Molecular Studies of *Treponema pallidum*

Craig Tipple

Imperial College London

Department of Medicine

Section of Infectious Diseases

Thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy of Imperial College London

2013

Abstract

Syphilis, caused by *Treponema pallidum* (*T. pallidum*), has re-emerged in the UK and globally. There are 11 million new cases annually. The WHO stated the urgent need for single-dose oral treatments for syphilis to replace penicillin injections. Azithromycin showed initial promise, but macrolide resistance-associated mutations are emerging. Response to treatment is monitored by serological assays that can take months to indicate treatment success, thus a new test for identifying treatment failure rapidly in future clinical trials is required. Molecular studies are key in syphilis research, as *T. pallidum* cannot be sustained in culture. The work presented in this thesis aimed to design and validate both a qPCR and a RT-qPCR to quantify *T. pallidum* in clinical samples and use these assays to characterise treatment responses to standard therapy by determining the rate of *T. pallidum* clearance from blood and ulcer exudates. Finally, using samples from three cross-sectional studies, it aimed to establish the prevalence of *T. pallidum* strains, including those with macrolide resistance in London and Colombo, Sri Lanka.

The sensitivity of *T. pallidum* detection in ulcers was significantly higher than in blood samples, the likely result of higher bacterial loads in ulcers. RNA detection during primary and latent disease was more sensitive than DNA and higher RNA quantities were detected at all stages. Bacteraemic patients most often had secondary disease and HIV-1 infected patients had higher bacterial loads in primary chancres. Treatment kinetics following benzathine penicillin injection were assessed in four men. The mean half-life for both blood and ulcer *T. pallidum* nucleic acid clearance was found to be short. All patients had serology consistent with cure at one month. Two *T. pallidum* strain types were found in Colombo, neither harbouring macrolide resistance. In London, several strain types were identified, the majority of which contained genetic determinants of macrolide resistance.

Declarations

The microbiology departments at St Mary's and Charing Cross Hospitals conducted laboratory tests used for the diagnosis of syphilis, including tests on cerebrospinal fluid. All clinical diagnoses referred to in this thesis were made by other physicians and not by myself.

Unless otherwise stated or referenced, all other work presented in this thesis is my own.

Craig Tipple

Acknowledgements

Firstly, I must thank my supervisors Dr Graham Taylor and Professor Myra McClure. Myra, I have valued highly your experience, support and your keen eye for detail (and grammar!). Graham, you have championed this project from almost the first time we met, and I am so very grateful for the time and the faith that you have invested me.

I would like to thank Professors Sheila Lukehart and Cathy Ison for sharing their expertise of laboratory-based and clinical syphilis research and for allowing me to spend time in their laboratories to learn *in vitro* and animal techniques relevant to this project.

I also thank my colleagues, many of whom are now friends, in the Jefferiss Trust Research Laboratories. Without your support, both practical and emotional, this work would not have been possible. I would like to thank especially Drs Steve Kaye and Maria-Antoinetta Demontis in this regard. Maria, you provided me an introduction to laboratory work with patience and kindness that I shall not forget. Steve, I cannot count the times that I wandered into your office with a question or query – you are a fount of knowledge, and I am most grateful for your support.

I must also thank the patients and staff of both the Jefferiss Wing, St Mary's Hospital and the sexual health clinics of Chelsea and Westminster Hospital for participating in the clinical studies presented in this thesis. Moreover, the staff of the McMichael-Wellcome Trust Clinical Research Facility at Hammersmith hospital and South Green ward at Charing Cross Hospital for hosting the inpatient element of the clinical study. I particularly thank Ms Lydia Hodson, Dr Rachael Jones, Sister Rosalind Nzeadi and Dr Karen Moseley for your kind assistance.

Finally, this work would not have been at all possible without the generous support of the Mason Medical Foundation, The National Institute for Health Research and the NIHR Imperial College Biomedical Research Centre.

Tim, Mum and Dad: I could not have done this without you.

Table of Contents

List of Tables

List of Figures and Appendices

List of Abbreviations

cDNA	Complementary DNA
CDTB	Communicable Diseases Tissue Biobank
DGM	Dark-ground Microscopy
EB	Elution Buffer
EDTA	Ethylenediaminetetraacetic Acid
EIA	Enzyme Immuno-assay
gDNA	Genomic DNA
IM	Intra-muscular
LB	Lysogeny Broth
LPS	Lipopolysaccharide
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	Messenger RNA
NHS	National Health Service
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
qPCR	Quantitative PCR
RIT	Rabbit Infectivity Test
RPR	Rapid Plasma Reagin
rRNA	Ribosomal RNA
RT	Reverse Transcriptase
RT-qPCR	Reverse Transcriptase qPCR
SDS	Sodium dodecyl sulfate
SOC	Super Optimal broth with Catabolite repression
TPPA	*Treponema pallidum* Particle Agglutination [assay]
TYLCV	Tomato Yellow Leaf Curl Virus
UDG	Uracil D-Glycosylase
UNG	Uracil-N-Glycosylase
VDRL	Venereal disease research laboratory [test]
β-ME	β-Mercaptoethanol

Chapter 1

Introduction

Introduction

Following four decades of low prevalence, syphilis is now resurgent in the UK and is posing both old and new challenges (1). Old problems include strategies for identifying and treating syphilis that are little changed since the 1940s and far from ideal. New obstacles are the emergence and rising prevalence of macrolide antibiotic resistance and the interaction between syphilis and HIV-1, which raise questions about the most appropriate management of the disease. In this introduction, the history and epidemiology of syphilis are explored in addition to the biology and pathogenesis of its causative organism. Current testing and treatment strategies are considered in the context of HIV-1 co-infection and in the light of new molecular techniques.

1.1 The bacterium: classification, origins and epidemiology

1.1.1 Classification

The spirochete bacterial phylum is one of the 40 major bacterial phyla and is sub-divided into three families: *Spirochetaecae*, *Brachyspiraecae* and *Leptospiraecae*. All three are thought to have a common ancestor and share, amongst other characteristics, an ultracellular structure with flagella and innate rifampicin resistance (2). The causative organism of syphilis, *Treponema pallidum* subsp. *pallidum* (*T. pallidum*) sits within the family *spirochetaecae* and the genus *Treponema* includes nine currently recognised members, often called treponemes, three of which are highly related to *T. pallidum* (3). The clinical disease caused by each of these three non-syphilis members (*Treponema pallidum* subsp. *pertenue*; *Treonema pallidum* subsp. *endemicum* and *Treponema carateum*) are summarised in Table 1.1. It is important to note that the exclusion of *T. carateum* from the *pallidum* subspecies was based only on the absence of a laboratory strain to enable its analysis and classification (4).

Table 1.1 Characteristics of three members of the genus *Treponema*

	T. pallidum* subsp. *pertenue	***T. pallidum* subsp. *endemicum***	***T. carateum***
Disease	Yaws	Endemic (non-venereal) syphilis	Pinta
Epidemiology	Central Africa, South America and Indonesia Tropical and desert climate Prevalence of 2 million cases worldwide	North Africa, Near East, South Europe Desert and temperate climates	Central and South America Tropical climates
Pathology	Mostly children under 15. Spread by close contact. Moderately invasive, skin and bone lesions	Moderately invasive. Early, latent and late stages	Non-invasive. Local dermal lesions. No systemic involvement

Table 1.1. The genus *Treponema* contains three members, in addition to *T. pallidum subsp. pallidum* (the causative organism of syphilis), which cause a distinct clinical disease (5, 6).

1.1.2 Historical origins of syphilis

The Milanese state kept detailed mortality records from 1452 and from 1503 began to note a new disease: *morbus gallicus* or the French disease. The disease was diagnosed on the basis of clinical observations such as ulcers and skin lesions and was distinct to descriptions of leprosy, which was common at that time (7). The French disease, of course, was syphilis although its modern description would not be made until 1837 by the (co-incidentally) French venereologist Phillipe Ricord and the discovery of its causative organism, *Treponema pallidum subsp. pallidum* (*T. pallidum*) would wait until Schaudinn and Hofmann's work in 1905 (8, 9). Whilst it is generally agreed that the syphilis epidemic was typical of a disease entering a non-immune population for the first time, its origin is still contested. A popular theory is that the disease was brought to the Italian peninsular with the return of Columbus from the New World. There is human skeletal evidence that venereal syphilis was present in Haiti prior to Columbus' arrival in 1492 (10). Moreover, there are descriptions of a disease characterised by 'buboes', which affected Columbus' crew and the disease appeared in the Milanese records shortly after his return to join the Spanish in the defense of Naples (11). The difficulty, however, is proving that venereal syphilis was not already present in Europe at the time of Columbus' return in 1493. The mortality records kept in Milan were not kept in other countries, thus earlier cases of the disease elsewhere may have gone un-recorded. Moreover, the fossil record contains some evidence to suggest the presence of treponematoses in Europe prior to 1492, but distinguishing the bone changes of non-venereal and venereal syphilis or even that of other conditions (rickets, Paget's disease etc) is difficult (12). An alternative theory is that modern syphilis evolved from Yaws, which was brought to Europe from Africa during the slave trade at around the time of Columbus' expeditions (13). Whilst the genetic distance between these strains appears to be at least several thousand years it is not inconceivable that *T. pallidum* could have adapted very quickly to its host (14). Wherever its origin, by the mid-16th century syphilis had reached epidemic proportions in Western Europe, its spread fuelled by the great Italian wars and the sexual relationships: consensual, transactional and forced, which are typical of conflict and social upheaval.

1.1.3 Modern epidemiology

By the 17th century, it was reported that 20% of 'patients' in London's poor houses were being treated for syphilis (15), but it was not until the Wasserman reaction (WR), the first serological test for syphilis, was described at the start of the 20th century that an accurate assessment of the prevalence of the disease was possible (12, 16). By the end of Queen Victoria's reign, 10% of British adults were reportedly infected and further increases during the First World War led, in part, to the venereal disease (VD) regulations of 1916 (17). The combination of this public health intervention together with the availability of effective treatment (Salvarsan (arsphenamine) in 1910 and penicillin in 1943) led to a significant fall in the prevalence of the disease and near complete-eradication (18-20). Today, however, syphilis has re-emerged as an important sexually transmitted disease. The annual incidence of early infection (primary, secondary and early latent) rose from 139 to a peak of 2874 in England between 1998 and 2008 and has remained high (2624 in 2010) (1, 21). Male diagnoses outnumber female in the order of 9 to 1, which reflects predominantly male-to-male transmission. One quarter of men who have sex with men (MSM) with syphilis in the UK are also HIV-1 infected, highlighting a close relationship between the epidemics (1). Rates of diagnosis have also risen in women (from 136 to 448 between 1999 and 2007) which has led to a worrying re-emergence of congenital syphilis, currently thought to be around 20 cases a year (22). Globally, syphilis reached an estimated prevalence of 36 million cases in 2005 (Figure 1.1) (23). In some populations, it has simply remained prevalent yet over-shadowed by the burden and importance of the HIV-1 epidemic. In others, there has been a recent large rise in the prevalence of the disease. China, for example, has seen a 30% annual rise in cases over the last five years resulting from worker migration and the rise of prostitution during their economic boom. Chinese MSM are also dis-proportionately affected, and a seroprevalence survey in 2008 found 11.9% to be infected with the disease. Another recent seroprevalence survey, this time in India's West Bengal region, found 8.2% of sexually transmitted disease (STD) clinic attendees to be syphilis seropositive and in neighbouring Sri Lanka diagnoses of early syphilis doubled between 2008 and 2012 (24, 25). Congenital syphilis is also a substantial problem worldwide and thought to cause 500,000 fetal and neonatal deaths each year (26).

Figure 1.1 Global prevalence and incidence of syphilis

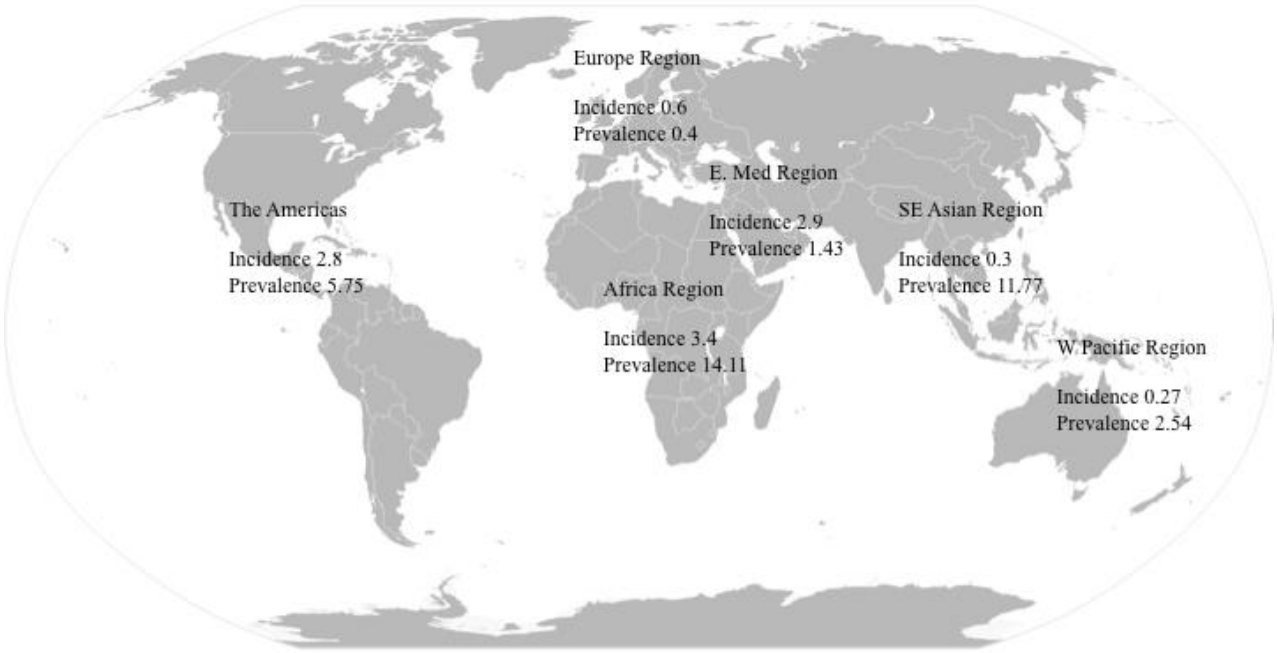


Fig. 1.1. The estimated global prevalence and incidence of syphilis in 2005 according to WHO estimates. Cases refer to men and women aged 15-49 and are presented in millions. The source data were extracted from (23).

1.2 Genome, proteome and metabolism

The *T. pallidum* genome is contained on a single circular chromosome comprising 1,014kb pairs and contains 1,041 open reading frames (representing 92.9% of the total genomic DNA (gDNA)) of which 55% codes for known proteins (Figure 1.2) (5, 27, 28). This is one of the smallest genomes of all bacteria, similar in size to that of *Borellia* and *Mycoplasma* and contains no prophage sequences, transposons or restriction-modification genes (29). *T. pallidum* sub-species have a high degree of sequence homology and through DNA hybridisation assays, subsp. *Pertenue* (yaws) and supsp. *pallidum* (syphilis) treponemes were judged so similar that they were re-classified as sub-species (30). However, Sanger sequencing and restriction fragment length polymorphism (RFLP) analysis have subsequently identified, single nucleotide differences between subsp. *pertenue* and subsp. *pallidum* in several genes, including the '*T. pallidum* repeat' (*tpr*) sub-family genes *tprI* and *tprC* (31). Most recently, next-generation genome sequencing compared seven *T. pallidum* strains comprising three subsp. *pertuenue* and four subsp. *pallidum* strains. Whilst 99.8% homology was observed, 97 genes had between one and six base differences, resulting in two or more amino acid changes and many of the proteins involved were known or putative virulence factors (32). Many pathogenic bacteria share virulence factors via the horizontal transfer of genes present on genomic islands. Other strategies include the transfer of plasmids, transposons and bacteriophage-mediated gene transfer (33). It is notable that bacteria lacking in these mechanisms, like treponemes, are usually well-adapted to life in their host, often unable to replicate outside it, and have a smaller genome. One of the purported sequelae of such adaptation and reliance on a host is that small genetic changes can have disproportionate effects on phenotype (34). This may explain, therefore, how the clinical diseases caused by such a genetically homogenous group are so different. Moreover, why continuous culture of *T. pallidum* outside a human or rabbit host remains elusive (35) and why penicillin resistance has not emerged, despite its widespread use to treat syphilis since 1943. Further exploration of these concepts requires a more detailed look at the 55% of the *T. pallidum* proteome that is homologous with proteins of known function.

T. pallidum is essentially a scavenger, relying on its host to provide energy, temperature regulation and most of the essential nutrients it requires. For energy production, *T. pallidum* seems entirely reliant on glycolysis as it lacks enzymes required for both the tricarboxylic

acid cycle and an electron transport chain (28). As a result, only two ATP can be produced from a molecule of glucose (compared with 38 ATP for *E. coli*). (28, 36) and partly explains the slow (30-33 hours) generation time of *T. pallidum* both *in vitro* and *in vivo* (35, 36). Another growth-limiting factor is the precarious relationship *T. pallidum* has with oxygen. The organism is micro-aerophilic, growing optimally in an environment containing 1.5-5% oxygen and surviving little beyond 4 hours at atmospheric oxygen levels (35, 37). This phenotype is conferred by an absence of the oxygen free-radical scavenging enzymes catalsase and superoxide dismutase and is characterised by the inability of *T. pallidum* to cope with oxidative stress (28). Of course, the organism does have some capacity for aerobic life, which is supported by an iron-containing enzyme with superoxide dismutase activity similar to the neelaredoxin enzymes typically found in thermophilic bacteria (38, 39). *T. pallidum* growth and survival is also dependent on temperature with the optimal environment being just below human body temperature at 32-36°C (35) and experimental temperatures above 45°C result in non-viability of the organism (5). This sensitivity was demonstrated in rabbits subjected to a unilateral sympathectomy. The animals developed one warm, vasodilated ear (29-34°C), which remained susceptible to *T. pallidum* infection and one cool ear (22-27°C) which could not be infected (40). Similarly in humans, larger numbers of organisms are found in the skin lesions of primary and secondary disease and far fewer in blood and deeper tissues, presumably because the skin is cooler (41). The absence of genes encoding heat-shock proteins makes *T. pallidum* unable to protect its enzymes from the denaturing effects of high temperatures and seems key to this temperature-sensitivity (42). Finally, this temperature intolerance was exploited as a syphilis treatment in the pre-penicillin era when patients were infected deliberately with malaria in order to produce fever. For the 90% that survived the malaria, the resulting high body temperatures led to an improvement in the signs and symptoms of their syphilis (4).

