could be investigated for diversity, in order to identify species-specific proteins or variants.

Next steps in *Strongyloides* diagnostic development

Regarding this study, the specific next steps towards developing an antigen capture assay are detailed below. This next phase of assay development will be carried out by a subsequent PhD student, under the supervision of the candidate and her supervisor:

1. Synthesis of the selected peptide coproantigen candidates (underway).

2. For larger candidate antigens, construction of gene clones and expression in *E. coli* would be preferable. Alternatively, to explore glycosylated antigens, a eukaryotic and preferably nematode, expression system could be used, potentially *C. elegans* or *S. ratti*. Novel primers would be needed and genes could be amplified and cloned from the *S. stercoralis* reference strain material available at LSHTM (kindly donated by Thomas Nolan, University of Pennsylvania, USA). Alternatively, gene sequences 'gene strings' of the relevant sequence could be purchased and inserted into plasmids for expression. Only small quantities of antigens would be needed for initial screening.

3. Screening, e.g. by ELISA or microarray, of the synthesised/recombinant proteins for antigenicity, using experimentally infected gerbil serum (kindly donated by Thomas Nolan, University of Pennsylvania, USA) and a non-infected control serum. Antigens well recognised by infected serum could indicate that they are likely to be strong antigens when raising antiserum.

4. Synthesis of a larger quantity of the antigenic candidates and at high purity for raising antibodies.

5. Raise antibodies to the candidate antigens in mice, rabbits or chickens via a commercial source. To reduce costs, antibodies could be raised to several of these candidates, e.g. 3, in a single animal. The antibodies would then be developed into a paired antigen capture assay initially using the original antigen which could be spiked into a *S. stercoralis* negative stool sample to assess interference and ideal sample pre-processing.

6. Perform *S. stercoralis* qPCR of ethanol-preserved stool samples collected in Guinea Bissau (ongoing). This would identify those that are positive for *S. stercoralis* and the corresponding formalin-preserved aliquots can be used in coproantigen assay development. Previous studies in Guinea Bissau have identified a *S. stercoralis* prevalence of 6.2%-7.9% in various locations throughout the country using various microscopic methods[276-278]. If these prevalence rates can be extrapolated, our sample size of 395 could contain about 31 positive samples.

7. Formalin antigen fixation and storage at -20°C has been reported to preserve coproantigens for months[279] and may contribute to lower background reactivity from stool samples[140]. The presently available stool samples in 10% formalin have been stored at -70°C to -80°C since October 2017. For assay development, the formalin stool should be homogenised and centrifuged to retain the supernatant which can then be

used in immunoassays to investigate sensitivity and specificity. In their *S. ratti* coproantigen assay development using antibodies against lysed whole worm antigen, Sykes and McCarthy (2011)[140] found that a dilution of at least 1:4 of the formalin stool extract was needed to ensure specificity.

8. Investigate antigenic diversity of the target sequences. In parallel to assay development, the amino acid sequence diversity of candidate or selected antigens should be investigated from different geographical regions and within an individual infection, preferably using DNA from worms collected by Baermann funnels or Koga agar plate cultures rather than stool DNA.

7.2 Visceral leishmaniasis

7.2.1. IgG1 serology

Background and success of this study

Previous studies have indicated that IgG1 may be a suitable indicator of cure and relapse after treatment for VL, currently a significant diagnostic gap[280]. In comparison with other IgG sub-classes, IgG1 provided the best discrimination between cure and relapse against a parasite lysate antigen in ELISA[280]. The current project, further investigating the utility of IgG1 diagnostics for VL, including in RDT format, used a larger sample size of paired sera than previously tested and showed that this test can differentiate between active VL and cure at 6 months in India, and potentially sooner in Sudan. This work also confirmed that IgG1, as measured by the RDT, was more likely to be elevated in relapsed individuals, where it was negative or reduced in those deemed cured. In addition, we found that the RDT, with further development, may have the potential to confirm PKDL, which currently lacks an empirical test.

rK39 and IgG1

The adaptation of the RDT to use rK39 antigen (at two concentrations), rather than parasite lysate, improved the detection of active VL (94.7-100% positive compared with 77.9% respectively), and relapse (85.7-90.5% and 84.8% positive respectively), as shown by other work in which the candidate was involved[281]. However, cured samples were more likely to be positive at 6 months with rK39 than with lysate (31.6-36.8% versus 24.0%), indicating the need for further development of the rK39 IgG1 RDT. Overall, the study showed that a simple adaptation, made to an existing diagnostic, could dramatically increase its capability.

Such an RDT, indicating cure or potential to relapse, would have profound impacts on VL disease control as it would facilitate targeted follow-up of individuals whose IgG1 remained elevated after treatment. However, earlier studies, and that presented here using paired sera from the same individuals before and after treatment, relied on a crude lysate antigen which required culturing of *L. donovani* in contained laboratory conditions and was not able to be standardised in its antigenic content. Therefore for large-scale disease control programs, it is preferable to use a defined antigen, such as rK39 or rK28, which can be produced in bulk as recombinant proteins and have no batch variation.

IgG1 western blots

We performed western blots with sera from a range of VL disease states to identify differences between VL and relapse IgG1 responses to *L. donovani* lysate. Certain bands were immunodominant across all clinical disease categories and were followed-up as potential sources of IgG1 seroantigens in the enclosed manuscript by Hinckel *et al*. It was not easily possible to correlate bands between blot strips due to slight differences in the gel separation. An electronic scanner could have assisted with matching bands. Distinguishing relapse from PKDL would be beneficial, particularly if a serological response could be used in advance to predict progression to either relapse or PKDL. The outcome we sought was the identification of individual proteins from these bands with a view to standardising the IgG1 assay to detect relapse or PKDL as distinct from each other and other VL states.

Mass spectrometry

Mass spectrometry identified several hundred proteins in each gel band, therefore it was not clear which ones were immunoreactive or immuodominant without further analysis. Computational filtering criteria were applied to these data to restrict the number of possible antigens. Alternative methods of revealing specific antigens in the crude lysate antigen include: 2-dimensional SDS PAGE to obtain better protein resolution; excising the gel region (unstained bands) of interest and re-separating them individually on an SDS PAGE gel of appropriate acrylamide percentage; probing a peptide library such as a phage display library, with sera or IgG1 from the different disease states.

Future development and use of IgG1 assays

Considering the performance of the IgG1 assays, the initial investigations detailed in this thesis relied on having paired serum/plasma samples available to observe a decrease in IgG1 titre between pre- and post-treatment. In a clinical situation, such samples may not be available. In addition, developing an assay that is required to be positive with active disease but negative 6 months later (in the case of India), is challenging. Existing immunoassays based on rK39 or rK28 with detection of total IgG have high sensitivity and specificity for active VL,

therefore, future development of the IgG1 RDT may be more appropriately focused on obtaining a negative result at 6 months follow-up to confirm cure. This single test would halve the cost per patient (compared with doing IgG1 tests before and after treatment). In addition it may require a lower antigen concentration to reduce sensitivity, such that it would only be positive with high titre sera, indicative of relapse. In a clinical context, the single time-point RDT would also avoid the need to train healthcare workers in interpreting the test line intensity of the RDT, which is subjective and goes against WHO efforts to standardise the interpretation of such tests as either positive or negative[282].

The RDT could be used with, or be replaced by, the quantitative IgG1 ELISA if paired samples were available or quantitative information was needed, for example in research. In the case of the IgG1 ELISA, the decline of titre could be quantified e.g., by a ratio of pre- to post-treatment results, in order to indicate cure without a sample becoming negative. This approach has been taken in serology for strongyloidiasis[67, 283]. A suitable ratio could be defined after analysis of a larger number of paired samples from known cure and relapse cases.

7.2.2. Urine antigen

Success of this study

This work successfully identified seven *L. donovani* parasite proteins in VL urine from India and Sudan, none of which had previously been found in urine. Analysis of these proteins for potential epitope regions indicated 22 with high specificity to *Leishmania*, that can be combined to increase the sensitivity and specificity of a subsequent assay. Several of the proteins identified in VL urine had a hypothetical identity due to incomplete annotation of the *L. donovani* genome. Better characterisation of these proteins could be useful to inform antigen production and assay development with these peptides in order determine the structure and surface exposure of predicted epitopes. Multiple online tools exist for protein identification and characterisation, as used by Shahbaaz *et al.* (2013)[284] to assign functions to over 70% of the previously hypothetical proteins of the bacterium, *Haemophilus influenzae*.

Potential for urine antigen diagnostics

In other studies, urine antigens have been found to decline in response to treatment, potentially offering more rapid monitoring than is possible with antibody detection[260]. Another potential utility of urine antigen diagnostics is in HIV-VL co-infection where urine antigen, as measured by the KAtex, indicated that those with a high antigenuria at diagnosis were significantly more likely to experience treatment failure than those with lower antigen levels. In addition, a high urine antigen result at follow-up, when deemed cured, indicated

higher probability of relapse[285, 286]. Thus, adaptation of a urine antigen capture assay to RDT format could be of significant benefit in VL endemic regions with high HIV coinfection.

Methods for urine antigen discovery

To identify parasite antigens in urine, mass spectrometry is less sensitive than immunoassay, which is why we combined these techniques, to first concentrate and then identify proteins. Alternative methods have included the use of nanoparticles to absorb small molecules including antigens, from urine, which after washing away the large urine volume, are released from the particles and immunoassayed. This technique has been applied to Chagas disease and Lyme borreliosis[287-289].

The alternative method, excising immunoreactive bands from a urine protein SDS-PAGE, had limitations as it involved 'wasting' the protein that was used to make the blot, and inefficient extraction of proteins from the gel. Alternatively, immunoreactive bands could be directly excised from a western blot, thus avoiding the need for a duplicate gel. However, all components would need to be optimised as there are differences between this and a standard western blot[290]. A higher throughput approach to urine biomarker discovery was presented by Balog *et al.* (2009)[291] in which urine protein was captured and fractionated from small volumes prior to mass spectrometry, yielding high sensitivity for low abundance peptides.

Mass spectrometry data analysis

Our simultaneous search of the mass spectra against two protein databases took into account the fact that a mass spectrum may have had a better match to a human protein peptide than to one from *L. donovani*. A more rigorous approach would have been to compile a database of human pathogens and commensals, particularly as VL patients may suffer from bacterial co-infections, thus the mass spectra could have been searched against these to exclude possible matches to peptides from other organisms. An additional process to strengthen our mass spectrometry data would have been to analyse EHC urine alongside the VL urine. Being able to make this comparison between VL and EHC samples would strengthen the identification of parasite peptides as any that apparently occurred in EHC urine could be excluded as false matches.

7.2.3. Strengths

This thesis describes the first RDT with the potential to indicate cure and relapse of VL, and to assist with diagnosis of PKDL. In addition, we showed that the already established serological assays using rK39 antigen could be adapted for determination of cure by incorporating anti-

IgG1 instead of anti-IgG. Separately, we identified *L. donovani* proteins in VL urine that could contribute to non-invasive diagnosis and test-of-cure development.

- Paired pre- and post-treatment samples showed decline of IgG1 at 6 months where relapse and PKDL cases were more likely to be positive.
- Assays worked equally well with strains of *L. donovani* from Indian and eastern African origins.
- Western blots followed by mass spectrometry allowed the most commonly recognised antigens to be identified and analysed further.
- Several potential urine antigens were identified by our immuno-panning method. The use of dishes coated at different antibody concentrations took into account possible steric hindrance- when antigens may not bind due to spatial molecular interactions. In addition, different urine concentration methods all resulted in parasite peptide identification.
- Criteria applied to mass spectrometry data ensured that only reliable peptide identifications resulted.

7.2.4. Limitations

For determination of the true utility of IgG1 in distinguishing cure from relapse and PKDL, and predicting progression to VL from asymptomatic infection, we would need access to a well-characterised panel of samples from a cohort that had ideally been followed actively over several years.

- There was limited follow-up of participants deemed cured in both India and Sudan that reduced our ability to measure the reliability of elevated IgG1 at 6 months in India and sooner in Sudan.
- Limited number of samples from other diseases involved in differential diagnosis such as infectious diseases causing fever. These would be needed to determine specificity of the diagnostic test if using lysate, but with rK39 or rK28 this would not be such an issue.
- Lack of sequential samples to determine the ability of IgG1 to predict VL relapse and PKDL, rather than confirm it only at the time of clinical presentation.

- Sensitivity of the lysate RDT for active VL could be improved by use of the rK39/rK28 instead, or focusing on a single 6-month measurement to remove this need.
- For urine antigen identification, it may be considered a limitation that none of the previously reported *Leishmania* proteins were identified and that there were no biological replicates.
- More generally, assessment of diagnostic tests is likely to give a more accurate measure of specificity if the samples are not pre-selected from defined groups of disease states i.e. cases and controls, which creates 'selection bias'[246]. Therefore, specificity should be re-evaluated in future studies using samples of unknown status.

7.2.5. Further work

IgG1 serology

The next step in validating IgG1 as a diagnostic tool for follow-up of VL would be its use on a large sample set from a cohort study. Specifically, there is a need to follow individuals for at least 2 years after treatment, to validate their cured status.

For rK39 as an antigen in IgG1-based diagnostics, the concentration of antigen on the rapid test requires optimising to maximise performance while minimising cost. This should be carried to on well-characterised samples from individuals followed for sufficient time to be considered definitively cured versus those who relapse.

In addition, optimisation of the RDT as a single-timepoint measure of cure e.g. at 6 months post-treatment, could allow it to use less antigen as only higher titres would be relevant to indicate relapse. This would remove the need for the test to detect pre-treatment VL, and would also reduce the cost per test.

The capacity of the RDT to indicate progression of asymptomatic seropositivity to VL could also be investigated using a larger number of samples of known clinical status.

If there are antigens that differ in IgG1 reactivity between VL, relapse and PKDL, these could be discovered by increasing the resolution of SDS-PAGE using 2-dimensional gel electrophoresis, peptide microarray[292] or genomic library screening.

An alternative version of the Ig1 RDT, with rK28 antigen, could be explored in order to increase sensitivity and specificity of the test, particularly in Africa. In addition, the IgG1 response should be investigated from the end of treatment (or even before) and the test

optimised to determine treatment outcome at this point to avoid loss-to-follow-up of patients.

Research into VL predictive and test-of-cure diagnostics could investigate the use of algorithms including levels other markers as well as IgG1. These markers could include IL-10 that has been linked to future development of PKDL in Sudan, and has been seen to decline in cure in Ethiopia[293, 294].

VL urine antigen detection

Development of a VL urine antigen assay from the peptides detailed in the enclosed manuscript would involve synthesis of those peptides and raising antibodies to them. This could initially be done with several peptides at a time, to create a polyclonal antiserum against 3 or more of the antigens. IgG could then be purified from the antiserum and used to coat ELISA plates. Urine of known VL status (VL disease and healthy controls) should be assayed at a low dilution e.g., neat or 1 in 2.

Additional development in this field could be computational analysis of the *Leishmania* proteins found by other studies, including tryparedoxin, nuclear transport factor 2 and iron speroxide dismutase[258] to discover specific epitopes that could be used to generate monoclonal or targeted polyclonal antibodies.

7.3 Conclusions

7.3.1. Summary of my findings

Strongyloidiasis

This study applied computational analyses to open access 'omic' data and identified multiple novel candidate coproantigens of *S. stercoralis.* These antigens were measured against three criteria: presence in host stool, specificity to *Strongyloides*; antigenicity. As such, the antigens have multiple types of evidence supporting their selection and should be tested as the targets of a rapid diagnostic test to indicate active infection. Ongoing developments are the synthesis of the short peptide candidate antigens for screening and raising antibodies.

Visceral leishmaniasis

For the neglected disease VL, this assessment of IgG1 serological assays provided additional evidence that high IgG1 is a marker for relapse after treatment, where low IgG1 could indicate cure, and that this could be measured with an RDT. Use of the existing seroantigen rK39 in a pilot RDT detecting IgG1, had promising results and removed the need for parasite culturing to produce the antigen.

Discovery of seven *L. donovani* proteins was made from urine of Indian and Sudanese VL patients using immunocapture and mass spectrometry methods. Computational analysis of these proteins revealed *Leishmania*-specific epitope regions that can be used to develop an antigen capture assay.

7.3.2. What's the need?

Both strongyloidiasis and VL are in need of diagnostics to indicate cure, as well as accurate primary diagnosis that is effective in HIV co-infected individuals. Both are potentially fatal diseases therefore there is a requirement for high sensitivity diagnostics. Strongyloidiasis in particular, lacks an agreed diagnostic method, leading to difficulty in comparing the results of prevalence surveys and assessments of cure.

7.3.3. Target product profiles and disease control

A combination of diagnostic tools is needed during the course of a disease control program, dependent on the goals of the program; morbidity control versus elimination/eradication. A target product profile (TPP) defines the characteristics of a product. In the case of diagnostics, the requirements of a product change over the course of a control program, depending on the decisions they support i.e., when to commence, continue, extend or end MDA, and to verify break in transmission. Lim *et al.* (2018)[295] described the TPP of STH diagnostics at these various stages of an MDA-based control program.

As there is currently no MDA specifically for strongyloidiasis control, a diagnostic would be used under the first 'use case'. It would ideally quantify the worm burden in those affected and could use stool, alongside STH surveys[295]. Results of such screening would inform the decision of which groups should receive MDA as well as its frequency and duration, and the frequency of follow-up testing. False positive results (low specificity of the test) in this scenario could lead to over-treatment. In the case of ivermectin this would be largely safe but could promote drug resistance in the worm. Low sensitivity could lead to underestimation of the true burden of the infection, as is currently the case, therefore a novel RDT should have high sensitivity to improve on the current diagnostic tools. During MDA, when prevalence is expected to be lower, a test must have very high sensitivity and be specific for active infection. This use case in particular, evaluating the progress of MDA, would favour antigen detection over serology for strongyloidiasis.

In the case of VL in India, Bangladesh and Nepal, where elimination is the target, diagnosis of VL is effective with the highly sensitive rK39 serological assay but MDA is impossible due to the toxicity of the drugs. Therefore, high specificity is a key requirement of VL diagnostics and is likely to continue to rely on aspirate microscopy. However, towards elimination or interruption of transmission, when prevalence is very low, confirmation of cure and the diagnosis of asymptomatic transmitters and PKDL cases becomes much more relevant. This is where the IgG1 assays could be integrated into the control program. A test to indicate cure would also enable more effective treatment regimens to be validated for the prevention of PKDL.

On an individual patient basis, both an *S. stercoralis* coproantigen assay and a VL test-of-cure would be beneficial to enable clinicians to make prompt and accurate treatment decisions. As RDTs, and with further development, these tests have the potential to fulfil the ASSURED[296] criteria of being:

- **A**ffordable by those at risk of infection
- **S**ensitive (few false negatives)
- **S**pecific (few false positives)
- **U**ser-friendly (simple to perform and requiring minimal training)
- **R**apid and **R**obust (to enable treatment at first visit, and not requiring refrigeration)
- **E**quipment-free
- **D**elivered to those who need it

Use and results of RDTs can be linked within a health system by mobile phone technology, termed 'mHealth'. For VL this could be implemented through healthcare workers, or the

population themselves, if phone ownership was common. For example, text or voice messages could prompt individuals to report symptoms that might indicate VL relapse or PKDL, to present to a clinic or be visited in their home, and tested with an RDT, the result of which would be reported to a data system to track real-time disease trends. Early warning systems for disease surveillance, particularly influenza and Dengue, have been tested based on internet searches[297]. However, these systems would have to be adapted to work for diseases like VL and other NTDs which frequently affect only the poorest and therefore, those least likely to own a mobile phone[298].

In 2018 the WHO launched the Essential Diagnostics List (EDL) to complement the Essential Medicines List with recommended *in vitro* diagnostics needed by a health system, along with advice to health ministries on the quality and implementation of these tests[299]. The initial list includes diagnostics for a range of non-infectious health conditions as well as prioritised infectious diseases such as HIV, TB, malaria and hepatitis viruses. This list will be updated annually and applications to include new diagnostics can be made. The impact of the EDL will hopefully be to increase the prominence of diagnostics equal to that of treatments, and contribute to the United Nations Sustainable Development Goal 3 related to health[300]. For strongyloidiasis in particular, the growing body of epidemiological data and diagnostic test assessments may facilitate diagnostics for this disease being added to the EDL, and thus becoming more widely available and used, as well as favouring a standardised method. For VL, the well-used rK39- and rK28-based serological tests would ideally be available in regions where the disease is highly endemic, including Brazil, eastern Africa and parts of India, thus a validated test-of-cure, predictor of relapse and confirmation of PKDL, would also be beneficial in these regions.

7.3.4. Potential impact of these diagnostics on disease control

Rapid diagnostic tests have revolutionised the care of individuals and have enabled large-scale control programs, including malaria and VL elimination campaigns. Antigen detection assays have great potential to fill diagnostic gaps left by serological tests, namely: accurate diagnosis of HIV positive patients who may have reduced serological responses; timely test of cure; indication/prediction of relapse; indication of active infection as opposed to previous exposure. These features facilitate the testing of new treatments and screening for antimicrobial resistance, which are held back by an absence of test-of-cure diagnostics. Antigen detection tests have the potential to be non-invasive, are generally cheaper than molecular tests, and can be unaffected by vaccine-induced immunity.

Strongyloides antigen detection

The impact of a stool antigen detection test on strongyloidiasis would be to increase the breadth and detail of epidemiological information, which in turn could support the implementation of MDA with ivermectin for this disease, as has been called for[43]. Such a test could easily be integrated into STH surveillance programmes when stool is collected for Kato Katz microscopy but is not tested for *Strongyloides* due to the logistical demands of the existing methods. As an indicator of active infection, an antigen detection assay would enable the true cure rates of existing and new drugs to be defined. It may also be useful as a tool to investigate the role of dogs as zoonotic reservoirs of *S. stercoralis*. At the clinical level, availability of an RDT would increase the profile of strongyloidiasis among clinicians when faced with cases of unexplained eosinophilia and prior to corticosteroid therapy or transplant, thus reducing the likelihood of unnecessary morbidity and mortality from the disease. In order to validate and implement a strongyloidiasis rapid test, partnerships must be created with study sites, industry, and control programmes. Diagnostic test accuracy must be assessed before and after treatment using a combination of diagnostic methods as the reference standard.

VL serology and antigen detection

As a serological test-of-cure or indicator of VL relapse, the IgG1 assays described here could be integrated into the existing health system in India where patients treated for VL return to the clinic for a follow-up 6 months later. Currently they are deemed cured if all clinical symptoms have resolved, however, some 10-20% go on to relapse and require re-treating. If the test indicated elevated levels of IgG1 at 6 months, that individual could be warned of their chance of relapse to avoid delay in the subsequent diagnosis, especially relevant in low-prevalence, disease surveillance, settings. With further validation, IgG1 assays may be applicable sooner than 6 months in Sudan. The *Leishmania* proteins identified in VL urine can be tested as the targets of a diagnostic test. Once validated, a highly accurate urine antigen test could replace blood collection for initial diagnosis of VL. Both types of test, IgG1 and urine antigen, could contribute to the diagnostic tools available for the long term control or elimination of VL.

8. References

1. Burbelo, P.D., Goldman, R., and Mattson, T.L., A simplified immunoprecipitation method for quantitatively measuring antibody responses in clinical sera samples by using mammalian-produced *Renilla* luciferase-antigen fusion proteins. BMC Biotechnol, 2005. **5**: p. 22.
2. Medgadget. Glow-in-the-dark paper test rapidly detects infectious diseases. 2018 [cited 2019 26/1/19]; Available from: https://www.medgadget.com/2018/10/glow-in-the-dark-paper-test-rapidly-detects-infectious-diseases.html.
3. Medgadget. Diagnostic blood tests in minutes: interview with Brianna Wronko, founder and CEO of Group K Diagnostics. 2019 [cited 2019 26/1/19]; Available from: https://www.medgadget.com/2019/01/diagnostic-blood-tests-in-minutes-interview-with-brianna-wronko-founder-and-ceo-of-group-k-diagnostics.html.
4. i-sense, the EPSRC IRC in early warning sensing systems for infectious diseases. 2017 [cited 2019 26/1/19]; Available from: https://www.i-sense.org.uk/.
5. Labes, E.M., et al., Genetic characterization of *Strongyloides* spp. from captive, semi-captive and wild Bornean orangutans (*Pongo pygmaeus*) in Central and East Kalimantan, Borneo, Indonesia. Parasitology, 2011. **138**(11): p. 1417-1422.
6. Hasegawa, H., et al., Molecular identification of the causative agent of human strongyloidiasis acquired in Tanzania: dispersal and diversity of *Strongyloides* spp. and their hosts. Parasitol Int, 2010. **59**(3): p. 407-13.
7. Nutman, T.B., Human infection with *Strongyloides stercoralis* and other related *Strongyloides* species. Parasitology, 2016: p. 1-11.
8. Neva, F.A., Biology and immunology of human strongyloidiasis. J Infect Dis, 1986. **153**(3): p. 397-406.
9. Seaman, J., Mercer, A.J., and Sondorp, E., The epidemic of visceral leishmaniasis in western Upper Nile, southern Sudan: course and impact from 1984 to 1994. Int J Epidemiol, 1996. **25**(4): p. 862-71.
10. Muniaraj, M., The lost hope of elimination of kala-azar (visceral leishmaniasis) by 2010 and cyclic occurrence of its outbreak in India, blame falls on vector control practices or co-infection with human immunodeficiency virus or therapeutic modalities? Trop Parasitol, 2014. **4**(1): p. 10-9.
11. Boelaert, M., et al., The poorest of the poor: a poverty appraisal of households affected by visceral leishmaniasis in Bihar, India. Trop Med Int Health, 2009. **14**(6): p. 639-44.
12. Harhay, M.O., et al., Who is a typical patient with visceral leishmaniasis? Characterizing the demographic and nutritional profile of patients in Brazil, East Africa, and South Asia. Am J Trop Med Hyg, 2011. **84**(4): p. 543-50.
13. Gilead Sciences Ltd. Press release: Gilead Sciences and the World Health Organization announce five-year visceral leishmaniasis collaboration. 2016 [cited 2018; Available from: http://www.gilead.com/news/press-releases/2016/9/gilead-sciences-and-the-world-health-organization-announce-fiveyear-visceral-leishmaniasis-collaboration.
14. National Vector Borne Disease Control Programme, M.o.H.a.F.W., Government of India,, Accelerated plan for kala-azar elimination 2017. 2017, Ministry of Health and Family Welfare, Government of India: India.
15. Singh, O.P., et al., Asymptomatic *Leishmania* infection: a new challenge for *Leishmania* control. Clin Infect Dis, 2014. **58**(10): p. 1424-9.
16. Burza, S., et al., Risk factors for visceral leishmaniasis relapse in immunocompetent patients following treatment with 20 mg/kg liposomal amphotericin B (AmBisome) in Bihar, India. PLoS Negl Trop Dis, 2014. **8**(1): p. e2536.
17. Zijlstra, E.E., et al., Post-kala-azar dermal leishmaniasis in the Indian subcontinent: a threat to the South-East Asia region kala-azar elimination programme. PLoS Negl Trop Dis, 2017. **11**(11): p. e0005877.

18. Patole, S., Burza, S., and Varghese, G.M., Multiple relapses of visceral leishmaniasis in a patient with HIV in India: a treatment challenge. Int J Infect Dis, 2014. **25**: p. 204-6.
19. Tefé-Silva, C., et al., Hyperinfection syndrome in strongyloidiasis, in Curr Top Trop Med. 2012.
20. Stoltzfus, J.D., et al., RNAseq analysis of the parasitic nematode *Strongyloides stercoralis* reveals divergent regulation of canonical dauer pathways. PLoS Negl Trop Dis, 2012. **6**(10): p. e1854.
21. Becker, S.L., et al., Diagnosis, clinical features, and self-reported morbidity of *Strongyloides stercoralis* and hookworm infection in a co-endemic setting. PLoS Negl Trop Dis, 2011. **5**(8): p. e1292.
22. Schar, F., et al., *Strongyloides stercoralis*: global distribution and risk factors. PLoS Negl Trop Dis, 2013. **7**(7): p. e2288.
23. Spinicci, M., et al., Seroepidemiological trend of strongyloidiasis in the Bolivian Chaco (1987-2013) in the absence of disease-specific control measures. Trop Med Int Health, 2017.
24. Buonfrate, D., et al., Epidemiology of *Strongyloides stercoralis* in northern Italy: results of a multicentre case-control study, February 2013 to July 2014. Euro Surveill, 2016. **21**(31).
25. Viney, M.E. and Lok, J.B. (2007) *Strongyloides* spp. WormBook,
26. Grant, W.N., et al., *Parastrongyloides trichosuri*, a nematode parasite of mammals that is uniquely suited to genetic analysis. Int J Parasitol, 2006. **36**(4): p. 453-66.
27. Ramachandran, S., Gam, A.A., and Neva, F.A., Molecular differences between several species of *Strongyloides* and comparison of selected isolates of *S. stercoralis* using a polymerase chain reaction-linked restriction fragment length polymorphism approach. Am J Trop Med Hyg, 1997. **56**(1): p. 61-5.
28. Hasegawa, H., et al., *Strongyloides* infections of humans and great apes in Dzanga-Sangha protected areas, Central African Republic and in degraded forest fragments in Bulindi, Uganda. Parasitol Int, 2016.
29. Ashford, R.W., Barnish, G., and Viney, M.E., *Strongyloides fuelleborni kellyi*: infection and disease in Papua New Guinea. Parasitol Today, 1992. **8**(9): p. 314-8.
30. Shield, J.M. and Kow, F., A comparative study of intestinal helminths in pre-school-age urban and rural children in Morobe Province, Papua New Guinea. P N G Med J, 2013. **56**(1-2): p. 14-31.
31. Dorris, M., Viney, M.E., and Blaxter, M.L., Molecular phylogenetic analysis of the genus *Strongyloides* and related nematodes. Int J Parasitol, 2002. **32**(12): p. 1507-17.
32. Centre for Tropical Diseases. Agar plate copro-culture method. [cited 2018 26/8/18]; Available from: http://www.tropicalmed.eu/Page/WebObjects/PageTropE.woa/wa/displayPage?name=ReadingCultureAgarMicro.
33. Hunt, V.L., et al., The genomic basis of parasitism in the *Strongyloides* clade of nematodes. Nat Genet, 2016.
34. Hunt, V.L., et al., The genome of *Strongyloides* spp. gives insights into protein families with a putative role in nematode parasitism. Parasitology, 2016: p. 1-16.
35. Lok, J.B., *Strongyloides stercoralis*: a model for translational research on parasitic nematode biology, in WormBook, T.*C.e.*R. Community, Editor. 2007.
36. Schad, G.A., Hellman, M.E., and Muncey, D.W., *Strongyloides stercoralis*: hyperinfection in immunosuppressed dogs. Exp Parasitol, 1984. **57**(3): p. 287-296.
37. Streit, A., Genetics: modes of reproduction and genetic analysis. Parasitology, 2016: p. 1-11.
38. Yamada, M., et al., Species-specific differences in heterogonic development of serially transferred free-living generations of *Strongyloides planiceps* and *Strongyloides stercoralis*. J Parasitol, 1991. **77**(4): p. 592-4.