Figure 1.2. Genetic map

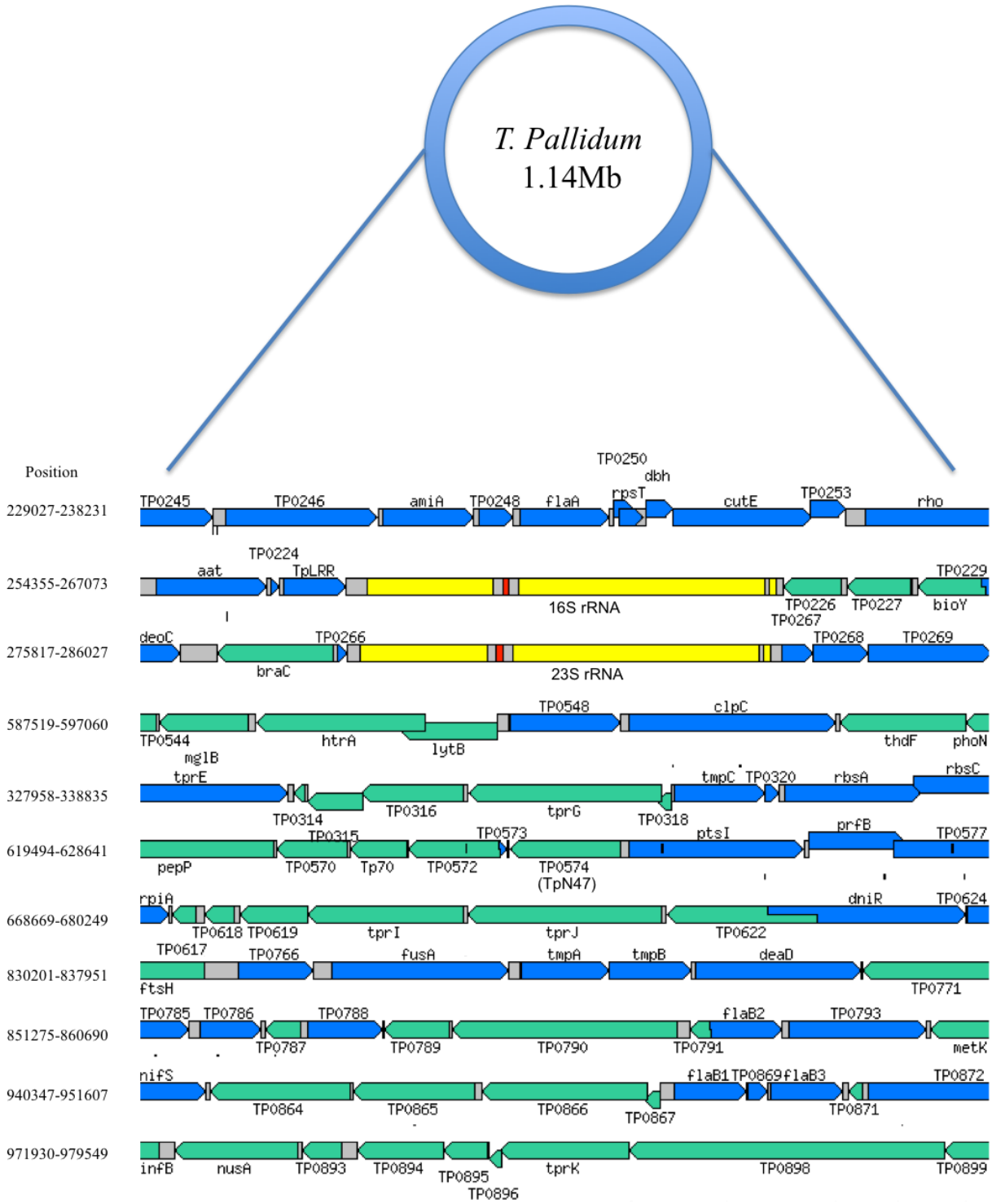


Fig. 1.2. The 1.14Mb *T. pallidum* genome is contained on a single circular chromosome. Genes of interest are displayed, but do not represent a continuous sequence. **Key** Blue: Plus strand gene; Green: Minus strand gene; Grey: Intergenic region; Red: transfer RNA; Yellow: ribosomal RNA

1.3 Culture and propagation

T. pallidum is a fastidious organism for which a continuous *in vitro* culture system has remained elusive. The best results to date were obtained with a tissue-culture system and serial passage of subcultures that maintained *T. pallidum* growth for 11 generations and resulted in a 2057-fold increase in the number of bacteria over 18 days (43). The conditions required to achieve this modest amplification are complex. Firstly, a static monolayer or microcarrier bead culture of Sf1Ep cottontail rabbit epithelial cells or RAB9 rabbit fibroblast cells is required. Culture medium is a modified BRMM with reduced essential (50%) and non-essential (25%) amino acids, which is enriched with superoxide dismutase, catalase, $CoCl_2$, co-carboxylase, histidine, mannitol and heat inactivated rabbit testis extract. Environmental conditions include three to four percent oxygen, five percent carbon dioxide and a temperature of 34-35°C (5, 35). In place of *in vitro* culture, *T. pallidum* research, therefore, relies on *in vivo* propagation in laboratory animals. Although venereal syphilis strains can be propagated in mice, guinea pigs and hamsters, rabbits are the preferred choice as the clinical primary and latent infection they develop closely resembles human disease (44, 45). Rabbits have been used to passage *T. pallidum* strains for over 100 years, beginning with the first isolated strain (Nichols strain) from the CSF of a patient with secondary disease in 1912 (46). This highly virulent strain has become the archetype for laboratory research, and despite a century of propagation in rabbits, remains infectious to humans (47, 48).

1.4 Invasion, evasion and pathogenesis of *T. pallidum*

T. pallidum is an well-adapted pathogen as what it lacks in capacity for host-free survival, it compensates for in virulence. In humans, it has an infectious dose of as few as 10 organisms (with an ID_{50} of 57 organisms); is able to infect almost any tissue type and can persist in its host for years (36, 48). In contrast to many bacterial infections, its success is better defined by invasion and immune evasion rather than a strong cytotoxic effect. When added to a variety of cell cultures or tissues, including rabbit capillaries, muscle and HEP2 cells, treponemes were shown to have a cytopathic effect but only when present in extremely large quantities (49, 50). By contrast, at low number, they are seen to cause little or no changes thus the tissue damage seen during clinical infection seems to be predominantly mediated by the host inflammatory response.

In rabbits, *T. pallidum flaA* DNA was detected in the plasma within 24 hours of intra-testicular injection of 10^6 organisms and then seen to rise steadily until a peak at 10 days (51). When organs were harvested from the same animals between 10 and 14 days post-infection, DNA was detected in some brains and kidneys and in most livers and spleens. The serum blood fraction was less likely to contain *T. pallidum* DNA than plasma or whole blood, although a recent meta-analysis found serum and plasma to be the best clinical samples in which to detect the organism (52). Whichever its preferred blood fraction, *T. pallidum* spread is predominantly haematogenous. It adheres readily to endothelial cells causing vasculitis and peri-vasculitis and has been shown to pass rapidly through endothelial inter-cellular junctions (53, 54). Once disseminated, treponemes are largely subject to immune clearance, however persistence in a variety of niches (collagen bundles, erector pili muscles, nerves) has been demonstrated and is thought to contribute to immune evasion and latency (55). There is evidence that *T. pallidum* bacteraemia can precede clinical signs and positive serological tests. For example, cases of congenital syphilis where the mother was infected in late pregnancy and was yet to develop positive serology (56). Moreover, while a recent study using PCR to detect *T. pallidum* in asymptomatic blood donors with positive syphilis serology failed to detect any bacteria there are documented cases of transfusion-related spread from individuals with incubating sero-negative disease (57). Invasion of the central nervous system is also well documented both in rabbits and humans and is considered separately in Chapter 1.6.

In order to invade, *T. pallidum* must first adhere to the surface of host cells. It binds to a wide variety of nucleated cells in a process mediated through proteins of the extra-cellular matrix, such as fibronectin, laminin and collagen I (58, 59). Laminin, a multi-domain glycoprotein is localised in the basement membrane of endothelial cells and may be important in *T. pallidum* gaining access to the blood (60). Fibronectin is a soluble dimer and appears to cross-link treponemes and mammalian cells (58). In the pre-genomic era, syphilitic serum and anti-fibronectin antibodies were found to prevent the bacteria binding to epithelial cells in culture systems (59). Subsequent knowledge of the *T. pallidum* genome allowed the '*in silico*' identification of potential adherence-associated proteins, three of which were subsequently cloned (61, 62). Two of these, *tp0155* and *tp0483* were identified as fibronectin binding proteins and the third *tp0751* as a laminin binding protein which has since been demonstrated to mediate *T. pallidum* adherence. The localisation of these binding proteins at the cell surface is yet to be fully elucidated, but once *T. pallidum* are bound to the host cell, the point of attachment seems to localise at the tip facing the organism's direction of travel (63).

The motility of *T. pallidum* is one of the major determinants of its virulence and 5% of the organism's genome has been found to encode motility and chemotactic proteins (28). Flagella, located in the periplasmic space (between the cytoplasmic and outer membrane), facilitate rotation around a longitudinal axis and result in a corkscrew motion, which can propel *T. pallidum* through viscous extra-cellular matrices at up to 19μM/second (64). The structure of the flagella, which contains a hook, basal body and filament is similar to that of other bacteria, except that the filament is made of four subunits (three class B core molecules and an outer class A sheath) in place of a single protein (65). These filaments are coded by the genes flaA (sheath) and flaB1, B2 and B3 (core) and are highly antigenic (66). The genes are coded in one large operon and knock-out of the first gene it contains (*tap1*) rendered *T. denticola* immotile (67). The ability of the unique flagellar structure to propel *T. pallidum* so effectively explains much of its invasive capacity and why a significant proportion of its small genome and limited metabolic capacity are given over to motility.

The cell surface of *T. pallidum* lacks the highly antigenic lipopolysaccharide (LPS) that is common to most gram negative bacteria and, overall, there are around 100 times fewer transmembrane proteins than are found in other spirochetes or *E. coli* (28, 68, 69). Antibodies in syphilitic serum only readily associate with *T. pallidum* if the outer membrane is degraded, following incubation at 4°C or treatment with detergent (70, 71). It is thought that proteins associated with the cell membrane are anchored predominantly in the cytoplasmic inner membrane and that few breach the outer membrane. Thus, only when they become exposed can antibody bind (72). Therefore, whilst the lack of outer membrane-associated proteins and LPS leads to increased fragility of the bacterium, it equally makes the organism much less immunogenic (4). One of the major, and best characterised, *T. pallidum* membrane-associated proteins is the penicillin-binding 47kDa surface lipoprotein (73). Whilst almost certainly anchored to the inner membrane, it is also surface-exposed and antibodies directed against it during infection have been identified (74-76). The gene sequence (*tpp047*) coding the protein is highly conserved among *T. pallidum* strains and subspecies, although it has only been found in pathogenic subspecies (77, 78). Another much-studied family of *T. pallidum* surface-exposed proteins are the *tpr* proteins. There is evidence that the in surface expression varies during the course of infection in a single host in a way that is thought to contribute to immune evasion (79-81).

1.5 T. pallidum strain-typing and molecular epidemiology

The United States Centres for Disease Control and Prevention (CDC) described the first and most commonly used typing system for *T. pallidum*, which relies on the molecular analysis of two genes (82). The first encodes an acidic repeat protein (*arp*) (*tp0433*), and contains a variable number of 60 base-pair repeats. The number of repeats for a given strain is determined by comparing the size of its *arp* amplicon with that of the Nichols strain *arp* gene, which is known to have 14. The second target, *tpr* subfamily II, has sequence heterogeneity that can be identified by comparing *Mse-I* restriction digestion patterns of *tpr* amplicons (each of which is given a separate letter) (82). The subtypes identified with this method are named according to the number of *arp* repeats and the *tpr* digest pattern, for example, 14d. This typing method was originally used to identify 12 different subtypes among 63 samples collected in the United States, Madagascar and South Africa. A recent meta-analysis revealed that 57 subtypes have now been identified using this method from 14 studies in eight geographic areas (83). Following its original description, a predominance of the subtype 14d was noted, leading some to conclude this was either the original circulating strain or had spread widely through linked transmission (83). Some felt, however, that the typing method may be relatively insensitive and three additional gene targets have been proposed to further sub-classify strains. The first enhancement involved determining the number of guanidine nucleotide repeats in the gene *rpsA* (*tp0297*) in a method similar to that used to identify the number of *arp* repeats (84). The authors proposed this number should be added to the end of the subtype, eg. '14d9'. The second method used *tp0136* gene sequence analysis to compare strains collected in the Czech republic with the macrolide resistant strain SS14, and has not been confirmed elsewhere (85). The third modification, now the most widely used, added gene sequence analysis of base pairs 131-215 of the *tpp0548* gene open reading frame (86). The authors assessed three different targets, and found *tpp0548* analysis, which was able to subdivide 173 samples from 14 CDC types into 25 types, to be the most discriminatory. Each *tpp0548* sequence identified was assigned a letter which was placed at the end of the CDC type, eg. '14d/g'.

Using the CDC typing method with or without *tpp0548* sequence analysis, several descriptions of syphilis epidemics in the USA (86), Scotland (87), South Africa (88, 89), China (90), Taiwan (91), Australia (92) and Canada (93) have now been made (Figure 1.3). The range of *arp* repeats identified was between two and 22; *tpr* patterns a to m and p have

been identified in addition tpp0548 sequences ‘c’ to ‘g’ and ‘I’ and finally eight, nine, 10 or 12 *rpsA* tandem repeats (83). Whilst certain types, notably ‘14d’ and ‘14f’ predominate, none of the epidemics characterised consisted of a single strain and have allowed analysis of population strain distribution. In this way, strain types have been associated with race, sexuality and macrolide resistance mutations (86, 93, 94).

Recent bioinformatic data derived from the complete genome sequence of nine *T. pallidum* strains has demonstrated the potential for further characterisation of *T. pallidum* molecular subtypes. In addition to *arp*, *tpr* and *tpp0548*, the genes showing the greatest sequence variation were *tp0138*, *tp0326* and *tp0488*. The optimal combination of molecular typing targets may be different depending on geographical location (32). In Madagascar, for example, all of the *tpp0548* sequences analysed were sequence ‘c’, thus the addition of this typing target added little to strain characterisation. Moreover, in Scotland all strain types identified contained 14 *arp* repeats, whereas using *tpp0548* or *tp0326* analysis may have sub-classified a number of strains (82, 87). Finally, a recent study found an association between neuro-syphilis and a *tpp0548*-enhanced strain type. Of 42 patients infected with type 14d/f, 50% had neurosyphilis compared with 10 of 41 patients with a non-14d/f type (86, 95). Whilst there is no biological basis for this association at present we have already seen how small genetic changes between *T. pallidum* subspecies are thought to account for the variation in the clinical disease they cause. It is entirely possible, therefore, that the phenotype associated with the *tpr*/*tpp0548* subtype d/f is indeed neurotropic and that in addition to molecular epidemiology, *T. pallidum* strain type analysis may have the power to predict clinical phenotypes.

Figure 1.3 Geographical distribution of *T. pallidum* subtypes

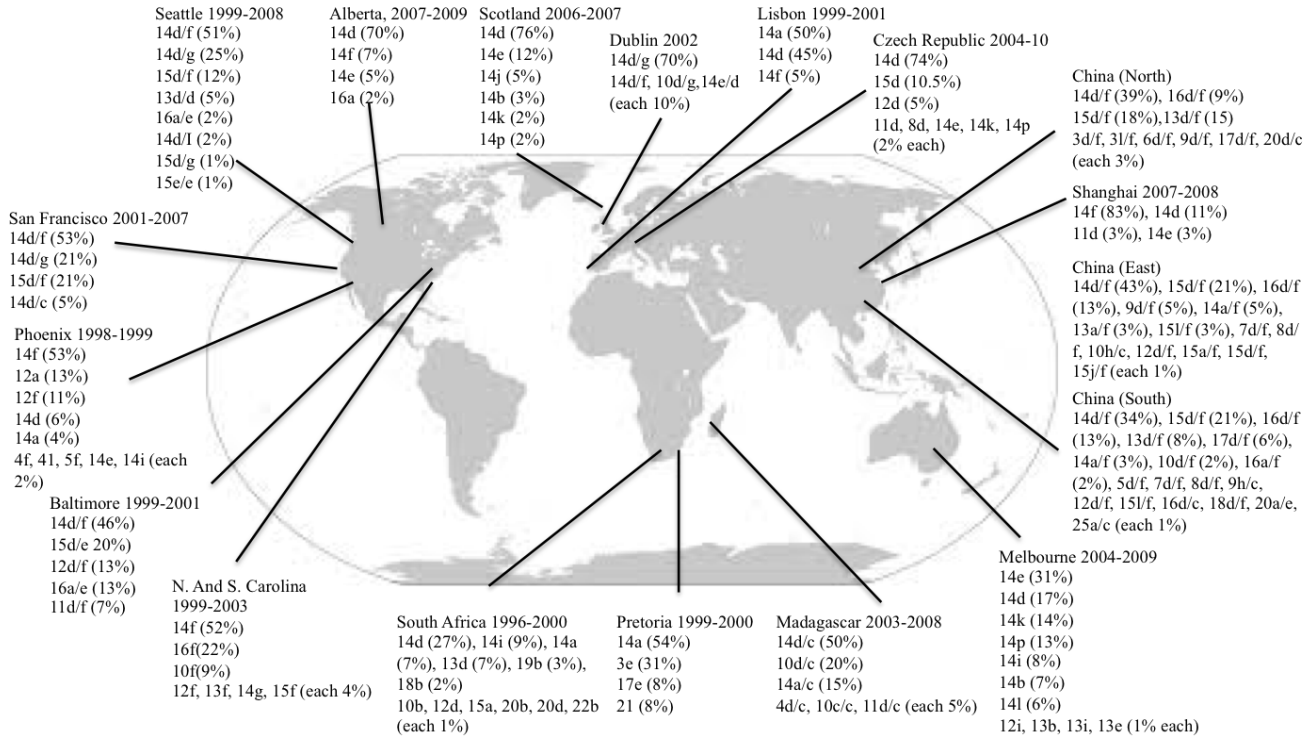


Fig. 1.3. Global distribution of *T. pallidum* strain types according to the Centres for Disease Control typing method with or without the addition of *tpp0548* gene sequence analysis. Source data were extracted from (85, 92, 96).

1.6 Clinical Disease

Syphilis is a multi-system disease with a variable course, which is divided into primary, secondary and latent stages (Figure 1.4). Transmission is both sexual, through direct skin or mucosal contact and vertical, as a result of trans-placental passage of *T. pallidum* during pregnancy. The infectious dose required for half of exposed patients to contract the disease is 57 organisms and approximately 50% of those who contact syphilis (higher in early disease) will become infected (36, 97, 98). Sexual transmission is said to be equally possible during the primary and secondary phases of infection, while transmission during latency is rare, though not impossible (97). Transmission during pregnancy is as high as 80% during early disease and then wanes as time since infection increases such that transmission risk is 30% during latent disease (99, 100).

Following sexual contact with an infectious partner, *T. pallidum* breach the skin through micro-abrasions and divide at the point of contact, commonly the mouth, genital or perianal skin (4). They enter the circulation quickly through blood and lymphatics while continuing to divide in the skin over a mean period of 24 days (range 10 to 90 days) before the initial papule of primary infection develops (48, 98). This papule then ulcerates and forms the chancre of primary disease, which is typically painless, with a firm margin and 0.5 – 2cm in diameter. Magnuson's work demonstrated that the size of the chancre developed by prisoners injected with *T. pallidum* was proportional to the infectious dose (48)[1]. Histological examination of the chancre reveals neovascularisation together with spirochetes, lymphocytes, macrophages and a variable number of plasma cells (101). Nerves are also infiltrated, via the vasa nervorum, which likely explains the absence of pain.

[1] NB. The prisoners taking part in this study were described as 'volunteers', but the use of prisoners in 1956 is likely to fall short of modern ethical standards and is not condoned.

Figure 1.4 The stages of syphilis

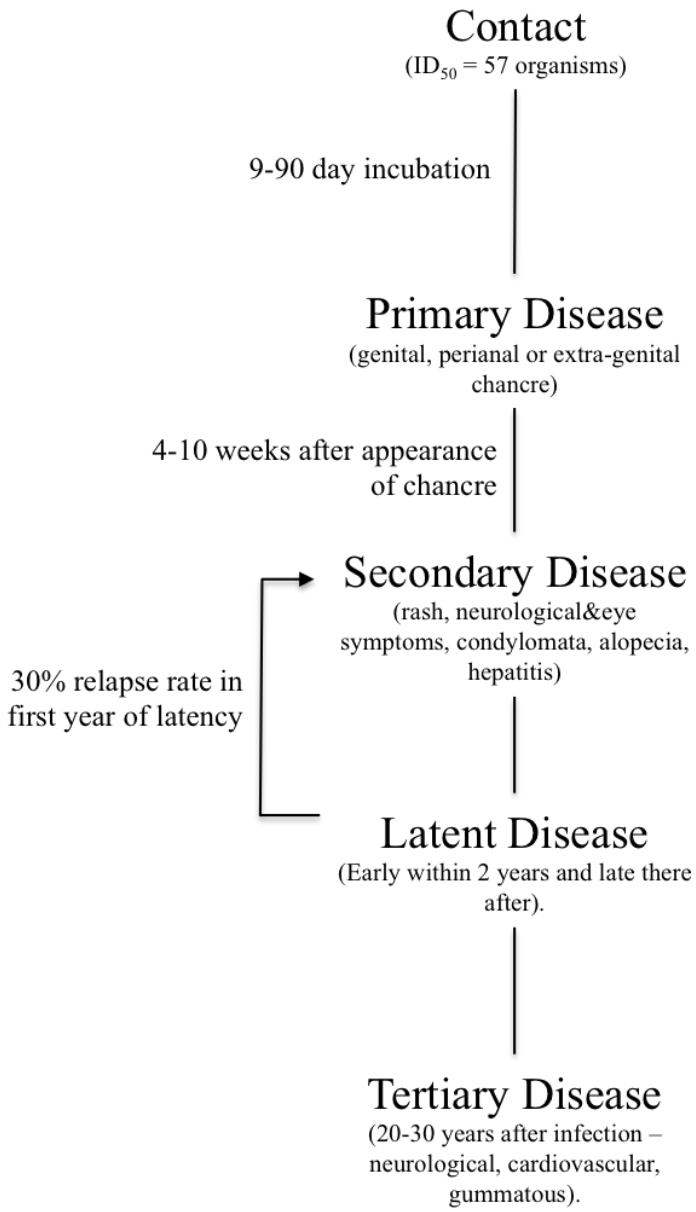


Fig 1.4. Syphilis is a multi-system disease which, if left untreated, progresses through early (primary and secondary), latent (early and late) and tertiary stages (4, 96).

T. pallidum enter the circulation from the first day of infection and continue to do so as the chancre heals over a period of 4-6 weeks (4). Interestingly, this does not result in the systemic inflammatory response that one generally associates with gram-negative bacteraemia, and is probably the result of *T. pallidum* not expressing LPS (28, 69). Moreover, the slow division time of the bacteria means that the symptoms and signs of secondary disease do not typically appear until four to 10 weeks after primary disease (100). In most cases, the chancre has healed before secondary disease is evident, but 15% (more in HIV-1 infected patients) will have concurrent signs of primary and secondary disease (102). The *sine qua non* of secondary disease is a rash, which will be macular-papular in 50-70% of patients with the unusual feature that it may affect the palms and soles (102). The histology of the rash is granulomatous and is thought to result from sufficient bacterial replication in the skin to provoke an immune response, hence the lag between bacterial dissemination and the development of clinically evident secondary disease (100). It is of note that, some patients never experience a rash, while others develop the rare and dramatic nodular, ulcerative cutaneous manifestation known as *lues maligna*. In addition to rash, a number of other manifestations both systemic and local are well documented. These are described in Table 1.2 and some are pictured in Figure 1.5.