39. Sudarshi, S., et al., Clinical presentation and diagnostic sensitivity of laboratory tests for *Strongyloides stercoralis* in travellers compared with immigrants in a non-endemic country. Trop Med Int Health, 2003. **8**(8): p. 728-732.
40. Nabeya, D., et al., Pulmonary strongyloidiasis: assessment between manifestation and radiological findings in 16 severe strongyloidiasis cases. BMC Infect Dis, 2017. **17**(1): p. 320.
41. Zaghlool, D.A., et al., A case of fatal gastrointestinal haemorrhage due to hyperinfection with *Strongyloides stercoralis*. J Parasit Dis, 2016. **40**(4): p. 1347-1350.
42. Buonfrate, D., et al., Severe strongyloidiasis: a systematic review of case reports. BMC Infect Dis, 2013. **13**: p. 78.
43. Bisoffi, Z., et al., *Strongyloides stercoralis*: a plea for action. PLoS Negl Trop Dis, 2013. **7**(5): p. e2214.
44. Albonico, M., et al., StrongNet: an international network to improve diagnostics and access to treatment for strongyloidiasis control. PLoS Negl Trop Dis, 2016. **10**(9): p. e0004898.
45. Maassen, W., et al., Microbiological screenings for infection control in unaccompanied minor refugees: the German Armed Forces Medical Service's experience. Mil Med Res, 2017. **4**: p. 13.
46. Ostera, G., et al., Strongyloidiasis in Latin American immigrants: a pilot study. J Helminthol, 2016. **FirstView**: p. 1-5.
47. Begona, M.-M., et al., Communitarian screening campaign for *Strongyloides stercoralis* among Latin American immigrants in Spain. Clin Microbiol Infect, 2018.
48. Cabezas-Fernandez, M.T., et al., Strongyloidiasis in immigrants in southern Spain. Enferm Infecc Microbiol Clin, 2015. **33**(1): p. 37-9.
49. ten Hove, R.J., et al., Molecular diagnostics of intestinal parasites in returning travellers. Eur J Clin Microbiol Infect Dis, 2009. **28**(9): p. 1045-53.
50. Barrett, J., et al., The changing aetiology of eosinophilia in migrants and returning travellers in the Hospital for Tropical Diseases, London 2002-2015: an observational study. J Infect, 2017.
51. Magnaval, J.F., et al., A retrospective study of autochthonous strongyloidiasis in Region Midi-Pyrenees (Southwestern France). Eur J Epidemiol, 2000. **16**(2): p. 179-82.
52. Russell, E.S., et al., Prevalence of *Strongyloides stercoralis* antibodies among a rural Appalachian population-Kentucky, 2013. Am J Trop Med Hyg, 2014. **91**(5): p. 1000-1.
53. Beknazarova, M., Whiley, H., and Ross, K., Strongyloidiasis: a disease of socioeconomic disadvantage. Int J Environ Res Public Health, 2016. **13**(5).
54. Khieu, V., et al., *Strongyloides stercoralis* infection and re-infection in a cohort of children in Cambodia. Parasitol Int, 2014. **63**(5): p. 708-12.
55. Khieu, V., et al., High prevalence and spatial distribution of *Strongyloides stercoralis* in rural Cambodia. PLoS Negl Trop Dis, 2014. **8**(6): p. e2854.
56. Forrer, A., et al., Ivermectin treatment and sanitation effectively reduce *Strongyloides stercoralis* infection risk in rural communities in Cambodia. PLoS Negl Trop Dis, 2016. **10**(8): p. e0004909.
57. Vargas, P., et al., Serologic monitoring of public health interventions against *Strongyloides stercoralis*. Am J Trop Med Hyg, 2017. **97**(1): p. 166-172.
58. Mobley, C.M., Dhala, A., and Ghobrial, R.M., *Strongyloides stercoralis* in solid organ transplantation: early diagnosis gets the worm. Curr Opin Organ Transplant, 2017. **22**(4): p. 336-344.
59. Galiano, A., et al., Donor-derived *Strongyloides stercoralis* hyperinfection syndrome after simultaneous kidney/pancreas transplant. Int J Infect Dis, 2016.
60. Carvalho, E.M. and Da Fonseca Porto, A., Epidemiological and clinical interaction between HTLV-1 and *Strongyloides stercoralis*. Parasite Immunol, 2004. **26**(11-12): p. 487-97.

61. Teixeira, M.C., et al., *Strongyloides stercoralis* infection in alcoholic patients. Biomed Res Int, 2016. **2016**: p. 4872473.
62. Abu Omar, M., et al., *Strongyloides* hyperinfection syndrome and VRE pneumonia. BMJ Case Rep, 2017. **2017**.
63. Myint, A., et al., *Strongyloides* hyperinfection syndrome in an immunocompetent host resulting in bandemia and death. BMJ Case Rep, 2017. **2017**.
64. Lam, C.S., et al., Disseminated strongyloidiasis: a retrospective study of clinical course and outcome. Eur J Clin Microbiol Infect Dis, 2006. **25**(1): p. 14-8.
65. Nutman, T.B., et al., Eosinophilia in Southeast Asian refugees: evaluation at a referral center. J Infect Dis, 1987. **155**(2): p. 309-13.
66. Osiro, S., et al., A case of *Strongyloides* hyperinfection syndrome in the setting of persistent eosinophilia but negative serology. Diagn Microbiol Infect Dis, 2017. **88**(2): p. 168-170.
67. Salvador, F., et al., Usefulness of *Strongyloides stercoralis* serology in the management of patients with eosinophilia. Am J Trop Med Hyg, 2014. **90**(5): p. 830-4.
68. Repetto, S.A., et al., High rate of strongyloidosis infection, out of endemic area, in patients with eosinophilia and without risk of exogenous reinfections. Am J Trop Med Hyg, 2010. **82**(6): p. 1088-93.
69. Rajamanickam, A., et al., Elevated systemic levels of eosinophil, neutrophil, and mast cell granular proteins in *Strongyloides stercoralis* infection that diminish following treatment. Front Immunol, 2018. **9**: p. 207.
70. Repetto, S.A., et al., Strongyloidiasis outside endemic areas: long-term parasitological and clinical follow-up after ivermectin treatment. Clin Infect Dis, 2018: p. cix1069-cix1069.
71. Ravi Acharya, K. and Ackerman, S., Eosinophil granule proteins: form and function. Vol. 289. 2014.
72. MacMicking, J., Xie, Q.W., and Nathan, C., Nitric oxide and macrophage function. Annu Rev Immunol, 1997. **15**: p. 323-50.
73. Montes, M., Sawhney, C., and Barros, N., *Strongyloides stercoralis*: there but not seen. Curr Opin Infect Dis, 2010. **23**(5): p. 500-4.
74. Vidarsson, G., Dekkers, G., and Rispens, T., IgG subclasses and allotypes: from structure to effector functions. Front Immunol, 2014. **5**: p. 520.
75. Quintana, J., Babayan, S., and Buck, A., Small RNAs and extracellular vesicles in filarial nematodes: from nematode development to diagnostics. Vol. 39. 2016.
76. Eberle, R., et al., Isolation, identification and functional profile of excretory-secretory peptides from *Onchocerca ochengi*. Acta Trop, 2015. **142**: p. 156-66.
77. Rebello, K.M., et al., Comprehensive proteomic profiling of adult *Angiostrongylus costaricensis*, a human parasitic nematode. J Proteomics, 2011. **74**(9): p. 1545-59.
78. Sotillo, J., et al., Secreted proteomes of different developmental stages of the gastrointestinal Nematode *Nippostrongylus brasiliensis*. Mol Cell Proteomics, 2014. **13**(10): p. 2736-2751.
79. Varatharajalu, R., et al., *Strongyloides stercoralis* excretory/secretory protein strongylastacin specifically recognized by IgE antibodies in infected human sera. Microbiol Immunol, 2011. **55**: p. 115-122.
80. World Health Organization. The top 10 causes of death. 2018 [cited; Available from: http://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death.
81. Kwan, J.L., et al., Seroepidemiology of helminths and the association with severe malaria among infants and young children in Tanzania. PLoS Negl Trop Dis, 2018. **12**(3): p. e0006345.
82. Nacher, M., et al., Intestinal helminth infections are associated with increased incidence of *Plasmodium falciparum* malaria in Thailand. J Parasitol, 2002. **88**(1): p. 55-8.

83. Stefani, A., et al., Interactions between gastrointestinal nematodes and malaria in a cohort of children in an amazonian village. J Trop Pediatr, 2017. **63**(2): p. 144-147.
84. Li, X.-X. and Zhou, X.-N., Co-infection of tuberculosis and parasitic diseases in humans: a systematic review. Parasit Vectors, 2013. **6**: p. 79-79.
85. Elias, D., Akuffo, H., and Britton, S., Helminthes could influence the outcome of vaccines against TB in the tropics. Parasite Immunol, 2006. **28**(10): p. 507-13.
86. Salvador, F., et al., Impact of helminth infection on the clinical and microbiological presentation of Chagas diseases in chronically infected patients. PLoS Negl Trop Dis, 2016. **10**(4): p. e0004663.
87. Azeredo-Coutinho, R.B., et al., Intestinal helminth coinfection is associated with mucosal lesions and poor response to therapy in American tegumentary leishmaniasis. Acta Trop, 2016. **154**: p. 42-9.
88. Blackwell, A.D., et al., Helminth infection, fecundity, and age of first pregnancy in women. Science, 2015. **350**(6263): p. 970-2.
89. Siegel, M.O. and Simon, G.L., Is human immunodeficiency virus infection a risk factor for *Strongyloides stercoralis* hyperinfection and dissemination. PLoS Negl Trop Dis, 2012. **6**(7): p. e1581.
90. Georas, S.N., Inhaled glucocorticoids, lymphocytes, and dendritic cells in asthma and obstructive lung diseases. Proc Am Thorac Soc, 2004. **1**(3): p. 215-21.
91. Smith, S.M. and Vale, W.W., The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. Dialogues Clin Neurosci, 2006. **8**(4): p. 383-395.
92. Luvira, V., Watthanakulpanich, D., and Pittisuttithum, P., Management of *Strongyloides stercoralis*: a puzzling parasite. Int Health, 2014. **6**(4): p. 273-81.
93. Wolstenholme, A.J. and Rogers, A.T., Glutamate-gated chloride channels and the mode of action of the avermectin/milbemycin anthelmintics. Parasitology, 2005. **131 Suppl**: p. S85-95.
94. Merck. The MECTIZAN donation program. [cited 2018 14/8/18]; Available from: http://www.merck.com/about/featured-stories/mectizan.html.
95. Barda, B., et al., Side benefits of mass drug administration for lymphatic filariasis on *Strongyloides stercoralis* prevalence on Pemba Island, Tanzania. Am J Trop Med Hyg, 2017.
96. Anselmi, M., et al., Mass administration of ivermectin for the elimination of onchocerciasis significantly reduced and maintained low the prevalence of *Strongyloides stercoralis* in Esmeraldas, Ecuador. PLoS Negl Trop Dis, 2015. **9**(11): p. e0004150.
97. Ménez, C., et al., Relative neurotoxicity of ivermectin and moxidectin in Mdr1ab (-/-) mice and effects on mammalian GABA(A) channel activity. PLoS Negl Trop Dis, 2012. **6**(11): p. e1883.
98. Wolstenholme, A., Glutamate-gated chloride channels. Vol. 287. 2012.
99. Ashraf, M., Gue, C.L., and Baddour, L.M., Strongyloidiasis refractory to treatment with ivermectin. Am J Med Sci, 1996. **311**(4): p. 178-179.
100. Verissimo, C.J., et al., Multidrug and multispecies resistance in sheep flocks from São Paulo state, Brazil. Vet Parasitol, 2012. **187**(1-2): p. 209-16.
101. Maroto, R., et al., First report of anthelmintic resistance in gastrointestinal nematodes of sheep from Costa Rica. Vet Med Int, 2011. **2011**: p. 145312.
102. Palmeirim, M.S., et al., Efficacy and safety of co-administered ivermectin plus albendazole for treating soil-transmitted helminths: a systematic review, meta-analysis and individual patient data analysis. PLoS Negl Trop Dis, 2018. **12**(4): p. e0006458.
103. Clarke, N.E., et al., Efficacy of anthelminthic drugs and drug combinations against soil-transmitted helminths: a systematic review and network meta-analysis. Clin Infect Dis, 2018.

104. Henriquez-Camacho, C., et al., Ivermectin versus albendazole or thiabendazole for *Strongyloides stercoralis* infection. Cochrane Database Syst Rev, 2016(1): p. Cd007745.
105. Ramanathan, R. and Nutman, T., *Strongyloides stercoralis* infection in the immunocompromised host. Curr Infect Dis Rep, 2008. **10**(2): p. 105-10.
106. Barda, B., et al., Efficacy of moxidectin versus ivermectin against *Strongyloides stercoralis* infections: a randomized controlled non-inferiority trial. Clin Infect Dis, 2017.
107. Weeks, J.C., et al., Sertraline, Paroxetine, and Chlorpromazine are rapidly acting anthelmintic drugs capable of clinical repurposing. Sci Rep, 2018. **8**(1): p. 975.
108. Ketzis, J.K., Limitations to the adoption of a standardized *Strongyloides stercoralis* diagnostic method: case study in the Caribbean. Acta Trop, 2017.
109. Luvira, V., et al., Comparative diagnosis of strongyloidiasis in immunocompromised patients. Am J Trop Med Hyg, 2016. **95**(2): p. 401-4.
110. Koga, K., et al., A modified agar plate method for detection of *Strongyloides stercoralis*. Am J Trop Med Hyg, 1991. **45**(4): p. 518-21.
111. Strkolcova, G., et al., The roundworm *Strongyloides stercoralis* in children, dogs, and soil inside and outside a segregated settlement in Eastern Slovakia: frequent but hardly detectable parasite. Parasitol Res, 2017. **116**(3): p. 891-900.
112. Baermann's larval technique, in Encyclopedia of Parasitology, H. Mehlhorn, Editor. 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 156-156.
113. Hirata, T., et al., Increased detection rate of *Strongyloides stercoralis* by repeated stool examinations using the agar plate culture method. Am J Trop Med Hyg, 2007. **77**(4): p. 683-4.
114. Buonfrate, D., et al., A retrospective study comparing agar plate culture, indirect immunofluorescence and real-time PCR for the diagnosis of *Strongyloides stercoralis* infection. Parasitology, 2017: p. 1-5.
115. Schar, F., et al., *Strongyloides stercoralis* larvae excretion patterns before and after treatment. Parasitology, 2014. **141**(7): p. 892-7.
116. Verweij, J.J., et al., Molecular diagnosis of *Strongyloides stercoralis* in faecal samples using real-time PCR. Trans R Soc Trop Med Hyg, 2009. **103**(4): p. 342-6.
117. Buonfrate, D., et al., Accuracy of molecular biology techniques for the diagnosis of *Strongyloides stercoralis* infection-A systematic review and meta-analysis. PLoS Negl Trop Dis, 2018. **12**(2): p. e0006229.
118. Barda, B., et al., Evaluation of two DNA extraction methods on the detection of *Strongyloides stercoralis* infection. J Clin Microbiol, 2018.
119. Schar, F., et al., Evaluation of real-time PCR for *Strongyloides stercoralis* and hookworm as diagnostic tool in asymptomatic schoolchildren in Cambodia. Acta Trop, 2013. **126**(2): p. 89-92.
120. Krolewiecki, A.J., et al., Transrenal DNA-based diagnosis of *Strongyloides stercoralis* (Grassi, 1879) infection: Bayesian latent class modeling of test accuracy. PLoS Negl Trop Dis, 2018. **12**(6): p. e0006550.
121. Goncalves, A.L., et al., Specific IgG and IgA to larvae, parthenogenetic females, and eggs of *Strongyloides venezuelensis* in the immunodiagnosis of human strongyloidiasis. Diagn Microbiol Infect Dis, 2012. **72**(1): p. 79-84.
122. Rodrigues, R.M., et al., IgG1, IgG4, and IgE antibody responses in human strongyloidiasis by ELISA using *Strongyloides ratti* saline extract as heterologous antigen. Parasitol Res, 2007. **101**(5): p. 1209-14.
123. Eamudomkarn, C., et al., Comparative evaluation of *Strongyloides ratti* and *S. stercoralis* larval antigen for diagnosis of strongyloidiasis in an endemic area of opisthorchiasis. Parasitol Res, 2015.
124. Bisoffi, Z., et al., Diagnostic accuracy of five serologic tests for *Strongyloides stercoralis* infection. PLoS Negl Trop Dis, 2014. **8**(1): p. e2640.

125. Rascoe, L.N., et al., Development of Ss-NIE-1 recombinant antigen based assays for immunodiagnosis of strongyloidiasis. PLoS Negl Trop Dis, 2015. **9**(4): p. e0003694.
126. Eamudomkarn, C., et al., Diagnostic performance of urinary IgG antibody detection: a novel approach for population screening of strongyloidiasis. PLoS One, 2018. **13**(7): p. e0192598.
127. Ruantip, S., et al., Accuracy of Urine and Serum Assays for the Diagnosis of Strongyloidiasis by Three Enzyme-Linked Immunosorbent Assay Protocols. Am J Trop Med Hyg, 2019. **100**(1): p. 127-129.
128. Soriano-Arandes, A., et al., Discordances between serology and culture for *Strongyloides* in an Ethiopian adopted child with multiple parasitic infections: a case report. Medicine (Baltimore), 2016. **95**(10): p. e3040.
129. Ravi, V., et al., Characterization of a recombinant immunodiagnostic antigen (NIE) from *Strongyloides stercoralis* L3-stage larvae. Mol Biochem Parasitol, 2002. **125**(1-2): p. 73-81.
130. Ravi, V., et al., *Strongyloides stercoralis* recombinant NIE antigen shares epitope with recombinant Ves v 5 and Pol a 5 allergens of insects. Am J Trop Med Hyg, 2005. **72**(5): p. 549-53.
131. Ramanathan, R., et al., A luciferase immunoprecipitation systems assay enhances the sensitivity and specificity of diagnosis of *Strongyloides stercoralis* infection. J Infect Dis, 2008. **198**(3): p. 444-51.
132. Bosqui, L.R., et al., Detection of parasite-specific IgG and IgA in paired serum and saliva samples for diagnosis of human strongyloidiasis in northern Parana state, Brazil. Acta Trop, 2015. **150**: p. 190-5.
133. Arifin, N., et al., Identification and preliminary evaluation of a novel recombinant protein for serodiagnosis of strongyloidiasis. Am J Trop Med Hyg, 2018.
134. Krolewiecki, A.J., et al., Improved diagnosis of *Strongyloides stercoralis* using recombinant antigen-based serologies in a community-wide study in northern Argentina. Clin Vaccine Immunol, 2010. **17**(10): p. 1624-30.
135. Buonfrate, D., et al., Accuracy of five serologic tests for the follow up of *Strongyloides stercoralis* infection. PLoS Negl Trop Dis, 2015. **9**(2): p. e0003491.
136. Mounsey, K., et al., Use of dried blood spots to define antibody response to the *Strongyloides stercoralis* recombinant antigen NIE. Acta Trop, 2014. **138**: p. 78-82.
137. Nabha, L., et al., Prevalence of *Strongyloides stercoralis* in an urban US AIDS cohort. Pathog Glob Health, 2012. **106**(4): p. 238-44.
138. Nageswaran, C., Craig, P.S., and Devaney, E., Coproantigen detection in rats experimentally infected with *Strongyloides ratti*. Parasitology, 1994. **108 (Pt 3)**: p. 335-42.
139. Goncalves, A.L., et al., A new faecal antigen detection system for *Strongyloides venezuelensis* diagnosis in immunosuppressed rats. Exp Parasitol, 2010. **125**(4): p. 338-41.
140. Sykes, A.M. and McCarthy, J.S., A coproantigen diagnostic test for *Strongyloides* infection. PLoS Negl Trop Dis, 2011. **5**(2): p. e955.
141. Goncalves, A.A.S., et al., Detection of immune complexes and evaluation of alcoholic individuals' serological profile in the diagnosis of strongyloidiasis. Parasitol Int, 2018.
142. Bosqui, L.R., et al., Immune complex detection in saliva samples: an innovative proposal for the diagnosis of human strongyloidiasis. Parasitology, 2017: p. 1-5.
143. Reitsma, J.B., et al., A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. J Clin Epidemiol, 2009. **62**(8): p. 797-806.
144. Knopp, S., et al., Diagnostic accuracy of Kato–Katz, FLOTAC, Baermann, and PCR methods for the detection of light-intensity hookworm and *Strongyloides stercoralis* infections in Tanzania. Am J Trop Med Hyg, 2014. **90**(3): p. 535-545.

145. Becker, S.L., et al., Real-time PCR for detection of *Strongyloides stercoralis* in human stool samples from Cote d'Ivoire: diagnostic accuracy, inter-laboratory comparison and patterns of hookworm co-infection. Acta Trop, 2015. **150**: p. 210-217.
146. Kristanti, H., et al., Diagnostic comparison of Baermann funnel, Koga agar plate culture and polymerase chain reaction for detection of human *Strongyloides stercoralis* infection in Maluku, Indonesia. Parasitol Res, 2018.
147. Fradejas, I., et al., Comparative study of two commercial tests for *Strongyloides stercoralis* serologic diagnosis. Trans R Soc Trop Med Hyg, 2018.
148. Levenhagen, M.A. and Costa-Cruz, J.M., Update on immunologic and molecular diagnosis of human strongyloidiasis. Acta Trop, 2014. **135c**: p. 33-43.
149. Real, F., et al., The genome sequence of *Leishmania (Leishmania) amazonensis*: functional annotation and extended analysis of gene models. DNA Res, 2013. **20**(6): p. 567-81.
150. Leishman, W.B., On the possibility of the occurrence of trypanosomiasis in India. Indian J Med Res, 1903. **123**(3): p. 1252-4; discussion 79.
151. Donovan, C., On the possibility of the occurence of trypanosomiasis in India. 1903.
152. Neave, S., "*Leishmania donovani*" in the Soudan. BMJ, 1904. **1**(2265): p. 1252-1252.
153. Lukeš, J., et al., Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. Proc Natl Acad Sci U S A, 2007. **104**(22): p. 9375.
154. Mauricio, I.L., Stothard, J.R., and Miles, M.A., The strange case of *Leishmania chagasi*. Parasitol Today, 2000. **16**(5): p. 188-9.
155. Harhay, M.O., et al., Urban parasitology: visceral leishmaniasis in Brazil. Trends Parasitol, 2011. **27**(9): p. 403-409.
156. Arango Duque, G. and Descoteaux, A., *Leishmania* survival in the macrophage: where the ends justify the means. Curr Opin Microbiol, 2015. **26**: p. 32-40.
157. Besteiro, S., et al., Protein turnover and differentiation in *Leishmania*. Int J Parasitol, 2007. **37**(10): p. 1063-1075.
158. Bern, C., et al., The epidemiology of visceral leishmaniasis and asymptomatic leishmanial infection in a highly endemic Bangladeshi village. Am J Trop Med Hyg, 2007. **76**(5): p. 909-14.
159. Hasker, E., et al., Latent infection with *Leishmania donovani* in highly endemic villages in Bihar, India. PLoS Negl Trop Dis, 2013. **7**(2): p. e2053.
160. Hirve, S., et al., Transmission dynamics of visceral leishmaniasis in the Indian subcontinent - a systematic literature review. PLoS Negl Trop Dis, 2016. **10**(8): p. e0004896.
161. Costa, C.H., et al., Asymptomatic human carriers of *Leishmania chagasi*. Am J Trop Med Hyg, 2002. **66**(4): p. 334-7.
162. le Fichoux, Y., et al., Occurrence of *Leishmania infantum* parasitemia in asymptomatic blood donors living in an area of endemicity in southern France. J Clin Microbiol, 1999. **37**(6): p. 1953-7.
163. Srivastava, P., et al., Molecular and serological markers of *Leishmania donovani* infection in healthy individuals from endemic areas of Bihar, India. Trop Med Int Health, 2013. **18**(5): p. 548-54.
164. Bhattarai, N.R., et al., PCR and direct agglutination as Leishmania infection markers among healthy Nepalese subjects living in areas endemic for kala-azar. Trop Med Int Health, 2009. **14**(4): p. 404-11.
165. Kirstein, O.D., et al., Minimally invasive microbiopsies: a novel sampling method for identifying asymptomatic, potentially infectious carriers of Leishmania donovani. Int J Parasitol, 2017.
166. Sarker, C.B., et al., Clinical profile of kala-azar in adults: as seen in Mymensingh Medical College Hospital, Mymensingh, Bangladesh. Mymensingh Med J, 2003. **12**(1): p. 41-4.

167. Dunning, N., Leishmania vaccines: from leishmanization to the era of DNA technology. Biosci Horizons, 2009. **2**(1): p. 73-82.
168. Oliveira, M.J., et al., Risk factors for acute kidney injury in visceral leishmaniasis (kala-azar). Am J Trop Med Hyg, 2010. **82**(3): p. 449-53.
169. Hasker, E., et al., Strong association between serological status and probability of progression to clinical visceral leishmaniasis in prospective cohort studies in India and Nepal. PLoS Negl Trop Dis, 2014. **8**(1): p. e2657.
170. Cota, G.F., et al., Exploring prognosis in chronic relapsing visceral leishmaniasis among HIV-infected patients: circulating *Leishmania* DNA. Acta Trop, 2017. **172**: p. 186-191.
171. Alvar, J., et al., The relationship between leishmaniasis and AIDS: the second 10 years. Clin Microbiol Rev, 2008. **21**(2): p. 334-59, table of contents.
172. Maciel, B.L., et al., Association of nutritional status with the response to infection with *Leishmania chagasi*. Am J Trop Med Hyg, 2008. **79**(4): p. 591-8.
173. Rink, L. and Gabriel, P., Zinc and the immune system. Proc Nutr Soc, 2000. **59**(4): p. 541-52.
174. Najim, R.A., Sharquie, K.E., and Farjou, I.B., Zinc sulphate in the treatment of cutaneous leishmaniasis: an in vitro and animal study. Mem Inst Oswaldo Cruz, 1998. **93**(6): p. 831-7.
175. Maciel, B.L., et al., Dual immune modulatory effect of vitamin A in human visceral leishmaniasis. PLoS One, 2014. **9**(9): p. e107564.
176. Custodio, E., et al., Nutritional supplements for patients being treated for active visceral leishmaniasis. Cochrane Database Syst Rev, 2018. **3**: p. Cd012261.
177. Mukhopadhyay, D., et al., Post kala-azar dermal leishmaniasis: an unresolved mystery. Trends Parasitol, 2014. **30**(2): p. 65-74.
178. World Health Organization, Leishmaniasis in high-burden countries: an epidemiological update based on data reported in 2014. Wkly Epidemiol Rec, 2016. **91**(22): p. 287-96.
179. Alvar, J., et al., Leishmaniasis worldwide and global estimates of its incidence. PLoS One, 2012. **7**(5): p. e35671.
180. World Health Organization. Status of endemicity of visceral leishmaniasis worldwide, 2016. Global Health Observatory Map Gallery 2018 [cited 2018 25/11/18]; Available from: http://gamapserver.who.int/mapLibrary/Files/Maps/Leish_VL_2016.png.
181. Bern, C., Courtenay, O., and Alvar, J., Of cattle, sand flies and men: a systematic review of risk factor analyses for South Asian visceral leishmaniasis and implications for elimination. PLoS Negl Trop Dis, 2010. **4**(2): p. e599.
182. Desjeux, P., et al., Report of the Post Kala-azar Dermal Leishmaniasis (PKDL) Consortium Meeting, New Delhi, India, 27-29 June 2012. Parasit Vectors, 2013. **6**: p. 196.
183. Dye, C. and Wolpert, D.M., Earthquakes, influenza and cycles of Indian kala-azar. Trans R Soc Trop Med Hyg, 1988. **82**(6): p. 843-50.
184. Islam, S., et al., Clinical and immunological aspects of post-kala-azar dermal leishmaniasis in Bangladesh. Am J Trop Med Hyg, 2013. **89**(2): p. 345-53.
185. Bern, C., et al., Loss of leishmanin skin test antigen sensitivity and potency in a longitudinal study of visceral leishmaniasis in Bangladesh. Am J Trop Med Hyg, 2006. **75**(4): p. 744-8.
186. Cameron, M.M., et al., Understanding the transmission dynamics of *Leishmania donovani* to provide robust evidence for interventions to eliminate visceral leishmaniasis in Bihar, India. Parasit Vectors, 2016. **9**: p. 25.
187. World Health Organization. Community-driven programme is key to defeating visceral leishmaniasis in Bangladesh. 2017 [cited 2018 22/09/18]; Available from: http://www.who.int/leishmaniasis/news/Defeating_visceral_leishmaniasis_in_Bangladesh/en/.

188. World Health Organization. Number of cases of visceral leishmaniasis reported, by country. Global Health Observatory Data Repository 2017 [cited 2018 15/9/18]; Available from: http://apps.who.int/gho/data/node.main.NTDLEISHVNUM?lang=en.
189. Mueller, Y.K., et al., Clinical epidemiology, diagnosis and treatment of visceral leishmaniasis in the Pokot endemic area of Uganda and Kenya. Am J Trop Med Hyg, 2014. **90**(1): p. 33-39.
190. Bourgeois, N., et al., 'Active chronic visceral leishmaniasis' in HIV-1-infected patients demonstrated by biological and clinical long-term follow-up of 10 patients. HIV Med, 2010. **11**(10): p. 670-673.
191. Burza, S., et al., One-year follow-up of immunocompetent male patients treated with miltefosine for primary visceral leishmaniasis in Bihar, India. Clin Infect Dis, 2013. **57**(9): p. 1363-4.
192. Gorski, S., et al., Visceral leishmaniasis relapse in Southern Sudan (1999-2007): a retrospective study of risk factors and trends. PLoS Negl Trop Dis, 2010. **4**(6): p. e705.
193. Lachaud, L., et al., Parasite susceptibility to amphotericin B in failures of treatment for visceral leishmaniasis in patients coinfected with HIV Type 1 and *Leishmania infantum*. Clin Infect Dis, 2009. **48**(2): p. e16-e22.
194. Das, V.N., et al., Clinical epidemiologic profile of a cohort of post-kala-azar dermal leishmaniasis patients in Bihar, India. Am J Trop Med Hyg, 2012. **86**(6): p. 959-61.
195. Molina, R., et al., Infectivity of post-kala-azar dermal leishmaniasis patients to sand flies: revisiting a proof of concept in the context of the kala-azar elimination program in the Indian subcontinent. Clin Infect Dis, 2017. **65**(1): p. 150-153.
196. Mondal, D., et al., Quantifying the infectiousness of post-kala-azar dermal leishmaniasis towards sandflies. Clin Infect Dis, 2018.
197. Musa, A.M., et al., The natural history of Sudanese post-kala-azar dermal leishmaniasis: clinical, immunological and prognostic features. Ann Trop Med Parasitol, 2002. **96**(8): p. 765-72.
198. Das, V.N.R., et al., Longitudinal study of transmission in households with visceral leishmaniasis, asymptomatic infections and PKDL in highly endemic villages in Bihar, India. PLoS Negl Trop Dis, 2016. **10**(12): p. e0005196.
199. Zijlstra, E.E., et al., Post-kala-azar dermal leishmaniasis. Lancet Infect Dis, 2003. **3**(2): p. 87-98.
200. Zijlstra, E.E., The immunology of post-kala-azar dermal leishmaniasis (PKDL). Parasit Vectors, 2016. **9**: p. 464.
201. Ostyn, B., et al., Incidence of symptomatic and asymptomatic *Leishmania donovani* infections in high-endemic foci in India and Nepal: a prospective study. PLoS Negl Trop Dis, 2011. **5**(10): p. e1284.
202. Gidwani, K., et al., Longitudinal seroepidemiologic study of visceral leishmaniasis in hyperendemic regions of Bihar, India. Am J Trop Med Hyg, 2009. **80**(3): p. 345-346.
203. Picado, A., et al., Risk factors for visceral leishmaniasis and asymptomatic *Leishmania donovani* infection in India and Nepal. PLoS One, 2014. **9**(1): p. e87641.
204. Ready, P.D., Epidemiology of visceral leishmaniasis. Clin Epidemiol, 2014. **6**: p. 147-154.
205. Baleela, R., et al., Leishmania donovani populations in Eastern Sudan: temporal structuring and a link between human and canine transmission. Parasit Vectors, 2014. **7**: p. 496.
206. Ribeiro, R.R., et al., Canine Leishmaniasis: An Overview of the Current Status and Strategies for Control. Biomed Res Int, 2018. **2018**: p. 3296893.
207. Grimaldi, G., Jr., et al., Field trial of efficacy of the Leish-tec(R) vaccine against canine leishmaniasis caused by *Leishmania infantum* in an endemic area with high transmission rates. PLoS One, 2017. **12**(9): p. e0185438.

208. Bezerra, J.M.T., et al., Burden of leishmaniasis in Brazil and federated units, 1990-2016: Findings from Global Burden of Disease Study 2016. PLoS Negl Trop Dis, 2018. **12**(9): p. e0006697.
209. Lima, I.D., et al., Changing demographics of visceral leishmaniasis in northeast Brazil: Lessons for the future. PLoS Negl Trop Dis, 2018. **12**(3): p. e0006164.
210. Pan American Health Organization, Plan of action to strengthen the surveillance and control of leishmaniasis in the Americas 2017-2022. 2017.
211. Romero, G.A.S. and Boelaert, M., Control of Visceral Leishmaniasis in Latin America—A Systematic Review. PLoS Negl Trop Dis, 2010. **4**(1): p. e584.
212. Amela, C., et al., Injecting drug use as risk factor for visceral leishmaniasis in AIDS patients. Eur J Epidemiol, 1996. **12**(1): p. 91-2.
213. Monge-Maillo, B., et al., Visceral leishmaniasis and HIV coinfection in the Mediterranean region. PLoS Negl Trop Dis, 2014. **8**(8): p. e3021.
214. Carrillo, E., Moreno, J., and Cruz, I., What is responsible for a large and unusual outbreak of leishmaniasis in Madrid? Trends Parasitol, 2013. **29**(12): p. 579-80.
215. Seva, A.D.P., et al., Efficacies of prevention and control measures applied during an outbreak in Southwest Madrid, Spain. PLoS One, 2017. **12**(10): p. e0186372.
216. Al-Salem, W., Herricks, J.R., and Hotez, P.J., A review of visceral leishmaniasis during the conflict in South Sudan and the consequences for East African countries. Parasit Vectors, 2016. **9**: p. 460.
217. Wasunna, M., et al., The Leishmaniasis East Africa Platform (LEAP): strengthening clinical trial capacity in resource-limited countries to deliver new treatments for visceral leishmaniasis. Trans R Soc Trop Med Hyg, 2016. **110**(6): p. 321-3.
218. World Health Organization, Visceral leishmaniasis control strategies and epidemiological situation update in East Africa. Report of a bi-regional consultation, Addis Ababa, Ethiopia 9-11 March 2015. 2015.
219. Sundar, S. and Chakravarty, J., Antimony toxicity. Int J Environ Res Public Health, 2010. **7**(12): p. 4267-4277.
220. South Sudan Ministry of Health, Guidelines for diagnosis, treatment and prevention of visceral leishmaniasis in South Sudan.
221. Sundar, S. and Singh, A., Recent developments and future prospects in the treatment of visceral leishmaniasis. Ther Adv Infect Dis, 2016. **3**(3-4): p. 98-109.
222. Trejo, W.H. and Bennett, R.E., *Steptomyces nodosus* sp. n., the amphotericin-producing organism. J Bacteriol, 1963. **85**(2): p. 436-439.
223. Saha, A.K., Mukherjee, T., and Bhaduri, A., Mechanism of action of amphotericin B on *Leishmania donovani* promastigotes. Mol Biochem Parasitol, 1986. **19**(3): p. 195-200.
224. Gray, K.C., et al., Amphotericin primarily kills yeast by simply binding ergosterol. Proc Natl Acad Sci U S A, 2012. **109**(7): p. 2234.
225. Sundar, S., et al., Efficacy of miltefosine in the treatment of visceral leishmaniasis in India after a decade of use. Clin Infect Dis, 2012. **55**(4): p. 543-50.
226. Burza, S., Croft, S.L., and Boelaert, M., Leishmaniasis. Lancet, 2018.
227. World Health Organization, Technical Report Series. 949. Control of the leishmaniases, in WHO Technical Report Series. 2010, World Health Organization: Geneva, Switzerland.
228. Singh, O.P. and Sundar, S., Developments in diagnosis of visceral leishmaniasis in the elimination era. J Parasitol Res, 2015. **2015**: p. 239469.
229. Burns, J.M., et al., Molecular characterization of a kinesin-related antigen of *Leishmania chagasi* that detects specific antibody in African and American visceral leishmaniasis. Proc Natl Acad Sci U S A, 1993. **90**(2): p. 775-779.
230. Sundar, S., et al., Rapid accurate field diagnosis of Indian visceral leishmaniasis. Lancet, 1998. **351**(9102): p. 563-5.

231. Boelaert, M., et al., Diagnostic tests for kala-azar: a multi-centre study of the freeze-dried DAT, rK39 strip test and KAtex in East Africa and the Indian subcontinent. Trans R Soc Trop Med Hyg, 2008. **102**(1): p. 32-40.
232. Bhattacharyya, T., Boelaert, M., and Miles, M.A., Comparison of visceral leishmaniasis diagnostic antigens in African and Asian *Leishmania donovani* reveals extensive diversity and region-specific polymorphisms. PLoS Negl Trop Dis, 2013. **7**(2): p. e2057.
233. Bhattacharyya, T., et al., Significantly lower anti-*Leishmania* IgG responses in Sudanese versus Indian visceral leishmaniasis. PLoS Negl Trop Dis, 2014. **8**(2): p. e2675.
234. Pattabhi, S., et al., Design, development and evaluation of rK28-based point-of-care tests for improving rapid diagnosis of visceral leishmaniasis. PLoS Negl Trop Dis, 2010. **4**(9).
235. Alce, T.M., et al., Expression of hydrophilic surface proteins in infective stages of *Leishmania donovani*. Mol Biochem Parasitol, 1999. **102**(1): p. 191-6.
236. Mukhtar, M., et al., Diagnostic accuracy of rK28-based immunochromatographic rapid diagnostic tests for visceral leishmaniasis: a prospective clinical cohort study in Sudan. Trans R Soc Trop Med Hyg, 2015.
237. Vaish, M., et al., Evaluation of rK28 antigen for serodiagnosis of visceral leishmaniasis in India. Clin Microbiol Infect, 2012. **18**(1): p. 81-5.
238. Dantas-Torres, F., et al., Level of agreement between two commercially available rapid serological tests and the official screening test used to detect Leishmania seropositive dogs in Brazil. Vet J, 2018. **234**: p. 102-104.
239. Molinet, F.J., et al., Specificity of the rapid rK39 antigen-based immunochromatographic test Kalazar Detect(r) in patients with cutaneous leishmaniasis in Brazil. Mem Inst Oswaldo Cruz, 2013. **108**(3).
240. Machado de Assis, T.S., Rabello, A., and Werneck, G.L., Latent class analysis of diagnostic tests for visceral leishmaniasis in Brazil. Trop Med Int Health, 2012. **17**(10): p. 1202-7.
241. da Silva, M.R.B., et al., Performance of two immunochromatographic tests for diagnosis of visceral leishmaniasis in patients coinfected with HIV. Parasitol Res, 2018. **117**(2): p. 419-427.
242. Islam, M.Z., et al., Enzyme-linked immunosorbent assay to detect urinary antibody against recombinant rKRP42 antigen made from *Leishmania donovani* for the diagnosis of visceral leishmaniasis. Am J Trop Med Hyg, 2008. **79**(4): p. 599-604.
243. Chakravarty, J., et al., Evaluation of rk39 immunochromatographic test with urine for diagnosis of visceral leishmaniasis. Trans R Soc Trop Med Hyg, 2011. **105**(9): p. 537-9.
244. Singh, D., et al., Evaluation of rK-39 strip test using urine for diagnosis of visceral leishmaniasis in an endemic region of India. Am J Trop Med Hyg, 2013. **88**(2): p. 222-6.
245. Ghosh, P., et al., Evaluation of diagnostic performance of rK28 ELISA using urine for diagnosis of visceral leishmaniasis. Parasit Vectors, 2016. **9**(1): p. 383.
246. de Ruiter, C.M., et al., Molecular tools for diagnosis of visceral leishmaniasis: systematic review and meta-analysis of diagnostic test accuracy. J Clin Microbiol, 2014. **52**(9): p. 3147-55.
247. Sakkas, H., Gartzonika, C., and Levidiotou, S., Laboratory diagnosis of human visceral leishmaniasis. J Vector Borne Dis, 2016. **53**(1): p. 8-16.
248. Sudarshan, M. and Sundar, S., Parasite load estimation by qPCR differentiates between asymptomatic and symptomatic infection in Indian visceral leishmaniasis. Diagn Microbiol Infect Dis, 2014. **80**(1): p. 40-2.
249. Pessoa, E.S.R., et al., Evaluation of urine for Leishmania infantum DNA detection by real-time quantitative PCR. J Microbiol Methods, 2016. **131**: p. 34-41.

250. Verma, S., et al., Development of a rapid loop-mediated isothermal amplification assay for diagnosis and assessment of cure of *Leishmania* infection. BMC Infect Dis, 2017. **17**(1): p. 223.
251. Mukhtar, M., et al., Sensitive and less invasive confirmatory diagnosis of visceral leishmaniasis in Sudan using loop-mediated isothermal amplification (LAMP). PLoS Negl Trop Dis, 2018. **12**(2): p. e0006264.
252. Castellanos-Gonzalez, A., et al., A Novel Molecular Test to Diagnose Canine Visceral Leishmaniasis at the Point of Care. Am J Trop Med Hyg, 2015. **93**(5): p. 970-5.
253. Mondal, D., et al., Mobile suitcase laboratory for rapid detection of *Leishmania donovani* using recombinase polymerase amplification assay. Parasit Vectors, 2016. **9**(1): p. 281.
254. Gunaratna, G., et al., Evaluation of rapid extraction and isothermal amplification techniques for the detection of Leishmania donovani DNA from skin lesions of suspected cases at the point of need in Sri Lanka. Parasit Vectors, 2018. **11**(1): p. 665.
255. Boelaert, M., et al., Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease. Cochrane Database Syst Rev, 2014. **6**: p. Cd009135.
256. Ahsan, M.M., et al., Evaluation of latex agglutination test (KAtex) for early diagnosis of kala-azar. Mymensingh Med J, 2010. **19**(3): p. 335-9.
257. Sarkari, B., Chance, M., and Hommel, M., A capture ELISA for the diagnosis of visceral leishmaniasis using a monoclonal antibody against a leishmanial urinary antigen. Iran Biomed J, 2005. **9**(3): p. 117-122.
258. Abeijon, C., et al., Identification and diagnostic utility of *Leishmania infantum* proteins found in urine samples from patients with visceral leishmaniasis. Clin Vaccine Immunol, 2012. **19**(6): p. 935-43.
259. Abeijon, C. and Campos-Neto, A., Potential non-invasive urine-based antigen (protein) detection assay to diagnose active visceral leishmaniasis. PLoS Negl Trop Dis, 2013. **7**(5): p. e2161.
260. Abeijon, C., et al., Novel antigen detection assay to monitor therapeutic efficacy of visceral leishmaniasis. Am J Trop Med Hyg, 2016. **95**(4): p. 800-802.
261. Abeijon, C., et al., Use of VHH antibodies for development of antigen detection test for visceral leishmaniasis. Parasite Immunol, 2018. **0**(ja): p. e12584.
262. Vallur, A.C., et al., Development and comparative evaluation of two antigen detection tests for visceral leishmaniasis. BMC Infect Dis, 2015. **15**(1): p. 384.
263. Bennuru, S., et al., Mining filarial genomes for diagnostic and therapeutic targets. Trends Parasitol, 2018. **34**(1): p. 80-90.
264. Saadi, M., Karkhah, A., and Nouri, H.R., Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. Infect Genet Evol, 2017. **51**: p. 227-234.
265. Repetto, S.A., et al., An improved DNA isolation technique for PCR detection of *Strongyloides stercoralis* in stool samples. Acta Trop, 2013. **126**(2): p. 110-4.
266. Cuesta-Astroz, Y., et al., Helminth secretomes reflect different lifestyles and parasitized hosts. Int J Parasitol, 2017. **47**(9): p. 529-544.
267. Worasith, C., et al., Advances in the diagnosis of human opisthorchiasis: development of *Opisthorchis viverrini* antigen detection in urine. PLoS Negl Trop Dis, 2015. **9**(10): p. e0004157.
268. Shirey, R., et al., Non-invasive urine biomarker lateral flow immunoassay for monitoring active onchocerciasis. ACS Infect Dis, 2018.
269. Fernandez-Soto, P., et al., Strong-LAMP: A LAMP Assay for Strongyloides spp. Detection in Stool and Urine Samples. Towards the Diagnosis of Human Strongyloidiasis Starting from a Rodent Model. PLoS Negl Trop Dis, 2016. **10**(7): p. e0004836.
270. Lodh, N., et al., Diagnosis of Strongyloides stercoralis: detection of parasite-derived DNA in urine. Acta Trop, 2016.

271. Formenti, F., et al., A diagnostic study comparing conventional and real-time PCR for Strongyloides stercoralis on urine and on faecal samples. Acta Trop, 2019. **190**: p. 284-287.
272. Guezala, M.C., et al., Development of a species-specific coproantigen ELISA for human *Taenia solium* taeniasis. Am J Trop Med Hyg, 2009. **81**(3): p. 433-7.
273. Casaravilla, C., et al., Production and characterization of monoclonal antibodies against excretory/secretory products of adult Echinococcus granulosus, and their application to coproantigen detection. Parasitol Int, 2005. **54**(1): p. 43-9.
274. Qian, M.B., et al., Clonorchiasis. Lancet, 2016. **387**(10020): p. 800-10.
275. Goodswen, S.J., Kennedy, P.J., and Ellis, J.T., Vacceed: a high-throughput *in silico* vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. Bioinformatics, 2014. **30**(16): p. 2381-3.
276. Carstensen, H., et al., The epidemiology of cryptosporidiosis and other intestinal parasitoses in children in southern Guinea-Bissau. Trans R Soc Trop Med Hyg, 1987. **81**(5): p. 860-4.
277. Pampiglione, S., et al., Human intestinal parasites in sub-Saharan Africa. Eastern Boé and Canhabaque Island (Guinea-Bissau). Parassitologia, 1987. **29**(1): p. 1-13.
278. Steenhard, N.R., Ornbjerg, N., and Molbak, K., Concurrent infections and socioeconomic determinants of geohelminth infection: a community study of schoolchildren in periurban Guinea-Bissau. Trans R Soc Trop Med Hyg, 2009. **103**(8): p. 839-45.
279. Allan, J.C. and Craig, P.S., Coproantigens in taeniasis and echinococcosis. Parasitol Int, 2006. **55 Suppl**: p. S75-80.
280. Bhattacharyya, T., et al., IgG1 as a potential biomarker of post-chemotherapeutic relapse in visceral leishmaniasis, and adaptation to a rapid diagnostic test. PLoS Negl Trop Dis, 2014. **8**(10): p. e3273.
281. Mollett, G., et al., Detection of IgG1 against rK39 improves monitoring of treatment outcome in visceral leishmaniasis. Clin Infect Dis, 2018.
282. World Health Organization. Malaria rapid diagnostic tests: RDT instructions and training. 2005 [cited 2018 2/9/18]; Available from: http://www2.wpro.who.int/sites/rdt/using_rdts/training/rdt_training_combination.htm.
283. Kobayashi, J., et al., Application of enzyme immunoassay for postchemotherapy evaluation of human strongyloidiasis. Diagn Microbiol Infect Dis, 1994. **18**(1): p. 19-23.
284. Shahbaaz, M., Md, I., and Ahmad, F., Functional annotation of conserved hypothetical Proteins from *Haemophilus influenzae* Rd KW20. PLoS One, 2013. **8**(12): p. e84263.
285. van Griensven, J., et al., *Leishmania* antigenuria to predict initial treatment failure and relapse in visceral leishmaniasis/HIV coinfected patients: an exploratory study nested within a clinical trial in Ethiopia. Front Cell Infect Microbiol, 2018. **8**: p. 94.
286. Riera, C., et al., Evaluation of a latex agglutination test (KAtex) for detection of *Leishmania* antigen in urine of patients with HIV-*Leishmania* coinfection: value in diagnosis and post-treatment follow-up. Eur J Clin Microbiol Infect Dis, 2004. **23**(12): p. 899-904.
287. Castro-Sesquen, Y.E., et al., Use of a novel Chagas urine nanoparticle test (Chunap) for diagnosis of congenital Chagas disease. PLoS Negl Trop Dis, 2014. **8**(10): p. e3211.
288. Douglas, T.A., et al., The use of hydrogel microparticles to sequester and concentrate bacterial antigens in a urine test for Lyme disease. Biomaterials, 2011. **32**(4): p. 1157-66.
289. Luchini, A., et al., Smart hydrogel particles: biomarker harvesting: one-step affinity purification, size exclusion, and protection against degradation. Nano letters, 2008. **8**(1): p. 350-361.

290. Luque-Garcia, J.L., et al., Analysis of electroblotted proteins by mass spectrometry: protein identification after Western blotting. Mol Cell Proteomics, 2008. **7**(2): p. 308-14.
291. Balog, C.I., et al., Novel automated biomarker discovery work flow for urinary peptidomics. Clin Chem, 2009. **55**(1): p. 117-25.
292. Driguez, P., et al., Protein microarrays for parasite antigen discovery. Methods Mol Biol, 2015. **1201**: p. 221-33.
293. Gadisa, E., et al., Serological signatures of clinical cure following successful treatment with sodium stibogluconate in Ethiopian visceral leishmaniasis. Cytokine, 2017. **91**: p. 6-9.
294. Gasim, S., et al., High levels of plasma IL-10 and expression of IL-10 by keratinocytes during visceral leishmaniasis predict subsequent development of post-kala-azar dermal leishmaniasis. Clin Exp Immunol, 1998. **111**(1): p. 64-9.
295. Lim, M.D., et al., Diagnostic tools for soil-transmitted helminths control and elimination programs: a pathway for diagnostic product development. PLoS Negl Trop Dis, 2018. **12**(3): p. e0006213.
296. Peeling, R.W., et al., Rapid tests for sexually transmitted infections (STIs): the way forward. Sex Transm Infect, 2006. **82**(suppl 5): p. v1.
297. Google. Google Flu Trends. 2014 [cited 2019 10/2/19]; Available from: https://www.google.org/flutrends/about/.
298. Wesolowski, A., et al., Heterogeneous mobile phone ownership and usage patterns in Kenya. PLoS One, 2012. **7**(4): p. e35319.
299. Bagcchi, S., WHO launches Essential Diagnostics List. Lancet Infect Dis, 2018. **18**(7): p. 724-725.
300. United Nations. Sustainable Development Goal 3: Ensure healthy lives and promote well-being for all at all ages. 2018 [cited 2019 10/2/19]; Available from: https://sustainabledevelopment.un.org/sdg3.

Appendix

Methods

Ethical statement

Collection of human stool samples carried out on Bubaque island, Guinea Bissau in October 2017 was approved by the London School of Hygiene & Tropical Medicine Ethical Committee, with reference number 9923-2, and as part of a separate LSHTM-hosted study by the National Committee of Ethics in Health, Ministry of Health, Guinea Bissau. All adult participants gave written informed consent for themselves and their dependent children. In addition, children aged 7-17 gave written informed assent to providing samples. Infants under 6 months old were excluded from the study.

Use of DNA extracted from human stool and kindly provided by Sören Becker was granted by the London School of Hygiene & Tropical Medicine Ethical Committee, reference number 9923-1. Original sample collection in Cote d'Ivoire, Nepal and Mali was approved as part of a separate study, 'NIDIAG', by all the respective national ethical committees and by the University of Antwerp Ethics Committee where the study was based. All participants gave informed consent.

Use of urine and serum samples from India and Sudan was ethically approved as a different component of the 'NIDIAG' study by the respective national ethics committees and by the London School of Hygiene & Tropical Medicine Ethics Committee where this part of the study was based. All participants gave informed consent.

Animal (rabbit) use for raising antibodies to *L. donovani* was carried out before the present study at the Royal Veterinary College, London, under conditions of animal husbandry and welfare, according to UK law.

Strongyloides methods

Coproantigen discovery methods are detailed in Chapter 2. Additional methods are described here.

Collection of stool samples and planned methodology for processing

Study site and participants

Human stool samples were collected as part of separate study which was a community-wide STH baseline prevalence survey from villages on the island of Bubaque in the Bijagos archipelago, Guinea Bissau. This was prior to a larger survey across the Bijagos, which would be followed by mass drug administration of anthelminthics. Participants were from

households chosen by closest proximity to a randomly-generated GPS point. Potential participants were informed in detail about the study and the use of their samples. Consenting individuals were provided with anonymously coded pots which were collected the following day and analysed by Kato Katz microscopy for STH. An aliquot of about 0.5 – 1 g of each sample was stored in 70% ethanol and shipped to LSHTM at ambient temperature and subsequently stored at -20°C for molecular analysis. An additional, larger, aliquot of some stool samples was stored in 10% formalin (HT5011, Sigma-Aldrich, UK) and shipped at ambient temperature then stored at -80°C for use in protein/antigen assay development.

Planned stool processing

The formalin-preserved stool samples were collected in order to provide *S. stercoralis* positive and negatives for coproantigen capture assay development. However, the Kato Katz technique carried out in the field did not detect *S. stercoralis*, therefore the paired ethanol stool samples were intended for use in qPCR, both the validate the microscopy results for the other STH, and also to test for *S. stercoralis*. These data would then inform the use of the formalin samples.

S. stercoralis gene sequencing- SCP/TAPS

Primer design and PCR optimisation

In order to investigate antigenic diversity across different geographic regions, a novel PCR assay was designed to amplify and sequence antigen genes of interest. The first selected gene was that of existing serological antigen NIE, for which there is currently no sequence diversity information, and which may also be suitable as a coproantigen. Amplification was of the SCP/TAPS protein family gene, SSTP_0001008900, in which our Set 1 primer amplicon corresponded to the NIE seroantigen region (Figure 1). The assay first required optimisation of the PCR reaction to amplify the region of interest, prior to sequencing (Figure 1). Stool DNA of different infection status used in this work originated from Cote d'Ivoire and was kindly provided by Sören Becker, Swiss TPH, Basel.

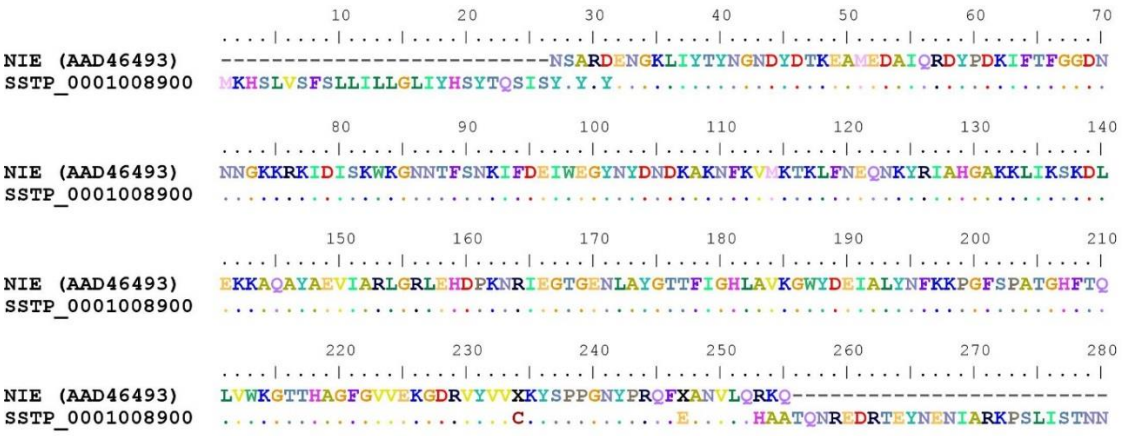


Figure 1. Very high amino acid sequence identity between seroantigen NIE and a region of the *S. stercoralis* protein encoded by gene SSTP_0001008900. Identity with the NIE sequence is indicated by a dot in the SSTP_0001008900 sequence.

Novel primers designed in this study and their binding positions in the SSTP_0001008900 gene are given in Table 1 and Figure 2. Two primer pairs, 'Set 1' and 'Set 2', were designed and used to avoid the large intron within the gene (Figure 2).

Table 1. Primers used to amplify regions of *S. stercoralis* gene SSTP_0001008900. Melting temperatures (Tm) were calculated using a pilot version of a custom-built program by Hai Le[1].