Secondary disease will usually regress spontaneously within three months. The disease then enters a latent stage, defined as ‘early’ up to one year after the presumed date of infection and late thereafter. This distinction is made as 25% of patients with latent disease followed up in the Oslo trial of untreated syphilis relapsed to secondary disease during the first year, and while relapses up to 5 years later were possible, they were rare (103). Patients are not generally considered infectious through sexual contact during the latent period, although they may still be bacteraemic and transmission of the disease from mother-to-child can occur (104). While it is important to note that withholding penicillin from patients following demonstration of its efficacy in 1943 cannot be condoned the Oslo and Tuskagee studies were of importance in determining the long-term outcome of patients with untreated syphilis (103, 105). Tertiary syphilis (the final stage of the disease) developed in approximately 30% of patients who were untreated for between 20 and 40 years, although it is now seldom seen due to partial treatment with antibiotics given for other indications (105). The most common presentation of tertiary disease was the gumma. These deep-seated destructive nodules in bone, skin and, occasionally, other organs affected around 15% of those with tertiary disease and are thought to result from a cellular hypersensitivity reaction (100). Cardiovascular

involvement caused important morbidity in around 10% of patients as *T. pallidum* has a predilection for the ascending aorta and can provoke aneurysmal change and aortic valve regurgitation (106). Neurological involvement can occur at any stage in infection, although it peaks at 12-15 months post-infection and is only seen in around a third of patients (107). During the tertiary stage, syphilis was recognised as causing meningo-vascular disease (proliferative enderarteritis), resulting in strokes at around 5-10 years and much later (20-30 years) patients developed the 'classic' signs of tabes dorsalis (spinal dorsal column damage), dementia and neuropathic 'Charcot' joints as a result of neuronal loss, demyelination and gliosis (107, 103).

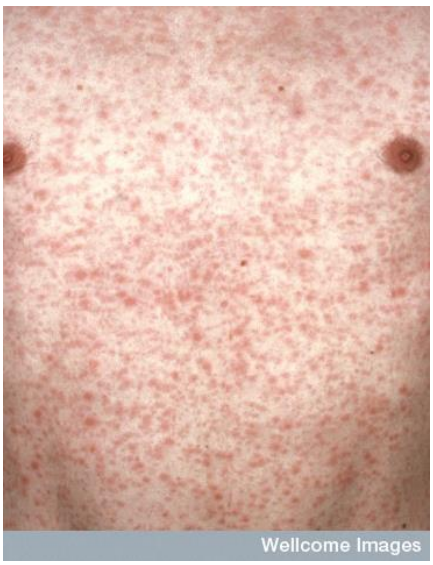
Table 1.2 Clinical manifestations of secondary syphilis

Systemic	• Myalgia
	• Fever
	• Weight Loss
	• Rash:
	  o Maculo-papular (50-70%)
	  o Papular (12%)
	  o Macular (10%)
	  o Annular (6-14%)
	  o Rarely: framboiseform, lues maligna
	• Generalised Lymphadenopathy (85%)
	• Alopecia (4-11%)

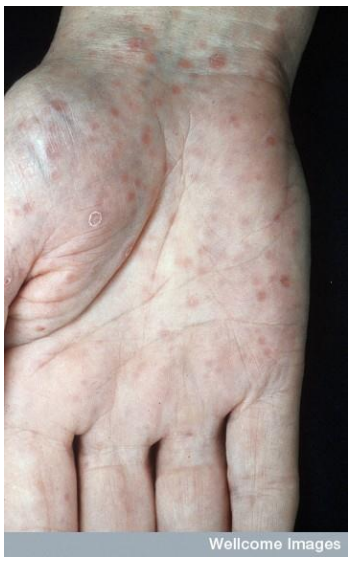
Local	• Condylomata lata (10%) Affect moist intertriginous areas (peri-anal, vulval, groin, scrotum .
	• Hepatitis
	• Meningo-vascular Neurological involvement:
	  o Stroke
	  o Ocular disease (uveitis, iritis)
	  o Meningitis
	• Renal (immune complex-mediated glomerulonephritis)

Table 1.2. A summary of the clinical manifestations of secondary syphilis (108, 109).

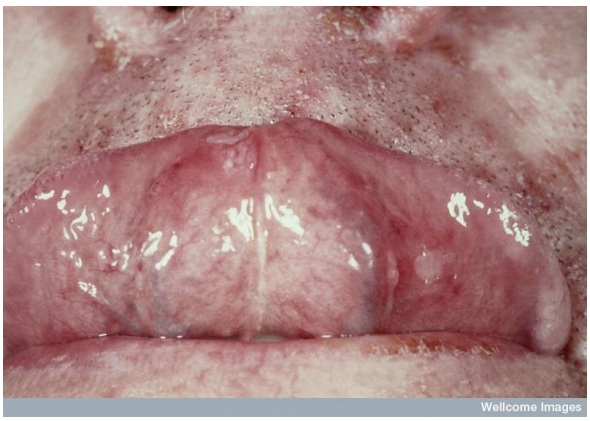
Figure 1.5 Features of secondary syphilis



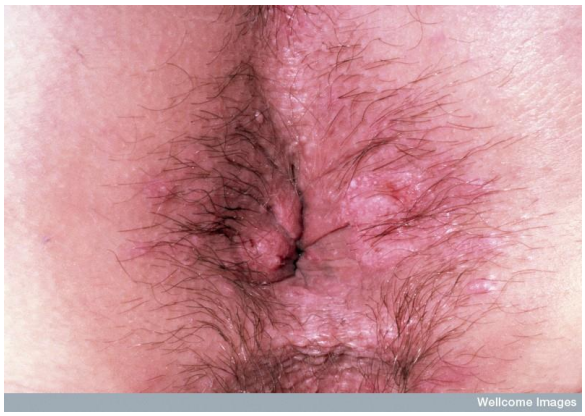
A. Acutely inflamed disseminated lesions



B. Pustular lesions on the palm



C. Oral mucosal ulceration



D. Condylomata lata

Fig. 1.5. Photographs of secondary syphilis. (Images courtesy of The Wellcome Trust Image Collection).

1.7 Diagnosing syphilis and its cure

The diagnosis of syphilis varies according to stage and relies on clinical findings, point of care tests and laboratory assays. The organism cannot be cultured, so testing has traditionally relied on direct visualisation and serological testing.

The chancre of primary disease and the moist lesions of secondary disease (condylomata lata, mucous patches, chancres) contain *T. pallidum* organisms that can be visualised directly with dark ground microscopy (DGM) (104). A sample of serous exudate is collected from the edge of a cleaned and abraded ulcer and placed immediately on a microscope slide. The treponemes are identified by their characteristic shape and motility and in the hands of a skilled operator the test provides an immediate diagnosis with a sensitivity of around 80% (110, 111). For incubating disease; DGM negative primary disease; secondary disease with no moist lesions and screening for later disease stages testing relies on a combination of serological tests.

The first serological test for syphilis was described in 1906 and became known as the Wasserman reaction (WR). Wasserman extracted what he thought to be treponemal antigens from the livers of newborns who succumbed to congenital syphilis and used them in a complement fixation test to identify reactivity with patients' sera (16). It was later discovered that the target of patient antibodies in the WR was not treponemal in origin, rather it was a phospholipid called cardiolipin (112). In a complex with lecithin and cholesterol this antigen, which could be extracted and standardized as a test reagent, became the basis for both the venereal disease research laboratory (VDRL) and rapid plasma reagin (RPR) tests, which are still in use (4, 113). These tests are often referred to as non-treponemal, because they are based on the detection of anti-cardiolipin antibodies as opposed to anti-treponemal antibodies and it is commonly asked why these anti-cardiolipin antibodies are produced in response to *T. pallidum* infection. One theory is that the antibodies are an auto-immune phenomenon and are raised against cardiolipin released from cells damaged by *T. pallidum*, although we have seen previously that the cytopathic effects of the bacteria are limited. Another is that they are raised against treponemal lipids directly. A third theory, and perhaps the most tangible, is that *T. pallidum* incorporates host lipids onto its surface during infection, which are subsequently presented to the immune system resulting in what is actually a *T. pallidum*-specific response

(4). Whatever the source, both the RPR and VDRL tests are sensitive for the diagnosis of primary disease (86% and 78%, respectively) and secondary disease (both said to be 100%) (111, 114). Of course, neither test is positive from the moment of infection nor has a specificity of 100%. A reported 30% of patients with primary disease will have a positive anti-cardiolipin test at presentation, and seroconversion is typical one to four weeks after the appearance of the chancre (115). False positive results are common in acute non-treponemal infections and in chronic auto-immune conditions, thus care must be taken to correlate serological results with clinical history and examination findings (104). As a final note, the difference between the two 'non-treponemal' tests, (VDRL and RPR) is essentially a practical one. The VDRL uses a single test slide per patient and requires a serum sample. The RPR is automatable, its antigen is more stable and, as the name suggests, requires plasma not serum. Both tests are quantifiable by using serial dilution to determine the strength of the antibody response, and the RPR is the most commonly used today (116). Antibody titres are low in primary infection, rise in secondary disease (typically to above 1:8) and then fall as disease passes into latency. As seen in Figure 1.6, they also fall following treatment and a four-fold drop in titre is taken to indicate treatment success. In most patients with early disease, the titre will decline, often to negativity in under a year, and is expected to have fallen four-fold within three to four months (117). For latent disease, the decline may be more gradual and may persist well beyond two years in 50% of patients (118). This state is often called 'serofast' and is not thought to indicate treatment failure, although the significance of a persistently positive RPR after treatment is not known and often causes clinical uncertainty (104). In addition to 'non-treponemal' antibodies, the humoral response to syphilis also comprises specific IgM and IgG antibodies. These can be detected from the sixth day of infection and are targeted against a broad range of *T. pallidum* surface-exposed lipids, lipoproteins (including the 47kDa and *tpr* proteins) and flagella proteins (4). In addition to opsonisation, these antibodies block the attachment of *T. pallidum* to host cells and some are immobilising (in conjunction with complement) (119). It is notable that while some antibodies are specific to the genus *Treponema* and others target only the pathogenic *T. pallidum* subspecies, none of the currently identified antibodies are specific to the syphilis-causing *T. pallidum* subspecies. Following their discovery, these 'specific' antibodies were harnessed as syphilis serological tests to improve the specificity of diagnosis. Immoblising antibodies were the first to be exploited as the *T. pallidum* immobilisation test (TPI) whereby Nichols organisms were immobilised by (syphilis-infected) patient serum and visualized by DGM (119). This laborious and insensitive technique was soon replaced with the fluorescent treponemal

antibody (FTA) test which involved coating slides with *T. phagadenis* antigen (from killed organisms), adding patient serum and then staining with fluorescein-labelled anti-human IgG. A subsequent modification with an absorption step gave rise the current to FTA-ABS (Figure 1.7) (120). Another technique, described in 1965, and also still used, is the *T. pallidum* haemagglutination (TPHA) test (121). In this assay, sheep erythrocytes are sensitised with antigens from a *T. pallidum* lysate and added to patient sera. A positive result is indicated by agglutination of the red blood cells into a 'mat' on the test slide. A modification of the test, the *T. pallidum* particle agglutination (TPPA) assay is also available where micro-particles replace the erythrocytes. One of the concerns with the TPHA/TPPA and FTA-ABS is that because they are produced from *T. pallidum* whole-organism lysates there is potential for cross-reactivity with other trepomenes and indeed other bacteria entirely. To overcome this, a number of tests have been developed using recombinant *T. pallidum* antigens, of which the enzyme immunoassays (EIA) are the most widely used and are commercially available. However, even this latest generation of tests fails to differentiate between infections caused by the four *T. pallidum* sub-species (111). Traditionally, syphilis screening began with the RPR and was then confirmed with a 'specific' test such as the TPPA or FTA-ABS. The high-throughput capability of the EIA, however, has led many laboratories to screen with a 'specific' test and then confirm results with the TPPA (which they were designed to replace). Either way, once two positive 'specific' test results are obtained, a VRDL or RPR is carried out to give an indication of the level of disease activity (122). A final consideration of serological tests is their use in the diagnosis of neurosyphilis. CSF analysis can help to predict risk of developing neurological disease, but results can be difficult to interpret. An abnormal CSF (raised protein, raised white cell count, positive VDRL test) portends a high risk of developing neurosyphilis if present two years or more after infection (123). However, in general only one to five percent of the 30-50% who have abnormal CSF results will go on to develop neurological disease and CSF abnormalities are frequently seen in neuro-asymptomatic early disease. Moreover, normal CSF results do not exclude neurosyphilis as a proportion of these patients will still have *T. pallidum* detectable by DGM or PCR and the sensitivity of VDRL in the CSF is only 30-70% (107, 123).

The detection of B-cells responsive to stimulation with *T. pallidum* protein in the peripheral blood of patients with syphilis has also been assessed (124). An Enzyme-Linked Immunoadsorbant Spot (ELISpot) assay was developed in which PBMCs were isolated from the blood of 35 patients with early and latent syphilis. Following separation, cells were

stimulated with recombinant Tpp17 protein and proliferation detected with a biotin-labelled anti-human IgG antibody. The assay had comparable sensitivity to RPR testing for the identification of primary and secondary disease and identified 46% of patients with latent disease. A total of six patients had a second blood sample taken following treatment and in five the ELISpot assay could no longer identify reactive B-cells. The sixth patient, whose B-cells remained reactive was deemed to be a treatment failure according to standard serological follow-up.

While molecular diagnostic tests are now gaining ground, most bacterial infections were traditionally diagnosed with culture. *T. pallidum*, being largely uncultivable, was one of the exceptions to this rule, hence our reliance over the years on direct visualisation and serological testing. It is possible, however, to use a modification of the method of propagating strains in rabbits as a diagnostic test, referred to as the rabbit infectivity test (RIT). The procedure, which can detect a single treponeme, remains the most sensitive method available for diagnosing syphilis but is expensive, can take up to 120 days to give a definitive result and requires the sacrifice of one or two rabbits. It is not widely used (125, 126).

Figure 1.6 Serum levels of syphilis antibodies following treatment

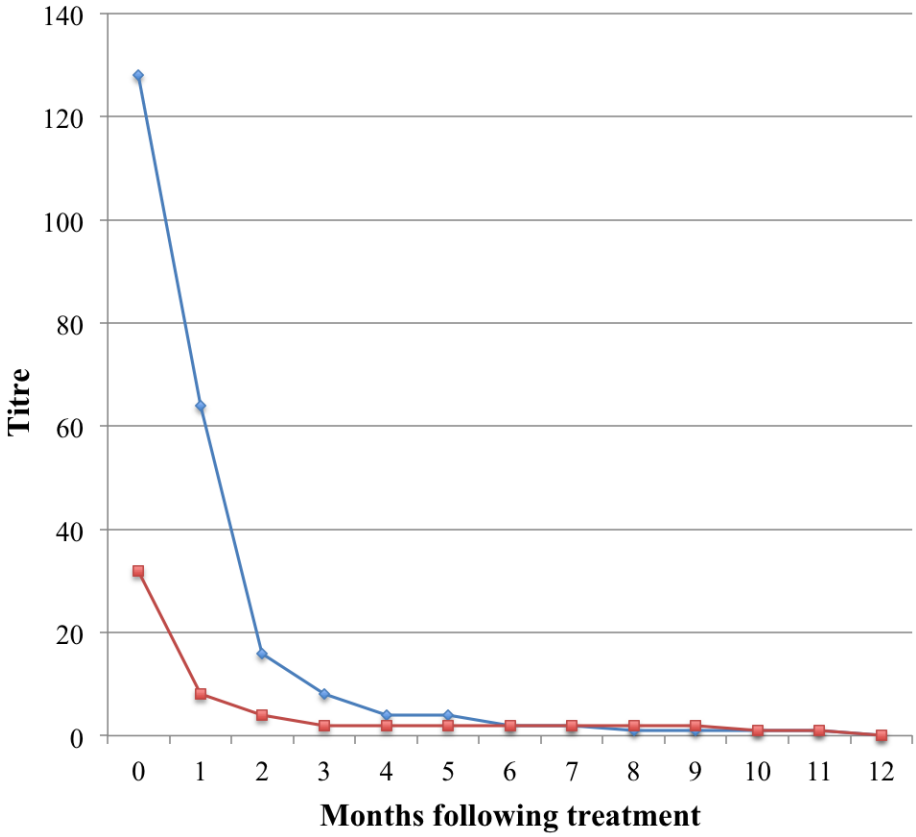


Fig 1.6. Following adequate treatment the non-treponemal antibody titre (RPR (blue diamonds) and VDRL (red squares) are pictured here) is expected to fall four-fold after approximately four months. It should then remain low or negative (111).

Figure 1.7 The FTA-Abs test for syphilis

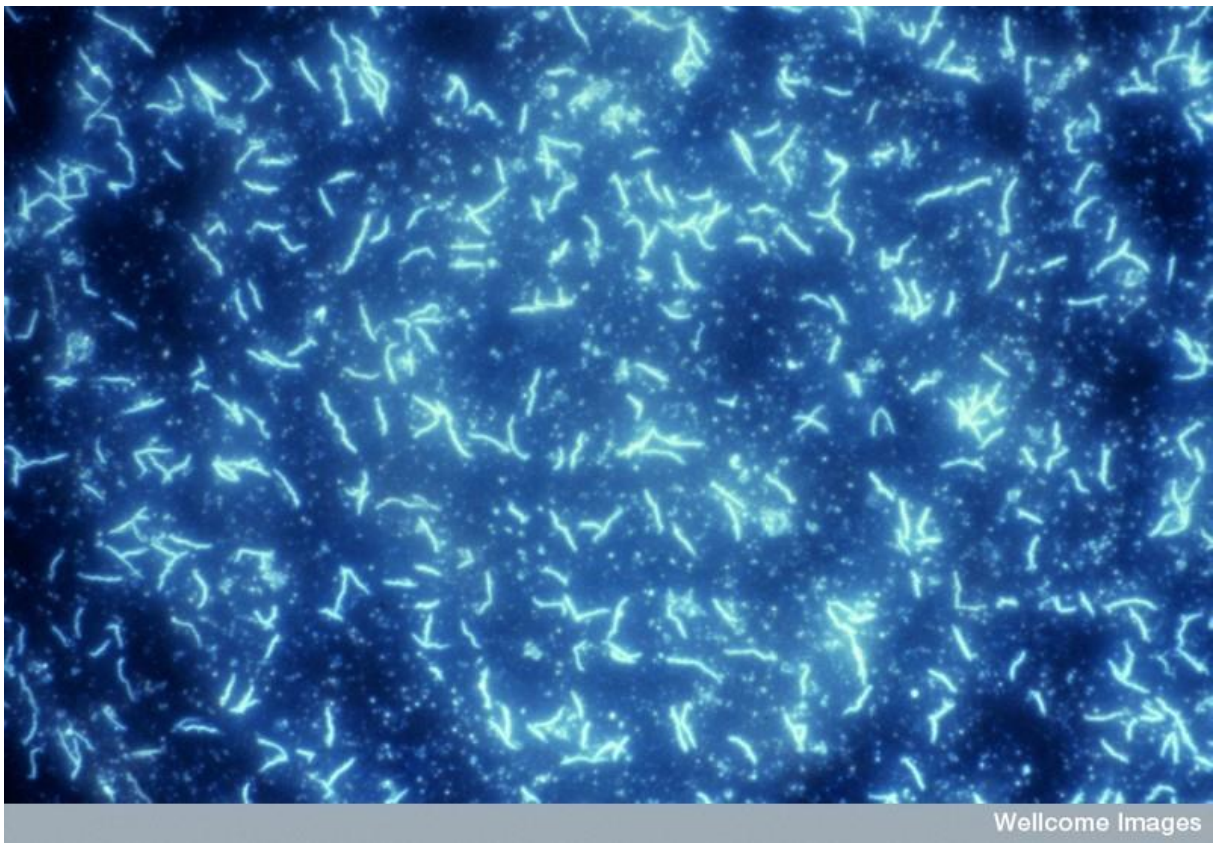


Fig 1.7. Following staining with a fluorophore-labelled antibody, *T. pallidum* in patient serum is viewed under a fluorescent microscope. This is the basis of the current FTA-Abs test for T. pallidum. Image courtesy of the Wellcome Trust Image Collection.

1.8 *T. pallidum* PCR

The prospect of PCR as a sensitive, specific and quick test for the diagnosis of syphilis was an inviting one. The first description of *T. pallidum* detection by PCR was made in 1990 using CSF samples collected from patients under-going lumbar puncture for the investigation of syphilis; HIV-related CNS disease or other non-STI related conditions (127). Overall, the clinical sensitivity of the assay was 47% (the assay sensitivity was 65 organisms/500μl sample) and specificity was 93% for the detection of *T. pallidum* DNA (*tmpA* and *4D* gene targets) in the CSF of patients with serologically confirmed syphilis. Of note, seven patients with latent disease and five patients who had previously received treatment were found to have positive CSF. The authors questioned the significance of *T. pallidum* DNA in the CSF i.e. whether the organisms detected were actively contributing to CNS disease.

It is notable that a previous study using RIT had found no viable *T. pallidum* in the CSF of 18 patients with latent disease compared with 30% positivity in early disease CSF (128). A subsequent study demonstrated that five patients with acute neurosyphilis who were symptomatically cured following treatment still had *T. pallidum* (*bmp* gene) detectable in their CSF following treatment (129).

This raised the question in both studies as to whether the DNA detected in the CSF was from viable or dead organisms, moreover if PCR inhibition or contamination were affecting results. A partial explanation came when a *tpp047* PCR was compared with RIT for the detection of *T. pallidum* in amniotic fluid, sera and CSF of neonates under investigation for congenital syphilis. The overall reported sensitivity was 80% (100% for amniotic fluid, 67% for sera and 71% for CSF) and specificity was 100% for all confirming the validity of PCR on CSF samples (130). However, further samples were not taken following treatment to assess the persistence of both DNA and viable organisms. This study also established that the method of sample preparation impacted on PCR sensitivity, namely that separating DNA from samples prior to PCR led to more reliable detection.

Following these initial successes, attention turned to a number of other sample types, DNA extraction methods and PCR targets. There is great appeal in being able to detect *T. pallidum* DNA in blood samples, especially in situations where serological tests can be ambiguous,

such as latent disease and treatment monitoring. In 1969, Turner and Hardy used RIT to demonstrate viable *T. pallidum* in 5/13 whole blood samples and 0/2 serum samples from patients with secondary syphilis. A single whole blood sample from a patient with primary disease was RIT negative (126). Wicher *et al* described a similar picture by employing a nested PCR to detect *T. pallidum* DNA (*bmp* gene) in rabbit whole blood and serum following intra-testicular injection. The only positive serum sample was partly haemolysed whereas 9/10 whole blood samples were PCR-positive at day eight (compared with 3/10 on day four).

In another study, whole blood from women with syphilis in pregnancy whose babies died as a result of the infection was analysed retrospectively with a *polA* PCR. Overall, 72% were found to have been bacteraemic at the time of syphilis screening. Many other studies have since examined the best blood fraction for *T. pallidum* PCR, which were summarised in a recent meta-analysis of 46 suitably controlled studies (52). The authors concluded that serum and plasma had greater sensitivity than whole blood, but not significantly so. Moreover, these data were a pooled analysis including different stages of disease, PCR targets and extraction techniques. If only whole blood is considered and detection compared with disease stage, however, a biologically plausible unity is found between studies that is supported by RIT data: Sensitivity is greatest in secondary disease (54.2%) compared with 36.1% for primary and 20.2% for latent disease. These results were confirmed in a second meta-analysis of 15 quantitative studies that found *T. pallidum* DNA in 55.8% of secondary disease blood samples compare with 34.1% for primary and 33.6% for latent (p=0.007) (83) It seems, therefore, that blood samples taken during highly active disease (secondary, congenital in the neonate, symptomatic neurosyphilis, fatal pregnancy outcomes) are more likely to be positive by PCR, and likely reflects the target DNA copy number.

Given the wide range of organs and tissues that *T. pallidum* has been show to invade, it is not surprising that PCR has also been employed to look for the organism in a number of sites of infection other than the bloodstream. The greatest success has been found in primary disease ulcer samples, with most studies reporting a sensitivity of 80-100% and with characteristically high specificity (131). For example, Orle *et al* developed a multiplex PCR for the diagnosis of genital ulcer disease (chancroid, genital herpes and syphilis), which demonstrated a sensitivity of 91% when multiplex PCR was compared with DGM (132). Moist secondary lesions have also proved rich in *T. pallidum* DNA, with a sensitivity of 72% in the recent meta-analysis

(52). Other samples examined include urine; placenta; gastric lesions; aorta and brain tissue and results of these investigations are summarised in Table 1.3. One final sample type which merits consideration is whole blood collected from ear lobe capillaries. In a study of 69 patients with serologically confirmed (RPR, TPHA and FTA-ABS positive), asymptomatic and untreated latent syphilis, 57.1% had detectable *T. pallidum* DNA (*tpp047* gene) in ear lobe capillary blood compared with 39.1% in whole blood samples from the same patients (133). The authors postulated that like Borellia spriochetes *T. pallidum* hide in capillary beds during latent infection and, thus, ear lobes, which are rich in capillaries, make an ideal sampling site. The sensitivity seen with this sample type far exceeds any other sample for the detection of *T. pallidum* DNA in latent syphilis, although it has yet to be replicated.

Table 1.3 Samples found to contain *T. pallidum* DNA

Sample (reference)	Sample provenance	Assay	Results
Urine (134)	Dermatology outpatient attendees. 126 samples from 76 patients VDRL and TPHA testing and stage according to CDC criteria (135)	Real-time PCR *tpp047* target	Primary disease (2/7, 29%) Secondary disease (4/9, 44%) Both 100% specific
Ejaculate (136)	4 samples (1 primary, 2 secondary, 1 early latent)	Nested PCR *tpp047* target	2/4 positive (50%). 1 primary and 1 secondary
Formalin-fixed paraffin embedded **placenta** (137)	49 placenta samples, 11 from mothers with negative serology (RPR and FTA-ABS) 38 from sero-positive mothers	*tpp047* PCR with southern blot target confirmation	0/11 from seronegative mothers 4/4 from positive mothers with syphilis consistent histology. 4/38 samples with inconsistent histology
Formalin-fixed **gastric biopsy specimens** (138)	2 patients with gastric syphilis thought to be lymphoma at endoscopy	*tmpA* gene confirmed with gene sequence and *SmaI* restriction digestion	Both samples were PCR positive.
Brain tissue (syphilitic gumma) (139)	Case report of a single patients with a brain lesion found to be a gumma	*bmp* gene Nested PCR	PCR-positive.
Post-mortem **aortic** biopsy (140)	Case report of an aortic aneurysm with typical morphological and histological appearance of tertiary syphilis	*tpp047* PCR with southern blot confirmation	PCR-positive

Table 1.3. A summary of clinical samples in which *T. pallidum* DNA has been identified by PCR.

The most commonly described *T. pallidum* PCR target is the *tpp047* gene, which encodes a cell-membrane associated 47kDa lipoprotein that is highly conserved, highly expressed and antigenically dominant in the human syphilis immune response (57, 76, 132, 141-143). Genes encoding other surface-associated proteins have also been targeted (*tmpA* (127); *bmp(144)*), in addition to other targets, such as *polA* which is involved in the regulation of genome duplication (133, 145, 146). Three of these targets (*tpp047*, *bmp* and *polA*) were compared in a study of 86 ulcer and blood specimens from patients with suspected syphilis and results (according to PCR target used) were found to be 100% concordant (147). However, when *tpp047* and *polA* were compared in blood samples from patients with latent disease, *tpp047* showed improved sensitivity of between three and 12 percent, depending on the blood fraction assessed (133).

As PCR technology has evolved, successive researchers have been quick to assess new techniques in the quest for sensitive and specific *T. pallidum* detection. As a qualitative diagnostic tool, real-time PCR (qPCR) can provide a result within two to three hours and, in one study, had a sensitivity comparable to nested PCR for the detection of *polA* in ulcer samples (148). Subsequent studies have validated qPCR with TaqMan probe detection of *polA* as a useful tool for the diagnosis of infectious syphilis by comparison with both serological results (sensitivity 80.38%) and DGM (sensitivity 72.8%) (149, 150). The true power of qPCR is in its ability to quantify the starting amount of the target DNA under investigation. The first study to describe *T. pallidum* qPCR was a cross-sectional study of secondary syphilis in Columbia. The authors were able to detect *polA* in 11/26 (46%) of whole blood samples (having previously determined that whole blood was the optimal sample type). By comparing the PCR cycle detection thresholds with an absolute quantification standard of purified *T. pallidum* DNA, they established a range of copies between 194.92 and 1954.2 copies/ml blood (151). Another study presented semi-quantitative real-time PCR data for the detection of *tpp047* demonstrating a greater number of PCR cycles required for the detection of *T. pallidum* in blood, urine and CSF compared with skin or mucosal swabs. They did not, however, compare these thresholds to those of a quantification standard thus their results are more difficult to interpret (152).

Reverse transcriptase PCR (RT-PCR) to amplify *T. pallidum* sequences is more sensitive than DNA detection for a variety of sample types. A 16S ribosomal RNA (16S rRNA) target was predicted to be present at 5000-10000 copies per organism, in contrast to the 2 known copies

of the gene (153). As sequence variation in the *16S rRNA* gene between bacterial species and within the *Treponema* genus is limited, primers specific for *T. pallidum* were identified and non-specific amplification was not observed. Messenger RNA (mRNA) targets may also be useful for sensitive and specific *T. pallidum* quantification. Smajs *et al* examined the *T. pallidum* transcriptome using a novel microarray and real-time PCR to discover that a number of genes, including those coding flagella proteins (*flaB3*, *flaB2* and *flaB1*) and the 47kDa protein (*tpp047*), were all highly transcribed (141). Lower stability of these mRNA targets compared with rRNA or DNA may impact on sensitivity of detection (154). However, it may make RNA a more attractive target for monitoring *T. pallidum* treatment response, as DNA from dead organisms persists for up to 15 days in rabbit skin lesions, and DNA appears to have persisted for up to 24 months in human CSF following neurosyphilis treatment (assuming that patients were not re-infected and did not fail treatment) (129, 155).

In summary, PCR is an effective tool for the diagnosis of early syphilis where swabbed lesions are rich in *T. pallidum* DNA. The sensitivity of *T. pallidum* bacteraemia detection is lower, but varies according to stage and is highest in early disease and whole blood appears to be the optimal sample. There is a wealth of evidence that the *tpp047* and *polA* genes are reliable targets for *T. pallidum* detection in both conventional and real-time assays and *tpp047* has been employed successfully for qPCR (52). Finally, RT-PCR may provide a more sensitive and responsive measure of spirochetaemia, but is yet to be used in a quantitative manner.

1.9 Treatment and antibiotic resistance

The earliest effective anti-syphilitic was arsphenamine (salvarsan), an arsenic-based drug developed by Paul Erlich in 1910 (20). Although toxic and difficult to administer, it was nonetheless an improvement on the inorganic and somewhat ineffective mercury-based drugs it superceded (15). It was the discovery and mass production of penicillin, however, which revolutionised the treatment of syphilis. Following the initial description of its efficacy, but without prior clinical trials, penicillin was widely used in mass treatment programmes in the 1940s (156). As a result, morbidity and mortality from syphilis declined in the USA between 1944 and 1954 by 75% and 98%, respectively and penicillin remains the mainstay of treatment today (104, 157). Initially, the drug was administered in its aqueous form as

multiple daily injections in order to maintain the stated MIC of >0.018mg/L. The subsequent co-formulation of penicillin with oil to make benzathine penicillin, meant that it could be given as a single dose and maintain a treponemocidal level for at least seven days (158). This long period is needed to cover the slow generation time of *T. pallidum* (30-33 hours) (4). Indeed, treatment is extended to 21 days (three weekly benzathine injections) in latent disease where *T. pallidum* are less metabolically active. For the treatment of neurosyphilis it is widely recommended that patients receive aqueous penicillin (procaine or benzylpenicillin), which, unlike benzathine, reaches adequate levels in the CSF (104, 158). For the treatment of patients with penicillin allergy, and those who decline parenteral treatment, oral second-line treatments are available and include amoxicillin with probenecid; doxycycline; erythromycin and azithromycin. With the exception of azithromycin the use of these alternatives, which all appear in treatment guidelines, has no basis in randomised controlled clinical trials (159-161).

It is remarkable that *T. pallidum* has developed no resistance to penicillin, despite the identification of several penicillin-binding proteins on the surface of the organism, which are likely to have a role in peptidoglycan and cell wall synthesis pathways (73, 162). Of particular interest is the abundant 47kDa lipoprotein which, in addition to penicillin-binding ability, has high beta-lactamase activity and can hydrolyse the beta-lactam bonds of penicillin (163). Fortunately, the activity of the enzyme is limited by strong product inhibition, thus, has not led to resistance, although the possibility of a mutant variant over-coming this inhibition is disconcerting. The only antibiotic class to which *T. pallidum* has become resistant is the macrolides, which includes the drugs erythromycin, spiramycin and azithromycin. In contrast to penicillin, macrolides are bacteriostatic antibiotics, which bind reversibly to the 23S rRNA in the 50S bacterial ribosomal subunit and block bacterial protein synthesis resulting in death (160). Azithromycin, which became available in the late 1980s, was of interest as a syphilis treatment due to its long half-life, good tissue penetration and good oral bioavailability (164). Following clinical trials showing comparable efficacy to benzathine penicillin, single-dose and orally-administered azithromycin (1g-2g) entered into clinical practice for the treatment of incubating and early syphilis in the United States (1999-2000), Canada (2000) and Uganda (mid-1990s) (164).

The story of macrolide resistance in *T. pallidum*, however, begins some time before the introduction of azithromycin. Erythromycin had already been available since the 1950s and was used for the treatment of syphilis in some patients with penicillin allergy, especially

pregnant women for whom tetracyclines are contra-indicated. Case reports during the 1970s of babies born with congenital syphilis to mothers who were treated in pregnancy with erythromycin were alarming, although it was not clear whether the cause of this treatment failure was resistance or transplacental transfer of the drug (165, 166). A more clear-cut case of resistance was reported in 1977 following treatment failure in a man with secondary syphilis who was given a prolonged course of erythromycin (167). The *T. pallidum* strain from this patient was isolated and became known as, 'street strain 14'. It was well-known from other bacterial species that A to G point mutations in the *23S rRNA* gene at positions 2058 and 2059 could confer macrolide resistance and in 1988 the A2058G point mutation was identified in the street strain 14 (167). Following this discovery, it was demonstrated *in vitro* that this strain was insensitive to the macrolide erythromycin. It was not until 2002, following the azithromycin treatment trials and the introduction of the drug into clinical practice, that the Centres for Disease Control (CDC) in the USA began to receive reports of syphilis treatment failures (168). These failures included patients who were contacts of infectious syphilis treated with azithromycin who went on to develop positive treponemal serology and patients with primary disease whose chancres continued to evolve following treatment. This prompted a study showing that rabbits infected with Nichols (wild-type) strain *T. pallidum* were cured with erythromycin and others infected with street 14 (mutant) *T. pallidum* were not (169). The same group then used restriction digestion and gene sequence analysis to describe the prevalence of mutant *T. pallidum* strains in Dublin (88%), Baltimore (11%), Seattle (13%) and San Francisco (22%) (169). This lead to the first *23S rRNA* A2059G mutation being discovered in a *T. pallidum* strain isolated from a Czech patient who had failed syphilis treatment with spiramycin (170). Treponemes harbouring both the A2058G and A2059G have since been found world-wide and appear to be increasingly prevalent, reaching 100% in one Shanghai centre (Figure 1.8) (164, 171, 172).

Finally, it is important to note the potential that exists for *T. pallidum* to develop resistance to other antibiotics. Despite the prior description of plasmids in *T. pallidum,* this was not confirmed when the genome sequence became available, nor were any other transposable genetic elements identified (14, 28, 173). This suggests that the only mechanism available to *T. pallidum* to develop drug resistance is through point mutations in genes coding for antibiotic targets. It is of concern that tetracycline resistance is easily conferred by *16S rRNA* gene point mutations and that the same mechanism has been predicted in *T. pallidum (160).* In

addition, *T. pallidum* has been shown to have innate resistance to rifampicin and clindamycin, which has a binding site on 23S rRNA overlapping that of the macrolides (160).

Figure 1.8 The worldwide prevalence of *T. pallidum* macrolide resistance

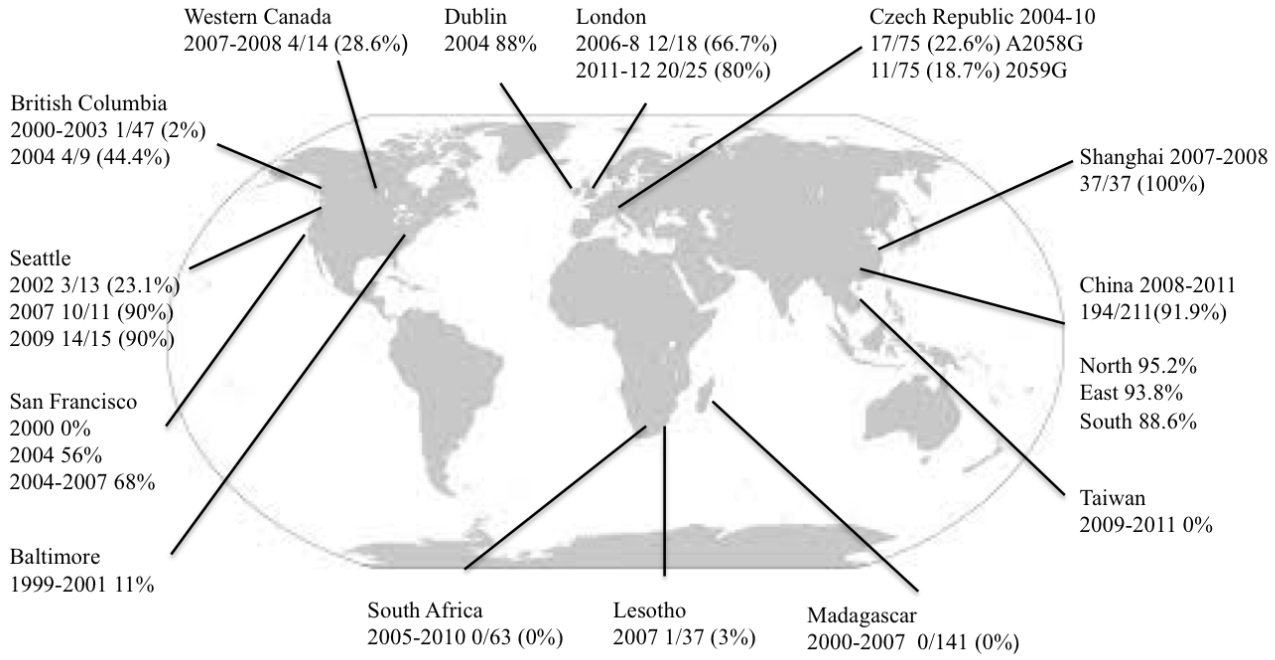


Fig 1.8. A summary of the prevalence of both A2058G and A2059G *23S rRNA* mutations detected in *T. pallidum* strains worldwide between 1999 and 2011 (84, 85, 91, 93, 169, 171, 174-180).

1.10 Inflammation and disease pathogenesis

Nearly 70 years ago, Stokes *et al* observed that the histological changes in early stage syphilitic lesions were essentially identical and deduced that the inflammatory response to *T. pallidum* was key in syphilis disease pathogenesis (15). Turner and Hollander made a similar observation in their review 10 years later, noting that the infiltration of inflammatory cells into rabbit skin following injection with *T. pallidum* was essential for lesion development (125). Another 40 years on, researchers noted that the development and resolution of early syphilis lesions is directly related to the immune response and concluded that the contribution of *T. pallidum* cytotoxicity to disease pathogenesis is minimal (181). This is confirmed in a study where rabbits were innoculated with *T. pallidum* antigens and were observed to develop lesions morphologically and histologically identical to those caused by living *T. pallidum* (182).

While not highly immunogenic, the outer membrane of *T. pallidum* contains lipoproteins, which following inoculation are stimulatory to dendritic cells and macrophages via CD14 and toll-like receptor (TLR) TLR 1 and 2 receptors (183, 184). Following antigen presentation and T-cell activation in draining lymph nodes, monocytes and T-cells are recruited to the site of inoculation by day three of infection and a predominantly Th1-type immune response develops (185-187). Secretion of IFN-γ, IL-2 and IL-12 by T-cells results in the activation of macrophages and the subsequent phagocytosis and clearance of the bacteria(187). In rabbits, a peak in the production of IFN-γ preceded the clearance of *T. pallidum* from lesions and marked the beginning of lesion resolution. The production of opsonising antibodies is also thought to be important, although the presence of B-cells in the inflammatory exudate is variable (187, 188). While CD4 cells predominate and IL-4 is absent, the number of CD8 cells present is seen to rise as lesions progress towards resolution although the significance of this observation is unclear (187).