		Sequence	Amplicon size (bp)	T_m (°C)
Set 1	FWD	5' CGTATGAATTTAAGAATGCTGTTAG 3'	1,308	56.4
	RVS	5' CTACAGTAATCCTTTTATTACGGTA 3'		59.7
Set 2	FWD	5' CTGGATTCAACCATCAATTTAGGCA 3'	1,232	59.3
	RVS	5' CCAAGGGTAGTAATGTATATGTGAG 3'		68.3

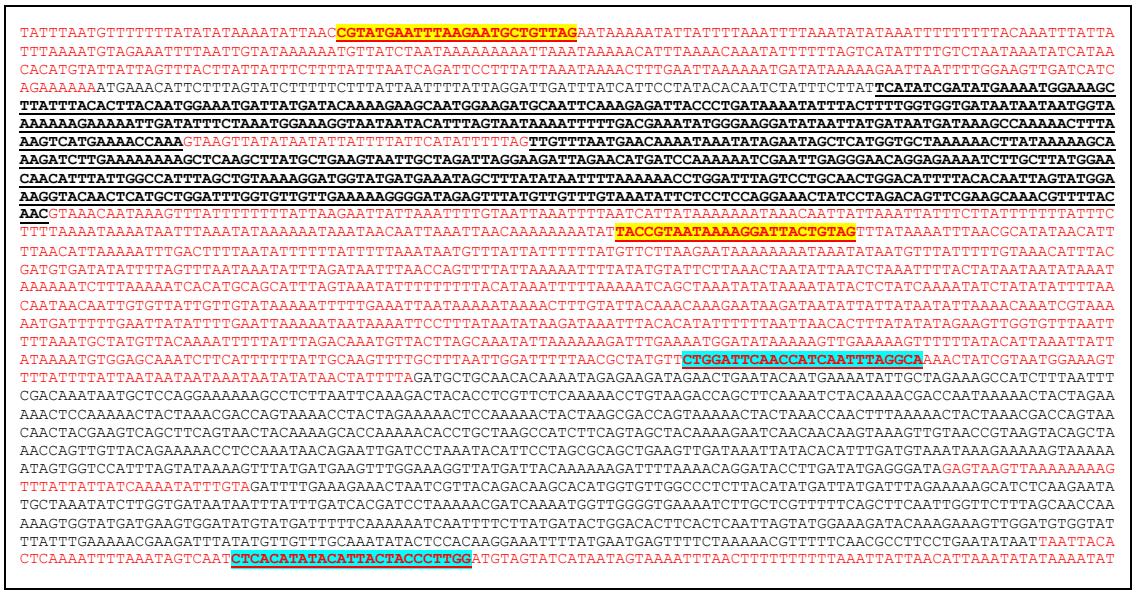


Figure 2. *Strongyloides stercoralis* SCP/TAPS gene SSTP_0001008900, a 602 aa protein encoded by 1809 base pairs in 4 exons on the antisense strand. Black and red indicate exons and introns respectively. Primer sequences are in bold underlined with Set 1 highlighted in yellow and Set 2 in blue. The NIE homologous coding region is indicated in bold underlined in exons one and two.

For the first reaction, a known *S. stercoralis* qPCR positive stool DNA sample from Cote d'Ivoire (#126), was used with Set 1 primers at three annealing temperatures and other

conditions as detailed in Tables 2 and 3. All reactions were run in a PTC-200 thermocycler (MJ Research, St Bruno, Canada). Reaction 2 used primers published by Verweij *et al.* (2009)[2], originally for qPCR detection of *S. stercoralis* in stool with a 101 bp amplicon of the 18S rRNA gene, which we used to ensure that *S. stercoralis* DNA was detectable by us in the stool DNA samples. The reaction used higher magnesium concentration and more sample DNA than reaction 1 (Table 2). Stool DNA sample IDs tested were: 10, 42, 126, 127 known *S. stercoralis* qPCR positives, and sample 2 which was known *S. stercoralis* negative. All samples were previously qPCR negative for both hookworm species. Reaction 3 used our novel Set 1 primers with three of the *S. stercoralis* positive samples (46, 126, 127) at the same annealing temperatures as reaction 1 but with increased Mg2+ and sample DNA. Reaction 4 was with our novel Set 2 primers, an intermediate level of magnesium and two annealing temperatures, 60°C and 62°C. In addition, primers described by Verweij *et al.* (2009)[2] were used under the same conditions alongside our primers as positive controls. Reaction 5 was to confirm the presence of *S. stercoralis* after DNA purification from larvae. Reaction 6 used both of our novel primer sets on larval DNA at various annealing temperatures and magnesium concentrations. 18S rRNA primers were used with the larval DNA as a positive control.

Table 2. PCR reaction conditions during optimisation of assay to amplify *S. stercoralis* SCP/TAPS gene SSTP_0001008900 prior to sequencing.

	Reaction 1 (ul/sample) [final conc.]	**Reaction 2**	**Reaction 3**	**Reaction 4**	**Reaction 5**	**Reaction 6**
Primers	Set 1	18S rRNA	Set 1	Set 2 & 18S rRNA	18S rRNA	Set 1, Set 2 and 18S rRNA
Sample	Stool DNA	Stool DNA	Stool DNA	Stool DNA	Larval DNA Stool DNA	Larval DNA
DNA[a]	1	5	5	5	5	5
10x buffer	2	2	2	2	2	2
Mg2+ @ 50 mM	[1.5 mM] 0.6	[2.5 mM] 1	[2.5 mM] 1	[2 mM] 0.8	[2.5 mM] 1	[1.5 and 2.0 mM]
dNTPs @ 10 mM	[0.2 mM] 0.4	[0.2 mM] 0.4	[0.2 mM] 0.4	[0.2 mM] 0.4	[0.2 mM] 0.4	[0.2 mM] 0.4
FWD primer @ 10 mM	[0.5 uM] 1	[0.5 pM/ul] 1	[0.5 uM] 1	1	[0.5 pM/ul] 1	1
RVS primer @ 10 mM	[0.5 uM] 1	[0.5 pM/ul] 1	[0.5 uM] 1	1	[0.5 pM/ul] 1	1
Taq @ 5 U/ul	(1 unit) 0.2	(1 unit) 0.2	(1 unit) 0.2	(1 unit) 0.2	(1 unit) 0.2	(1 unit) 0.2
dH_2O	13.8	9.4	9.4	9.6	9.4	as needed
Total	20	20	20	20	20	20

[a] All reactions also included a template-free negative control with water instead of DNA.

Cycling conditions for the different experimental reactions are given in Table 3.

Table 3. Cycling conditions used during PCR with 18S rRNA gene primers as per Verweij (2009), and PCR development with novel primer sets for *S. stercoralis* gene SSTP_0001008900.

Reaction	1 and 3	2	4	5	6
1. Denaturation/ Activation	95°C, 2 mins	95°C, 15 min	95°C, 2 mins	95°C, 15 min	95°C, 2 mins
2. Denaturation	95°C, 30 sec	95°C, 30 sec	95°C, 30 sec	95°C, 30 sec	95°C, 30 sec
3. Annealing	55, 58, 60°C, 30 sec	60°C, 30 sec	60, 62°C, 30 sec	60°C, 30 sec	55, 58, 60, 62°C, 30 sec
4. Extension	72°C, 90 sec	72°C, 30 sec	72°C, 90 sec	72°C, 30 sec	72°C, 90 sec
5. Cycles (steps 2-4)	30	30	30	30	30
6. Final extension	72°C, 10 mins	72°C, 10 mins	72°C, 10 mins	72°C, 10 mins	72°C, 10 mins

PCR products were visualised on a 1% agarose gel with GelRed (Biotium, USA) and the appropriate Hyperladder (Bioline, UK).

DNA purification from S. stercoralis larvae

Larval material of the *S. stercoralis* reference strain PV0001 was kindly provided by Thomas Nolan, University of Pennsylvania, USA. There, infection was maintained in an immunosuppressed dog under laboratory conditions and larvae were isolated from a stool culture before being allowed to migrate out of soft agar in order to clean them of bacteria. Larvae were subsequently washed in distilled water and stored frozen at -80°C. We purified DNA from 24 mg of these whole larvae using a QIAamp DNA mini kit (51304, Qiagen) and performed three elutions of the DNA from the column. The first elution was used for all experiments described here.

Sequencing and diversity analysis of NIE

In order to investigate heterogeneity in the SCP/TAPS candidate coproantigen of interest (SSTP_0001008900), we sequenced regions of the gene from *S. stercoralis* larval DNA of the reference strain PV0001. This would reveal potential sequence diversity hidden in the consensus that forms the reference genome. Sequencing reactions were carried out on PCR products from both Set 1 and Set 2 primers. Capillary sequencing was performed by central services at LSHTM. Sequence data were analysed using MEGA 6.0 (Megasoftware.net) and BioEdit[3].

S. stercoralis qPCR

Quantitative PCR (qPCR) was performed on DNA extracted from *S. stercoralis* larvae of the PV0001 reference strain, to confirm the correct species and to briefly compare pipetting accuracy of two technicians. We used a published species-specific assay for a 101 base pair (bp) region of the *S. stercoralis* 18S rRNA gene (AF279916)[2].

Strongyloides comparative genomics

The genomes of *S. stercoralis*, *S. papillosus*, *S. venezuelensis* and *Parastrongyloides trichosuri* were aligned to that of the reference genome, *S. ratti*, in order to detect genetic differences (variants) between them. In particular, with the aim of investigating variants which lead to amino acid differences between the species, to indicate possible *S. stercoralis* species-specific candidate antigens.

Genome mapping

Paired-end genome reads of *S. stercoralis*, *S. papillosus*, *S. venezuelensis* and *Parastrongyloides trichosuri* were downloaded from NCBI SRA (accessions ERR422414, ERR422409, ERR696964 respectively for the *Strongyloides* species and ERR422413, ERR422412 and ERR539713 for *P. trichosuri*) using Aspera Connect (Aspera, USA) and checked for completeness. Fermikit[4] was used to map the reads to the *S. ratti* reference genome (PRJEB125) which was obtained from WBPS release 7 along with the annotation file. The reference genome was first indexed with bwa[5]. Initial attempts at mapping *S. venezuelensis* using Fermikit were unsuccessful due to the large number of reads (136 million, versus about 24-48 million for the other species). Therefore we made a random subset of 20 million reads for this species using seqtk[6] which were then mapped against the reference in the same way as the other species.

Variant calling

Variant calling within the mapped reads for each species was done with Fermikit. Resulting variant call format (vcf) files comprised: structural variants (sv.vcf) which included long deletions, novel sequence insertions and complex structural variations; filtered variants (flt.vcf) comprising single nucleotide polymorphisms (SNPs) and short insertions or deletions (INDELs); binary sequence alignments (bam). Mapped genomes and variants were viewed in Integrative Genome Viewer (IGV)[7] using the *S. ratti* genome and annotation as reference. SnpEff[8] was used to obtain statistics on the effects of the variants and any effect on protein sequences.

Supplementary material for Chapter 4

Capture and identification of Leishmania donovani protein antigens in human urine during visceral leishmaniasis

Table S1. Number of *L. donovani* peptides and proteins identified by mass spectrometry of antigens captured from Indian VL urine with anti-1S2D antibody.

Indian VL urine immuno-captured material		**Dish 8**	**Dish 35**	**Found in both dishes**	**Total unique**
Total number of proteins		3	3	1	5
Total number of peptides		12	12	4	20
Peptides which:	Led to protein IDs	7	6	2	11
	Solo peptides	5	6	2	9

Table S2. Solo peptides of *Leishmania donovani* identified in urine of Indian and Sudanese patients with visceral leishmaniasis.

Sample	Experimental origin of peptide	Peptide	Parent protein	UniProtKB / GenBank accession
India VL urine	Dish 8 and 35	VTALEENIEAALR	Hypothetical protein, conserved	LdBPK_160420 / XP_003859730.1
	Dish 8	VTLLDR	3-oxoacyl-(acyl-carrier protein) reductase, putative	LdBPK_242110 / XP_003861301.1
	Dish 8 and 35	LSLEPR	Hypothetical protein, conserved	LdBPK_354400 / XP_003865018.1
	Dish 8	LVEEFHFSK	MP44, putative	LdBPK_270350 / XP_003861892.1
	Dish 8	ESPSPWVR	Hypothetical protein, conserved	LdBPK_281920 / XP_003862291.1
	Dish 35	SELDARK	Hypothetical protein, conserved	LdBPK_100630 / XP_003858899.1
	Dish 35	ITQLVQLMK	Hypothetical protein, unknown function	LdBPK_312110 / XP_003863299.1
	Dish 35	CFNDDIQGTGAVIAAGFLNAVK	Malic enzyme	LdBPK_240780 / XP_003861168.1
	Dish 35	YDAASQIAILSMER	Enoyl-CoA hydratase/isomerase family protein, conserved	LdBPK_350360 / XP_003864629.1
Sudan VL urine	VLu	AGNVSINQHEGQR	RNA binding protein, putative	LdBPK_190290 / XP_003860208.1
	VLu/uAg*	MEEYLHSKDSAEQR	Hypothetical protein, conserved	LdBPK_220520 / XP_003860802.1
	VLu/uAg*	NSSSRAKADYKPSSSR	Hypothetical protein, conserved	LdBPK_333140 / XP_003864138.1

	uAg	SPASASVEAVGAAAFAR ****	Hypothetical protein, conserved	LdBPK_091050 / XP_003858784.1
	VLu	VWTKDLSQMK	Hypothetical protein, conserved	LdBPK_353990 / XP_003864977.1
	uAg	IAASVPSLR	Coatomer gamma subunit, putative	LdBPK_282820 / XP_003862380.1
	VLu/uAg*	NHAKQLYMR	Protein kinase, putative	LdBPK_210190 / XP_003860536.1
	uAg	GTYEVICR ****	Hypothetical protein, conserved	LdBPK_151420 / XP_003859665.1
	uAg	QAETALVNR ****	Hypothetical protein, conserved	LdBPK_303110 / XP_003863022.1

**** These peptides were assigned to *L. donovani* proteins from mass spectrometry data after a Mascot search of an *L. donovani*-only database. All other peptides were assigned to *Leishmania* proteins after a simultaneous *L. donovani* and human database search. However, these indicated peptides had very high homology to *Leishmania* proteins in subsequent homology searches and were therefore retained.

*Some gel fragments analysed by mass spectrometry spanned two sample types.

Table S3. Number of epitope peptides of ≥8 amino acids in *Leishmania donovani* proteins identified from VL patients' urine.

VL urine origin and method used	**Protein name (UniProtKB accession)**	**Number of predicted epitopes (length range in amino acids) at two epitope cut-off scores**		**Number of N-linked glycosylation sites**
		0.65	**0.55**	
India (immunopanning)	Hypothetical protein (LdBPK_191140)	6 (8-10 aa)	48 (8-86 aa)	1
	40S ribosomal protein S9 (LdBPK_070760)	4 (8-17 aa)	25 (8-142 aa)	0
	Hypothetical protein (LdBPK_363030)	2 (12-14 aa)	28 (9-42 aa)	0
	Hypothetical protein (LdBPK_323250)	8 (8-24 aa)	18 (8-64 aa)	1
	Protein kinase (LdBPK_262110)	12 (8-14 aa)	23 (9-284 aa)	3
Sudan (immunocapture and western blot)	Hypothetical protein ‡ (LdBPK_160110)	0	1 (29 aa)	0
	Protein kinase (LdBPK_351070)	3 (9-25 aa)	4 (15-149 aa)	1
Total		35	147	4 proteins

‡ The lower epitope score threshold was used for this protein due to absence of epitopes >8 aa at the higher score threshold.

Oral presentations by the candidate

- London Centre for Neglected Tropical Disease Research (LCNTDR), Genetic Diversity of NTD Pathogens research afternoon, London, 23/5/18. 'Exploiting genetic diversity for development of strongyloidiasis diagnostics'.
- International Society for Neglected Tropical Diseases (ISNTD) d^3 conference, London, 26/6/18. 'Using 'omics' data for *Strongyloides stercoralis* diagnostic antigen discovery'. Video available here [9].
- Three Minute Thesis, LSHTM Finalist, 7/6/18. 'A Tricky Little Worm'.
- LSHTM, work-in-progress seminar, 29/1/16. 'Identifying diagnostic antigens for visceral leishmaniasis'.

Poster presentations by the candidate

World Leish 6, Toledo, May 2017.

- Marlais et al. 'Non-invasive diagnosis and test-of-cure of VL: Using mass spectrometry to identify the antigen detected in urine by a commercial capture ELISA.'
- Marlais et al. 'Cure or Relapse? Further studies of an IgG1 rapid test and western blot in assessing treatment outcome and progression to VL in India.'

British Society for Parasitology, Aberystwyth, April 2018.

- Marlais et al. 'Comparative 'Omics' identification of coproantigens for diagnosis of *Strongyloides stercoralis* infection'. (Student poster prize winner).

References within the Appendix

1. Le, H. primer-design. Primer design program for visualising and calculating secondary structure stability. 2018 [cited; Available from: https://github.com/lehaibg/primer-design/blob/master/06072018%20Primer%20design.py.
2. Verweij, J.J., et al., Molecular diagnosis of *Strongyloides stercoralis* in faecal samples using real-time PCR. Trans R Soc Trop Med Hyg, 2009. **103**(4): p. 342-6.
3. Hall, T., BioEdit version 7.0.0 manual. 2004.
4. Li, H., FermiKit: assembly-based variant calling for Illumina resequencing data. Bioinformatics, 2015. **31**(22): p. 3694-3696.
5. Li, H. and Durbin, R., Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009. **25**(14): p. 1754-60.
6. https://github.com/lh3/seqtk. Seqtk: a fast and lightweight tool for processing sequences in the FASTA or FASTQ format. [cited.
7. Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P., Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 2013. **14**(2): p. 178-92.
8. Cingolani, P., et al., A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin), 2012. **6**(2): p. 80-92.
9. International Society for Neglected Tropical Diseases. Tegwen Marlais (LSHTM): 'Omics' data for Strongyloides stercoralis diagnostic antigen discovery. 2018 [cited 2019 10/2/19]; Available from: https://www.youtube.com/watch?v=fyRIatj0BkQ&index=29&list=PLdJabcLRtQRqUO9Wc8wrzuwAvbNU1PRMo&t=0s.

frontiers in Cellular and Infection Microbiology

ORIGINAL RESEARCH
published: 13 December 2018
doi: 10.3389/fcimb.2018.00427

Visceral Leishmaniasis IgG1 Rapid Monitoring of Cure vs. Relapse, and Potential for Diagnosis of Post Kala-Azar Dermal Leishmaniasis

Tegwen Marlais[1†], Tapan Bhattacharyya[1†], Om Prakash Singh[2], Pascal Mertens[3], Quentin Gilleman[3], Caroline Thunissen[3], Bruno C. Bremer Hinckel[3,4], Callum Pearson[1], Bathsheba L. Gardner[1], Stephanie Airs[1], Marianne de la Roche[1], Kiera Hayes[1], Hannah Hafezi[1], Andrew K. Falconar[5], Osama Eisa[6], Alfarazdeg Saad[6], Basudha Khanal[7], Narayan Raj Bhattarai[7], Suman Rijal[8], Marleen Boelaert[9], Sayda El-Safi[6], Shyam Sundar[2] and Michael A. Miles[1]*

[1] Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, [2] Department of Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi, India, [3] Coris BioConcept, Gembloux, Belgium, [4] Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium, [5] Departamento de Medicina, Universidad del Norte, Barranquilla, Colombia, [6] Faculty of Medicine, University of Khartoum, Khartoum, Sudan, [7] Department of Microbiology, B.P. Koirala Institute of Health Sciences, Dharan, Nepal, [8] Department of Internal Medicine, B.P. Koirala Institute of Health Sciences, Dharan, Nepal, [9] Department of Public Health, Institute of Tropical Medicine, Antwerp, Belgium

Background: There is a recognized need for an improved diagnostic test to assess post-chemotherapeutic treatment outcome in visceral leishmaniasis (VL) and to diagnose post kala-azar dermal leishmaniasis (PKDL). We previously demonstrated by ELISA and a prototype novel rapid diagnostic test (RDT), that high anti-*Leishmania* IgG1 is associated with post-treatment relapse versus cure in VL.

Methodology: Here, we further evaluate this novel, low-cost RDT, named VL Sero K-SeT, and ELISA for monitoring IgG1 levels in VL patients after treatment. IgG1 levels against *L. donovani* lysate were determined. We applied these assays to Indian sera from cured VL at 6 months post treatment as well as to relapse and PKDL patients. Sudanese sera from pre- and post-treatment and relapse were also tested.

Results: Of 104 paired Indian sera taken before and after treatment for VL, when deemed clinically cured, 81 (77.9%) were positive by VL Sero K-SeT before treatment; by 6 months, 68 of these 81 (84.0%) had a negative or reduced RDT test line intensity. ELISAs differed in positivity rate between pre- and post-treatment ($p = 0.0162$). Twenty eight of 33 (84.8%) Indian samples taken at diagnosis of relapse were RDT positive. A comparison of Indian VL Sero K-SeT data from patients deemed cured and relapsed confirmed that there was a significant difference ($p < 0.0001$) in positivity rate for the two groups using this RDT. Ten of 17 (58.8%) Sudanese sera went from positive to negative or decreased VL Sero K-SeT at the end of 11–30 days of treatment. Forty nine of 63 (77.8%) PKDL samples from India were positive by VL Sero K-SeT.

OPEN ACCESS

Edited by:
Javier Moreno, Instituto de Salud Carlos III, Spain

Reviewed by:
Suresh Kumar Kalangi, Indrashil University, India
Kevin M. Tyler, University of East Anglia, United Kingdom

***Correspondence:**
Tegwen Marlais
tegwen.marlais@lshtm.ac.uk

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Parasite and Host, a section of the journal Frontiers in Cellular and Infection Microbiology

Received: *09 May 2018*
Accepted: *28 November 2018*
Published: *13 December 2018*

Citation:
Marlais T, Bhattacharyya T, Singh OP, Mertens P, Gilleman Q, Thunissen C, Hinckel BCB, Pearson C, Gardner BL, Airs S, de la Roche M, Hayes K, Hafezi H, Falconar AK, Eisa O, Saad A, Khanal B, Bhattarai NR, Rijal S, Boelaert M, El-Safi S, Sundar S and Miles MA (2018) Visceral Leishmaniasis IgG1 Rapid Monitoring of Cure vs. Relapse, and Potential for Diagnosis of Post Kala-Azar Dermal Leishmaniasis. Front. Cell. Infect. Microbiol. 8:427. doi: 10.3389/fcimb.2018.00427

Conclusion: We have further shown the relevance of IgG1 in determining clinical status in VL patients. A positive VL Sero K-SeT may also be helpful in supporting diagnosis of PKDL. With further refinement, such as the use of specific antigens, the VL Sero K-SeT and/or IgG1 ELISA may be adjuncts to current VL control programmes.

Keywords: visceral leishmaniasis, serology, treatment, relapse, cure, IgG1, RDT, PKDL

INTRODUCTION

Visceral leishmaniasis (VL; kala-azar), is caused by the protozoan parasites *Leishmania donovani* in Asia, Africa and the Middle East and *Leishmania infantum* in Europe and South America. These parasites are transmitted by blood-feeding female phlebotomine sand flies. Symptomatic VL is usually fatal if untreated. Symptoms include prolonged fever >14 days, wasting, splenomegaly, hepatomegaly and anemia (Sundar and Rai, 2002). While VL is present in about 75 countries, the majority (90%) of cases in 2015 occurred in India, Sudan, South Sudan, Ethiopia, Somalia, Kenya, and Brazil (World Health Organization, 2017), where it is closely linked to poverty, both as cause and effect (Boelaert et al., 2009; Sarnoff et al., 2010).

Following clinical suspicion of VL, serology is used for diagnosis. Techniques vary by region and include the immunofluorescence antibody test (IFAT), direct agglutination test (DAT), enzyme linked immunosorbent assay (ELISA), and detection of IgG against recombinant antigens rK39 or rK28 (Singh and Sundar, 2015). In India the DAT and rK39 serology are used, with a positive result in either test indicative of exposure to infection with *L. donovani*. For confirmatory parasitological diagnosis, seropositive individuals undergo spleen, bone marrow or lymph node biopsy to search for the intracellular amastigote stage in films of Giemsa-stained aspirates. These are invasive, costly and potentially hazardous techniques with low and variable sensitivities ranging from 53 to 99% (Singh and Sundar, 2015).

VL is treated with antimonials, miltefosine, paromomycin, amphotericin B, liposomal amphotericin (AmBisome) or drug combinations (World Health Organization, 2010). Currently, post-treatment outcome is determined by assessment of clinical signs and symptoms, initially on the last day of drug treatment and, in India, again 6 months after administration of the last dose (World Health Organization, 2010). Possible outcomes are: cure; relapse; death (by VL or not); post kala-azar dermal leishmaniasis (PKDL); loss to follow up. However, recent studies from India and Nepal have reported relapse rates of between 1.4 and 20%, including up to and beyond 12 months following the end of treatment (Burza et al., 2013, 2014; Rijal et al., 2013). In Sudan, relapse rates around 6% have been reported (Gorski et al., 2010; Atia et al., 2015). Patients who relapse face a further biopsy procedure to confirm the presence of parasites.

PKDL is a non-painful sequela of VL occurring in over 50% of cases in Sudan (Zijlstra et al., 2003) but is far less prevalent in South Asia (Zijlstra et al., 2003; Uranw et al., 2011). PKDL is less reported in *L. infantum* endemic regions where cases have mostly been associated with HIV/AIDS (Ridolfo et al., 2000; Bittencourt et al., 2003; Celesia et al., 2014), other co-infections (Trindade et al., 2015) or immune suppression (Roustan et al., 1998). PKDL manifests between 0.5 months to one or more years after apparently successful VL treatment (Musa et al., 2002; Uranw et al., 2011; Singh et al., 2012; Moulik et al., 2017) and may occasionally occur without a prior episode of VL (el Hassan et al., 1992; Zijlstra et al., 2003; Das et al., 2012, 2016). PKDL is suspected based on dermal manifestations that are non-specific and diagnosis is made on previous VL treatment history and confirmed parasitologically by microscopy of slit skin smear or biopsy or PCR (Zijlstra et al., 2017). Conventional serology is likely to remain positive from the earlier VL and there is no test in use to predict PKDL (Gidwani et al., 2011). The high parasite density in PKDL skin provides a source of infection to sand flies and thus sustains long term transmission and endemicity (Molina et al., 2017; Mondal et al., 2018).

An unresolved crucial question is how to identify asymptomatic infected individuals simply and reliably (as defined by seropositivity, lack of clinical symptoms and no prior history of VL) who will progress to active VL. High DAT and/or rK39 ELISA titres have been associated with increased risk of progression in the Indian subcontinent but as yet there is no single rapid test in use for this purpose (Hasker et al., 2014; Chapman et al., 2015). To improve outcome monitoring of VL and disease control, the World Health Organization has identified the vital need for a marker of post-chemotherapeutic cure, and the high priority incorporation of such a biomarker into a point-of-care rapid diagnostic test (RDT) (World Health Organization, 2012). Such a test should meet the "ASSURED" criteria of being: affordable; sensitive (few false negatives); specific (few false positives); user-friendly, requiring minimal training; rapid; robust, not requiring cold-storage; equipment-free, and deliverable to those who need it (Peeling et al., 2006).

We have previously shown that high anti-*Leishmania* IgG1 ELISA titers are associated with treatment failure, whereas, in cases deemed to be cured following chemotherapy, IgG1 levels diminish significantly by 6 months post-treatment and only IgG1 gave this level of discrimination (Bhattacharyya et al., 2014a). We demonstrated this by ELISA against *L. donovani* whole cell lysate, and then adapted the assay to a prototype lateral flow immunochromatographic RDT. Here, we present further evaluation of this RDT, called VL Sero K-SeT, to indicate cure after VL treatment in a larger, paired, sample set and to confirm relapse. We also performed western blot on the same sample set. Additionally we show the potential utility of VL Sero K-SeT and other IgG1 assays to confirm PKDL.

METHODS

Ethics Statement

In India, the collection of samples was approved by the Ethics Committee of Banaras Hindu University, Varanasi. In Sudan approval was by the Ethical Research Committee, Faculty of Medicine, University of Khartoum and the National Health Research Ethics Committee, Federal Ministry of Health, Sudan. Written informed consent was obtained from adult subjects included in the study or from the parents or guardians of individuals <18 years of age. In Nepal, informed consent was obtained from all the participants and the ethical committee of the B.P. Koirala Institute of Health Sciences (BPKIHS) approved the study. This research was also approved, as part of the EC NIDIAG project, by the London School of Hygiene and Tropical Medicine Ethics Committee.

Sources of Sera/Plasma

We retrospectively selected sera or plasma from an archive of different VL disease states. Samples had been collected in VL endemic regions, namely Muzaffarpur in Bihar, India after 2007 and in 2013 in Gedaref, Sudan. Sample sizes used during this evaluation were dependent on availability of appropriate samples and reagents.

In India, cases of VL had been diagnosed by positive rK39 serology and/or parasitologically by microscopy of splenic aspirates. In Sudan active cases of VL had been diagnosed by microscopy of bone marrow or lymph node aspirates in conjunction with serological assays. These diagnoses were made according to their respective national procedures, prior to the present study. Sera/plasma were stored at −80°C until use. All patients were HIV negative. We have previously observed that serum and plasma derived from the same sample show no difference in IgG titer in ELISA against *L. donovani* lysate (unpublished observations), although we have not specifically assessed IgG1 with both sample types.

India

Indian sample types are described in **Table 1**. We have previously found that in Indian VL, IgG1 titer up to day 30 post-treatment initiation is not statistically significantly different from pre-treatment (Bhattacharyya et al., 2014a) and therefore we consider these as "pre-treatment" in paired samples for the purposes of this study. Treatment of VL was with single-dose AmBisome alone or with 10 days of miltefosine. PKDL was treated with miltefosine for 84 days. DAT and rK39 ELISA were conducted prior to the present study as part of standard diagnostic procedures in India.