A previously discussed (Chapter 1.6), haematogenous spread of *T. pallidum* is key to disease pathogenesis and by the time the primary chancre has resolved, the organism is widely disseminated. The patient, however, may be mostly unaware of this dissemination until the lesions of secondary disease appear as a result of local inflammation at the sites of bacterial replication, which are predominantly peri-vascular. It is often shortly after treatment that

patients feel most unwell, due to an inflammatory reaction described by Jarisch in 1895 and again in 1902 by Herxheimer and Krause (189). The Jarish-Herxmeimer reaction (JHR), as it became known, affects approximately 30% of patients following the treatment of syphilis (190). It occurs predominantly in early disease (although can certainly occur in late disease) and results in fever, myalgia, headache and increased prominence of skin lesions typically 4-6 hours after treatment (190). It is widely thought to result from rapid bacterial killing following treatment and the release of bacterial antigens into the circulation. The JHR has also been noted in other spirochetal diseases. For example, following the treatment of tick-bourne relapsing fever (seconday to infection with *Borellia spp*.) there is an inflammatory reaction characterised immunologically by increased levels of the cytokines IL-6, IL-8 and TNFα (191). Cytokine release during the syphilis JHR is yet to be measured, but is predicted to be similar. Histologically, capillaries and small veins in the skin become congested, the vascular endothelium swells and an inflammatory exudate appears. These changes are maximal by 14-18 hours and are typically resolved by 72 hours (190). Bryceson *et al* studied physiological parameters of the JHR in humans by admitting 15 patients with early syphilis to hospital for intensive monitoring following the administration of penicillin. Patients experienced fever at around two hours following treatment which peaked between four and eight hours at an average of 1.5°C above baseline (39.5°C max) (189).

In another study, the acute phase response protein C-reactive protein (CRP) was measured before and after penicillin injection. Increases were noted at 12 hours and were maximal at 24 hours. Those patients with the highest CRP levels prior to treatment had the more severe JHRs although no association was found between VDRL titre and strength of JHR. Furthermore, the single patient not to develop a JHR was the only patient included in the study with early latent as opposed to primary or secondary disease (192). In a 1948 review the percentage of patients with neurosyphilis experiencing the JHR was between 64% and 79%. It often resulted in worsening of neurological symptoms; was more common when the results of CSF analysis were abnormal and has been associated with death (193). For this reason, it is recommended that patients with neurosyphilis receive corticosteroid therapy before and after treponemocidal antibiotics (104).

In summary, it is the host inflammatory response to *T. pallidum* infection rather than a direct pathogen effect which charactersises clinical signs and symptoms of syphilis and its

treatment. The immune response, predominantly Th1 in nature, results in the rapid clearance of *T. pallidum* from both primary and secondary lesions, although the highly invasive nature of the bacterium means that the immune system is often one step behind.

1.11 HIV-1 and syphilis co-infection

The clinico-pathological and epidemiological interactions between syphilis and HIV-1 are significant. HIV-1 impacts on the clinical presentation of syphilis and potentially on the response of the disease to treatment. Despite active interest and research in the field, a consensus on the management of syphilis in the context of HIV-1 co-infection has still not been reached.

HIV-1 and syphilis co-infection is common. In 1991, 515 patients attending an STD outpatient clinic in Chicago were tested for both diseases. Among the 8% of patients found to be HIV-1 infected, 31.7% were co-infected with syphilis compared with 5.9% of HIV-1 uninfected patients (194). A subsequent literature review identified 10 case-control studies comparing the risk of syphilis in patients with and without HIV-1 infection and nine showed statistically significant relationships with odds ratios between 1.7 and 8.6. The tenth study found an association, but this was not significant when adjusted for other factors in a logistic regression (195). In the United Kingdom, 27% of syphilis diagnoses are made in the context of HIV-1 co-infection and 34% are among HIV-1-infected MSM (1). On an epidemiological level, high rates of co-infection have been associated with sex between men; sero-sorting, the practice of choosing sexual partners with the same HIV-1 infection status; recreational drug use and unprotected oral sex (196-198). In a recent review, the authors cite a rising incidence of early syphilis and falling HIV-1 incidence in San Francisco and suggest that by sero-sorting and engaging only in oral sex the transmission of syphilis is enhanced and that of HIV-1 decreased as syphilis is readily transmitted by the oral route (perhaps 30% of cases among MSM) and HIV-1 rarely so (<1%) (199). This does not, however, take note of the significant reduction on HIV-1 transmission of anti-retroviral therapy nor the body of evidence that demonstrates sero-sorting increases the transmission of other sexually transmitted infections (STIs) whilst only modestly decreasing HIV-1 transmission (200, 201). Moreover, a mathematical model suggested that syphilis is responsible for an extra 1000 cases of heterosexual HIV-1 transmission in the Unites States annually (202).

It seems unlikely that high rates of HIV-1 and syphilis co-infection can be explained by epidemiological factors alone and a number of plausible biological factors have been identified. In terms of HIV-1 transmission risk, STIs that cause genital ulcers have been associated with increased shedding of HIV-1 in the genital tract; higher HIV viral loads in semen and ulcers may bleed during sexual intercourse. Moreover, *T. pallidum* lipoproteins can upregulate HIV-1 gene expression and multiplication in tissue culture models using monocytes (203).

In all these scenarios an HIV-1 uninfected sexual partner is exposed to increased levels of HIV-1 (204, 205). Mechanisms to explain enhanced HIV-1-acquisition in patients with syphilis have also been identified. Firstly, genital ulcers damage mucosal integrity and provide a point of entry for HIV. Secondly, both living treponemes and *T. pallidum* lipoproteins have been shown to enhance the expression of CCR5, to which HIV-1 can bind, on the surface of macrophages and dendritic cells which are present in primary chancres (182). The potential for syphilis to increase the risk of contracting HIV is supported in a number of clinical studies. The risk ratio of HIV-1 acquisition was found to be 4.7 in a study of 149 cases of chancroid or primary syphilis and ranged between 2.2 and 13.5 in six other studies of genital ulcer disease (not necessarily caused by syphilis) (206, 207). Furthermore, in a retrospective cohort of 5164 patients in Miami the incidence ratio between patients diagnosed with syphilis prior to a positive HIV test and those who developed syphilis in the interval between a negative and a positive HIV test was 2.9 (95% CI 1.9-4.3) and compares with that of several other studies (207, 208).

Co-infection with HIV-1 has been shown to impact of the clinical course of syphilis. This interaction was first described early in the HIV epidemic through a number of case reports of patients experiencing increased neurological complications in addition to unusually aggressive disease. Subsequently, two prospective, multi-centre clinical trials aimed to assess the differences in the presentation and outcomes of syphilis treatment according to HIV-1 infection status. In the first, HIV-1 co-infected patients were more likely to present with multiple and deeper chancres than those without HIV-1 infection. Moreover, HIV-1 was associated with an increased risk of the primary chancre persisting into the secondary stage (209).

The second study was a randomised, controlled trial comparing standard benzathine therapy with enhanced therapy in patients with early syphilis both with and without HIV-1 co-infection. No significant differences were seen in the systemic manifestations of syphilis, although the study did not examine inpatients, was probably underpowered (recruiting 590 patients of a planned 1200) and included few patients with advanced HIV-1 related disease. Despite these limitations it was noted that HIV-1 infected patients have significantly higher RPR titres (1:128 vs 1:64, $p<0.05$) at presentation; that the JHR was more common in co-infected patients ($p=0.02$) and chancre resolution was slower in the context of HIV-1 infection (16.5 days vs 13 days, $p=0.08$). Finally, on multivariate logistic regression more serological treatment failures (failure of RPR to fall four-fold) were identified at six and 12 months in co-infected patients but this did not translate into clinical failure, nor was it seen at any other stage of disease.

A major concern regarding HIV-1 and syphilis co-infection is the potential for more severe *T. pallidum*-associated neurological disease, which is recalcitrant to treatment. In a study by Lukehart *et al*, patients with early syphilis underwent lumbar puncture before and after syphilis treatment. CSF analysis in 40 patients included RIT to identify viable *T. pallidum* and seven samples from patients with secondary disease were positive. Of those seven, three patients received two or three doses of IM benzathine penicillin and were cured (according to a repeat RIT result). Four patients received a single dose of IM benzathine and three, all HIV-1 infected, were found to have failed treatment while the fourth (HIV-1 uninfected) was cured (128). This study led to the adoption of more rigorous treatment regimes for patients with early syphilis and HIV-1 co-infection.

Subsequently, during the randomised treatment trials of azithromycin vs benzathine penicillin discussed in Chapter 1.9, a proportion of patients underwent lumbar puncture following treatment and *T. pallidum* detection by PCR or RIT was found to be no more common in those with HIV-1 infection despite a higher baseline rate of CSF abnormalities (210). Subsequent researchers have attempted to identify risk factors associated with neurosyphilis in the context of HIV-1 co-infection and have identified association with baseline CD4 counts below 350 and RPR titres of 1:32 or more (123, 211). Additionally, those patients with a CD4 count of less than 200 were 3.7 times less likely to become CSF VDRL negative following treatment (212). There is, however, no evidence from longitudinal clinical trials to support the use of enhanced syphilis therapy in HIV-1 infected patients. Accordingly, the UK syphilis

treatment guideline now advises a single penicillin injection for the treatment of early syphilis regardless of HIV-1 seropositivity, although the need for close monitoring for treatment failure and CSF examination in the case of suspected treatment failure is stated (104).

The diagnosis of syphilis can also be more difficult in HIV-1 infected patients. There can be a delay, seemingly unrelated to CD4 counts, in seroconversion for both treponemal specific and non-specific tests; false-positive non-treponemal serological results may be more common (213). A new diagnosis of syphilis has also been shown to impact on CD4 counts and viral loads of patients with HIV-1 infection. In a study of 41 patients with early syphilis, CD4 count decreased by an average of 106 cells/mm^3 ($p=0.03$) and rose by an average of 66 cells/mm^3 following treatment (214). In a larger study comparing HIV-1 infected men with (77 patients) and without (205 patients) an incident case of early syphilis the odds ratio of an HIV-1 viral load increase was 1.87. Even in those with suppressed HIV-1 replication (HIV-1 RNA <500 copies/ml blood) on anti-retroviral treatment, a syphilis diagnosis was associated with risk of HIV viral load becoming detectable (OR 1.52). CD4 cell counts were also seen to fall by an average of 28 cells/mm^3 following syphilis diagnosis, but returned to previous levels after treatment (215). The longer-term impact of these changes, if anything, is unknown (216). As a final note, many of the observations made about syphilis treatment failures and atypical serological and CSF findings were made in the pre-highly active anti-retroviral therapy (HAART) era. It is important to note that HAART reduced the risk of serological failure in one study by 60% and thus the impact of HIV-1 on syphilis and *vice versa* may follow a different pattern in the context of effective anti-retroviral treatment (217).

1.12 Measuring the efficacy of treatment

A problem that often faces clinicians is how to ensure that syphilis has been adequately treated. The UK syphilis management guideline recommends that serological tests are followed for 12 months, or 24 if the patient is HIV-1 infected (104). Treatment failure is defined serologically as less than a four-fold fall in the patient's RPR titre, but is complicated by the non-specific nature of the RPR and by patients with latent disease who may have a low RPR titre at baseline (104, 218). Moreover, in the context of a clinical trial of a new syphilis treatment, such as the recent azithromycin trials, it is necessary to wait for a minimum of six

months following the administration of the trial drug to identify treatment success or failure (161, 219, 220).

In the early post-penicillin era, silver staining or dark field microscopy of lymph nodes and RIT were used to assess *T. pallidum* persistence both in humans and animals. In one study, lymph nodes from 45 patients with penicillin-treated syphilis were examined. Five were found to contain treponemal forms on microscopy and two were shown to harbor virulent organisms by RIT (221). The authors subsequently re-confirmed their findings but this time identifying *T. pallidum* in 11 of the 45 lymph node samples (222). In a similar study, 40 patients with treated syphilis and three with untreated disease were investigated for evidence of *T. pallidum* persistence. Of the three untreated patients, one was positive by RIT and another had *T. pallidum* identified by DGM in a sample of aqueous humour. *T. pallidum* was identified by microscopy in samples from 13 of the treated patients, but all were negative by RIT (223). These may have been false positive results, as the identification of treponemal forms on microscopy is difficult, however, it may also represent the presence of attenuated organisms with greatly reduced virulence (224). More recently, researchers have explored the use of PCR to replace RIT for direct detection of *T. pallidum* and for the evaluation of patients following treatment. Sensitivity in all specimens except ulcer swabs precludes the routine use of PCR in diagnosis, however the persistence of *T. pallidum* DNA has been noted in the CSF of patients following neurosyphilis treatment. In rabbits, PCR has also successfully demonstrated the kinetics of *T pallidum* dissemination following inoculation where DNA was reliably detected in the animal's blood during the active early stage of disease (51, 155). No study, however, has used PCR either qualitatively or quantitatively to detect *T. pallidum* in animal or human samples before and immediately following therapy in order to measure the treatment response directly. It remains to be seen, therefore, how quickly *T. pallidum* DNA levels fall following the administration of effective treatment and how this correlates with standard serological testing.

1.13 Summary

Syphilis is a systemic bacterial infection caused by *T. pallidum*, which is transmitted both sexually and from mother-to-child and is currently resurgent in the UK (1). The primary stage consists of bacterial invasion with subsequent haematogenous spread and is characterised by a

self-limiting mucosal ulcer. The disease then progresses to the secondary stage and, if untreated, to tertiary disease after two to four decades (44). The humoral response to *T. pallidum* infection can be divided into 'non-treponemal' antibodies, directed against cardiolipin-lecithin-cholesterol complex, and treponemal antibodies directed against bacterial proteins. Both antibody types have been exploited for diagnostic tests, the former for screening and as a measure of disease activity and the latter for confirmatory testing. An arbitrary two-dilution (four-fold) fall in non-treponemal titre is considered to represent an adequate treatment response and it is recommended that titres be followed for a year after treatment (two years in HIV-positive patients) to diagnose cure. Treatment for syphilis became widely available from the mid-20th century following the discovery of penicillin, and this remains the preferred treatment option for those without allergy to the antibiotic. Second-line treatment is usually tetracycline antibiotics, which have been shown in small, uncontrolled trials to be effective, but which are contra-indicated in pregnancy. A second alternative, the oral macrolide antibiotic, azithromycin, has equivalent efficacy to penicillin for the treatment of incubating and early disease but resistance is emerging.

It has been shown using PCR in a rabbit model that the dissemination of *T. pallidum* after inoculation is rapid and wide-spread. Studies exploring PCR for syphilis diagnosis in humans find sensitivities of 72-91% in primary chancres, and 24-73% in blood samples and detection of treponemal bacteraemia varies according to stage, being greatest in early disease and pregnancy (for mother's with adverse neonatal outcomes) and lowest in latent disease. RT-PCR has also been described for *T. pallidum* detection *in vitro* and has the potential to improve detection sensitivity. The use of qPCR has been described in a cross-sectional study, but *T. pallidum* bacterial loads have never been measured longitudinally as a measure of treatment response. Syphilis and HIV-1 co-infection is common in the UK and those with early syphilis are more likely to both acquire and transmit HIV. It is not known, however, whether co-infected patients are more likely to transmit syphilis than those who are HIV-1 un-infected. The publication of the *T. pallidum* genome, and that of other treponemal species has provided insight into the origins of syphilis and enabled the detailed characterisation of strain types. Molecular techniques for strain type analysis have been described and are providing useful information for epidemiologists but no data are currently available in England. Finally, the genetic identification of macrolide antibiotic resistance mutations is able to inform the use of these drugs in specific populations but prevalence of these mutations is currently unknown in many countries including the United Kingdom and Sri Lanka.

1.14 Hypotheses and aims of the thesis

1.14.1 Hypotheses

1. *T. pallidum* nucleic acids can be quantified in blood and ulcer exudates taken from patients with syphilis.
2. Bacterial load will correlate with stage of disease and disease activity.
3. The efficacy of single-dose benzathine penicillin for the treatment of early syphilis can be monitored by molecular means.
4. Macrolide-resistant *T. pallidum* strains are increasingly prevalent in London as a result of antibiotic selection pressure and are likely to be unusual in Colombo, Sri Lanka.
5. *T. pallidum* strain types identified in London are distinct from those in Colombo and in both populations, strains are associated with demographic and behavioural factors.

1.14.2 Aims

1. To establish the most efficient methods for *T. pallidum* DNA and RNA extraction from blood and ulcer samples.
2. To select the most appropriate molecular targets for quantification of *T. pallidum* in clinical samples (blood and ulcer exudates).
3. To design and validate, using samples from human subjects, two qPCR assays for the quantification of *T. pallidum* DNA and RNA.
4. To establish the blood compartment (whole blood, plasma, PBMCs) in which *T. pallidum* bacteria are most numerous.
5. To measure *T. pallidum* bacterial load in samples from patients with different stages of syphilis and to correlate this load with standard serological tests.
6. To design an internal control for monitoring the efficiency of *T. pallidum* DNA extraction from clinical samples and its subsequent amplification by qPCR.
7. To monitor the clearance of *T. pallidum* DNA and RNA from blood and ulcer exudates collected from patients with primary and secondary syphilis and to correlate these data with standard serological outcomes.

8. To establish the best clinical samples for strain-type analysis; the utility of including *tpp0548* sequence analysis in the strain type and to optimise the PCR amplification of the typing target gene *arp*.
9. To determine the prevalence of macrolide resistant *T. pallidum* strains at St Mary's Hospital, London, UK and the Colombo district clinics of the National STD/AIDS control programme in Sri Lanka.
10. To determine whether the prevalence of *T. pallidum* macrolide resistant strains has increased over time.
11. To compare *T. pallidum* strain types in London and Colombo and to make associations between strains and patient demographics.

Chapter 2

Materials and Methods

2.1 Controls

2.1.1 T. pallidum organisms

Nichols strain *T. pallidum* organisms were required both for the construction of nucleic acid quantification standards and as a source of syphilis DNA for the positive control of PCRs.
Bacteria were obtained from both Newmarket Laboratories (Newmarket, UK) and the University of Washington (Seattle, USA). In both cases the bacteria were passaged in male, adult white New Zealand rabbit testes according to a well-defined method (225). Prior to suspension in either Phosphate buffered saline (PBS) and 20% glycerol or RNAlater® solution (Life technologies, UK), organisms were quantified with darkfield microscopy according to the formula:

$$treponemes/ml = \frac{(total\ treponemes\ counted\ in\ 100\ fields)}{(100\ fields\ \times\ field\ volume\ in\ ml)}$$

2.1.2 T. pallidum nucleic acids

RNA, to be used as a positive control for reverse transcription and PCR, was extracted from 10^6 *T. pallidum*, stored in RNAlater®, using the Roche High Pure RNA Isolation Kit (Roche, Mannheim, Germany). The bacteria were first lysed during a 10-minute, 30°C incubation with 4µl of 50mg/ml lysozyme solution (Sigma, Gillingham, UK). Following the addition of 400µl Lysis/Binding buffer the lysed bacteria were applied to a spin column and centrifuged for 15s at 8,000g. On column DNAse treatment was then performed for 60 minutes at room temperature following the addition of 100µl DNAse I (Roche). The column was then washed twice, first with 500µl of Wash Buffer I, then Wash Buffer II. Following the second wash, the column was dried by centrifuging at 13000g for two minutes. Finally, RNA was eluted into 100µL of Roche elution buffer (EB) by passing two 50µl aliquots of EB through the column each following a one-minute incubation at room temperature. The eluted RNA was stored in 20µL aliquots at -80°C.

DNA was also extracted from Nichols strain *T. pallidum* for use as a positive control. DNA from $6x10^6$ organisms in PBS and glycerol was extracted with a Qiagen DNA mini kit (Qiagen, Crawley, UK). Nichols organisms were added to Qiagen tissue lysis buffer 'ATL' to

give a total volume of 180µl. Next, 20µl Qiagen proteinase K was added and the mixture incubated for three hours at 56°C. A second lysis step was then performed by adding 200µL of Qiagen lysis buffer (AL) and incubating for 10 minutes at 70°C. Once lysis was complete the mixture was added to 200µl molecular grade ethanol (Sigma, Gillingham, UK) and applied to a spin column. The column was centrifuged at 6,000g for one minute to capture the bacterial DNA. Following capture, DNA was washed with 500µl of buffer AW1 and then 500µl of buffer AW2. Following the second wash, the column was dried by centrifuging at 14,000g for three minutes and DNA eluted into 60µL of proprietary elution buffer to give a concentration of $1x10^6$ genome-equivalents per 10µL.

2.1.3 Other pathogenic organisms

DNA from a number of other bacteria, viruses and fungi, obtained from the Department of Microbiology at Imperial College Healthcare NHS Trust, was extracted with the same Qiagen method as detailed for *T. pallidum* above. A full list of organisms is available in Appendix 1.

2.1.4 Tomato Yellow Leaf Curl Virus (TYLCV)

TYLCV DNA to be used as an internal control for nucleic acid extraction and amplification from clinical samples was obtained which had been extracted from approximately 0.15g of leaf tissue from Tomato plants (*Lycopersicon esculentum)* with TYLCV symptoms (Consiglio Nazionale delle Ricerche, Istituto per lo Studio degli Ecosistemi, Li Punti, Italy) (226).

2.1.5 Syphilis-negative human blood

Blood was collected into EDTA or Tempus™ collection tubes from healthy human volunteers. It was deemed syphilis-free using the Newmarket IgM/IgG enzyme immunoassay (EIA), (Alere, Stockport, UK). Additional syphilis-free human whole blood was obtained from the NHS Blood and Transplant Service. Blood from both sources was stored in 500µL aliquots at -80°C.