Sudan

Sudanese paired serum samples (n = 17 pairs) were taken on day of diagnosis of VL and at the end of treatment at 11 days (AmBisome), 17 days (sodium stibogluconate (SSG) + paromomycin), or 30 days (SSG only). These samples were previously tested for IgG1 by ELISA (Bhattacharyya et al., 2014a). Additional Sudanese serum samples used in the present study were unpaired treated individuals (n = 2) taken an unknown time after treatment, and relapse (n = 1). Sudanese Endemic Healthy Control (EHC) samples had previously been tested by the IgG1 ELISA using the same antigen and were negative (Bhattacharyya et al., 2014a) but were not retested here.

TABLE 1 | Indian sample types and total numbers tested by IgG1 assays.

Sample type	Definition	*n*
Pre- and post-treatment pairs, deemed cured	Treated for VL, with improvement in clinical symptoms and no evidence of relapse at any time 6 months after treatment. Samples were taken at or around the start of treatment and at 6 months.	105 pairs
Relapse	VL treated and subsequently relapsed to active disease. Sampled at the time of relapse diagnosis.	33
PKDL	Samples taken at or up to 30 days after diagnosis of PKDL. Parasite infection was confirmed by PCR or a slit-skin smear or biopsy.	63
Asymptomatic	Asymptomatic seropositive, on the basis of DAT and/or rK39 ELISA, without symptoms or history of VL. Progressors (n = 4) developed VL after sampling. Non-progressors (n = 4) did not develop VL during follow-up of at least 3 years.	8
Other diseases	Malaria (n = 5); tuberculosis (n = 5), rheumatoid arthritis (n = 1); dengue (n = 1); multiple myeloma (n = 1).	13
Endemic healthy control	Resident in VL endemic area, seronegative by DAT and rK39 ELISA, no history of VL, healthy.	30

Antigen Production

Whole cell lysate of *L. donovani* strain MHOM/IN/80/DD8 isolated from India, and MHOM/SD/97/LEM3458 isolated from Sudan, was prepared as described previously (Bhattacharyya et al., 2014a). Lysate antigen was used for VL Sero K-SeT development (strain LEM3458), ELISA and western blot (strain DD8). Antigen preparation for western blot strips contained 50 μl of protease inhibitor cocktail (P8340, Sigma, UK) per 1 ml of *L. donovani* cells; centrifugation after sonication was 16,160 × g for 45 min at 4°C.

ELISA for IgG1 Anti *L. donovani*

Duplicate ELISA plates were coated overnight at 4°C with *L. donovani* DD8 strain antigen prepared as above, at 2 μg/ml in coating buffer (35 mM $NaHCO_3$, 15 mM $NaCO_3$, pH 9.6), 100 μl/well. Plates were washed 3 times with phosphate buffered saline (PBS) + 0.05% Tween 20 (PBST) prior to blocking with 200 μl/well PBS + 2% w/v non-fat milk powder (Premier International Foods, UK) (PBSM) for 2 h at 37°C, followed by three PBST washes. Sera/plasma were diluted 1/100 in PBST+ 2% w/v non-fat milk powder (PBSTM) and applied at 100 μl/well, incubated for 1 h at 37°C then washed 6 times with PBST. Mouse anti human IgG1-horse radish peroxidase (HRP)

(ab99774, Abcam, UK) was diluted 1:5,000 in PBSTM and incubated at 100 µl/well, 37°C for 1 h. Plates were washed 6 times with PBST before the addition of 100 µl/well of freshly prepared substrate solution (50 mM citric acid, 50 mM Na_2HPO_4, 2 mM *o*-phenylenediamine HCl, 0.009% H_2O_2). Plates with substrate were incubated in the dark at room temperature for 10–15 min when the reaction was stopped with 100 µl/well of 1 M H_2SO_4 and absorbance read at 490 nm. Each plate contained an EHC sample as a negative serological control for determining the positivity cut-off and a known seropositive VL sample as positive control. All ELISA results reported are the mean A_{490} of duplicate plates.

RDT Production and Use

Whole cell lysate was prepared as described previously (Bhattacharyya et al., 2014a) from *L. donovani* strain MHOM/SD/97/LEM3458. The VL Sero K-SeT lateral flow immunochromatographic tests were developed at Coris BioConcept and consisted of a cassette with a nitrocellulose membrane, a sample pad, a conjugate pad and an absorbent pad, backed with a plastic strip. The nitrocellulose membranes were sensitized with the *L. donovani* lysate antigen and anti-human IgG1-specific antibody labeled with colloidal gold was dried on the conjugate pad. This strip was housed in a plastic cassette with two windows: the smaller buffer well and the long central test window.

To perform the test, 3.5 µl of serum/plasma was applied to the test window at the point indicated by a dot (·) on the cassette, followed immediately by 120 µl of supplied running buffer to the buffer well (**Figure 1**). Devices were incubated flat, at ambient temperature for 15 min before being assessed visually. Any test line at position T was considered a positive result if a control line was also present at position C. Positive test line intensity was assessed visually for samples from pre- and post-treatment VL (**Figure 1**). A subset of samples was tested on different batches of the VL Sero K-SeT. Readers of the RDTs were blinded to all the corresponding ELISA results.

Western Blotting

Western blots were performed to visualize antigen recognition in patients from the different clinical groups, as described in **Supplementary Material S1**. Briefly, tricine SDS-PAGE gels were made as per Schägger (2006). *L. donovani* DD8 lysate was used as antigen with sera/plasma diluted 1 in 300 (Sudan) or 1 in 400 (India) and detection was by mouse anti human IgG1-HRP.

Statistical Analysis

We performed a two-tailed Fisher's exact test on Indian VL Sero K-SeT and IgG1 ELISA data to calculate *p*-values between samples from pre- and matched 6 months post-treatment (deemed cured), separately between post-treatment and relapse, and between post-treatment and PKDL. Cut-off for the IgG1 ELISA was calculated as the mean absorbance of the EHC samples plus 3 standard deviations.

RESULTS

IgG1 Diminishes by 6 Months in Cured VL Patients

Samples taken from Indian patients before or at the outset of therapy, were compared by VL Sero K-SeT and IgG1 ELISA with paired samples taken 6 months later when the individuals were deemed cured. Both IgG1 assay methods showed a statistically significant difference in positivity rate between pre- and post-treatment samples ($p = 0.0162$ and $p < 0.0001$ for ELISA and RDT, respectively) (**Figure 2**). A consistent and strongly significant difference was also observed between cured vs. relapsed samples ($p < 0.0001$), again with both IgG1 assay methods (**Figure 2**).

A subset of pre- and post-treatment of the cured pairs samples ($n = 87$) was tested on different batches of the VL Sero K-SeT, with agreement between individual RDTs of 92.0% (80/87).

Changes in IgG1 Levels by ELISA and VL Sero K-SeT

ELISA absorbance and corresponding VL Sero K-SeT results for individual samples are given in **Supplementary Material S2**. Of the 80 Indian paired samples tested for anti *L. donovani* IgG1 by ELISA, 54 (67.5%) were positive before treatment. Of these, 51/54 declined in titer: 21/51 (representing 26.3% of the total 80) went from positive to negative and 30/51 (representing 37.5% of the total 80) had reduced IgG1 at 6 months when deemed cured (**Figure 2** and **Table 2**). Twenty one (26.3%) paired cured sera were negative by IgG1 ELISA before treatment and remained so at 6 months.

Overall, including those negative at the start, at 6 months after treatment 79/104 (76.0%) were negative by VL Sero K-SeT (**Table 2**). VL Sero K-SeT results were additionally assessed according to whether the Indian 6 month post-treatment (cure) sample had a decreased or not decreased test line intensity compared to the paired pre-treatment sample. Of the 104 paired samples tested by VL Sero K-SeT from deemed cured Indian VL patients, 81 (77.9%) were positive at start of treatment (**Table 2**); of these, 68/81 (84.0%) had either become negative or had a visibly reduced test line intensity at 6 months when deemed cured. Thirteen (12.5%) initially RDT positive individuals showed no visible decrease in RDT band intensity at 6 months, despite being deemed cured, and none became positive from negative.

Ninety four percent of samples positive by ELISA at pre-treatment, decreased in seropositivity; for VL Sero K-SeT, this proportion was 84%. However, at 6 months post-treatment, the ELISA was more likely to remain positive than the RDT, using the cut-off value established for the IgG1 ELISA.

Seventy nine Indian samples were tested by both VL Sero K-SeT and ELISA. Of these samples, the RDT was more likely than the ELISA to be positive with pre-treatment samples (78.5 vs. 67.1%) and negative with 6 month samples (78.5 vs. 53.2%). Of samples which remained positive at 6 months by both methods ($n = 14$), the change in intensity of RDT test line generally mirrored the change in ELISA absorbance value for the same sample. Three of the Indian samples increased markedly in IgG1

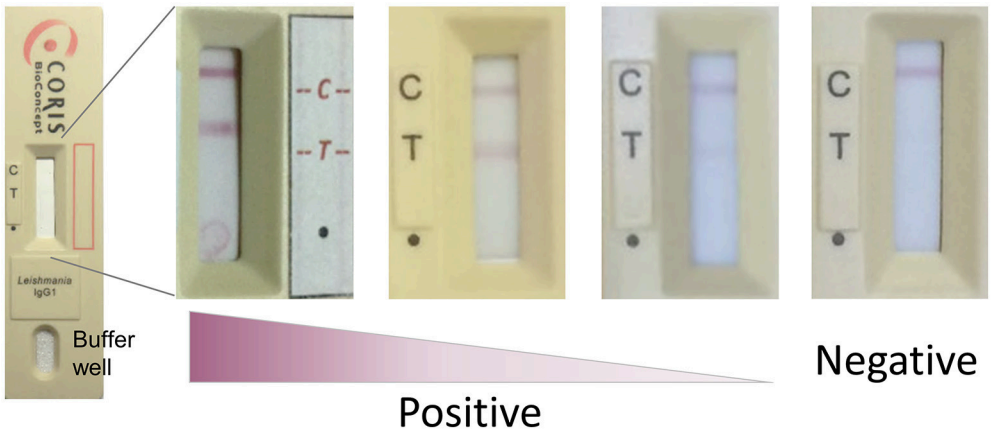


FIGURE 1 | Representative examples of VL Sero K-SeT test line intensity. C, control line; T, test line; dot (•) indicates where sample is applied. Test strip manufacture was identical despite being housed in different cassettes. Image is best viewed in digital, color format.

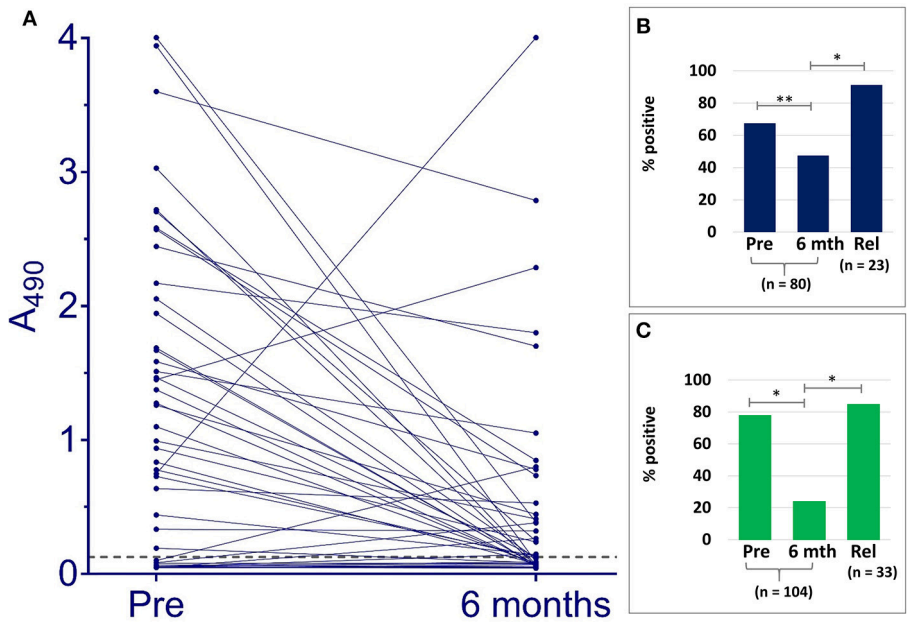


FIGURE 2 | IgG1 anti *L. donovani* assays with Indian VL samples detect differences according to treatment outcome. **(A)** ELISA A_{490} change between paired pre-treatment samples and at 6 months post-treatment when deemed cured. Dashed line indicates cut-off ($A_{490} = 0.128$). Positivity rates with paired pre-treatment and cured samples at 6 months (6 mth), and non-paired relapse (Rel) for **(B)** ELISA, **(C)** VL Sero K-SeT. $^{*}p < 0.0001$, $^{**}p = 0.0162$.

titer by ELISA at 6 months (**Figure 2**). Two of these accorded with a corresponding rise in VL Sero K-SeT test line intensity; for the third sample, both pre-treatment and 6 month RDTs were negative (**Figure 2A**).

Sudanese paired samples taken before and immediately after treatment (11–30 days later) were similarly assessed (**Table 2**). For Sudanese paired samples prior to treatment, 13/17 (76.5%) were positive by VL Sero K-SeT and at completion of treatment, 10/13 (76.9%) had a negative or reduced test line intensity (**Table 2**). If taken as a single time point at the end of treatment, 10/17 (58.8%) Sudanese VL patients had negative VL Sero K-SeT result. Four (23.5%) of the Sudanese treated individuals were negative pre-treatment, similar to the proportion of Indian samples (22.1%). Two additional un-paired treated Sudanese samples were negative by RDT (not shown).

IgG1 Western Blot Confirmed Negative/Declined RDT in Cure

For a subset of 25 of the paired Indian samples, western blots mirrored the VL Sero K-SeT RDT findings, in that IgG1 declined dramatically in all but one VL patient at 6 months follow up after treatment (**Supplementary Material S3**). As with the RDT, the blots showed that samples that were positive and detecting many antigens before treatment had become negative or reduced

TABLE 2 | Change in IgG1 response of pre- and post-treatment paired samples from India and Sudan.

Change in IgG1 response	ELISA A_{490}	VL Sero K-SeT test line intensity	
	India *n* (%)	India *n* (%)	Sudan *n* (%)
Positive to negative	21 (26.3%)	56 (53.8%)	7 (41.2%)
Positive clear decrease	30 (37.5%)	12 (11.5%)	3 (17.6%)
Positive no clear decrease	3 (3.8%)	13 (12.5%)	3 (17.6%)
Negative no change	21 (26.3%)	23 (22.1%)	3 (17.6%)
Negative to positive	5 (6.3%)	0 (0%)	1 (5.9%)
Total	80 (100%)	104 (100%)	17 (100%)

in intensity by 6 months. Corresponding RDT images are shown in **Supplementary Material S4**.

Elevated IgG1 in VL Relapse

For 33 Indian patients for whom we had unpaired samples at the time of relapse, the VL Sero K-SeT was 84.8% (28/33) positive and ELISA 91.3% (21/23) positive, confirming relapse. Of the 23 samples tested by ELISA that were also tested by RDT, 19 gave the same result by both assays (**Supplementary Material S2**). The single available Sudanese relapse sample was IgG1 positive (**Supplementary Material S5**). Twenty five of the Indian samples and the single Sudanese sample were also tested on western blot for IgG1 against *L. donovani* lysate antigen and showed concordance between the RDT and blots (**Supplementary Materials S2**, **S5**). For two of the 33 Indian relapse samples, a paired pre-treatment sample was available. Both individuals were VL Sero K-SeT positive at both time points.

All samples from other diseases, namely malaria, tuberculosis, dengue fever, rheumatoid arthritis, and multiple myeloma were negative by VL Sero K-SeT, as were all samples from endemic healthy controls.

VL Sero K-SeT Can Provide Evidence for PKDL but Not for Its Cure

Of the 63 PKDL samples tested, 49 (77.8%) were positive by VL Sero K-SeT and of the subset of 45 tested by IgG1 ELISA, 43 (95.6%) were positive (**Supplementary Material S2**). A subset of 10 VL Sero K-SeT-positive PKDL samples were tested by western blot, of which 9 showed discernible bands. Images of the blots and their corresponding VL Sero K-SeT RDTs are shown in **Supplementary Material S6**. There was a highly statistically significant difference between post-treatment cured samples at 6 months and PKDL by both VL Sero K-SeT and IgG1 ELISA (Fisher's exact $p < 0.0001$ for both assays).

Seventeen of the 63 individuals with PKDL provided between 1 and 5 additional sequential follow-up samples over intervals ranging from 15 to 365 days post-treatment. These PKDL post-treatment sequential samples retained the initial RDT result in 12/17 (70.6%) cases, decreased in 3/17 (17.6%), increased slightly in one case (5.9%) and varied between positive and negative over time in one case (5.9%).

IgG1 Can Indicate Progression From Asymptomatic Status

When samples from asymptomatic seropositive individuals who later progressed to symptomatic disease (progressors, $n = 4$) were tested on the VL Sero K-SeT, all gave a positive test result (**Supplementary Material S7**). In contrast, 4 individuals who were seropositive but did not develop symptomatic VL were negative by VL Sero K-SeT. Thus, in our limited sample size, elevated IgG1 levels, as detected by VL Sero K-SeT, were associated with progression to symptomatic disease. This result was corroborated by ELISA and western blot (**Supplementary Material S7**).

DISCUSSION

Conventional serology for VL diagnosis relies on detecting the overall IgG response. This has been reported to remain elevated, often for years, after treatment (Bhattarai et al., 2009; Gidwani et al., 2011; Srivastava et al., 2013). This makes current serology unsuitable for timely monitoring of treatment outcome. We have previously found using ELISA that a decreased or negative anti *Leishmania* IgG1 titer at 6 months post-treatment can be indicative of VL cure, whereas elevated IgG1 levels are associated with post-chemotherapeutic relapse (Bhattacharyya et al., 2014a).

Monitoring of Post-treatment Outcomes

Here we used a larger panel of paired samples to assess the IgG1 response as detected by the rapid test, VL Sero K-SeT, where 77.9% of Indian samples were positive before treatment and of these 69.1% had become negative 6 months later when deemed cured (**Table 2**). In total, 76% of 6 month samples were negative, a significant difference from pre-treatment ($p < 0.0001$). Of those still positive at 6 months using this RDT, we found that a diminished test line intensity was also consistent with cure. This decline was corroborated by ELISAs, and despite slight differences in the antigen preparations. We have found no difference in performance of the VL Sero K-SeT when DD8 strain antigen is used instead of LEM3458 (unpublished observations). Thus, the VL Sero K-SeT is a promising innovation, although there is a need to improve further its discriminative capacity.

Sudanese samples declined from positive to negative or decreased VL Sero K-SeT test line intensity in 76.9% of patients immediately after treatment, no more than 30 days after the first sample. This apparently rapid drop in IgG1 was not seen in Indian samples and could be due to the overall lower IgG titer observed in Sudanese samples (Bhattacharyya et al., 2014b; Abass et al., 2015). Thus, a small drop in IgG1 titer could have taken these samples below the detection limit of the VL Sero K-SeT. This may suggest that the VL Sero K-SeT can be used before 6 months to indicate cure or relapse in eastern Africa. The unexpectedly low sensitivity of the VL Sero K-SeT at the start of treatment for both Indian (77.9%) and Sudanese (76.5%) samples does not hinder the subsequent assessment of cure at 6 months, because a negative IgG1 result at 6 months can indicate cure. In addition, we do not propose to use IgG1 assays as a diagnostic for active VL but rather to assist with confirming cure, relapse and

PKDL, all of which currently lack an appropriate diagnostic test. With Indian samples, where there was discrepancy between VL Sero K-SeT and ELISA, the RDT was generally more accurate, being positive with pre-treatment and negative with 6 month samples (**Supplementary Material S2**). As for the Indian sera, the strength of RDT test line intensity broadly corresponded with ELISA signal for an individual sample.

Elevated levels of IgG1 were associated with VL relapse in both assays here for Indian samples. Likewise, the single Sudanese relapse patient was positive by VL Sero K-SeT, whilst 2 treated individuals were negative. We do not know the length of time between treatment and relapse for relapsed individuals (India and Sudan), or the outcome of treated Sudanese individuals. Burza et al. (2014) advised that patient follow-up should be extended from 6 to 12 months as 50–85% of relapses have been found to occur 6 to ≥12 months post-treatment (Rijal et al., 2013; Burza et al., 2014). Our evaluations of a limited number Nepalese relapse samples eluted from filter paper indicated that, although encouraging, elution volumes and conditions need further optimisation before they can be more extensively used with VL Sero K-SeT (data not shown).

We found that in Indian cases who relapsed, the RDT positivity rate was significantly different from 6 month samples from patients deemed cured ($p < 0.0001$). Thus, the VL Sero K-SeT, with Indian samples, can contribute to distinguishing patients deemed cured from those who have relapsed. Of the 13 Indian patients deemed cured at 6 months but who had no clear decrease in VL Sero K-SeT test line intensity (**Table 2**), none is known to have relapsed with VL. However, the quantitative ELISA did detect an IgG1 decrease in these samples, consistent with cure. Apparent relapses might however, occasionally include re-infections given the highly endemic locations (Morales et al., 2002). Although beyond the scope of the present study, the inclusion of parasite genotyping in a future study would be an advantage.

Cases co-infected with HIV and VL were not included in the present study. Serological diagnosis is less reliable in HIV/VL co-infection (Cota et al., 2012; Abass et al., 2015) and the dynamics of IgG1 response in HIV/VL co-infections need to be determined. Other techniques such as a loop mediated isothermal amplification (LAMP) or qPCR detecting parasite DNA might have the potential to discriminate cure from relapse in HIV/VL patients but are currently less accessible than immunological tests (Mukhtar et al., 2018).

PKDL

Indian individuals with PKDL tested here were defined as being with or without a previous history of VL, presenting with a dermal macular, papular or nodular rash often starting on the face with further spread to other parts of the body without loss of sensation. VL Sero K-SeT and IgG1 ELISA results suggest that these assays might contribute to PKDL case detection, as found by a study by Saha et al. (2005), whereas conventional serology may be of limited utility (Gidwani et al., 2011). Our data did not assess the predictive value of IgG1 for development of PKDL.

Where the information was available with our sample set, we did not observe an association between elevated IgG1 and macular vs. polymorphic PKDL presentation, this is in contrast to the report of Mukhopadhyay et al. (2012). For a subset of these PKDL samples, we also tested sequential samples taken up to 1 year after the initial sample. We did not observe a consistent decrease in IgG1 after PKDL treatment.

Progression From Asymptomatic to Active VL

Asymptomatic, seropositive cases outnumber active VL cases (Bern et al., 2007; Ostyn et al., 2011; Hasker et al., 2013; Hirve et al., 2016; Saha et al., 2017) but a proportion are at elevated risk of progressing to active VL (Gidwani et al., 2009; Topno et al., 2010; Ostyn et al., 2011). Asymptomatics have been reported to occasionally have detectable parasites by PCR or culture of blood (le Fichoux et al., 1999; Costa et al., 2002; Bhattarai et al., 2009; Srivastava et al., 2013). Therefore, neither standard seropositivity nor parasitaemia are indicators of progression to clinical disease. Gidwani et al. (2009) found that this progression to VL occurred up to 2 years after serological positivity.

Our limited sample size of seropositive asymptomatic individuals were identified during a community serological screening study, before the present study. Those who later progressed to clinical VL were positive by VL Sero K-SeT and ELISA, whilst those who did not progress were negative by both assays. High titres in both DAT and rK39 ELISA have been indicative of progression in larger studies (Ostyn et al., 2011; Hasker et al., 2014). However, this combination of tests requires laboratory facilities, therefore it would be desirable to have an RDT that could predict progression.

Additional validation of the VL Sero K-SeT should compare larger cohorts who do and do not progress to VL.

Potential Clinical Application of IgG1 Tests

On the basis of the IgG1 responses reported here by VL Sero K-SeT and ELISA, we propose that IgG1 levels may contribute to monitoring the therapeutic outcome of VL, irrespective of whether there is a pre-treatment sample or result. With further development and validation, IgG1 assays, including the VL Sero K-SeT, which can be produced in large-scale at a cost of a few Euros per test, can be used as an adjunct to the clinical assessment of VL status following treatment. A negative, or defined decrease in IgG1 result at 6 months post treatment in India could be supportive of the clinical assessment of cure. Conversely, an un-paired positive or non-decreased paired positive result at 6 months could indicate the need for additional follow-up. In Sudan, the test may be applicable for defining cure before 6 months. A positive IgG1 result in suspected PKDL or relapse could support the presence of leishmaniasis compared to differential diagnoses. Although western blots were supportive of the use of IgG1, we did not specifically assess banding patterns, and do not propose their use in VL diagnosis. However, we are investigating the discriminative diagnostic potential of antigens separated on acrylamide gels.

Recommendations for Further Validation of IgG1 Assays

We propose that a prospective study, with extended follow-up of a larger cohort of treated VL patients, should be used to validate the use of IgG1 ELISA and the VL Sero K-SeT for confirming cure in all endemic areas and defining the optimal time for testing, which may differ between regions. This longer follow-up would also indicate the potential of elevated IgG1 to predict relapse and PKDL and in turn, link these with different treatment regimens. A more extensive study of PKDL is required to determine the potential role of IgG1 in identifying PKDL as distinct from leprosy and fungal skin diseases (Saha et al., 2005; Mondal and Khan, 2011). In addition, use of the IgG1 assays on a much larger panel of seropositive asymptomatic individuals would help to define its role in predicting progression to VL. In all cases, comparison with existing diagnostics, including definitive parasitological methods, would directly assess the advantage of IgG1 assays.

Technical refinement of the VL Sero K-SeT should consider the use of electronic RDT readers to give an objective assessment of test band intensity. In addition, the identification of specific antigens suitable to replace the use of parasite lysate would obviate issues regarding batch-to-batch variation. These developments could improve precision of IgG1 readings and reproducibility. A comparison of whole blood and serum/plasma is also required for point-of-care use, although a study in Bangladesh on various VL RDTs did find high agreement between the two sample types (Ghosh et al., 2015).

CONCLUSION

IgG1 assays, particularly in the VL Sero K-SeT RDT format, may be a useful adjunct in the assessment of VL treatment outcome and diagnosis of PKDL, which have been identified as research priorities for VL (World Health Organization, 2012). With additional refinement and validation, the VL Sero K-SeT and IgG1 ELISA could contribute to life-saving follow-up of treated patients and to control programme monitoring, surveillance, and targeting of strategies for long-term control of VL.

AUTHOR CONTRIBUTIONS

MM conceived and designed the study. TM, TB, and MM wrote the manuscript. TM, TB, CP, BG, SA, MdlR, KH, and HH performed the experiments, collected data. TM, TB, CP, SA, MdlR, KH, and HH analyzed data. OS, PM, QG, CT, BH, AF, OE, AS, BK, NB, SR, SE-S, and SS provided materials. AF, MM, OS, and PM provided feedback on final draft. MM, TB, and TM supervised the project. MB and MM obtained funding.

FUNDING

This work was part of the NIDIAG network research partnership supported by the European Commission under the Health Cooperation Work Programme of the 7th Framework Programme (Grant agreement no. 260260, http://cordis.europa.eu/fp7/home_en.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. TM is funded by the Sir Halley Stewart Trust (http://www.sirhalleystewart.org.uk/). The views expressed within this report are those of the authors and not necessarily those of the Trust. TM was additionally supported by the John Henry Memorial Fund (http://www.johnhenrymf.org/). BH is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 642609. (https://ec.europa.eu/programmes/horizon2020/).

ACKNOWLEDGMENTS

We thank Abhishek Kumar Singh and Keshav Rai for their kind assistance during the laboratory work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2018.00427/full#supplementary-material

Supplementary Material S1 | Detailed SDS-PAGE and western blotting methods.

Supplementary Material S2 | Spreadsheet with ELISA, RDT & blot results for Indian cured paired, PKDL and relapse samples.

Supplementary Material S3 | Images of western blots for Indian paired cured samples.

Supplementary Material S4 | Images of VL Sero K-SeT for Indian paired cured samples, corresponding to those in S3.

Supplementary Material S5 | Images of VL Sero K-SeT and western blots for Indian and Sudanese relapsed samples.

Supplementary Material S6 | Images of VL Sero K-SeT and western blots for Indian PKDL samples.

Supplementary Material S7 | Images of VL Sero K-SeT and western blots for Indian asymptomatic progressors and non-progressors.

REFERENCES

Abass, E., Kang, C., Martinkovic, F., Semiao-Santos, S. J., Sundar, S., Walden, P., et al. (2015). Heterogeneity of *Leishmania donovani* parasites complicates diagnosis of visceral leishmaniasis: comparison of different serological tests in three endemic regions. *PLoS ONE* 10:e0116408. doi: 10.1371/journal.pone.0116408

Atia, A. M., Mumina, A., Tayler-Smith, K., Boulle, P., Alcoba, G., Elhag, M. S., et al. (2015). Sodium stibogluconate and paromomycin for treating visceral leishmaniasis under routine conditions in eastern Sudan. *Trop. Med. Int. Health* 20, 1674–1684. doi: 10.1111/tmi.12603

Bern, C., Haque, R., Chowdhury, R., Ali, M., Kurkjian, K. M., Vaz, L., et al. (2007). The epidemiology of visceral leishmaniasis and asymptomatic leishmanial infection in a highly endemic Bangladeshi village. *Am. J. Trop. Med. Hyg.* 76, 909–914. doi: 10.4269/ajtmh.2007.76.909

Bhattacharyya, T., Ayandeh, A., Falconar, A. K., Sundar, S., El-Safi, S., Gripenberg, M. A., et al. (2014a). IgG1 as a potential biomarker of post-chemotherapeutic

relapse in visceral leishmaniasis, and adaptation to a rapid diagnostic test. *PLoS Negl. Trop. Dis.* 8:e3273. doi: 10.1371/journal.pntd.0003273

Bhattacharyya, T., Bowes, D. E., El-Safi, S., Sundar, S., Falconar, A. K., Singh, O. P., et al. (2014b). Significantly lower anti-Leishmania IgG responses in Sudanese versus Indian visceral leishmaniasis. *PLoS Negl. Trop. Dis.* 8:e2675. doi: 10.1371/journal.pntd.0002675

Bhattarai, N. R., Van der Auwera, G., Khanal, B., De Doncker, S., Rijal, S., Das, M. L., et al. (2009). PCR and direct agglutination as Leishmania infection markers among healthy Nepalese subjects living in areas endemic for kala-azar. *Trop. Med. Int. Health* 14, 404–411. doi: 10.1111/j.1365-3156.2009.02242.x

Bittencourt, A., Silva, N., Straatmann, A., Nunes, V. L., Follador, I., and Badaro, R. (2003). Post-kala-azar dermal leishmaniasis associated with AIDS. *Braz. J. Infect. Dis.* 7, 229–233. doi: 10.1590/S1413-86702003000300009

Boelaert, M., Meheus, F., Sanchez, A., Singh, S. P., Vanlerberghe, V., Picado, A., et al. (2009). The poorest of the poor: a poverty appraisal of households affected by visceral leishmaniasis in Bihar, India. *Trop. Med. Int. Health* 14, 639–644. doi: 10.1111/j.1365-3156.2009.02279.x

Burza, S., Nabi, E., Mahajan, R., Mitra, G., and Lima, M. A. (2013). One-year follow-up of immunocompetent male patients treated with miltefosine for primary visceral leishmaniasis in Bihar, India. *Clin. Infect. Dis.* 57, 1363–1364. doi: 10.1093/cid/cit508

Burza, S., Sinha, P. K., Mahajan, R., Lima, M. A., Mitra, G., Verma, N., et al. (2014). Risk factors for visceral leishmaniasis relapse in immunocompetent patients following treatment with 20 mg/kg liposomal amphotericin B (Ambisome) in Bihar, India. *PLoS Negl. Trop. Dis.* 8:e2536. doi: 10.1371/journal.pntd.0002536

Celesia, B. M., Cacopardo, B., Massimino, D., Gussio, M., Tosto, S., Nunnari, G., et al. (2014). Atypical presentation of PKDL due to Leishmania infantum in an HIV-infected patient with relapsing visceral leishmaniasis. *Case Rep. Infect. Dis.* 2014:370286. doi: 10.1155/2014/370286

Chapman, L. A. C., Dyson, L., Courtenay, O., Chowdhury, R., Bern, C., Medley, G. F., et al. (2015). Quantification of the natural history of visceral leishmaniasis and consequences for control. *Parasit. Vectors* 8:521. doi: 10.1186/s13071-015-1136-3

Costa, C. H., Stewart, J. M., Gomes, R. B., Garcez, L. M., Ramos, P. K., Bozza, M., et al. (2002). Asymptomatic human carriers of *Leishmania chagasi*. *Am. J. Trop. Med. Hyg.* 66, 334–337. doi: 10.4269/ajtmh.2002.66.334

Cota, G. F., de Sousa, M. R., Demarqui, F. N., and Rabello, A. (2012). The diagnostic accuracy of serologic and molecular methods for detecting visceral leishmaniasis in HIV infected patients: meta-analysis. *PLoS Negl. Trop. Dis.* 6:e1665. doi: 10.1371/journal.pntd.0001665

Das, V. N., Pandey, R. N., Siddiqui, N. A., Chapman, L. A. C., Kumar, V., Pandey, K., et al. (2016). Longitudinal study of transmission in households with visceral leishmaniasis, asymptomatic infections and PKDL in highly endemic villages in Bihar, India. *PLoS Negl. Trop. Dis.* 10:e0005196. doi: 10.1371/journal.pntd.0005196

Das, V. N., Ranjan, A., Pandey, K., Singh, D., Verma, N., Das, S., et al. (2012). Clinical epidemiologic profile of a cohort of post-kala-azar dermal leishmaniasis patients in Bihar, India. *Am. J. Trop. Med. Hyg.* 86, 959–961. doi: 10.4269/ajtmh.2012.11-0467

el Hassan, A. M., Ghalib, H. W., Zijlstra, E. E., Eltoum, I. A., Satti, M., Ali, M. S., et al. (1992). Post kala-azar dermal leishmaniasis in the Sudan: clinical features, pathology and treatment. *Trans. R. Soc. Trop. Med. Hyg.* 86, 245–248.

Ghosh, P., Hasnain, M. G., Ghosh, D., Hossain, F., Baker, J., Boelaert, M., et al. (2015). A comparative evaluation of the performance of commercially available rapid immunochromatographic tests for the diagnosis of visceral leishmaniasis in Bangladesh. *Parasit. Vectors* 8:331. doi: 10.1186/s13071-015-0935-x

Gidwani, K., Kumar, R., Rai, M., and Sundar, S. (2009). Longitudinal seroepidemiologic study of visceral leishmaniasis in hyperendemic regions of Bihar, India. *Am. J. Trop. Med. Hyg.* 80, 345–346. doi: 10.4269/ajtmh.2009.80.345

Gidwani, K., Picado, A., Ostyn, B., Singh, S. P., Kumar, R., Khanal, B., et al. (2011). Persistence of *Leishmania donovani* antibodies in past visceral leishmaniasis cases in India. *Clin. Vaccine Immunol.* 18, 346–348. doi: 10.1128/cvi.00473-10

Gorski, S., Collin, S. M., Ritmeijer, K., Keus, K., Gatluak, F., Mueller, M., et al. (2010). Visceral leishmaniasis relapse in Southern Sudan (1999-2007): a retrospective study of risk factors and trends. *PLoS Negl. Trop. Dis.* 4:e705. doi: 10.1371/journal.pntd.0000705

Hasker, E., Kansal, S., Malaviya, P., Gidwani, K., Picado, A., Singh, R. P., et al. (2013). Latent infection with *Leishmania donovani* in highly endemic villages in Bihar, India. *PLoS Negl. Trop. Dis.* 7:e2053. doi: 10.1371/journal.pntd.0002053

Hasker, E., Malaviya, P., Gidwani, K., Picado, A., Ostyn, B., Kansal, S., et al. (2014). Strong association between serological status and probability of progression to clinical visceral leishmaniasis in prospective cohort studies in India and Nepal. *PLoS Negl. Trop. Dis.* 8:e2657. doi: 10.1371/journal.pntd.0002657

Hirve, S., Boelaert, M., Matlashewski, G., Mondal, D., Arana, B., Kroeger, A., et al. (2016). Transmission dynamics of visceral leishmaniasis in the Indian subcontinent - a systematic literature review. *PLoS Negl. Trop. Dis.* 10:e0004896. doi: 10.1371/journal.pntd.0004896

le Fichoux, Y., Quaranta, J. F., Aufeuvre, J. P., Lelievre, A., Marty, P., Suffia, I., et al. (1999). Occurrence of Leishmania infantum parasitemia in asymptomatic blood donors living in an area of endemicity in southern France. *J. Clin. Microbiol.* 37, 1953–1957.

Molina, R., Ghosh, D., Carrillo, E., Monnerat, S., Bern, C., Mondal, D., et al. (2017). Infectivity of post-kala-azar dermal leishmaniasis patients to sand flies: revisiting a proof of concept in the context of the kala-azar elimination program in the Indian subcontinent. *Clin. Infect. Dis.* 65, 150–153. doi: 10.1093/cid/cix245

Mondal, D., Bern, C., Ghosh, D., Rashid, M., Molina, R., and Chowdhury, R., el al. (2018). Quantifying the infectiousness of post-kala-azar dermal leishmaniasis towards sandflies. *Clin Infect Dis.* 891. doi: 10.1093/cid/ciy891. [Epub ahead of print].

Mondal, D., and Khan, M. G. M. (2011). Recent advances in post-kala-azar dermal leishmaniasis. *Curr. Opin. Infect. Dis.* 24, 418–422. doi: 10.1097/QCO.0b013e32834a8ba1

Morales, M. A., Cruz, I., Rubio, J. M., Chicharro, C., Cañavate, C., Laguna, F., et al. (2002). Relapses versus reinfections in patients coinfected with *Leishmania infantum* and Human Immunodeficiency Virus Type 1. *J. Infect. Dis.* 185, 1533–1537. doi: 10.1086/340219

Moulik, S., Chaudhuri, S. J., Sardar, B., Ghosh, M., Saha, B., Das, N. K., et al. (2017). Monitoring of parasite kinetics in Indian post kala azar dermal leishmaniasis. *Clin. Infect. Dis.* 66, 404–410. doi: 10.1093/cid/cix808

Mukhopadhyay, D., Das, N. K., De Sarkar, S., Manna, A., Ganguly, D. N., Barbhuiya, J. N., et al. (2012). Evaluation of serological markers to monitor the disease status of Indian post kala-azar dermal leishmaniasis. *Trans. R. Soc. Trop. Med. Hyg.* 106, 668–676. doi: 10.1016/j.trstmh.2012.07.005

Mukhtar, M., Ali, S. S., Boshara, S. A., Albertini, A., Monnerat, S., Bessell, P., et al. (2018). Sensitive and less invasive confirmatory diagnosis of visceral leishmaniasis in Sudan using loop-mediated isothermal amplification (LAMP). *PLoS Negl. Trop. Dis.* 12:e0006264. doi: 10.1371/journal.pntd.0006264

Musa, A. M., Khalil, E. A., Raheem, M. A., Zijlstra, E. E., Ibrahim, M. E., Elhassan, I. M., et al. (2002). The natural history of Sudanese post-kala-azar dermal leishmaniasis: clinical, immunological and prognostic features. *Ann. Trop. Med. Parasitol.* 96, 765–772. doi: 10.1179/000349802125002211

Ostyn, B., Gidwani, K., Khanal, B., Picado, A., Chappuis, F., Singh, S. P., et al. (2011). Incidence of symptomatic and asymptomatic *Leishmania donovani* infections in high-endemic foci in India and Nepal: a prospective study. *PLoS Negl. Trop. Dis.* 5:e1284. doi: 10.1371/journal.pntd.0001284

Peeling, R. W., Holmes, K. K., Mabey, D., and Ronald, A. (2006). Rapid tests for sexually transmitted infections (STIs): the way forward. *Sex. Transm. Infect.* 82(Suppl. 5) v1–6. doi: 10.1136/sti.2006.024265

Ridolfo, A. L., Gervasoni, C., Antinori, S., Pizzuto, M., Santambrogio, S., Trabattoni, D., et al. (2000). Post-kala-azar dermal leishmaniasis during highly active antiretroviral therapy in an AIDS patient infected with *Leishmania infantum*. *J. Infect.* 40, 199–202. doi: 10.1053/jinf.1999.0630

Rijal, S., Ostyn, B., Uranw, S., Rai, K., Bhattarai, N. R., Dorlo, T. P. C., et al. (2013). Increasing failure of miltefosine in the treatment of kala-azar in Nepal and the potential role of parasite drug resistance, reinfection, or noncompliance. *Clin. Infect. Dis.* 56, 1530–1538. doi: 10.1093/cid/cit102

Roustan, G., Jiménez, J. A., Gutiérrez-Solar, B., Gallego, J. L., Alvar, J., and Patron, M. (1998). Post-kala-azar dermal leishmaniasis with mucosal involvement in a kidney transplant recipient: treatment with liposomal amphotericin B. *Br. J. Dermatol.* 138, 526–528.

Saha, P., Ganguly, S., Chatterjee, M., Das, S. B., Kundu, P. K., Guha, S. K., et al. (2017). Asymptomatic leishmaniasis in kala-azar endemic areas

of Malda district, West Bengal, India. *PLoS Negl. Trop. Dis.* 11:e0005391. doi: 10.1371/journal.pntd.0005391

Saha, S., Mazumdar, T., Anam, K., Ravindran, R., Bairagi, B., Saha, B., et al. (2005). *Leishmania promastigote* membrane antigen-based enzyme-linked immunosorbent assay and immunoblotting for differential diagnosis of Indian post-kala-azar dermal leishmaniasis. *J. Clin. Microbiol.* 43, 1269–1277. doi: 10.1128/jcm.43.3.1269-1277.2005

Sarnoff, R., Desai, J., Desjeux, P., Mittal, A., Topno, R., Siddiqui, N. A., et al. (2010). The economic impact of visceral leishmaniasis on rural households in one endemic district of Bihar, India. *Trop Med Int Health* 15(Suppl. 2), 42–49. doi: 10.1111/j.1365-3156.2010.02516.x

Schägger, H. (2006). Tricine-SDS-PAGE. *Nat. Protoc.* 1, 16–22. doi: 10.1038/nprot.2006.4

Singh, O. P., and Sundar, S. (2015). Developments in diagnosis of visceral leishmaniasis in the elimination era. *J. Parasitol. Res.* 2015:239469. doi: 10.1155/2015/239469

Singh, R. P., Picado, A., Alam, S., Hasker, E., Singh, S. P., Ostyn, B., et al. (2012). Post-kala-azar dermal leishmaniasis in visceral leishmaniasis-endemic communities in Bihar, India. *Trop. Med. Int. Health* 17, 1345–1348. doi: 10.1111/j.1365-3156.2012.03067.x

Srivastava, P., Gidwani, K., Picado, A., Van der Auwera, G., Tiwary, P., Ostyn, B., et al. (2013). Molecular and serological markers of *Leishmania donovani* infection in healthy individuals from endemic areas of Bihar, India. *Trop. Med. Int. Health* 18, 548–554. doi: 10.1111/tmi.12085

Sundar, S., and Rai, M. (2002). Laboratory diagnosis of visceral leishmaniasis. *Clin. Diagn. Lab. Immunol.* 9, 951–958. doi: 10.1128/CDLI.9.5.951-958.2002

Topno, R. K., Das, V. N., Ranjan, A., Pandey, K., Singh, D., Kumar, N., et al. (2010). Asymptomatic infection with visceral leishmaniasis in a disease-endemic area in Bihar, India. *Am. J. Trop. Med. Hyg.* 83, 502–506. doi: 10.4269/ajtmh.2010.09-0345

Trindade, M. A., Silva, L. L., Braz, L. M., Amato, V. S., Naafs, B., and Sotto, M. N. (2015). Post-kala-azar dermal leishmaniasis and leprosy: case report and literature review. *BMC Infect. Dis.* 15:543. doi: 10.1186/s12879-015-1260-x

Uranw, S., Ostyn, B., Rijal, A., Devkota, S., Khanal, B., Menten, J., et al. (2011). Post-kala-azar dermal leishmaniasis in Nepal: a retrospective cohort study (2000-2010). *PLoS Negl. Trop. Dis.* 5:e1433. doi: 10.1371/journal.pntd.0001433

World Health Organization (2010). *Technical Report Series. 949. Control of the Leishmaniases*. WHO Technical Report Series. Geneva: World Health Organization.

World Health Organization (2012). *Technical Report Series. 975. Research Priorities for Chagas Disease, Human African Trypanosomiasis and Leishmaniasis*. WHO Technical Report Series. Geneva: World Health Organization.

World Health Organization (2017). *Leishmaniasis Fact Sheet*. Available online at: http://www.who.int/mediacentre/factsheets/fs375/en/ (Accessed October 1, 2018).

Zijlstra, E. E., Alves, F., Rijal, S., Arana, B., and Alvar, J. (2017). Post-kala-azar dermal leishmaniasis in the Indian subcontinent: a threat to the South-East Asia region kala-azar elimination programme. *PLoS Negl. Trop. Dis.* 11:e0005877. doi: 10.1371/journal.pntd.0005877

Zijlstra, E. E., Musa, A. M., Khalil, E. A., el-Hassan, I. M., and el-Hassan, A. M. (2003). Post-kala-azar dermal leishmaniasis. *Lancet Infect. Dis.* 3, 87–98. doi: 10.1016/S1473-3099(03)00517-6

Conflict of Interest Statement: PM, QG, and CT are employed by Coris BioConcept which developed the VL Sero K-SeT.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Refining wet lab experiments with *in silico* searches: A rational quest for diagnostic peptides in visceral leishmaniasis

Bruno Cesar Bremer Hinckel[1,2]*, Tegwen Marlais[3], Stephanie Airs[3], Tapan Bhattacharyya[3], Hideo Imamura[2], Jean-Claude Dujardin[2], Sayda El-Safi[4], Om Prakash Singh[5], Shyam Sundar[5], Andrew Keith Falconar[6], Bjorn Andersson[7], Sergey Litvinov[8], Michael A. Miles[3], Pascal Mertens[1]

1 Coris BioConcept, Gembloux, Belgium, **2** Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium, **3** Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **4** Faculty of Medicine, University of Khartoum, Khartoum, Sudan, **5** Department of Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India, **6** Departamento de Medicina, Universidad del Norte, Barranquilla, Colombia, **7** Department of Cell- and Molecular Biology, Karolinska Institutet, Stockholm, Sweden, **8** Aptum Biologics Ltd., Southampton, United Kingdom

* bruno.hinckel@corisbio.com

OPEN ACCESS

Citation: Bremer Hinckel BC, Marlais T, Airs S, Bhattacharyya T, Imamura H, Dujardin J-C, et al. (2019) Refining wet lab experiments with *in silico* searches: A rational quest for diagnostic peptides in visceral leishmaniasis. PLoS Negl Trop Dis 13(5): e0007353. https://doi.org/10.1371/journal.pntd.0007353

Editor: Sitara SR Ajjampur, Christian Medical College, Vellore, INDIA

Received: June 28, 2018

Accepted: April 1, 2019

Published: May 6, 2019

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: BCBH was part of the Euroleish.net training network (http://www.euroleish.net/) and has received funding from the European Commission Horizon 2020 Research and Innovation Programme under Marie-Sklodowska Curie grant agreement number 642609. TM is funded by the Sir Halley Stewart Trust

Abstract

Background

The search for diagnostic biomarkers has been profiting from a growing number of high quality sequenced genomes and freely available bioinformatic tools. These can be combined with wet lab experiments for a rational search. Improved, point-of-care diagnostic tests for visceral leishmaniasis (VL), early case detection and surveillance are required. Previous investigations demonstrated the potential of IgG1 as a biomarker for monitoring clinical status in rapid diagnostic tests (RDTs), although using a crude lysate antigen (CLA) as capturing antigen. Replacing the CLA by specific antigens would lead to more robust RDTs.

Methodology

Immunoblots revealed *L. donovani* protein bands detected by IgG1 from VL patients. Upon confident identification of these antigens by mass spectrometry (MS), we searched for evidence of constitutive protein expression and presence of antigenic domains or high accessibility to B-cells. Selected candidates had their linear epitopes mapped with *in silico* algorithms. Multiple high-scoring predicted epitopes from the shortlisted proteins were screened in peptide arrays. The most promising candidate was tested in RDT prototypes using VL and nonendemic healthy control (NEHC) patient sera.

Results

Over 90% of the proteins identified from the immunoblots did not satisfy the selection criteria and were excluded from the downstream epitope mapping. Screening of predicted epitope peptides from the shortlisted proteins identified the most reactive, for which the sensitivity

(http://www.sirhalleystewart.org.uk/). The views expressed within this report are those of the authors and not necessarily those of the Trust. TM was additionally supported by the John Henry Memorial Fund (http://www.johnhenrymf.org/). This work was part of the NIDIAG network research partnership supported by the European Commission under the Health Cooperation Work Programme of the 7th Framework Programme (Grant agreement no. 60260, http://cordis.europa.eu/fp7/home_en.html). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We are grateful to Albert Picado, coordinator of Euroleish.net (http://www.euroleish.net/) and to Marleen Boelaert coordinator of EC NIDIAG, for their initiative in securing funding to establish those research networks.

Competing interests: BCBH was and PM is employee of Coris BioConcept and they do not have any share in the company. Sergey Litvinov is an employee of Aptum Biologics Ltd. which is an official partner institution in the Consortium that produced this work. Aptum Biologics Ltd. has and provided to the Consortium the software for epitope prediction, and beyond this has no consultancy, employment or any other forms of financial benefits from this work, nor affiliation with whatever potential intellectual property that may arise from this particular work. All other authors have declared that no competing interests exist.

for IgG1 was 84% (95% CI 60—97%) with Sudanese VL sera on RDT prototypes. None of the sera from NEHCs were positive.

Conclusion

We employed *in silico* searches to reduce drastically the output of wet lab experiments, focusing on promising candidates containing selected protein features. By predicting epitopes *in silico* we screened a large number of peptides using arrays, identifying the most promising one, for which IgG1 sensitivity and specificity, with limited sample size, supported this proof of concept strategy for diagnostics discovery, which can be applied to the development of more robust IgG1 RDTs for monitoring clinical status in VL.

Author summary

Visceral leishmaniasis (VL) is a neglected tropical disease caused by protozoan parasites of the *Leishmania donovani* complex. Without treatment, VL is fatal. Although diagnostic techniques, mainly based on the detection of anti-*Leishmania* antibodies are available, invasive procedures such as microscopy from spleen or bone marrow aspirates are still required for the diagnosis of seronegative VL suspects, for the detection of recurrent cases and to confirm cure after successful treatment. Previous investigations showed the potential of IgG1 as a biomarker of post-chemotherapeutic relapse for VL in rapid diagnostic tests (RDTs) sensitised with crude lysate antigen (CLA). Here we employed *in silico* tools to search for desired protein features in a large number of *L. donovani* antigens detected by human IgG1 in western blots. We then employed prediction algorithms to profile epitopes from the shortlisted proteins. We screened a panel of high-scoring peptides in a high-throughput manner using arrays, with low reagent consumption. The most reactive peptide was adapted to RDTs, showing promising results of both sensitivity and specificity. This peptide has the potential of replacing the CLAs in IgG1 RDTs. Thus we believe that *in silico* tools can be used to optimise wet lab experiments for a rational search of biomarkers.

Introduction

The leishmaniases comprise a group of vector-borne diseases caused by parasites of the genus *Leishmania*. The visceral form—visceral leishmaniasis (VL), also known as kala-azar (Hindi for 'black fever') affects internal organs such as liver, spleen and bone marrow, and leads to death if left untreated. VL is caused by parasites of the *Leishmania donovani* complex and in 2015 over 23.000 new cases were reported to the world health organization (WHO) worldwide [1] while an overall case-fatality rate of 10% has been estimated [2]. VL is diagnosed by a combination of clinical symptoms, including prolonged fever, weight loss, hepatosplenomegaly and malaise, and the detection of parasite-specific immunoglobulins (Igs). The recombinant protein rK39, a fragment of a *Leishmania infantum* kinesin-like gene, was described in 1993 [3] and remains the most widely used antigen for VL serodiagnosis. Nonetheless, novel and improved antigens are still required to complement the use of the rK39, to improve sensitivity in eastern Africa [4, 5] and to determine cure after successful chemotherapy (versus relapse).

The recombinant chimeric protein rK28, which is a derivative of the rK39, incorporating the first two repeats of a Sudanese *L. donovani* kinesin flanked by proteins HASPB1 and HASPB2, was engineered to address low sensitivity values reported from eastern Africa. Although a slight better performance compared to the rK39, both in terms of sensitivity and specificity, was reported with Sudanese VL patients [6], important variations in sensitivity amongst different rK39 rapid diagnostic test (RDT) manufacturers were also reported from that region [7]. Moreover, false positive rates as high as 19.5% in eastern African patients were described [8], meaning that further investigations are still required to confirm the relevance of the rK28 as well as its superiority over the rK39-based diagnostic tests for VL in eastern African patients. Due to Igs persistence even after complete parasite clearance [9–13], neither the rK39 nor the rK28 commercial diagnostic kits can be employed to determine cure after successful chemotherapy.

There has been a dramatic reduction in genome sequencing costs, accompanied by an exponential increase in the number of available sequences in public repositories [14, 15]. The first *Leishmania* spp. genome sequencing was completed in 2005 [16]. More recently, the advent of high-throughput technologies has enabled the completion of the whole genome sequencing of *L. donovani*, the causative agent of VL in the Indian subcontinent [17]. Moreover, the availability of next-generation sequencing made it possible to perform whole transcriptome sequencing. RNA sequencing (RNA-seq) generates data on the transcriptome at a specific stage of a pathogen life cycle or in a specific culture condition of an organism. The growing number of available, high quality, whole genome sequences has become a central element in the area of comparative genomics, which has contributed greatly to a better understanding of multiple aspects of the leishmaniases, including determinants of disease phenotype [18], mode of action of drugs [19] and parasite biology [20].

Igs are a major component of the immune system. They bind specific regions of pathogens' proteins (epitopes), tagging them for clearance by the immune system. Epitopes can be divided into linear (a continuous stretch of amino acids (AAs)) and discontinuous (where non-proximal residues are brought together by protein folding) and can be identified by functional and structural studies (e.g. X-ray crystallography), while *in silico* epitope prediction algorithms are gaining popularity. The early prediction methods of linear B-cell epitopes were mainly based on propensity scales [21]. More recently, machine learning methods have been employed to improve prediction performance [22–24]. The prediction of discontinuous epitopes still depends on the availability of 3D structures. *In silico* tools can also be used to locate antigenic domains from DNA sequences or peptide sequences influencing protein localisation within cells.

Diagnostic research can incorporate genomics, transcriptomics, proteomics as well as bioinformatic prediction algorithms for the discovery of new biomarkers. Such a systematic 'omics' approach has been applied alone for the discovery of vaccine candidates [25–27] as well as for diagnostic biomarkers [28]. These *in silico* searches can also be applied downstream of wet lab experiments, refining their output for a rationalised search.

Mass spectrometry (MS) can be used to identify proteins from wet lab experiments (e.g. immunoblots). Comparative genomics enables identification of species-specific genes, while levels of life-stage specific proteins can be estimated using publicly available RNA-seq data. *In silico* prediction algorithms can be employed to infer protein localisation, search for antigenic domains as well as to predict linear B-cell epitopes. Synthetic peptides can then be incorporated into arrays, enabling the screening of a large number of candidates in pilot serological assays. Promising candidates can be adapted and tested in RDTs, a format suitable for field use.

Here we employed comparative genomics as well as *in silico* algorithms to reduce drastically the excessive number of *L. donovani* protein candidates recognised by human IgG1, which has been shown to be a potential indicator of post-chemotherapeutic relapse in VL, using RDT prototypes sensitised with a *L. donovani* crude lysate antigen (CLA) [29]. By screening a large number of predicted epitopes from selected candidate antigens containing desired protein features, we identified one peptide specifically recognised by IgG1 in arrays. This peptide was also tested in prototype RDTs with VL patient sera as well as sera from nonendemic healthy controls (NEHCs), showing promising results of both sensitivity and specificity on a small sample size. Thus we propose that this approach is a valid proof of concept for the discovery of diagnostic peptides, which can be used to replace the CLA in IgG1 RDTs, leading to the production of cheaper and more robust RDTs for monitoring clinical status in VL.

Methods

Ethics statement

In India, the collection of samples was approved by the Ethics Committee of Banaras Hindu University, Varanasi while in Sudan the approval for collection and research was granted by the Ethical Research Committee, Faculty of Medicine, University of Khartoum and the National Health Research Ethics Committee, Federal Ministry of Health, Sudan. Written informed consent was obtained from all adult subjects included in the study or from the parents or guardians of individuals less than 18 years of age. This research was also approved, as part of the EC NIDIAG project, by the London School of Hygiene and Tropical Medicine Ethics Committee as well as by the Ethics Committee of the Antwerp University Hospital.

Sources of sera

A detailed description of all serum samples used in this work can be found below while a summary is shown in Table 1.

India. Indian sera were selected from archived samples, collected after 2007 from active VL, relapsed and endemic healthy controls (EHCs), all from the endemic region of Muzaffarpur, Bihar. Active VL cases had been diagnosed by positive rK39 and/or DAT serology and parasitologically by microscopy of splenic aspirates prior to the present study. Relapses were diagnosed clinically. All Indian samples were HIV negative.

Sudan. Sudanese serum samples were collected in 2011 and 2012 from the VL-endemic region of Gedaref in eastern Sudan. VL cases had been diagnosed by microscopy of bone marrow or lymph node aspirates in conjunction with serological assays (rK39 or rK28) and were all HIV negative. These diagnoses were carried out according to their respective national procedures, prior to the present study.

Table 1. Serum samples used in this study. Details of all serum samples used in the western blots, peptide arrays and RDT prototypes.

Sample type	Region	Definitions	n
Active VL	India	Samples were taken at or around onset of treatment (D0, D7 or D15)	25
Relapse		VL treated and subsequently relapsed to active disease. Sampled at the time of relapse diagnosis	26
		Total India	51
Active VL	Sudan	Samples were taken at or around onset of treatment (D0, D7 or D15)	19
NEHC	France	Samples from healthy donors living in region where no transmission of VL is documented	10

https://doi.org/10.1371/journal.pntd.0007353.t001

NEHC. Sera was obtained from whole blood collected from *Etablissement Français du Sang*, France. All donors were informed of the use of the blood for research purpose and gave their informed consent for the purpose of scientific research use and that all national laws and ethical principles were fulfilled. The samples were fully anonymised. All samples were certified to be negative for the following transmissible diseases: HIV-1, HIV-2, HCV, HBV, HTLV I, HTLV II and syphilis.