2.2 Clinical sample handing

2.2.1 Whole blood samples

Following the withdrawal of two or three 500µl aliquots of whole EDTA blood, the remainder of the sample was added gently to a 15ml Falcon tube containing histopaque (Sigma, UK). (The volume of Histopaque was equivalent to 60% that of the blood volume). The sample was then centrifuged at 1000g for 20 minutes at room temperature. The plasma and buffy coat were removed carefully with Pasteur pipettes and stored separately at -80°C prior to DNA extraction.

For samples received after Dec 2006, EDTA blood was no longer separated with Histopaque. Instead, either two 500µl aliquots of blood were added to 500µl of 2x lysis buffer (20mM Tris pH 8, 0.2M EDTA, 1% SDS) and stored at -80°C, or whole blood was simply stored at -80°C in 400µl aliquots. From Jan 2011, a further 1ml aliquot was centrifuged for five minutes at 8,000g and room temperature to separate plasma, which was then stored at -80°C.

2.2.2 Snostrip samples

Sterile filter paper Snostrips (Chauvin, France) were used to collect the exudate from ulcers (genital, perianal, oral) and condylomata lata. Ulcers were first cleaned and abraded with sterile gauze and water, as though preparing for dark ground microscopy. Exudate from the ulcer was then absorbed onto the snostrip until it reached the notch (see Figure 2.1), or until the ulcer was dry (whichever was sooner). The strip was then cut transversely along line two using a sterile, disposable scalpel blade and section ‘A’ placed either into 1.3mL of RNAlater® solution or into an empty, sterile 1.5ml microcentrifuge tube.

Figure 2.1 Method for ulcer exudate collection

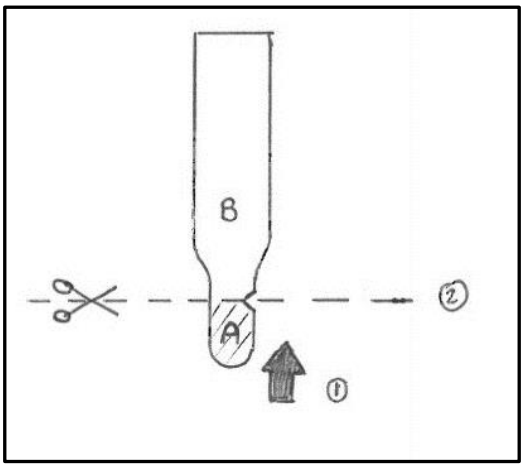


Fig 2.1 Exudate was absorbed onto a filter paper snotrip in the direction of the arrow (1) until it reached the notch. The strip was then cut transversely along line two with a sterile blade and section ‘A’ retained.

2.2.3 Tempus™ tubes for blood collection and RNA stabilisation

Three millilitres of blood were collected directly from patients into Tempus™ vacutainer tubes (Applied Biosystems, Carlsbad, CA). The tubes were supplied pre-filled with 9ml of a proprietary solution containing guanidine and detergent which simultaneously inhibited RNAses; stablilised RNA; lysed blood cells and selectively precipitated RNA in micelles. If RNA was not being extracted immediately, tubes were frozen at -80°C.

2.3 Nucleic acid extraction techniques

2.3.1 DNA extraction from whole blood, plasma and PBMCs

T. pallidum DNA was extracted from whole blood, plasma and PMBCs using one of two methods. The first method, for samples prior to Jan 2011, was based on the Qiagen QIAamp DNA Blood Mini Kit, protocol version Feb 2003 as follows: 400µl of whole blood, buffy coat or plasma were added to 40µl of Proteinase K (Qiagen, Crawley, UK). Following the addition of 400µl of lysis buffer (AL) the sample was incubated at 56°C for three hours. Next, 200µl of molecular grade ethanol (Sigma, UK) were added and the mixture applied to a QIAamp spin column. The DNA was then washed with ethanol-containing buffers AW1 and AW2 prior to elution into 120µl of PCR-grade water (Sigma, UK).

The second method, for DNA extraction from whole blood in 2x lysis buffer, used the Qiagen QIAamp DNA blood midi kit. A positive control of 500µL syphilis seronegative blood spiked with 5000 *T. pallidum* organisms in RNAlater® and a negative control of 1mL 2x lysis buffer were included with each extraction. A maximum of eight samples were extracted simultaneously.

Firstly, 1ml of blood and lysis buffer mixture was added to 100µL proteinase K (Qiagen, Crawley, UK) and 1.2ml AL buffer, then vortexed and incubated at 65°C for three hours. Next, 1ml molecular grade ethanol was added to the sample and the voretexed mixture applied to a midi prep glass-fibre column. Following centrifugation at 1850g and 4°C for three minutes, the filtrate was discarded and the column washed first with 2ml of buffer AW1, then 2ml of buffer AW2. For elution of extracted DNA, the column was placed into a clean

tube and 200µl of water applied to the membrane. Following a five-minute incubation at room temperature, the column was centrifuged at 4500g for five minutes. The process is then repeated with a second 200µl of EB. The eluted DNA was then concentrated by ethanol precipitation.

2.3.2 Ethanol precipitation of DNA

Extracted DNA in water was added to 2.5 volumes of 100% molecular-grade ethanol; 0.1 volumes of 3M sodium acetate; 2µl glycogen from mussels (Roche, Paisley, UK) and mixed by inverting the tube. Glycogen is a carrier that binds to DNA, making it visible and easier to handle. Following precipitation at -20°C for two hours (or overnight), samples were centrifuged for 30 minutes at 16,000g and 4°C. Two washes with 500µl 75% ethanol and 5 minute centrifuges of 9,500g at room temperature produced a DNA-glycogen pellet, which was dried at 62°C for 1-2 minutes, then re-suspended in 60µl of EB and stored at -80°C in four 15µl aliquots.

2.3.3 RNA extraction from blood

2.3.3.1 Tempus™ collection and extraction system

RNA from blood collected into Tempus™ tubes was extracted according to the manufacturer's protocol. Positive (Tempus™ tubes filled with 3ml of syphilis-negative blood and 5000 *T. pallidum* organisms in RNAlater®) and Negative (3ml PBS) controls were extracted together with up to 8 clinical samples in each batch. Following storage at -80°C, samples and controls were thawed, then added to 3ml of PBS in 50ml Falcon tubes to give a total volume of 12ml and vortexed for 30 seconds. The mixture was then centrifuged at 3000g and 4°C for 30 minutes. Next, the lysed blood was decanted carefully, the tube allowed to dry partially, and the remaining RNA pellet re-suspended in 400µl of resuspension solution. The RNA was adsorbed onto silica contained in a spin column and DNA digested enzymatically by adding 100µl of AbsoluteRNA wash solution (DNAse) (Applied Biosystems, Carlsbad, CA) and incubating at room temperature for 15 minutes. The captured RNA was washed twice with wash solutions one and two and, finally, eluted into 100µL of elution solution following a two-minute incubation at 70°C.

2.3.3.2 Alternative methods assessed

Ambion Ribopure™ Kit

The Ribopure™ kit (Ambion, UK) comprised cell lysis in a guanidinium-based solution, followed by purification with phenol/chloroform and solid-phase extraction on a glass fibre filter. To assess suitability of the method, 500μL of EDTA whole blood in 1.3ml RNAlater® solution were spiked with 2,000 *T. pallidum* organisms. Following centrifugation at 14,000g for 1 minute, the pellet was resuspended in a proprietary lysis solution and 3M sodium acetate. Acid-Phenol Chloroform was added and the RNA, which dissolved in the upper aqueous layer. This RNA-containing layer was then removed and applied to a column-based glass fibre filter. Now bound to the filter, the RNA was washed and finally eluted into 50μL of proprietary elution solution. Next, contaminating gDNA was removed by the addition of 2.5μl of DNAse buffer (Ambion, UK) and eight units of DNAseI (Ambion, UK). Following incubation at 37°C for 30 minutes, 10μl of DNAse inactivation agent (Ambion, UK) were added, the mixture vortexed and incubated at room temperature for one minute. The inactivation agent was then removed by pelleting it in a centrifgue at 12,000g for one minute. The DNAse-treated RNA was then transferred to a clean RNAse-free micro-centrifuge tube.

Magnetic-bead capture of human RNA species (Ambion MicrobEnrich™ kit)

Human mRNA was depleted from blood-extracted RNA samples (Tempus™ extraction system) in order to reduce *T. pallidum* PCR inhibition. The MicrobEnrich™ system was designed to enrich bacterial RNA from mixtures also containing human RNA. Human mRNA is polyadenylated at the 3' end and, in this procedure, becomes bound to magnetic beads bearing polythymidylated capture oligonucleotides. The beads are then pulled out of the sample using a magnet and discarded, leaving only bacterial RNA in solution. Briefly, in the final step of Tempus™ RNA extraction, EDTA (Sigma, Gillingham, UK) was added to the elution buffer at a final concentration of 1mM, to chelate any divalent cations present. The RNA was mixed with RNA binding buffer and capture oligonucleotide mix, heated to 70°C for 5 minutes to denature any RNA secondary structures, then incubated for an hour at 37°C to allow hybridization. Prepared and equilibrated magnetic beads were added to the RNA-oligonucleotide mixture and incubated at 37°C for 15 minutes to allow binding. Finally, the

mixture was placed on a magnetic stand, the beads were captured and bacterial RNA in the resultant supernatant was precipitated and concentrated with ethanol-glycogen precipitation.

2.3.4 Nucleic acid extraction from Snostrip samples

In the *T. pallidum* qPCR validation study, ulcer exudates were absorbed onto snostrips as described in Figure 2.1, then stored 'dry' in screw-cap microcentrifuge tubes. Those samples collected for the Communciable Diseases Tissue Biobank (CDTB) were added to 500µl 1x lysis buffer (10mM Tris pH 8, 0.1M EDTA, 0.5% SDS) prior to storage. Finally, samples collected for the *in* vivo treatment kinetics study (Chapter four) and the cross-sectional study of *T. pallidum* macrolide resistance in Sri Lanka (Chapter five) were stored in 1.3mL of RNAlater® solution.

2.3.4.1 DNA extraction from ulcer samples (snostrips)

Snostrips stored in lysis buffer were first vortexed, then a 200µl aliquot was added to 25µl Proteinase K and incubated at 56°C for one hour. An aliquot of 200µl of buffer AL was then added, the sample vortexed and incubated for a further 10 minutes at 70°C. Next, 200µl molecular grade ethanol were added and the mixture applied to a QIAamp spin column. Following one minute of centrifugation at 6,000g, the filtrate was discarded and the sample washed with buffers AW1 and AW2 before a final centrifugation (to dry the column) at 20,000g for one minute. The DNA was eluted using 200 µl of EB, passed through the column in two 100µl aliquots, each with a five-minute incubation prior to a one-minute centrifuge at 6,000g.

Those strips stored dry were first added to 400µl PBS and 20µl Proteinase K prior to incubation at 56°C for three hours. Following this enzymatic digestion, 400µl ethanol were added and the mixture applied to a QIAamp mini-prep spin column (Qiagen, Crawley, UK). The column-adsorbed DNA was then washed with buffers AW1 and AW2 prior to elution into 120µl of molecular grade water.

Finally, DNA from strips stored in RNA later solution was extracted in combination with RNA using an all-prep DNA/RNA mini kit (Qiagen, Crawley, UK).

2.3.4.2 DNA/RNA co-extraction from ulcer snostrip samples using a Qiagen all-prep mini kit

The manufacturers of RNAlater® solution, designed to protect RNA from degradation during storage, state that DNA can also be efficiently extracted from samples stored in the solution. For the study of clearance kinetics, one ulcer sample was to be taken at each time point, but both RNA and DNA needed to be quantified in that sample. During this process, the DNA and total RNA from samples were extracted concurrently. A positive control (snostrip spiked with 5,000 *T. pallidum* organisms in RNAlater®) and a negative control (an unspiked snostrip in RNAlater®) were extracted together with up to eight clinical samples in each batch. Samples were kept on ice whenever possible, and RNAse ZAP solution (Ambion, UK) was used to decontaminate thoroughly pipettes; racks; centrifuges and work-surfaces prior to handling RNA.

Snostrips in 1.3ml RNAlater® (stored at -80°C) were first thawed thoroughly, then centrifuged at 14,000g and 4°C for 30 minutes. Next, 1.1ml of RNAlater® were carefully removed and discarded. The remaining sample (200µl) was added to 600µl of buffer RLT plus which contains a highly denaturing guanidine-isothiocyanate buffer and 0.143M β-mercaptoethanol (β-ME), both to disrupt cells and to inactivate DNAses and RNAses. The mixture and the paper strip, were placed into a QIAshredder column (Qiagen, Crawley, UK) and centrifuged at 14,000g and 4°C for one minute to homogenise the lysate. The lysate was passed through a DNA all-prep spin column, which due to the high salt buffer selectively bound DNA, leaving the unbound RNA from the sample in the eluate. This RNA, once added to an equal volume of ethanol to allow efficient RNA binding, was captured on an RNeasy spin column. Leaving the DNA columns on ice, the column-captured RNA was washed with 700µl buffer RW1, then two 500µl aliquots of buffer RPE. Finally, the RNA was eluted into 60µl of RNAse-free water and stored in 4x15µl aliquots and DNA extraction was completed with two wash steps, as described in 2.3.4.1, and elution into 100µl of water (stored as 5x20µl aliquots). Both DNA and RNA aliquots were stored at -80°C prior to qPCR (or RT-qPCR).

2.4 Extraction of DNA from agarose gels

The following methods for gel extraction and PCR purification were carried out according to the Qiagen QIAquick spin protocol, March 2008. The DNA fragment to be purified was separated and identified according to size by 2% agarose gel electrophoresis. The fragment was then excised with a clean scalpel and added to three volumes of buffer QG prior to incubation at 50°C for 10 minutes. Once fully dissolved, the agarose was added to one volume of isopropanol, mixed and applied to a Qiaquick column. The column-adsorbed DNA was then washed with 500µl of buffer QG, to remove any remaining agarose, then 750µl of ethanol-based buffer PE to remove unwanted salts. The column was dried by full-speed centrifugation and DNA eluted with 50µl of the basic and low-salt buffer EB.

2.5 PCR product purification

PCR product was added to 5 volumes of buffer PB and 10µl of 3M Na acetate (pH 5), as a pH <7.5 is required for efficient DNA binding to the purification column. Once applied to the column, the DNA was washed with 750µl of buffer PE and then eluted into 50µl of buffer EB.

2.6 Gene sequence analysis

Chain-termination or 'Sanger' gene sequence analysis confirmed the presence of macrolide resistance-associated point mutations; to confirm the cloned plasmid sequences and to identify the requisite *tpp0548* sequence for *T. pallidum* strain type analysis. In this method, a DNA polymerase was used to amplify the target sequence of either the sense or antisense strand of the target DNA in a primer-specific reaction. (227) The sequencing reaction mix contained additional modified nucleotides which lack a 3'-hydroxyl group and are fluorescently labeled – each nucleotide (A,C, T or G) emitting light at a different wavelength. During the sequencing reaction, DNA polymerase was blocked at random each time a modified nucleotide was incorporated resulting in different length oligonucleotides. Subsequently, when separated electrophoretically, the fluorescence of each oligonucleotide length was read in ascending order of nucleotide size to produce a DNA sequence.

Sequencing reactions and electrophoresis were achieved using the Applied Biosystems dRhodamine kit and 3100 capillary sequencer.

Target DNA was generally sequenced in both the forward and reverse directions by addition of sense or antisense primers, respectively. A consensus sequence was then established by comparing the chromatograms of both sequences.

2.7 Reverse transcription of RNA to cDNA

This method used the enzyme reverse transcriptase to convert RNA into complementary DNA (cDNA). It was performed either in a 'one-tube' reaction, where sequence specific primers both reverse transcribed and subsequently PCR-amplified the target sequence, or in two separate reactions.

Two-step reverse transcription began with elimination of any contaminating gDNA by DNAse treatment. gDNA wipeout buffer (Qiagen, Crawley, UK) was added to RNA and water to give a total volume of 14µl. The reaction was then incubated at 42°C for two minutes before being placed on ice. Following DNAse treatment, all 14µl of the reaction were added to 1µl of reverse transcriptase (Qiagen, Crawley, UK), 4µl 5x reaction buffer (containing magnesium and an RNAse inhibitor) and 1µl of dNTPs and the resulting mixture incubated at 42°C for 15 minutes. The reverse transcriptase was then inactivated during a final incubation for 3 minutes at 95°C.

2.8 Site-directed mutagenesis

A splice-overlap PCR created a mutated allele of the *T. pallidum tpp047* gene in which a 25 base-pair sequence is replaced with an alternative probing sequence. (228) Ultimately, a known amount of the mutated sequence was added to clinical samples prior to processing to serve as an internal control of the efficiency of DNA extraction and PCR amplification of *tpp047*. Co-extraction and amplification of the internal control from samples with wild-type sequence thus controlled for PCR inhibition and the efficiency of extraction, amplification, and quantification of *tpp047*.

Three rounds of PCR were required for the creation of the internal control sequence (Figure 2). In round one, primers TP3 and TP4 were used to amplify a 1357 base-pair section of the Nichols *T. pallidum* strain *tpp047* gene. PCR product was identified by agarose gel electrophoresis and gel-extracted as before.

In the second round of PCR, the gel-extracted PCR product from round one was added to two reactions. The first contained the sense primer from round one (TP3) and an antisense primer (TP5) complementary to 19bp of the negative strand immediately downstream of the probe-binding site. The second contained the antisense primer from round one (TP4) and a sense primer (TP6) complementary to 19bp of the positive strand upstream of the probe-binding site. Both primers, TP5 and TP6, were 5'-tailed with the internal control probe sequence. Reactions conditions were as for round one. The two reactions of this round of PCR created two amplicons, both tailed with the internal control probe sequence.

In the third round, equimolar concentrations of the round two amplicons were added to a single PCR with conditions and primers otherwise identical to round one. During this reaction, the complementary internal control sequences from the two amplicons anneal and are extended by the TP3 and TP4 primers to produce DNA where the internal control sequence is flanked on either side by the *tpp047* sequence from round one.

Figure 2.2 Construction of an internal control

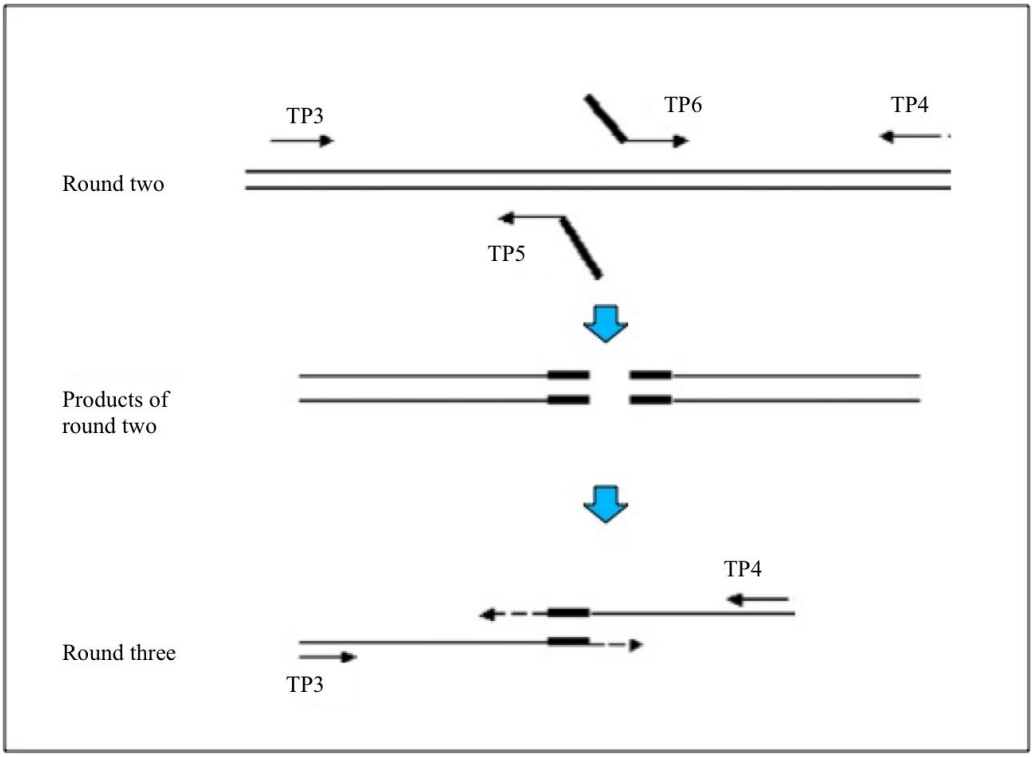


Fig 2.2. Three rounds of PCR were performed. In the first (not shown), primers TP3 and TP4 (Appendix 2) amplified the target section of the *tpp047* gene. During round two the primer pairs TP3&TP5 and TP4&TP6 produced two tpp047 amplicons, during two separate reactions, which were 5'-tailed with the 19bp internal control sequence. In the third round, the round two products annealed and were extended using primers TP3&TP4. Thick lines represent internal control sequence and dashed lines show extension of annealed complementary internal control probe sequence (228).