Wet lab identification of *L. donovani* antigens detected by IgG1

L. donovani promastigote antigens recognised by human IgG1 were profiled using western blots as described in the repository protocols.io (dx.doi.org/10.17504/protocols.io.u8rezv6). Briefly, the cytosolic proteins from a *L. donovani* (strain MHOM/IN/80/DD8) whole cell lysate were separated by SDS-PAGE, blotted onto a nitrocellulose (NC) membrane and sliced into individual strips, to be immunoassayed with sera from individual patients from India. Upon visual identification of bands of interest on the individual strips using HRP-labelled anti-human IgG1 as secondary antibody, new gels were run and corresponding immunogenic bands were excised and analysed by mass spectrometry (MS) (syn. liquid chromatography tandem mass spectrometry (LC-MS/MS)), according to the methods described in S1 Text. Protein hits were identified by matching peptide fragments against the *L. donovani* reference genome-derived proteome only (ENA accession nos. FR799588-FR799623 [17]), henceforth referred to as LdBPK v1 genome. These proteins were submitted to the *in silico* filter detailed below.

In silico refinement of wet lab output

Desired protein features were searched *in silico* in order to decrease the number of candidates from the MS output while shortlisting the proteins more likely to be antigenic, to have their B-cell epitopes mapped with *in silico* prediction algorithms. Some protein features were searched in series while others in parallel, meaning that all shortlisted proteins satisfy the criteria described in either branch as shown in Fig 1. These shortlisted v1 IDs were then merged with the IDs from an improved LdBPK282 reference genome (ENA accession number ERP022358, henceforth referred to as LdBPK v2) according to the correspondence in Data Set 2 from [30]. Hypothetical proteins were excluded, unless containing tandem repeats (TRs). Additional information on each step of this *in silico* filter are detailed in the following subsections.

Gene expression analysis—RNA-seq. Transcriptomic data on seven promastigotes and four amastigotes, representing different parasite life stages was obtained. In the present study the improved LdBPK282 reference genome was used [30]. *Leishmania* chromosome copy number is known to be variable, especially in cultivated promastigotes, which can potentially change within a few passages. To quantify the RNA-seq depth levels, which were not affected by somy variability, we used normalised haploid depth (HAP). The average normalised HAP of all genes were ranked for promastigotes and amastigotes separately and only those genes whose values were greater than the first quartile (HAP>Q1) in both parasite life stages were considered as constitutively expressed and and kept for downstream analysis.

Confident MS identification. We matched the peptides identified by MS against the LdBPK v1 genome only. Each protein hit identified by MS with the software MASCOT [31] has its own score defined by:

$$\text{MASCOT score} = -10 \times \log_{10}(P)$$

The score converted the probability that the observed match was not a random event (P) into an ascending scale where the lowest score is the most unreliable match, and higher scores

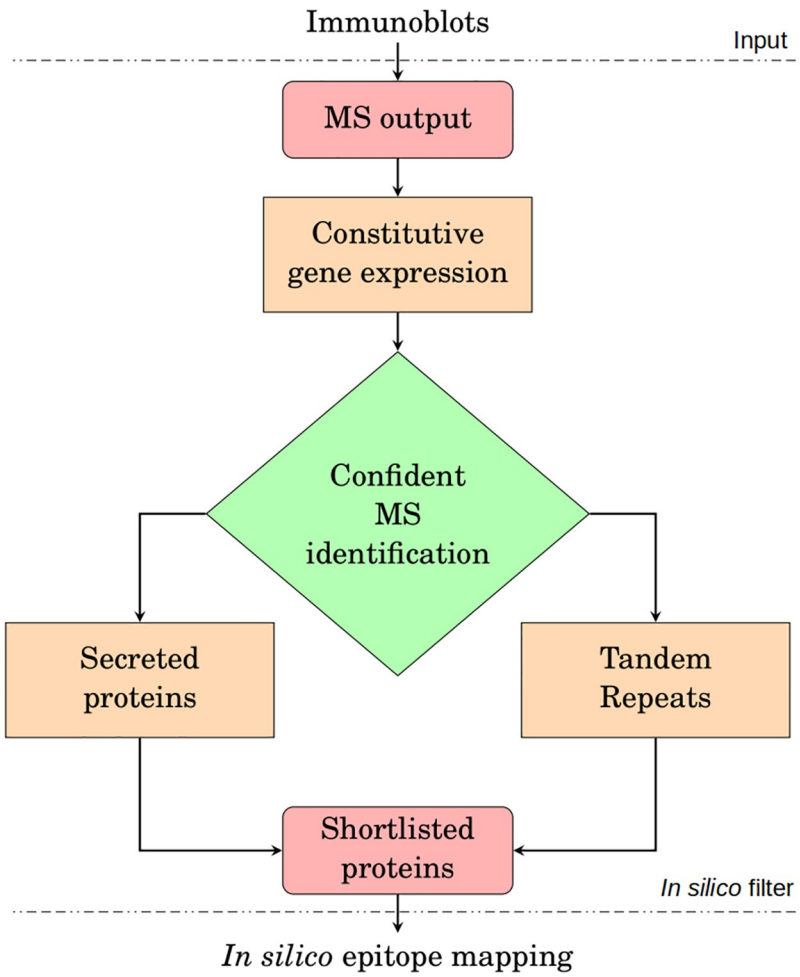


Fig 1. *In silico* filter applied to select desired protein features. Selected protein properties (yellow and green nodes) were searched *in silico* on all protein hits identified by MS from selected IgG1 sero-reactive western blot bands (*MS output*—top node), with the ultimate goal of shortlisting protein candidates most likely to be antigenic (*Shortlisted proteins*—bottom node).

https://doi.org/10.1371/journal.pntd.0007353.g001

indicate more reliability [32]. A score of 100 was chosen as the cut-off value to minimise the chances that any match could be found by chance.

Protein localisation. Exported *L. donovani* proteins, identified as part of exosomes by experimental approaches [33] and/or by the presence of a signal peptide were considered highly accessible to B-cells and thus selected for further analysis. The presence of a signal peptide sequence was searched with the stand-alone version of the software SignalP v4.0 [34] with the option 'noTM'. All other optional parameters were default.

Antigenic domain selection. *L. donovani* proteins harbouring TRs include known B-cell antigens such as rK39 and rK28, therefore the presence of this domain was included as a criterion correlating with antigenicity. The protein coding DNA sequences were scanned for the presence of TRs using the stand-alone version of the program Tandem Repeats Finder [35] (v4.09). All program parameters were default.

Epitope mapping from shortlisted proteins

Four *in silico* algorithms were used independently to define B-cell epitopes: ABCPred [23], Bepipred1.0 [22], EpiQuest-B [36] and lbtope [37]. For all epitopes predicted from each algorithm individually (generally a 15 or 16mer), a 'core sequence', stretching from the fifth to the eleventh AA residue was defined and scanned against the rest of the predicted antigenic peptides, in order to identify peptides of similar sequences. These peptides sharing core sequences with other predicted peptides at any position in the AA sequence are henceforth referred as partially overlapping peptides.

While ABCPred, lbtope and EpiQuest-B were run on default values, predicting antigenic sequences of 16, 15 and 16 AA residues, respectively, shorter sequences with high antigenic scores were also considered for EpiQuest-B. The highest scoring peptides predicted from Bepipred were obtained by calculating the area under the curve (AUC) from plots of AA residue position vs. individual residue score for all possible combinations of 16mers with one AA offset, considering all contiguous predicted antigenic sequences from all submitted proteins. Additional details about the prediction algorithms employed can be found in Table 2.

Peptide synthesis

Desalted peptides i.e. of varying purity grades were synthesised with an N-terminal biotin molecule linked via AAG spacer so that they could bind to NC membranes (1620215, Bio-Rad, USA), which had been previously soaked with neutralite avidin (NLA) (NLA30, e-proteins, Belgium) and dried at 50˚C. Lyophilised peptides were dissolved according to standard protocols [38]. Where the solvent mixture in which the peptides were dissolved exhibited stable background values, the peptide concentration was calculated based on their molar absorbances, measured either at 280 nm, for tyrosine (Y) and tryptophan (W) containing peptides [39], or at 205 nm [40], for those without W or Y. For the remaining peptides, the concentration was calculated by dividing the peptide amount reported by the manufacturer by the volume of the solvent mixture in which they were dissolved.

Any promising candidate identified from the pilot peptide screening was synthesised with higher purity grade (>90%) and with a N-terminal biotin—polyethylene glycol-glycine spacer in order to improve water solubility as well as to increase rotation and ensure that the full amino acid sequence could freely interact with Igs, as opposed to being adsorbed onto the solid support and therefore unavailable for recognition by Ig [41].

Table 2. Information on the B-cell epitope prediction algorithms employed. A total of 80 peptides with high antigenic scores, independently predicted from four different B-cell prediction algorithms, were selected for synthesis. AA: amino acid; AUC: area under the curve.

Algorithm	Details	n
ABCPred	Predicts antigenic regions of fixed even length and assigns a score, ranging from zero to one, for the whole predicted k-mer.	20
lbtope	Scans all possible 15mers with one AA offset on all the submitted sequences. A numeric score as well as the probability of correctnesses of each prediction are outputted.	20
EpiQuest-B	Calculates individual scores per AA residue (AGR) and the final peptides scores are ranked by the relative immunogenicity index, which corresponds to the curve (AUC) in a plot residue position vs. AGR.	20
Bepipred	Assigns a score for each AA residue on the submitted sequences. Predicted epitope regions can be of any length $\geq$ 1 AA with a score greater than the set cut-off (0.35 for default values). The antigenicity of the 16mers was calculated from the AUC in a plot residue position vs. score.	20
	Total	80

https://doi.org/10.1371/journal.pntd.0007353.t002

Pilot screening of peptides using arrays

Array production. sciFLEXARRAYER (Scienion AG, Germany) with a PDC 70 piezo dispense capillary (type 3 coating, P-2030/ S-6051, Scienion AG, Germany) was used to spot 7 nl of the selected peptides at the required concentrations in multiple replicates onto NC membranes previously soaked with NLA and dried at 50˚C. rK39 (RAG0061, Rekom Biotech, Spain) and a whole *L. donovani* lysate, obtained as described in [29], were both spotted on the diagonals as positive controls as well as for orientation purposes. As negative controls, we spotted a peptide specific for *T. cruzi* [42]. All controls were spotted at 0.1 mg/ml. Arrays were incubated overnight (ON) at 50˚C upon completion of the spotting, to be hybridised with serum the next day.

Hybridisation with serum/ image acquisition. Peptide arrays were blocked with phosphate-buffered saline (PBS) + 3% bovine serum albumin (BSA) (PBSB) ON at 4˚C or for 2h at room temperature (RT), followed by three 5 minute washes of PBS + 0.05% Tween 20 (PBST). In order to assess the peptide recognition by sera from VL patients and NEHCs, separate arrays were hybridised with pooled Sudanese serum samples positive for VL and with commercial pooled NEHC sera (S1-100ML, EMD Millipore Corporation, USA), respectively, diluted in PBST + 3% BSA (PBSTB), for 1h. After five 5 minute washes with PBST, they were incubated with fluorescent mouse anti-Human IgG1 Fc—Alexa Fluor 488 antibody (AB) (A-10631, ThermoFisher) diluted 1:1000 in PBSTB, for 1h at RT. Followed by five 5 minute washes with PBST, the arrays were incubated at 50˚C until completely dry.

Images were acquired at 500ms and 20 dB gain with a digital CCD camera (ORCA-R^2, Hamamatsu, Japan) coupled to a fluorescence microscope (model BX53, Olympus, Germany) equipped with filter cube U-FGFP (N271350, Olympus, Germany). The fluorescence of the spots was quantified using the software cellSens Dimensions v.1.7 (Olympus GmbH, Germany)

Adaptation to an RDT

RDTs were composed of a NC strip sensitised with NLA at 3.5 mg/ml and a conjugate pad, impregnated with anti-human IgG1-specific antibody conjugated to 40nm gold beads (nanoQ, Belgium). The strip was either housed within a plastic cassette, with a buffer application well and a test/reading window (cassette), or not (dipstick). Prior to application on the RDT, equal volumes of serum and the biotinylated peptide at stock concentration were mixed and incubated at 37˚C for 15 minutes. 3.5 μl of the mix was then pipetted onto the sample application zone, just above the top of the sample pad and at the bottom end of the NC strip. 150 μl of buffer solution was dispensed into the buffer application well (cassette) or dipsticks were dipped into a recipient vessel filled with same volume of running buffer. After 15 minute incubation, a test was deemed valid if a clear red control band was present in line with the 'C', and deemed positive if a second band was present in line with the 'T'. If no band was visible at the 'T' and a clear control band at 'C' developed, then the test was deemed 'negative'. Invalid tests were assessed by the absence of the red control band 'C' or by any other clear migration problem.

Statistical analysis

All statistical analyses were performed using the computing environment R [43]. The final fluorescence of each spot on the peptide arrays (spot fluorescence) was expressed in arbitrary units and calculated by subtracting the mean background value of a given acquired image from the individual fluorescence values. Mean background value was calculated from the mean of six-eight fluorescence measures in random positions on the arrays where no

fluorescence signal was detected. The peptide specificity was expressed as 95% Fiellers confidence intervals (CIs) of the ratio between the mean spot fluorescence of eight replicates from the arrays hybridised with VL serum and that of arrays hybridised with NEHC serum (ratio VL/NEHC). CIs were calculated with the mratios R package [44].

RDT results were compared with defined clinical status to establish sensitivity and 95% CIs with Clopper-Pearson exact method using the PropCIs R package [45].

Results

MS of selected antigenic bands identified over 1300 hits

The development of the western blot strips immunoassayed with active Indian VL serum samples and relapsed patients revealed the protein bands detected by human IgG1. New gels were run in order to get three selected antigenic bands (B1, B2 and B3) excised (Fig 2), to ultimately have their constituent proteins revealed by MS. Criteria to select the antigenic bands were (I) strong recognition, visually assessed by their intensity, (II) ubiquitous recognition i.e. recognised by the majority of the patients and (III) confident match between the bands from the western blots and the new gels.

A total of 1357 putative *L. donovani* proteins were identified from the three selected antigenic bands, with some identified in multiple bands (redundant matches). All these candidate antigen proteins were submitted to the *in silico* filter (Fig 1) to reduce the number of promising candidates while restricting to those possessing protein features that correlate with antigenicity.

The *in silico* filter excluded over 90% of the uniquely identified proteins

The initial 1357 IDs from the three selected antigenic bands (Fig 2) contained 678 unique IDs that matched the LdBPK v1 genome (we only matched the peptides identified by MS against the LdBPK v1 genome, such that the analysis returned *L. donovani* protein hits only). Out of these, 538 were considered to be constitutively expressed in both amastigote and promastigote

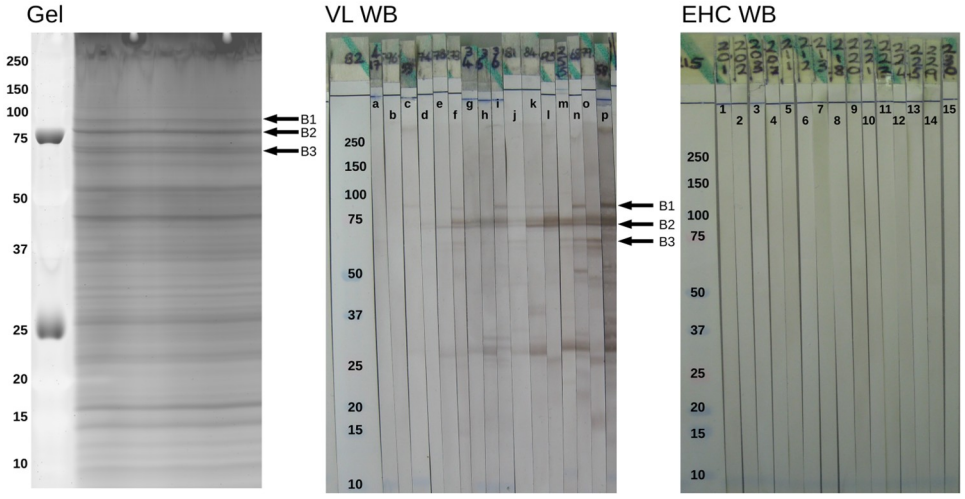


Fig 2. Gel and western blot strips immunoassayed with sera from Indian VL and EHC patients. The western blot strips incubated with individual active VL sera revealed the *L. donovani* protein bands reacting with IgG1 from Indian patients (VL WB). Three antigenic bands, B1, B2 and B3 were excised from gels to have their constituent proteins identified by MS (Gel). The strips incubated with EHC sera did not develop any band (EHC WB). Molecular mass markers are given in kDa.

https://doi.org/10.1371/journal.pntd.0007353.g002

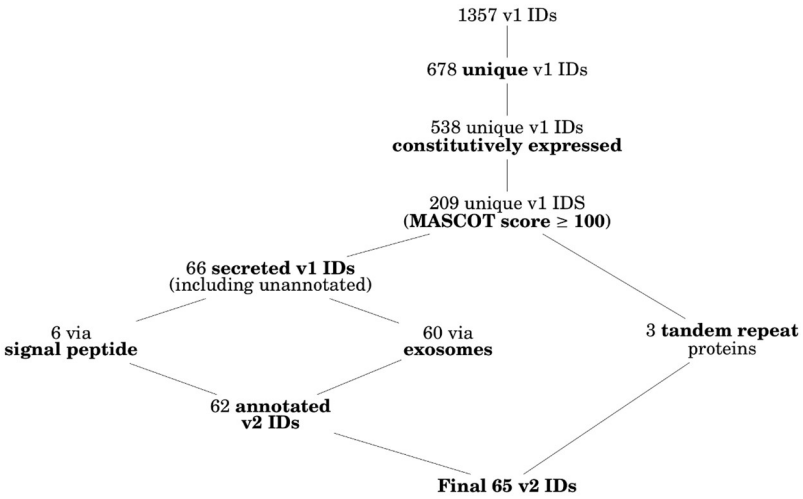


Fig 3. Selected proteins for *in silico* epitope mapping. 65 proteins ('Final 65 v2 IDs') satisfied the criteria in either branch shown in Fig 1 and were shortlisted to have their epitopes mapped with multiple *in silico* B-cell prediction algorithms. In **bold** are the features added after each step of the filter.

https://doi.org/10.1371/journal.pntd.0007353.g003

life stages following gene expression analysis, of which 209 had a MASCOT score $\geq$ 100. Sixty six of these candidates were found to be secretory, 60 of which as part of exosomes while 6 where actively secreted to the outside of the cell via the presence of a signal peptide. These 66 unique v1 IDs were merged with the corresponding IDs from the improved LdBPK v2 reference genome, making 62 annotated extracellular and constitutively expressed proteins, identified with confidence by MS, as shown in Fig 3.

Further three unique candidates harbouring TR domains were found to be constitutively expressed and were identified with confidence by MS from the excised bands. Although none of the three was found to be extracellular, they were kept for downstream epitope mapping as such proteins containing TR domains are often B-cell antigens [46]. A single TR protein satisfying all the criteria described remained unannotated in the improved LdBPK v2 reference genome. Results including the number of hits satisfying all the criteria applied are schematically shown in Fig 3 while the detailed list with all the 65 shortlisted proteins can be found in S1 Table.

In silico selection of B-cell epitopes

All 65 shortlisted proteins were profiled with the B-cell prediction algorithms as detailed in the methods section. Considering partially overlapping peptides, only that of highest score was shortlisted. The top 20 scoring peptides, with no shared core sequences within a given algorithm, were selected for synthesis, making a total of 80 high scoring peptides from four different prediction algorithms, as shown in S2 Table.

Peptide antigenicity screening

Initial hybridisation revealed the most antigenic peptide. A first batch of peptide arrays was produced in order to obtain insights about the reactivity of the peptides with human IgG1. The visual inspection of the imaged arrays hybridised with both pooled serum groups (Sudanese VL and NEHC) diluted 1:100 revealed that a high number of peptides spotted at their stock concentrations reacted with both serum pools (Fig 4).

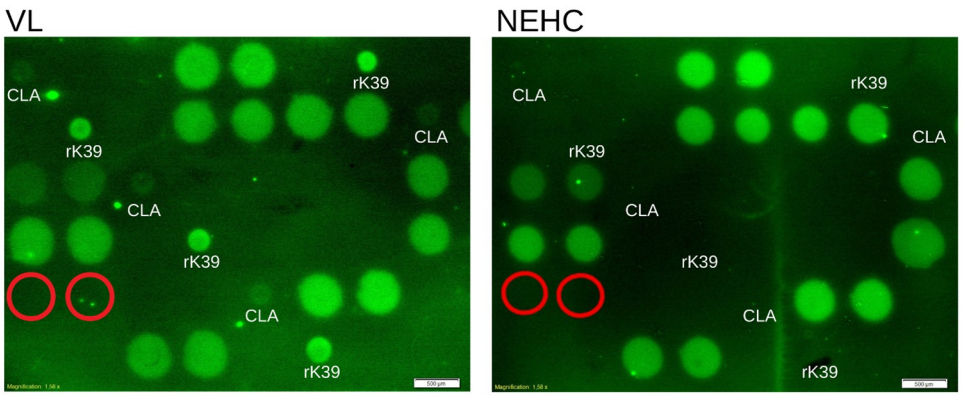


Fig 4. Peptides reactivity with a pool of Sudanese VL and NEHC sera diluted at 1:100 showed to be unspecific. [VL]: Section of an array hybridised with pooled Sudanese VL serum at 1:100 dilution. rk39 and a whole *L.donovani* lysate (CLA) were spotted on the diagonal as positive controls as well as for orientation purposes. [NEHC]: the same array section hybridised with a pool of NEHC sera at same dilution. The red circles indicate the spotting position of a peptide specific for *T. cruzi* [42], employed here as negative control. Both images indicate the array section where the peptide epitopes predicted from the lbtope algorithm were spotted.

https://doi.org/10.1371/journal.pntd.0007353.g004

Hybridisation of the same batch of arrays with pooled serum at a higher dilution (1:200) showed that a reduced number of peptides reacted with both serum groups, visually revealing the most reactive peptide, as shown in Fig 5. Details of the most reactive peptide found, henceforth referred as EpQ11, are shown in Table 3.

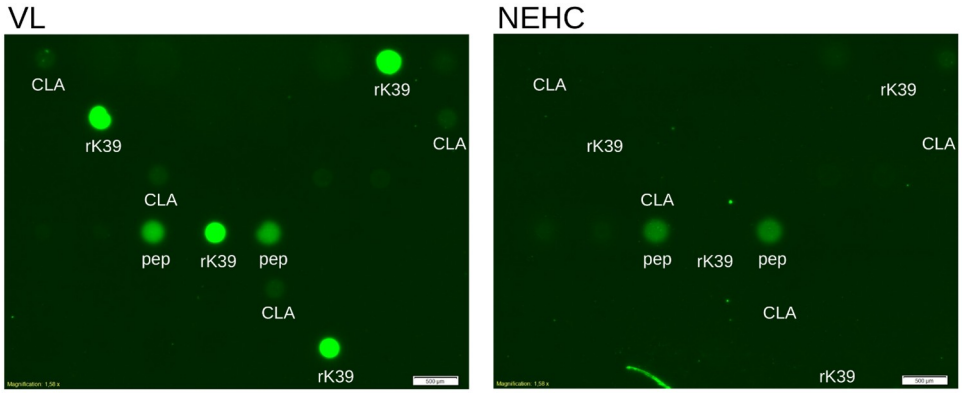


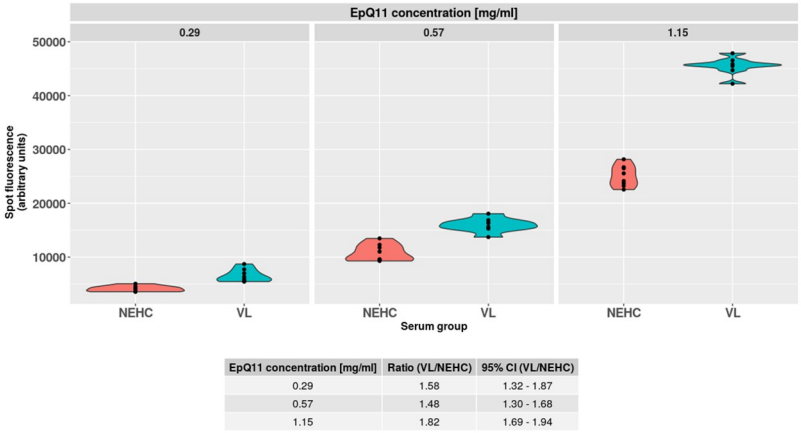
Fig 5. Hybridisation of the peptide arrays with a pool of serum samples diluted 1:200 revealed the most sensitive peptide. [VL]: Section of a peptide array hybridised with pooled Sudanese VL serum at 1:200 dilution. The most reactive peptide (pep) was spotted in duplicates. rk39 and CLA spotting positions are shown. [NEHC]: the same array hybridised with a pool of NEHC sera at same dilution. Both images are from the array section where the peptide epitopes predicted from the EpiQuest algorithm were spotted.

https://doi.org/10.1371/journal.pntd.0007353.g005

Table 3. Information on the most antigenic peptide identified visually from the pilot screening with desalted peptides. Column 'v2 ID' and 'v2 gene product' refer to the improved LdBPK reference genome.

Peptide ID	Short ID	Sequence	AA residues	v2 ID	v2 gene product
EpQ_11_NIRI	EpQ11	NIRIHLGDTIRIAPCK	82—97	LdBPK_360019900.1	Transitional endoplasmic reticulum ATPase, putative

https://doi.org/10.1371/journal.pntd.0007353.t003



EpQ11 concentration [mg/ml]	Ratio (VL/NEHC)	95% CI (VL/NEHC)
0.29	1.58	1.32 - 1.87
0.57	1.48	1.30 - 1.68
1.15	1.82	1.69 - 1.94

Fig 6. Peptide specificity for VL IgG1 was highly concentration dependent. Peptide EpQ11 was spotted at three different concentrations in eight replicates each onto dry NC membranes, previously soaked with NLA. The highest discrimination between VL and NEHC IgG1 (Ratio VL/NEHC) was obtained when the peptide was spotted at 1.15 mg/ml.

https://doi.org/10.1371/journal.pntd.0007353.g006

Specificity of peptide EpQ11 was highly concentration dependent. The most reactive peptide from the initial pilot screening, EpQ11, was spotted at multiple concentrations in eight replicates each onto a NC membrane pre-soaked and dried with NLA at 1 mg/ml and dried. The hybridisation of two separate arrays with pooled serum samples diluted 1:200 showed that the discrimination between the VL and the NEHC pooled sera was highly dependent on peptide concentration, with the highest discrimination displayed when the peptide was spotted at 1.15 mg/ml, as shown in Fig 6.

Peptide EpQ11 bound specifically to IgG1 from VL patients in RDT prototypes. RDT prototypes in two formats, cassettes and dipsticks, were tested with individual Sudanese VL serum samples as well as NEHC from Europe, in order to confirm the specificity of the EpQ11 peptide in a point-of-care format. All of the NEHC sera (n = 10) were negative by both RDT formats. Sensitivity values, assessed with VL serum samples (n = 19) varied between 79% (54—94%) and 84% (60—97%) depending on the format tested (Table 4). Representative results for these RDTs are shown in Fig 7.

Discussion

Identifying a biomarker of post-chemotherapeutic relapse remains a key element for VL control worldwide. The majority of relapses occur between 6 and 12 months after therapy

Table 4. EpQ11 performance in RDTs with Sudanese VL and NEHC patient sera. Summary of the RDT prototypes tested with individual VL sera from Sudan or NEHC serum samples from Europe.