2.9 Construction and quantification of plasmids

TA-cloning was employed to create six plasmid constructs, each containing one of the six sequences described in Table 2.1. Following PCR-amplification of the relevant sequence, ligation into 1μl of pcr4-TOPO cloning vector (Invitrogen, Paisley, UK) (Appendix 4) was carried out at room temperature for five minutes using 4μL of the relevant PCR product and a salt solution containing 200mM NaCl and 10mM $MgCl_2$. The cloning vector is supplied in a linearised form with a single thymidine 3' overhang and the enzyme topoisomerase bound covalently to the vector in an activated form. *Taq* polymerase automatically adds an adenine base to the 3' end of the DNA it synthesizes, in a non-template dependent process, thus the complementary overhangs of vector and PCR product allow efficient combination. TOP-10 chemically competent *E. coli* (Invitrogen, Paisley, UK) were then transformed with the plasmids produced. An aliquot of 2μl of the ligated cloning vector was added to the *E. coli* and mixed gently prior to a 30-minute incubation on ice. The bacteria were then heat-shocked by placing in a 42°C water bath for 45 seconds followed immediately by a further two-minute incubation on ice. During this process, the bacteria are transformed by internalising ligated cloning vector. Transformed cells were incubated for one hour at 37°C in 250μl SOC medium, to allow expression of the plasmid-encoded ampicillin (*Bla*) resistance gene, prior to inoculation onto ampicillin-containing agar plates and incubation at 37°C overnight. Successful transformants produced colonies from which one was chosen for further culture and added to 100ml of lysogeny broth (LB) containing 100mg/ml ampicillin for incubation in a 37°C shaking incubator overnight.

Plasmid DNA was extracted from the above culture using a Qiagen Plasmid mini kit (Qiagen, Crawley, UK). Bacteria were harvested from the 100ml LB culture by centrifugation at 6000g and 4°C for 15 minutes. Pelleted bacteria were then resuspended in 0.3ml of buffer P1 (50mM Tris.Cl pH 8.0;10mM EDTA). Next, 0.3ml of alkaline Buffer P2 (200mM NaOH; 1% SDS) was added and the bacteria lysed during a 5-minute incubation. Following the addition of 300μl of chilled buffer P3 (3.0M Na acetate, pH 5.5) and incubation on ice for 5 minutes, the precipitated gDNA, proteins and cell debris are removed by centrifugation at full-speed for 10 minutes. A Qiagen-tip column was next equilibrated with 1ml buffer QBT (750 mM NaCl; 50mM MOPS, 15% isopropanol, 0.15% Triton X-100) and the supernatant, containing

plasmid DNA, applied to the column. Captured in the column resin, the plasmid DNA was then washed with 2 x 2ml aliquots of buffer QC (1 M NaCl; 50mM MOPS; 15% isopropanol). The DNA was finally eluted, by passing 800µl buffer QF (1.25 M NaCl; 50mM Tris.Cl, pH 8.5; 15% isopropanol) through the column, then precipitated by adding 560µl isopropanol, mixing and centrifuging at 15,000g for 30 minutes at room temperature. The DNA pellet was then washed by adding 1ml of 70% ethanol and centrifuging for another 10 minutes at 15,000g. Once washed, the pellet was air-dried and re-dissolved in 200µl of Qiagen buffer TE (10mM Tris.Cl, pH 8.0; 1mM EDTA).

The concentration of extracted plasmid DNA (µg/µl) was determined by spectrophotometry using a P360 UV/visible nanophotemeter (Implen, Munich, Germany) The number of plasmid copies was then calculated by first multiplying this by Avogadro's constant and subsequently dividing this concentration by plasmid weight. The weight of one plasmid is equal to its length (in base pairs) multiplied by the molar mass (g/mol) of a single base pair.

$$Plasmids/\mu l = \frac{plasmid\ concentration\ (\mu g/\mu l) \times 6.022 \log_{10} 23\ (molecules/mole)}{plasmid\ length\ (bp) \times 650 g/mol\ of\ bp}$$

Confirmation that the plasmids contained the desired *T. pallidum* or TYLCV sequences was achieved with gene sequence analysis.

Following construction, plasmids were linearised in order to improve the efficiency of downstream reactions and assays. Linearisation was achieved either with the restriction enzyme *nco-I* (New England Biolabs, Ipswich, MA), which makes a single cut in the plasmid, or by excision of the requisite *T. pallidum* sequence from the plasmid by restriction digestion with *Eco-R1 (*New England Biolabs, Ipswich, MA) and subsequent gel-extraction.
Nco-I restriction digestion was performed on plasmids pTYLCV, pTp47, pTpQC and p16SRNA. The remainder were Eco-RI digested. The conditions used for endonuclease restriction digestion are detailed in Appendix 3.

Following linearisation and quantification by spectrophotometry, plasmids pTp47; p16SRNA and pTYLCV were re-analysed with real-time PCR in a series of limiting dilution assays in

order to construct standard curves for *T. pallidum* and TYLCV quantification. By knowing the starting quantity of DNA and the probability of a negative reaction at a given dilution it is possible, using the following equation, to calculate gene copy number. (229)

$$Gene\ copy\ number/reaction = \ln[frequency\ of\ a\ negative\ reaction] \times dilution\ factor$$

Table 2.1 Plasmids constructed for the project

1. pTpQc Partial *tpp047* sequence (1357bp) mutated to contain the internal control sequence.
2. pTp47 Partial (326bp from 409 to 715) *tpp047* sequence from *T. pallidum* Nichols strain (Genbank ref: AE000520). Sequence amplified with primers Tpnest4F and Tpnest4R (Appendix 2).
3. pTp2058 Partial (629bp) *23S rRNA* gene sequence from Street 14 strain (Genbank ref: AF200365). This section of the gene is identical for both copies of the gene (located 231950 to 234850 and 280394 to 283294). Sequence amplified with primers A2058G_F and A2058G_R (Appendix 2).
4. pTp2059 Partial (629bp) *23S rRNA* gene from UW228B, an A2059G mutant. Sequence amplified with primers A2058G_F and A2058G_R (Appendix 2).
5. p16SRNA Partial (131bp) *16SrRNA* gene from *T. pallidum* nichols strain (Genbank ref: AE000520, bases 231464 to 231593). Sequence amplified with primers 16S_F(seattle) and 16S_R7.
6. pTYLCV Partial (100bp) sequence of TYLCV virus (Genbank ref: X61153.1). Sequence amplified with the primers TYLCV_F3 and TYLCV_R3 (Appendix 2).

Table 2.1. A total of seven plasmids were constructed for use as quantification standards and positive controls.

2.10 Point mutation analysis for macrolide resistance

The identification of two base substitutions, from adenine to guanine at positions 2058 and 2059 of the *T. pallidum 23S rRNA* gene, was achieved with PCR-amplification followed by restriction digestion and gene sequence analysis. (169)

Two 629 base-pair sections, one from each copy of the *T. pallidum 23srRNA* gene (Genbank ref. AF200367), were PCR-amplified from extracted DNA using two rounds of PCR. In the first round, the primers TP0226_F and TP0226_R amplified the first allele, and TP0267_F and TP0267_R the second allele (Appendix 2). Each reaction contained: 1 unit FastStart Taq DNA polymerase (Roche, Mannheim, Germany), 0.2 μM of each primer, 1x reaction buffer containing 20mM $MgCl_2$ (Roche, Mannheim, Germany), 0.2mM PCR grade nucleotide mix (Roche, Mannheim, Germany) 5μL of target DNA and PCR-grade water to a total volume of 50μL. Cycling conditions were as follows: three minutes at 93°C followed by 35 cycles of 93°C for one minute, 62°C for 2 minutes and 72°C for one minute and a final 72°C extension step for 10 minutes.

Conditions for the second round were identical, except for an annealing temperature of 63°C; the primers A2058G_F and A2058G_R (Appendix 2) for both alleles and the addition on 1μL of product from round one in place of 5μL DNA.

Separate aliquots of each amplicon were subject to restriction digestion using the endonuclease *Mbo-II* (New England Biolabs, Ipswich, MA) at 37°C for 1 hour, or *Bsa-I* (New England Biolabs, Ipswich, MA) at 50°C for one hour, as previously described (169, 170) (Appendix 3). The resultant DNA fragments were analysed by two percent agarose gel electrophoresis, which revealed bands of 440bp and 188bp for a positive *Mbo-II* digest (ie. an A2058G mutant) and 432bp and197bp for a positive *Bsa-I* digest (ie. an A2059G mutant). Additionally, the gene sequence of a third, undigested, aliquot was determined using the primer A2058G_F. The plasmids pTp2058 and pTp2059, containing DNA from Street 14 strain (A2058G mutant) or UW228B (A2059G mutant) respectively, constituted positive controls for restriction digestion and Nichols strain *T. pallidum* was the negative (wild-type) control.

2.11 *T. pallidum* strain type analysis

T. pallidum strain identification was based on the analysis of the genes *arp*, *tpr* and *tpp0548*, as described previously. (82, 86) A no-template control (molecular-grade water) and a positive control (Nichols strain *T. pallidum*) were included in each PCR.

2.11.1 arp analysis by PCR

arp analysis began with PCR-amplification of the *arp* gene by adding 0.4μM each of the primers ARP_F and ARP_R (Appendix 2); 2 units of FastStart Taq polymerase; 200μM PCR grade nucleotide mix; 1mM $MgCl_2$ and 1x PCR reaction buffer to 5μL DNA and making up to 50μL with molecular-grade water. Cycling conditions were as follows: 94°C for 4 minutes; 40 cycles of 94°C, 60°C and 72°C, each for one minute, and finally 10 minutes at 72°C. The amplicons were of varying size depending on the number of identical 60bp repeats they contained. Determination of the number of repeats gave the *arp* type.

For size-determination, samples were applied to a 2% agarose gel and run, together with a 100bp ladder (Invitrogen, paisley, UK), at 70V for around six hours. The gel was analysed using LabWorks Image Acquisiton and Analysis Software version 4.6 (Media Cybernetics Inc, Bethesda, USA). The size of each amplicon was compared to that of the Nichols strain, known to have 14 60bp repeats. Each 60bp difference represented a different type, and was numbered accordingly.

2.11.2 *tpr* analysis

tpr analysis also begins with PCR amplification by adding 0.4μM each of the primers EGJ_F and EGJ_R (Appendix 2); 1 unit of FastStart Taq polymerase; 200μM Roche PCR grade nucleotide mix; 1x Roche PCR reaction buffer containing 20mM $MgCl_{2)}$ to 5μL DNA and making up to 50μL with molecular-grade water. Cycling conditions were as for *arp* amplification. Amplicons were subsequently PCR-purified using the QIAquick purification system, as described before, and digested for one hour at 37°C with the endonuclease *Mse-I* (New England Biolabs, Ipswich, MA) (Appendix 3). In order to compare restriction digestion patterns, DNA fragments were separated by two percent agarose gel electrophoresis run at

100v for 2-5 hours with a 1kb ladder. The resultant band compositions were compared with standard profiles (82).

2.11.3 tpp0548 typing

tpp0548 typing requires amplification and sequence analysis of a section of the *tpp0548* gene. (86) The PCR used the primers 548_S and 548_AS3 (Appendix 2) in a reaction with the same reagents and concentrations as for *tpr* amplification described previously. Cycling conditions were as follows: 95°C for 4 minutes; 40 cycles of 95°C, 62°C and 72°C each for one minute and finally 10 minutes at 72°C. The amplified DNA was then PCR purified, as before, and sequenced in both directions using the primers 548_S and 548_AS (Appendix 2). The sequence is matched to one of eight established sequences and assigned a letter (Figure 2.3).
The three components of the strain type are then assembled and presented in the order *arp, tpr, tpp0548.* For example: '14d/g' which has 14 60bp repeats in *arp, MseI tpr* digestion pattern 'd' and *tpp0548* gene sequence 'g'.

Figure 2.3 *T. pallidum* subtypes identified by gene sequence analysis

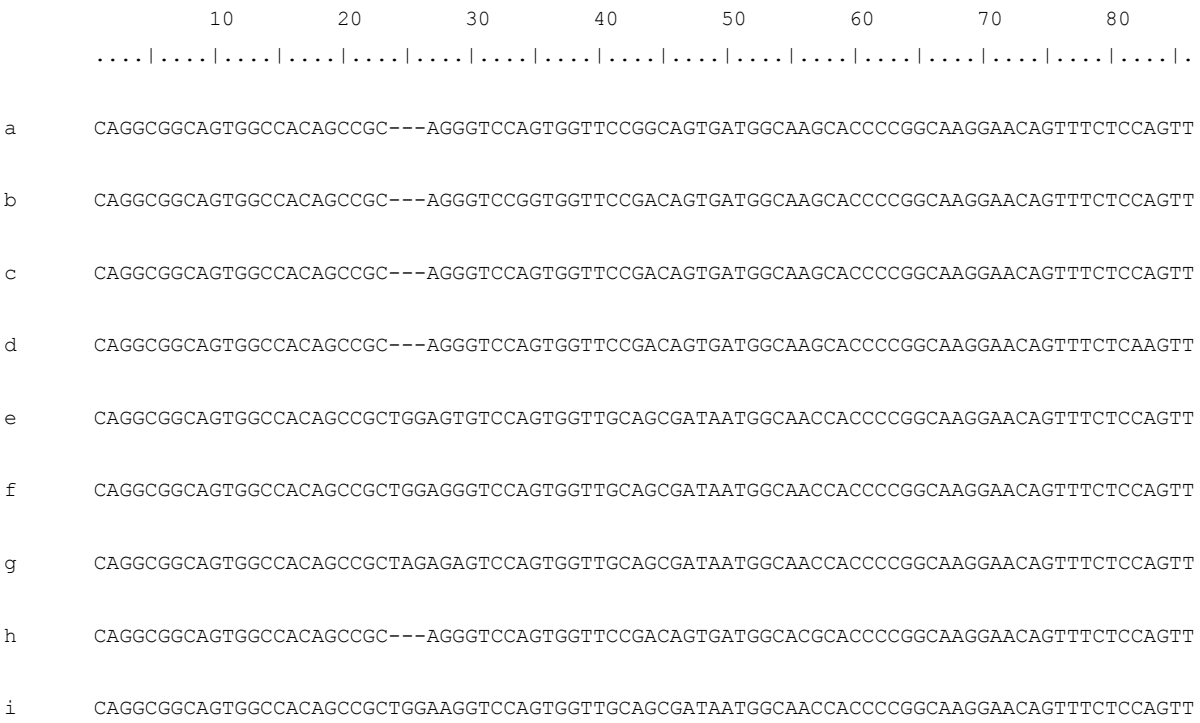


Fig. 2.4. In the enhanced method for strain-type analysis, a third *T. pallidum* gene, *tpp0548*, is PCR-amplified and its sequence determined in both direction with dye-termination sequencing. The consensus sequence is then compared with published sequences a-i (86).

2.12 *T. pallidum* quantification by polymerase chain reaction

Real-time, quantitative PCR (qPCR) was used to quantify *T. pallidum* nucleic acids in clinical samples. In a qPCR, fluorescent dyes included in the reaction are used to measure the amount of DNA present at the end of each cycle. Changes in fluorescence are used to calculate the amount of DNA present in samples at the beginning of the qPCR. The LightCycler 2.0 real-time system (Roche, Mannheim, Germany) and SYBR-green I dye were used initially, but subsequently replaced with the CFX96 real-time system, C1000 thermal cycler (Biorad, Hercules, USA) and fluorescently labeled TaqMan probes.

2.12.1 DNA quantification

2.12.1.1 *Quantification of tpp047 in vitro by Roche LightCycler*

A 178bp segment of the 1377bp *T. pallidum tpp047* gene (Genbank accession number M88769.1) was amplified using the primer pair TP1 and TP2 (Appendix 2) (nucleotides 525 to 547 and 702 to 680 respectively). Each reaction contained: 10µL extracted DNA; 0.6 units LightCycler Uracil-DNA Glycosylase (UDG) (Roche, Mannheim, Germany); 0.5µM of each primer; 4µL FastStart Plus reaction mixture (Roche, Mannheim, Germany) (containing FastStart Taq DNA polymerase, reaction buffer, $MgCl_2$, SYBR Green I dye and dNTP mix), and PCR-grade water to a final 20µL volume.

Real-time PCR was performed using a Roche LightCycler 2.0, software version 4.0, with the following conditions: UDG activation at 40°C for 10 minutes, pre-incubation at 95°C for 10 minutes and then 46 cycles of: 95°C for 10 seconds; 60°C for five seconds; 72°C for eight seconds, and 81°C for 10 seconds (acquisition). Two no template controls (water) were included in each experiment and standard precautions to avoid contamination were employed. Samples were run in duplicate and discordant results were repeated. Negative samples were diluted 1:10 and repeated if inhibition suspected. Melting curve analysis was performed on all amplicons from clinical samples and compared to an in-run standard in order to confirm amplification of the correct target. Fluorescence emission in the 530nm channel at a linear temperature transition rate of 0.1º/s from 60ºC to 95ºC was monitored with continuous acquisition. The melting temperature (Tm) of the correct 178bp amplicon was 82.5ºC.

The *T. pallidum* DNA used for both the in-run standard curve and positive PCR control was derived from Nichols organisms as described in Chapter 2.1.2.

2.12.1.2 *tpp047 quantification with the Biorad CFX Real-time system*

This method replaced SYBR green I (above) with a fluorescent TaqMan probe. The probe was a 25bp oligonucleotide with the dye Carboxyfluorescein (6-FAM) added to the 5' end and the 'quencher' tetramethylrhodamine (TAMRA) capping the 3' end. When in proximity (ie. bound to the oligonucleotide) the fluorescence of 6-FAM is quenched by TAMRA. During the PCR, the probe binds to the single-stranded DNA template. Taq polymerase exhibits 5'-3' exonuclease activity and when it meets the probe during the extension phase it cleaves the fluorophore such that its fluorescence is no longer quenched (Figure 2.4). As the PCR proceeds, the level of fluorescence increases in parallel with the amount of PCR product. As with the SYBR green I assay, a crossing point is established which represents the starting copy number. In contrast, the quantification standard used with this method is the plasmid pTp47.

Each reaction contained 12.5µl 2x Quantitect PCR mix (Qiagen, UK) (reaction buffer, 8mM $MgCl_2$, dNTPs and HotStarTaq DNA polymerase); 0.4µM each of the primers TP1 and TP2; 0.2µM of TPP probe; 5µl DNA and water to a total volume of 25µl. Cycling conditions consisted of 15 minutes at 95°C for Taq activation, then 44 cycles of 95°C for 15 seconds and 80 seconds at 60°C (combined annealing and extension step). Samples were run in triplicate and those with only a single positive replica were repeated. Each run contained a quantification standard of pTp47 diluted 1:10 from 10^5 to 10 copies and two no-template (water) controls.

Figure 2.4 Mechanism of action of 'TaqMan' PCR probes

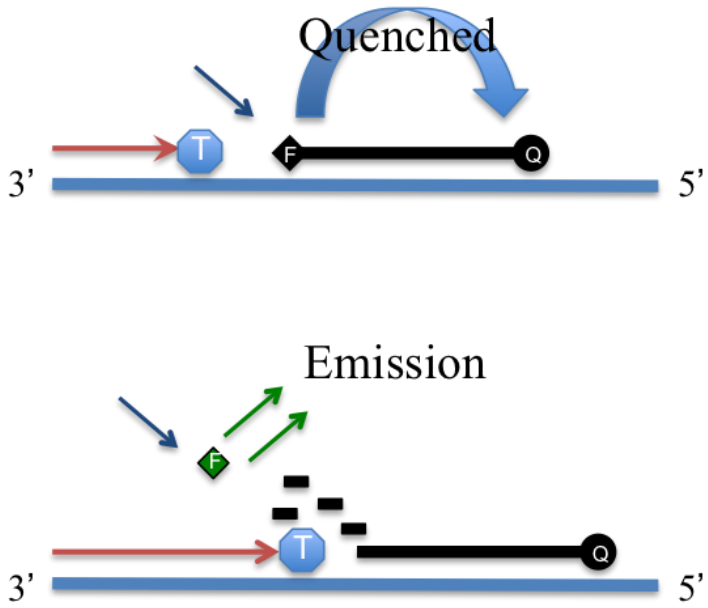


Fig. 2.5. Initially, the 'TaqMan' probe (black line) 'fluorophore' (F) is in close proximity to the 'quencher' (Q) and does not fluoresce. Following sequence-specific binding to template DNA the fluorophore and quencher are separated when taq polymerase (T), which possesses 5'-3' exonuclease activity, arrives at the probe binding site from the downstream primer (red arrow) cleaving the fluorophore and probe nucleotides (black blocks). Following its release, the fluorophore fluorescence (green arrows) is no longer quenched and can be detected by the real-time PCR instrument.

2.12.1.3 Multiplex detection T. pallidum and an internal control

This multiplex PCR uses two distinct primer pairs and two probes, each with a different fluorophore, to co-amplify *tpp047* and pTYLCV DNA. Spiking each blood sample with the same quantity of pTYLCV (5000 copies) and then quantifying *T. pallidum* and plasmid DNA simultaneously allows determination of the efficiency of DNA extraction and PCR amplification.