Serum Group	n	Format	RDT results	
			Positive	Negative
VL positive	19	Cassette	84,2% (16/19)	15.8% (3/19)
		Dipstick	78.9% (15/19)	21.1% (4/19)
NEHC	10	Cassette[a]	0% (0/8)	100% (8/8)
		Dipstick	0% (0/10)	100% (10/10)

[a]: 2/10 tests were deemed invalid (control line did not develop due to leakage of running buffer from the buffer application well) and were therefore excluded from the sensitivity calculation.

https://doi.org/10.1371/journal.pntd.0007353.t004

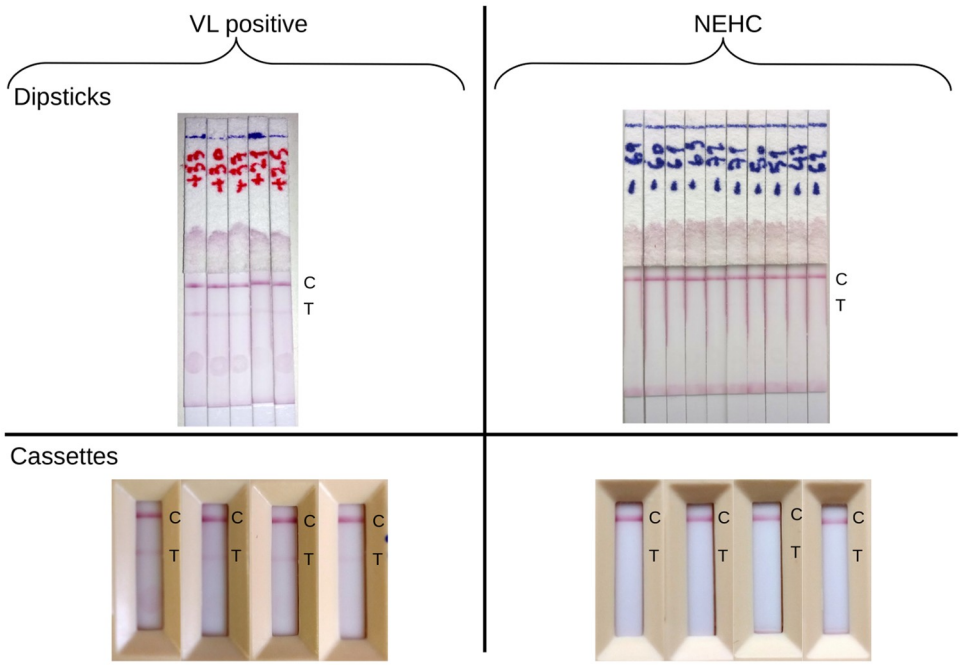


Fig 7. Peptide EpQ11 is specifically detected by Sudanese VL IgG1 in cassette as well as in dipstick format. 'T' indicates the location of the NLA at which, in a positive test, there is a red coloured line due to the presence of peptide-IgG1 complex. Successful migration was ensured by the development of the control line at 'C'.

https://doi.org/10.1371/journal.pntd.0007353.g007

completion [47], while rates are drug dependant and can reach levels up to 20% at 12 months after completion of therapy in immunocompetent patients [48]. The situation is even more dramatic in HIV-coinfected patients, where relapse rates as high as 70% have been reported [49]. Considering that most of the patients affected by VL live in poor and remote villages in developing countries [50], the development of easy-to-use diagnostic tools is of utmost importance in order to follow up patients up to 12 months upon treatment completion in a primary care setting. Currently there is no point-of-care serodiagnostic tool for detection of relapses in VL. Despite unquestionable value of the rK39-based RDTs for active cases identification, they cannot be used as a tool to detect relapses due to antibody persistence even after complete disease clearance [9–13].

We have identified a peptide detected by human IgG1 from VL patients by mapping epitopes from multiple proteins, which were selected from immunoblots by searching antigenic features using comparative genomics and *in silico* algorithms. The peptide was adapted to RDT prototypes and showed promising results for VL serology in terms of both sensitivity and specificity with limited sample size. Given the free online availability and growing number of various *in silico* algorithms to predict or scan multiple protein features from DNA or AA sequences and the growing number of high quality sequenced genomes, the adopted strategy can also be employed for the search of diagnostic biomarkers or vaccine candidates for other infectious diseases [51, 52].

The decision to make use of comparative genomics and *in silico* algorithms to refine the output from immunoblots came from the excessive number of putative hits identified by MS, and we believe that the strength of our approach lies in this combination. Nonetheless, we consider that technical issues might have been related to the excessive number of hits initially identified by MS from the western blots. The immunogenic bands excised from the acrylamide gels that were analysed by MS were excised from gels with wide lanes (59 mm). We believe

that excising such wide lanes from acrylamide gels contributed to the excessive number of hits: by mechanically cutting along very wide lanes, we have probably included proteins that would not have been present had we excised narrower (e.g. 4 mm wide) lanes, although multiple hits would have been probably identified anyway.

We have identified antigens from *L. donovani* promastigotes due to ease of culturing. In the host, promastigotes are phagocytosed by macrophages or other types of mononuclear phagocytic cells, where they differentiate into amastigotes, remaining in this form until transmission to a new vector [53]. Therefore we believe that amastigote proteins are more likely to be targeted by host Igs. We have chosen a non-stringent HAP cut-off for both promastigotes and amastigotes (HAP>Q1) due to the generally nonstrong correlation between mRNA and protein expression levels observed in eukaryotes [54, 55]. By doing so, we avoided missing antigens that might be constitutively expressed, despite low mRNA levels. Given the excessive number of protein candidates identified by MS we could set a stricter MASCOT score cut-off, ruling out proteins that could have been identified by chance in each antigenic band. Extracellular proteins are more accessible for binding with host Igs, thus we have included this criterion in our filter. Finally, because intracellular proteins harbouring TRs are amongst the most widely used antigens for VL diagnosis [56], we decided to select such candidates in parallel to extracellular proteins. Although the vast majority of epitopes founds in proteins are discontinuous [57], the use of such prediction algorithms is only meaningful in case the native 3D structure of the proteins is retained, which is not the case upon completion of an SDS-PAGE gel. In addition, the prediction of discontinuous epitopes relies on the scarce availability of experimentally validated epitopes.

Even though the *in silico* filter that we applied ruled out over 90% of the proteins initially identified by MS from the immunoblots, it remained impracticable to test the antigenicity of all 65 selected proteins. By screening synthetic peptides instead of recombinant proteins we circumvented the lengthy and costly protein expression and purification steps and could therefore test a large number of top-scoring candidates predicted from multiple algorithms with various prediction methods. Moreover, the production of peptide arrays allowed the screening of multiple peptides at various concentrations using low volumes of reagents of limited availability to us (e.g. VL serum).

The pilot screening with peptide arrays revealed that the EpQ11 was the most sensitive peptide, despite cross-reacting with NEHC serum (Fig 5). By showing the clear effect of the hybridisation conditions on the specificity of the EpQ11 in arrays, especially that of the spotting concentration (Fig 6), we sought to test the EpQ11 in RDTs, taking into account that the format of the test (i.e. peptide array vs. RDTs) could have an even larger impact on the specificity/sensitivity of the peptide. We believe that the long hybridisation periods employed in our peptide arrays played a crucial role on the higher sensitivities observed when compared to the RDT prototypes. On the other hand, by using a format of lower sensitivity, we managed to avoid the false positive signals generated by the NEHC samples in peptide arrays. The intensity of the test line of our RDT prototypes might be improved by using a larger peptide composed of multiple copies in series ('peptide trains') or in parallel (multiple antigen peptides), a strategy already used in vaccine development [58, 59], to boost the recognition of synthetic peptides by Igs, although we are aware that this improvement in sensitivity might have a negative effect on the specificity of the test.

In spite of these advantages of screening peptide epitopes in a high-throughput manner, the most important antigens for VL diagnosis (e.g. rK39, rK28) are still recombinant proteins. It would have been an interesting alternative to apply the *in silico* filter described here to proteins of lower molecular weight (e.g. those between 25 and 37 kDa—Fig 2), to test the antigenicity of a few of them. Low molecular weight proteins are generally easier to express and purify when

compared to large antigens. We decided, however, to focus our search in predicted epitopes from larger proteins due to (I) the large number of possible epitopes that proteins of high molecular weight have and (II) we were not entirely convinced about the matching between the antigenic bands in the western blots and those in the gels, mainly due to the high density of proteins in that region of the gel (Fig 2).

Another limitation of our study lies in the restricted range of serum samples available to us. Our ultimate goal was to use the EpQ11 peptide to detect IgG1 from pre-treatment VL patients paired with post-treatment (deemed cured), such that we could assess its use as a replacement for the *Leishmania* CLA in IgG1 RDTs, as described in [29]. Replacing the CLA by synthetic peptides would allow the manufacturing of more robust RDTs as a non-invasive tool to help in supporting the confirmation of cure after successful chemotherapy. Advantages of employing synthetic peptides (vs. whole-cell lysates) include unproblematic and low-cost synthesis and the production of more standardised (and potentially more sensitive and specific) diagnostic tests. Due to local regulations, however, it is not possible to export human sera from India, where the cured paired serum samples from our partner institutions are available (western blots were carried out on site in 2015), such that we could only carry out the peptide pilot screening as well as the sensitivity tests in RDTs with Sudanese active VL sera. As, however, Sudanese patients show generally a weaker immune response when compared to Indian VL patients [60], the serum samples that we tested can be considered 'worst-case scenario' samples such that we would expect the sensitivity in India to be at least as high as the values reported in this study (Table 4). The EpQ11 was also not tested with EHCs. However, EHCs used in both western blots (Fig 2) and ELISAs were negative with *L. donovani* CLA [29]. Finally, even though it is important to test the EpQ11 with serum samples from patients presenting with potentially cross-reacting diseases (e.g. malaria, human African trypanosomiasis), we would not expect a high rate of false positives due to low protein sequence conservation between the LdBPK_360019900.1. and proteins from organisms causing potentially cross-reacting diseases, especially in the region harbouring the EpQ11 sequence (Fig 8).

Our work was conducted as a proof of concept that comparative genomics and *in silico* algorithms can be employed downstream of wet lab experiments for a rational search for diagnostic peptides. Further investigations, using a larger sample size of various clinical status, including cured paired samples, EHCs as well as sera from patients presenting with potentially cross-reacting diseases, are still required to confirm the potential of the EpQ11 (with possible structural optimisations described above) as a stand-alone antigen for VL serodiagnosis. Meanwhile, given the evidence from the present work that the described peptide specifically

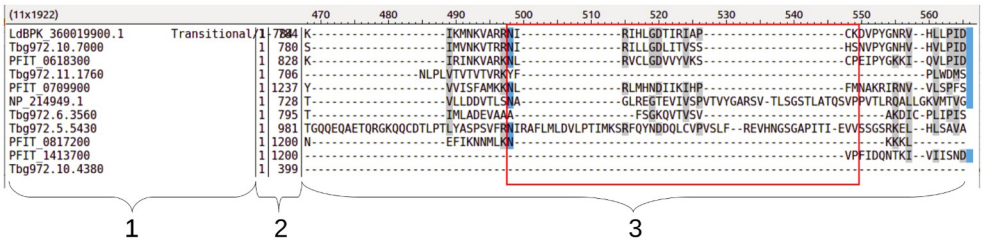


Fig 8. The *L. donovani* protein containing the EpQ11 sequence shows low similarity to proteins from organisms causing potentially cross-reacting diseases. The *L. donovani* protein harbouring the EpQ11 peptide (LdBPK_360019900.1—v2 ID) was aligned against the proteome of organisms causing potentially cross-reacting diseases with VL using MAFFT v7.222 [61]. Column 1: GenBank accession number; column 2: protein length (position of first amino acid | position of last amino acid, 784 for LdBPK_360019900.1), column 3: alignment around the region where the EpQ11 peptide is located. Numbers on top (470-560) indicate the alignment position. Screenshot from belvu, SeqTools—4.44.1 [62] showing the most similar proteins to LdBPK_360019900.1.

https://doi.org/10.1371/journal.pntd.0007353.g008

binds to IgG1 from Sudanese VL patient sera, the sequence of the EpQ11 could be synthetically coupled to that of the rK39, potentially improving the sensitivity of the latter antigen, including in IgG1 RDTs, whose potential in monitoring clinical status in VL has recently been shown [63].

Conclusion

Based on the interpretation of the results of our experiments we conclude that:

1. Comparative genomics as well as *in silico* algorithms are useful tools for refining large output from wet lab experiments for a rational search for diagnostic peptides.
2. B-cell epitopes prediction algorithms represent an interesting option for epitope mapping, enabling the screening of peptides in a high-throughput manner with minimal reagent consumption.
3. We have identified a peptide that specifically binds to human IgG1 from Sudanese VL patients by refining wet lab experiments with *in silico* searches. Replacing the CLA by specific antigens would enable the manufacturing of more robust IgG1 RDTs for monitoring clinical status in VL.
4. While further investigations are still required to confirm the potential of the EpQ11 peptide as a stand-alone antigen for VL serodiagnosis, incorporation of its sequence into that of rK39-based assays (i.e. rK39-EpQ11) might boost the sensitivity of this antigen in eastern African patients as well as in IgG1 RDTs.

Supporting information

S1 Text. LC-MS/MS method employed for protein identification.
(PDF)

S1 Table. Shortlisted proteins for *in silico* epitope mapping. List of 65 proteins identified by immunoblots and mass spectrometry (MS), then shortlisted by *in silico* algorithms. NCBI GenBank accession numbers are given for v1 IDs. Mass is given in kDa.
(PDF)

S2 Table. Synthesised peptides for antigenicity screening in arrays. The top 20 scoring peptides from each of the four prediction algorithms were shortlisted and synthesised for antigenicity screening with Sudanese VL serum samples. 'C.reported' was calculated by dividing the reported amount by the volume of the solvent mixture in which the peptides were dissolved, while 'C.280' and 'C.205' were calculated based on the molar absorbances at 280 and 205 nm, respectively. All concentrations are given in mg/ml. MW is given in g/mol. Peptides that could not be dissolved and were therefore not included in the pilot peptide screening are marked with an asterisk.
(PDF)

Acknowledgments

Mass spectrometry and subsequent protein identification was carried out by the Advanced Proteomics Facility at the University of Oxford, UK.

We would like to thank to Thomas Simon, Laetitia Avrain and Isabelle Ote from Coris BioConcept for producing the peptide arrays.

Author Contributions

Conceptualization: Bruno Cesar Bremer Hinckel, Tegwen Marlais, Tapan Bhattacharyya, Andrew Keith Falconar, Michael A. Miles, Pascal Mertens.

Data curation: Bruno Cesar Bremer Hinckel.

Formal analysis: Bruno Cesar Bremer Hinckel, Tegwen Marlais, Stephanie Airs, Hideo Imamura, Andrew Keith Falconar, Michael A. Miles, Pascal Mertens.

Funding acquisition: Jean-Claude Dujardin, Michael A. Miles, Pascal Mertens.

Investigation: Bruno Cesar Bremer Hinckel, Tegwen Marlais, Stephanie Airs, Tapan Bhattacharyya, Andrew Keith Falconar, Michael A. Miles, Pascal Mertens.

Methodology: Bruno Cesar Bremer Hinckel, Tegwen Marlais, Stephanie Airs, Andrew Keith Falconar, Michael A. Miles, Pascal Mertens.

Project administration: Andrew Keith Falconar, Michael A. Miles, Pascal Mertens.

Resources: Hideo Imamura, Jean-Claude Dujardin, Sayda El-Safi, Om Prakash Singh, Shyam Sundar, Bjorn Andersson, Michael A. Miles, Pascal Mertens.

Software: Bjorn Andersson, Sergey Litvinov.

Supervision: Bjorn Andersson, Michael A. Miles, Pascal Mertens.

Visualization: Hideo Imamura, Pascal Mertens.

Writing – original draft: Bruno Cesar Bremer Hinckel.

Writing – review & editing: Tegwen Marlais, Tapan Bhattacharyya, Jean-Claude Dujardin, Michael A. Miles, Pascal Mertens.

References

1. World Health Organization. Global leishmaniasis update, 2006–2015: a turning point in leishmaniasis surveillance. Weekly Epidemiological Record. 2017; 92(38):557–565. PMID: 28945057
2. Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, et al. Leishmaniasis worldwide and global estimates of its incidence. PloS one. 2012; 7(5):e35671. https://doi.org/10.1371/journal.pone.0035671 PMID: 22693548
3. Burns JM, Shreffler WG, Benson DR, Ghalib HW, Badaro R, Reed SG. Molecular characterization of a kinesin-related antigen of Leishmania chagasi that detects specific antibody in African and American visceral leishmaniasis. Proceedings of the National Academy of Sciences. 1993; 90(2):775–779. https://doi.org/10.1073/pnas.90.2.775
4. Cunningham J, Hasker E, Das P, El Safi S, Goto H, Mondal D, et al. A global comparative evaluation of commercial immunochromatographic rapid diagnostic tests for visceral leishmaniasis. Clinical infectious diseases. 2012; 55(10):1312–1319. https://doi.org/10.1093/cid/cis716 PMID: 22942208
5. Boelaert M, El-Safi S, Hailu A, Mukhtar M, Rijal S, Sundar S, et al. Diagnostic tests for kala-azar: a multi-centre study of the freeze-dried DAT, rK39 strip test and KAtex in East Africa and the Indian subcontinent. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2008; 102(1):32–40. https://doi.org/10.1016/j.trstmh.2007.09.003 PMID: 17942129
6. Pattabhi S, Whittle J, Mohamath R, El-Safi S, Moulton GG, Guderian JA, et al. Design, development and evaluation of rK28-based point-of-care tests for improving rapid diagnosis of visceral leishmaniasis. PLoS neglected tropical diseases. 2010; 4. https://doi.org/10.1371/journal.pntd.0000822 PMID: 20856856
7. World Health Organization. Visceral leishmaniasis rapid diagnostic test performance. Diagnostics Evaluation Series No 4. 2011.
8. Bezuneh A, Mukhtar M, Abdoun A, Teferi T, Takele Y, Diro E, et al. Comparison of point-of-care tests for the rapid diagnosis of visceral leishmaniasis in East African patients. The American journal of tropical medicine and hygiene. 2014; 91(6):1109–1115. https://doi.org/10.4269/ajtmh.13-0759 PMID: 25311696

9. Gidwani K, Picado A, Ostyn B, Singh SP, Kumar R, Khanal B, et al. Persistence of Leishmania donovani antibodies in past visceral leishmaniasis cases in India. Clinical and vaccine immunology: CVI. 2011; 18:346–348. https://doi.org/10.1128/CVI.00473-10 PMID: 21159922

10. Patil RR, Muliyil JP, Nandy A, Addy A, Maji AK, Chatterjee P. Dynamics of the antibodies in cohorts of cured cases of visceral leishmaniasis: its implication on the validity of serological test, value in prognosis and in post therapeutic assessment. Human vaccines & immunotherapeutics. 2012; 8:725–730. https://doi.org/10.4161/hv.19542

11. De Almeida Silva L, Romero HD, Prata A, Costa RT, Nascimento E, Carvalho SFG, et al. Immunologic tests in patients after clinical cure of visceral leishmaniasis. The American journal of tropical medicine and hygiene. 2006; 75:739–743. https://doi.org/10.4269/ajtmh.2006.75.739 PMID: 17038704

12. Zijlstra EE, Nur Y, Desjeux P, Khalil EA, El-Hassan AM, Groen J. Diagnosing visceral leishmaniasis with the recombinant K39 strip test: experience from the Sudan. Tropical medicine & international health: TM & IH. 2001; 6:108–113. https://doi.org/10.1046/j.1365-3156.2001.00680.x

13. Zijlstra EE, Daifalla NS, Kager PA, Khalil EA, El-Hassan AM, Reed SG, et al. rK39 enzyme-linked immunosorbent assay for diagnosis of Leishmania donovani infection. Clinical and diagnostic laboratory immunology. 1998; 5:717–720. PMID: 9729541

14. National Center for Biotechnology Information. GenBank and WGS Statistics;. Available from: https://www.ncbi.nlm.nih.gov/genbank/statistics/.

15. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP);. Available from: https://www.genome.gov/sequencingcostsdata/.

16. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, Leishmania major. Science (New York, NY). 2005; 309:436–442. https://doi.org/10.1126/science.1112680

17. Downing T, Imamura H, Decuypere S, Clark TG, Coombs GH, Cotton JA, et al. Whole genome sequencing of multiple Leishmania donovani clinical isolates provides insights into population structure and mechanisms of drug resistance. Genome research. 2011; 21:2143–2156. https://doi.org/10.1101/gr.123430.111 PMID: 22038251

18. McCall LI, McKerrow JH. Determinants of disease phenotype in trypanosomatid parasites. Trends in parasitology. 2014; 30(7):342–349. https://doi.org/10.1016/j.pt.2014.05.001 PMID: 24946952

19. Kaur G, Rajput B. Comparative analysis of the omics technologies used to study antimonial, amphotericin B, and pentamidine resistance in leishmania. Journal of parasitology research. 2014; 2014. https://doi.org/10.1155/2014/726328 PMID: 24900912

20. Pawar H, Renuse S, Khobragade SN, Chavan S, Sathe G, Kumar P, et al. Neglected tropical diseases and omics science: proteogenomics analysis of the promastigote stage of Leishmania major parasite. Omics: a journal of integrative biology. 2014; 18(8):499–512. https://doi.org/10.1089/omi.2013.0159 PMID: 24937402

21. Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS letters. 1990; 276(1-2):172–174. https://doi.org/10.1016/0014-5793(90)80535-Q PMID: 1702393

22. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome research. 2006; 2(1):2. https://doi.org/10.1186/1745-7580-2-2 PMID: 16635264

23. Saha S, Raghava G. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins: Structure, Function, and Bioinformatics. 2006; 65(1):40–48. https://doi.org/10.1002/prot.21078

24. Chen Z, Li J, Wei L. A multiple kernel support vector machine scheme for feature selection and rule extraction from gene expression data of cancer tissue. Artificial Intelligence in Medicine. 2007; 41 (2):161–175. https://doi.org/10.1016/j.artmed.2007.07.008 PMID: 17851055

25. Capecchi B, Serruto D, Adu-Bobie J, Rappuoli R, Pizza M. The genome revolution in vaccine research. Current issues in molecular biology. 2004; 6:17–28. PMID: 14632256

26. Bai X, Borrow R. Genetic shifts of Neisseria meningitidis serogroup B antigens and the quest for a broadly cross-protective vaccine. Expert review of vaccines. 2010; 9(10):1203–1217. https://doi.org/10.1586/erv.10.116 PMID: 20923270

27. Zhao BP, Chen L, Zhang YL, Yang JM, Jia K, Sui CY, et al. In silico prediction of binding of promiscuous peptides to multiple MHC class-II molecules identifies the Th1 cell epitopes from secreted and transmembrane proteins of Schistosoma japonicum in BALB/c mice. Microbes and infection. 2011; 13 (7):709–719. https://doi.org/10.1016/j.micinf.2011.03.005 PMID: 21458582

28. Carvalho GBF, Silva-Pereira RAd, Pacífico LGG, Fonseca CT. Identification of Schistosoma mansoni candidate antigens for diagnosis of schistosomiasis. Memorias do Instituto Oswaldo Cruz. 2011; 106 (7):837–843. https://doi.org/10.1590/S0074-02762011000700009 PMID: 22124556

29. Bhattacharyya T, Ayandeh A, Falconar AK, Sundar S, El-Safi S, Gripenberg MA, et al. IgG1 as a potential biomarker of post-chemotherapeutic relapse in visceral leishmaniasis, and adaptation to a rapid diagnostic test. PLoS neglected tropical diseases. 2014; 8(10):e3273. https://doi.org/10.1371/journal.pntd.0003273 PMID: 25340782

30. Dumetz F, Imamura H, Sanders M, Seblova V, Myskova J, Pescher P, et al. Modulation of aneuploidy in Leishmania donovani during adaptation to different in vitro and in vivo environments and its impact on gene expression. MBio. 2017; 8(3):e00599–17. https://doi.org/10.1128/mBio.00599-17 PMID: 28536289

31. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999; 20:3551–3567. https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2 PMID: 10612281

32. Matrix Sciences. Mascot database search: Scoring;. Available from: http://www.matrixscience.com/help/scoring_help.html.

33. Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ, et al. Proteomic analysis of the secretome of Leishmania donovani. Genome biology. 2008; 9:R35. https://doi.org/10.1186/gb-2008-9-2-r35 PMID: 18282296

34. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods. 2011; 8:785–786. https://doi.org/10.1038/nmeth.1701 PMID: 21959131

35. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic acids research. 1999; 27:573–580. https://doi.org/10.1093/nar/27.2.573 PMID: 9862982

36. Aptum Biologics LTD. EpioQuest B—find immunodominant epitopes;. Available from: https://www.epiquest.co.uk/epiquest-b.

37. Singh H, Ansari HR, Raghava GP. Improved method for linear B-cell epitope prediction using antigen's primary sequence. PloS one. 2013; 8(5):e62216. https://doi.org/10.1371/journal.pone.0062216 PMID: 23667458

38. Genscript. Peptide solubility and stability;. Available from: https://www.genscript.com/peptide_solubility_and_stablity.html.

39. Layne E. [73] Spectrophotometric and turbidimetric methods for measuring proteins. In: Methods in Enzymology. Elsevier; 1957. p. 447–454.

40. Anthis NJ, Clore GM. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. Protein science: a publication of the Protein Society. 2013; 22:851–858. https://doi.org/10.1002/pro.2253

41. Falconar AK. Epitope reactions can be gauged by relative antibody discriminating specificity (RADS) values supported by deletion, substitution and cysteine bridge formation analyses: potential uses in pathogenesis studies. BMC research notes. 2012; 5(1):208. https://doi.org/10.1186/1756-0500-5-208 PMID: 22546090

42. Bhattacharyya T, Falconar AK, Luquetti AO, Costales JA, Grijalva MJ, Lewis MD, et al. Development of peptide-based lineage-specific serology for chronic Chagas disease: geographical and clinical distribution of epitope recognition. PLoS neglected tropical diseases. 2014; 8:e2892. https://doi.org/10.1371/journal.pntd.0002892 PMID: 24852444

43. R Core Team. R: A Language and Environment for Statistical Computing; 2018. Available from: https://www.R-project.org/.

44. Djira GD, Hasler M, Gerhard D, Schaarschmidt F. mratios: Inferences for ratios of coefficients in the general linear model; 2012. Available from: https://CRAN.R-project.org/package=mratios.

45. Scherer R. PropCIs: Various Confidence Interval Methods for Proportions; 2018. Available from: https://CRAN.R-project.org/package=PropCIs.

46. Goto Y, Coler RN, Reed SG. Bioinformatic identification of tandem repeat antigens of the Leishmania donovani complex. Infection and immunity. 2007; 75:846–851. https://doi.org/10.1128/IAI.01205-06 PMID: 17088350

47. Burza S, Sinha PK, Mahajan R, Lima MA, Mitra G, Verma N, et al. Risk factors for visceral leishmaniasis relapse in immunocompetent patients following treatment with 20 mg/kg liposomal amphotericin B (Ambisome) in Bihar, India. PLoS neglected tropical diseases. 2014; 8:e2536. https://doi.org/10.1371/journal.pntd.0002536 PMID: 24392166

48. Rijal S, Ostyn B, Uranw S, Rai K, Bhattarai NR, Dorlo TP, et al. Increasing failure of miltefosine in the treatment of kala-azar in Nepal and the potential role of parasite drug resistance, reinfection, or non-compliance. Clinical Infectious Diseases. 2013; 56(11):1530–1538. https://doi.org/10.1093/cid/cit102 PMID: 23425958

49. Cota GF, De Sousa MR, Rabello A. Predictors of visceral leishmaniasis relapse in HIV-infected patients: a systematic review. PLoS neglected tropical diseases. 2011; 5(6):e1153. https://doi.org/10.1371/journal.pntd.0001153 PMID: 21666786

50. Boelaert M, Meheus F, Sanchez A, Singh SP, Vanlerberghe V, Picado A, et al. The poorest of the poor: a poverty appraisal of households affected by visceral leishmaniasis in Bihar, India. Tropical medicine & international health: TM & IH. 2009; 14:639–644. https://doi.org/10.1111/j.1365-3156.2009.02279.x

51. List C, Qi W, Maag E, Gottstein B, Müller N, Felger I. Serodiagnosis of Echinococcus spp. infection: explorative selection of diagnostic antigens by peptide microarray. PLoS neglected tropical diseases. 2010; 4:e771. https://doi.org/10.1371/journal.pntd.0000771 PMID: 20689813

52. Carvalho GBF, Silva-Pereira RAd, Pacífico LGG, Fonseca CT. Identification of Schistosoma mansoni candidate antigens for diagnosis of schistosomiasis. Memorias do Instituto Oswaldo Cruz. 2011; 106:837–843. https://doi.org/10.1590/S0074-02762011000700009 PMID: 22124556

53. Liu D, Uzonna JE. The early interaction of Leishmania with macrophages and dendritic cells and its influence on the host immune response. Frontiers in cellular and infection microbiology. 2012; 2:83. https://doi.org/10.3389/fcimb.2012.00083 PMID: 22919674

54. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. Molecular BioSystems. 2009; 5(12):1512–1526. https://doi.org/10.1039/b908315d PMID: 20023718

55. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS letters. 2009; 583(24):3966–3973. https://doi.org/10.1016/j.febslet.2009.10.036 PMID: 19850042

56. Bhattacharyya T, Marlais T, Miles MA. Diagnostic antigens for visceral leishmaniasis: clarification of nomenclatures. Parasites & vectors. 2017; 10:178. https://doi.org/10.1186/s13071-017-2120-x

57. Van Regenmortel MHV. What is a B-cell epitope? Methods in molecular biology (Clifton, NJ). 2009; 524:3–20. https://doi.org/10.1007/978-1-59745-450-6_1

58. Tam JP. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. Proceedings of the National Academy of Sciences of the United States of America. 1988; 85:5409–5413. https://doi.org/10.1073/pnas.85.15.5409 PMID: 3399498

59. Fujita Y, Taguchi H. Current status of multiple antigen-presenting peptide vaccine systems: Application of organic and inorganic nanoparticles. Chemistry Central journal. 2011; 5:48. https://doi.org/10.1186/1752-153X-5-48 PMID: 21861904

60. Bhattacharyya T, Bowes DE, El-Safi S, Sundar S, Falconar AK, Singh OP, et al. Significantly lower anti-Leishmania IgG responses in Sudanese versus Indian visceral leishmaniasis. PLoS neglected tropical diseases. 2014; 8(2):e2675. https://doi.org/10.1371/journal.pntd.0002675 PMID: 24587456

61. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution. 2013; 30(4):772–780. https://doi.org/10.1093/molbev/mst010 PMID: 23329690

62. Barson G, Griffiths E. SeqTools: visual tools for manual analysis of sequence alignments. BMC research notes. 2016; 9:39. https://doi.org/10.1186/s13104-016-1847-3 PMID: 26801397

63. Mollet G, Bremer Hinckel BC, Bhattacharyya T, Marlais T, Singh OP, Mertens P, et al. Detection of IgG1 against rK39 improves monitoring of treatment outcome in visceral leishmaniasis. Clinical Infectious Diseases. 2018. https://doi.org/10.1093/cid/ciy1062