The reaction conditions for the multiplex PCR included a Uracil-N-Glycosylase (UNG) step. This hydrolytic enzyme removed uracil from dUMP in any contaminating molecules in the PCR. The resultant DNA, with apyrimidinic sites, was subsequently destroyed by hydrolysis of the phosphate backbone at these sites during the Taq activation step at 95°C. The final 25µl reaction thus contained: 12.5µl of 1x Qiagen Quantitect Multiplex PCR reaction mix; 0.5units UNG; 0.4 µM of each of the four primers (TP1; TP2; TYLCV_F3 and TYLCV_R3); 0.2µM each of the probes TPP and PTYCLV; 5µl of DNA and water to a total volume of 25µl. Cycling conditions were as follows: 50°C for 2 minutes (UNG step); 95°C for 15 minutes (Taq activation and UNG inactivation) then 45 cycles of 94°C for 1 minute; 60°C for 75 seconds (annealing, extension and plate read). Two standard curves were included in every run, each comprising a 1:10 dilution series from 10^5 to 10 copies per reaction. The standards were the plasmids pTp47 and pTYLCV, constructed as described previously. A no-template (water) control and extraction controls, both positive and negative, were also run simultaneously.

2.12.2 16SrRNA quantification with Biorad CFX Real-time system

This one-step RT-qPCR assay was selected for *T. pallidum* RNA quantification in blood and ulcer samples using the *16S rRNA* target. Each 25µl reaction contained: 1µl enzyme mix (Omniscript reverse transcriptase; Sensiscript reverse transcriptase and HotStarTaq DNA polymerase) (Qiagen); 1µl DNTP mix (0.4mM final concentration of each nucleotide); 5µl of reaction buffer (including 12.5mM $MgCl_2$); 0.6µM each of the primers 16S_F_Seattle and 16S_R7; 0.2µM of the probe 16S_P3; 5µl RNA and water. Cycling conditions began with a 30-minute step at 50°C for reverse transcription followed by 15 minutes at 95°C both to inactivate the two RT enzymes and to activate the Taq polymerase. Next were 44 cycles of

95°C for 15 seconds then 60°C for 90 seconds. Fluorescence at 530nm was measured at the end of each round of PCR. Samples were run in triplicate and runs contained two no-template (water controls). A 1:10 dilution series of the plasmid p16SRNA was included in each run as an absolute quantification standard.

2.13 Development of an RNA standard for *T. pallidum* quantification

2.13.1 RiboMAX T7 RNA polymerase

Several methods of producing *T. pallidum* RNA in sufficient quantities for use as quantification standard were assessed and ultimately the RiboMAX Express T7 RNA polymerase (Promega, Southampton, UK) was selected. This DNA-dependent phage RNA polymerase binds with high specificity to its priming site (T7 promoter sequence) and subsequently produces RNA from a DNA transcriptor vector downstream of the T7 promoter. The pCR4-TOPO cloning vector, used to produce the *T. pallidum* DNA quantification plasmid (pTp47), contains a T7 priming site 31 bases upstream from the cloning site which enabled a T7 polymerase to produce RNA copies of the plasmid p16SRNA. The T7 Ribomax express system contains the polymerase T7 Express, 2µl of which were added to $8x10^5$ copies of *Nco-1* linearised p16SRNA. The reaction also contained the recombinant enzymes Ribonuclease inhibitor (RNAseIn) and Inorganic Pyrophosphatase (in the T7 polymerase enzyme mixture) and 10µl of reaction buffer (HEPES, pH7.5; NTPs and Magnesium). The reaction volume was made up to 20µl with water and incubated at 37°C for 30 minutes. To remove contaminating template DNA, 2µl of RQ1 DNAse (Promega, Southampton, UK) were added, then the reaction stirred and incubated for 15 minutes at 37°C. To purify the resultant RNA and remove any unincorporated rNTPs, the entire reaction volume was applied to a Roche Highpure extraction column (as described previously), washed and eluted into 100µl water.

2.13.2 Invitrogen T7 polymerase

Invitrogen T7 polymerase was assessed as an alternative to RiboMAX T7 polymerase. The basic manufacturer's protocol was as follows: 10µL (1x final concentration) trans-optimised buffer were added to 5mM DTT; 50 Units RNAseIn; 1mM rNTP mix; 50 Units bacteriophage

RNA polymerase; 10µl of plasmid DNA (10^6 plasmid copies) and water to a total volume of 50µl. The reaction was then mixed and incubated at 37°C for one hour. Following the T7 polymerase reaction, template plasmid DNA was removed with on-column DNAse treatment and RNA was eluted into 100µl water using the Roche highpure method described in Chapter 2.1.2.

2.13.3 Incorporating T7 promoter sequence into TYLCV target sequence

This modification of the splice-overlap PCR described previously was designed to add the T7 promoter sequence to the 5' end of the TYLCV target sequence. New primers were synthesised: The primer TYLCV_F3 was modified by adding the 20bp T7 promoter sequence (shown here in bold) to the 5'end (**TAA TAC GAC TCA CTA TAG GGC** GTC CGT CGA) and the primer TYLCV_R3 was elongated (shown in bold) to produce a comparable melting temperature (**AGG TCA GCA CAT TTC C**AT CCG AAC ATT CAG GGA GCTA). During the first cycle of PCR, the primers annealed to the TYLCV target sequence and the T7 promoter sequence, which was initially a 5' overhang, was reverse complemented. During the second cycle, the TYLCV-T7 hybrid sequence, was replicated and soon became the dominant species in the PCR. The predicted product size of 136bp was confirmed by agarose gel electrophoresis. The PCR product was purified and used as a template for T7 polymerase, with the standard conditions described previously (Chapter 2.5).

2.14 Clinical Studies

The data from four separate clinical studies are described in this thesis. Recruitment criteria and details of ethical approvals for each are detailed here and copies of the associated approvals and protocols are included in the appendices.

2.14.1 Validation of *T. pallidum* DNA quantification in blood and ulcer samples

Patients were recruited for the initial validation of *T. pallidum* detection and quantification by SYBR-green base qPCR in a prospective, cross-sectional study, conducted at St Mary's Hospital London, following ethical approval from The Hounslow and Hillingdon LREC (06/Q0407/29, 26th May 2006). All patients being investigated for syphilis (including routine

screens) were eligible. A sample of 6ml of whole blood was collected into EDTA and for patients 25-99, an additional 3ml of blood was collected into a Tempus™ tube. Samples collected between July and December 2006 were processed immediately and DNA extracted as described previously. Samples collected after this period were stored at -80°C until DNA and RNA were extracted. Ulcer exudate was collected with snostrips according to the method in Chapter 2.3.4. The strips were stored dry in microcentrifuge tubes until DNA extraction. DGM was performed on genital ulcers according to clinic policy prior to the collection of study samples.

Baseline and convalescent syphilis serology were documented for all patients in addition to the results of investigations for other STIs. The Newmarket IgM/IgG enzyme immunoassay (EIA), (Alere, Stockport, UK) was used to screen for *T. pallidum* infection. Fujirebo *T. pallidum* particle agglutination (TPPA) (Mast Diagnostics, Bootle, UK) and Bio-kit rapid plasma reagin (RPR) (Launch Diagnostics, Longfield, UK) were used to confirm and quantify the serological response respectively.

Patients' final clinico-microbiological diagnoses were recorded as one of the following: *Herpes simplex* virus (HSV); primary syphilis; secondary syphilis; other syphilis (comprising latent and symptomatic late disease); neuro-syphilis; asymptomatic contacts of syphilis and non-specific genital ulceration. Case definitions are given in Appendix 5.

2.14.2 Cross-sectional study of *T. pallidum* quantification

A second set of cross-sectional samples was collected prospectively from patients presenting to St Mary's Hospital, London with untreated syphilis of any stage between January 2011 and December 2012. Ethical approval for sample collection was granted from the Communicable Diseases Tissue Biobank (CDTB), Imperial College London.
Whole blood (3-5ml) was collected into EDTA from each patient in addition to 3ml of blood into a Tempus™ collection tube. Ulcer swabs were taken from patients with a positive DGM result and stored in microcentrifuge tubes containing 1.3ml RNA later. CSF was also collected from patients undergoing Lumbar Puncture examination for suspected neurosyphilis.

Additional EDTA blood samples were also collected from patients who were admitted to hospital for their syphilis treatment. These samples comprised EDTA blood surplus to requirement for full blood count measurement, which had been stored at 4°C after collection. The same demographic data, clinical details and serological results were documented for patients in this study as for the validation study. In addition, information on previous syphilis treatment and recent treponemocidal antibiotics was collected. All samples were stored at -80°C prior to nucleic acid extraction.

2.14.3 Molecular characterisation of *T. pallidum* in Colombo, Sri Lanka

Patients attending the Central STD clinic, Colombo between January and October 2012 with confirmed symptomatic early syphilis (primary chancre or moist lesion of secondary syphilis) were considered eligible for the study. Ethical approval was granted by the Ethics Review Committee, Faculty of Medicine, University of Colombo (Approval No EC-11-194, 19th Jan 2012) to recruit up to 40 patients. Following informed consent, ulcer exudate was collected using filter paper snostrips as described previously. Samples were stored in 1.3ml of RNAlater® solution at -80°C until they were shipped to the UK, at ambient temperature. Approval for the transfer of samples to the Communicable Disease Tissue Biobank at Imperial College London was granted by way of a Material Transfer Agreement between Imperial College London and The National STD/AIDS Control Programme of Sri Lanka (24th November 2011). Demographic data were recorded for each patient in addition to the results of standard serological tests (*T. pallidum* EIA and RPR) and dark ground microscopy. No patient identifiable information was transferred from Sri Lanka to London.

2.14.4 Study of treatment kinetics

Ethical approval was gained to recruit 4 patients with confirmed late primary or early secondary syphilis to this longitudinal study. Primary Disease was defined as a typical looking genital ulcer, which was beginning to heal, confirmed as syphilis by either dark ground microscopy (DGM) OR a positive ulcer multiplex PCR result from the HPA and/or serology consistent with a new infection. Positive serology was not an absolute requirement for study entry as it may not have been available on the day of recruitment. Secondary Disease was defined as a new positive EIA result with an RPR of 1:16 or more with or

without any relevant symptom or sign of secondary syphilis (listed in the national guideline for the management of syphilis and including deranged liver function tests) (104). If symptoms or signs were present and of recent onset then the case could have been classified as early secondary disease. In the absence of symptoms or signs the patient may still have been eligible for study entry if the RPR titre was high (>1:128). Previous syphilis did not exclude a patient from study entry. However, they must have had serology showing either a negative RPR or a 'serofast' RPR titre (≤ 1:4) before the current new infection. All patients had to have the ability to give informed consent for participation in the study and needed to be available for a 7-day period, including the first three potentially as an inpatient. Participating patients then needed to be available for serological follow-up, to a maximum of 6 months. Every effort was made to give patients the maximum amount of time to consider their decision to enter the study without unduly delaying their treatment or prolonging their stay in clinic. Potential subjects were given preliminary information about the study by their doctor or nurse prior to meeting the study team and posters were displayed in the waiting rooms detailing the study.

Exclusion criteria for study entry were as follows:

1. Exposure to any class of anti-treponemocidal antibiotic within 90 days of enrolment (macrolides, tetracyclines, penicillins, cephalosporins, other b-lactams eg. carbopenems)
2. Presence or suspicion of concomitant bacterial sexually transmitted infections (STI) or other bacterial infections (skin, chest etc) that require treatment with treponemocidal antibiotics.
3. Pregnancy. Potential female participants of child-bearing age who are not sterilized were tested for pregnancy prior to study entry.
4. Following full discussion with patients about the study, those who seemed insufficiently committed were not recruited.
5. Patients with significant drug or alcohol dependence and/or mental health issues were excluded.

Once recruited and once the treating physician had established the management plan, the patient was transferred immediately to an inpatient study unit. An intravenous cannula was sited (23g) to facilitate frequent venesection and samples were collected from patients according to the sample schedule in Appendix 7. Ulcer samples were collected using filter

paper snostrips, as described previously. Prior to collecting the first sample, however, the ulcers were photographed and measured with a disposable paper ruler. EDTA and Tempus™ blood samples were taken from a peripheral vein using an intravenous cannula and either a 4ml EDTA vacutainer tube pre-marked to 2mL or a Tempus™ vaccutainer collection tube. Once collected, EDTA samples were stored on dry ice until DNA extraction was performed, typically in batches of 6-8 samples. Tempus™ tubes were placed immediately in dry ice, then stored at -80°C for up to two weeks prior to RNA extraction

Chapter 3

Cross-sectional studies of *T. pallidum* nucleic acid quantification *in vivo*

3.1 Introduction

3.1.1 Background

The primary stage of syphilis consists of a self-limiting mucosal ulcer. If untreated, the bacteria invade and spread widely through the bloodstream, giving rise to the secondary stage (rash; alopecia; hepatitis; eye and neurological symptoms). Again, these symptoms will resolve spontaneously if left untreated, but a small number of slowly dividing bacteria lodged in the brain; spinal cord; cardiovascular system or elsewhere may eventually cause destructive tertiary disease over two-three decades. The disease is called 'early' in the primary and secondary stage and 'latent' when asymptomatic between stages (4, 104). Using PCR in a rabbit model the dissemination of *T. pallidum* after inoculation has been shown to be quick and wide-spread (51). Studies exploring the sensitivity of PCR for syphilis diagnosis in humans find sensitivity of 72-91% in primary chancres, and 24-73% in blood samples (133, 145, 151). RT-PCR was described as having the potential to enhance the qualitative detection of *T. pallidum in vitro*, but failed to do so in one subsequent *ex vivo* study (57, 153). No previous description of *T. pallidum* RT-qPCR has been made and only a single study of secondary syphilis has presented any *T. pallidum* qPCR (151). The humoral response to *T. pallidum* infection can be divided into nontreponemal antibodies and treponemal antibodies (100). Both antibody types have been exploited for diagnostic tests, the former for screening and as a measure of disease activity and the latter for confirmatory testing (230). Non-treponemal antibody titres are low in incubating disease, elevated in early disease and fall as disease passes into latency or after treatment.

The current study describes the development of novel qPCR and RT-qPCR assays for the measurement of *T. pallidum* bacterial load in clinical samples. This includes a determination of the most efficient methods for extracting nucleic acids and finding optimal PCR targets. It subsequently describes the validation of these assays using clinical samples collected in three separate cross-sectional studies; compares qPCR data with standard serological results; establishes the optimal blood compartment (whole blood, plasma, PBMCs) for *T. pallidum* detection and correlates bacterial load with disease stage and HIV-1 co-infection.

3.2 Results: assay development

In order to quantify *T. pallidum* nucleic acids in clinical samples it was paramount both to identify robust methods for sample collection and storage and to optimise techniques for nucleic acid extraction and quantification. DNA and RNA extraction from blood and ulcer samples is described in Chapter 2.3 and was largely based on previously published methods. However, the preservation of *T. pallidum* RNA in ulcer samples using RNAlater® solution and in blood with Tempus™ vaccutainer tubes was novel as were qPCR assays for both DNA and RNA quantification. Each of the steps taken in the development of a qPCR and an RT-qCPR capable of quantifying *T. pallidum* in clinical samples are now described.

3.2.1 *T. pallidum* nucleic acid extraction from snostrips.

The ability to co-extract DNA and RNA from a single ulcer sample was important as only one sample would be available at each time-point during the treatment monitoring study (Chapter 4) and both nucleic acids were to be quantified. Of equal importance was the need for the DNA extraction method to be compatible with RNA later solution in which snostrip-adsorbed ulcer exudates were to be stored. Two DNA extraction methods were assessed for use with snostrips stored in RNAlater® solution. In the first, an aliquot of RNAlater® was withdrawn and DNA extracted leaving a second aliquot for subsequent RNA extraction. The second method lysed and homogenised the entire sample and was followed by DNA and RNA extraction in parallel. Table 3.1 demonstrates that DNA extracted from 5000 *T. pallidum* organisms with the second method was detected an average of 1.08 PCR cycles sooner ($p=0.0058$) than DNA extracted with the first method. Moreover, by extracting DNA and RNA in parallel, the second method allowed subsequent RNA extraction both without the need for a separate process or the potential for RNA degradation as a result of re-freezing the sample.

Table 3.1 A comparison of two methods for *T. pallidum* DNA extraction from Snostrips

	Mean† cycle threshold					
	Experiment 1		Experiment 2			
Method	**Sample 1**	**Sample 2**	**Sample 1**	**Sample 2**	**Mean**	
All-prep	34.62	35.07	34.19	34.68	34.64	p=0.0058*
QIAamp	36.26	35.67	35.53	35.43	35.72	

Table 3.1. *T. pallidum* organisms (n=5,000) were added to Snostrips and stored in 1.3ml RNA later solution. DNA was extracted with either a Qiagen 'all-prep' mini kit (DNA/RNA co-extraction) or a Qiagen 'QIAamp' kit (DNA extraction alone). Two samples were extracted in duplicate (samples one and two) and the experiment repeated to ensure reproducibility (experiments one and two).

*A two-sided student's t-test was used to determine statistical significance.

†Three technical replicates of each sample were performed.

3.2.2 The Tempus™ blood tubes are suitable for *T. pallidum* RNA preservation and extraction.

Maximising the efficiency of RNA extraction was key in order to improve *T. pallidum* detection in blood, which had previously been reported to be low (52). Two techniques for RNA extraction from whole blood samples were assessed. The first was the Tempus™ system, where blood was drawn from the patient directly into a tube containing a proprietary detergent (for sample lysis) and guanidine thiocyanate (for RNAse inactivation). The RNA extraction process (Chapter 2.3.3.1) subsequently took around 60 minutes. The second technique, using Ambion's Ribopure™ system, involved an aliquot of blood being stored in RNAlater® solution and then extracted using a phenol:chloroform-based method taking 50 minutes. Both techniques included a DNAse step to remove any contaminating gDNA. A comparison of the two techniques was made by spiking whole blood with *T. pallidum* at a final concentration of 4000 bacteria/ml, extracting RNA according to the method described in Chapter 2.3.3.2 and quantifying extracted RNA by RT-qPCR (Chapter 2.12.2). Table 3.2 compares the mean Cts of an RT-qPCR amplifying the extracted RNA and shows that Tempus™-extracted RNA was detected 0.5 cycles earlier at 40.53 cycles. The mean cycle thresholds for positive control RNA ($5x10^4$ *T. pallidum* organisms per reaction) was 31.60 cycles and a no template (water) control was negative. In addition to this marginal improvement in extraction efficiency, the Tempus™ system immediately preserved RNA without the need for a laboratory (important in the clinical studies). Moreover, whilst the total time taken for extraction was similar, the Tempus™ protocol was less complex as it included a 30-minute centrifugation step.

Tempus™ tubes are designed to extract RNA from human lymphocytes, not bacteria in blood samples, thus two potential optimisation steps were assessed. The first was additional proteinase K digestion to enhance the bacterial lysis achieved with Tempus™ detergent alone. Two Tempus™ samples containing 3mL of blood drawn from syphilis negative controls were spiked with 5000 RNAlater®-preserved *T. pallidum* control organisms. Following the initial 30-minute centrifugation one sample was resuspended in 200µL PBS and 10µL Proteinase K (Qiagen, Crawley, UK), mixed and incubated at 56°C for 3 hours. A second sample was resuspended in 400µL resuspension solution (as per the Tempus™ extraction protocol). RNA

extraction from both samples was then completed according to the Tempus™ protocol (Chapter 2.3.3.1).

Figure 3.1 demonstrates that this extra step increased the cycle threshold for *T. pallidum* detection by over 10 cycles and it was not adopted. The second was selective removal of human mRNA from samples prior to RT-qPCR with magnetic bead oligonucleotide capture. As Table 3.3 demonstrates, the extraction of 5000 *T. pallidum* organisms with the Tempus™ system alone (condition 'a') resulted in a mean cycle threshold of 31.70 (std. dev. 0.59) (measured by SYBR-green RT-qPCR). Following human mRNA depletion, the cycle threshold for *T. pallidum* detection was later at 34.20 (std. dev. 0.73) cycles. A control step with the omission of the capture oligonucleotide mix (condition 'b') also resulted in a reduction of RNA starting quantity (cycle threshold 32.31 (std. dev. 0.49) cycles).

Table 3.2 *T. pallidum* RNA extraction from blood

Method	Mean* cycle threshold (std. dev.) for *tpp047* gDNA detection	Cost of extraction (£/sample)	Time taken for extraction (minutes)
Tempus™ with on column DNAse treatment[1]	40.53 (0.87)	8.80	60
Ribopure™ RNA extraction with Ambion DNAse treatment[2]	41.04 (1.1)	8.70	50
Positive Control[3]	31.60 (0.29)	-	-
NTC (water)	Negative		

Table 3.2. Whole human blood was spiked with 5,000 *T. pallidum* which had previously been stored in RNAlater®. Following RNA extraction with either the Tempus™ or Ambion Ribopure systems, *tpp047* was quantified by RT-qPCR and the cycle thresholds of detection were compared. The cost and time taken for each method were also compared.

[1]3ml blood collected into a Tempus™ tube and spiked with 4000 *T. pallidum*/ml). [2]500µl of blood spiked with 4000 *T. pallidum*/ml and added to 1.3ml RNAlater®. [3]RNA extracted from $5x10^4$ *T. pallidum*/reaction.

* Mean of three replicas in a single experiment.

Figure 3.1 Comparison of *T. pallidum* RNA extraction methods

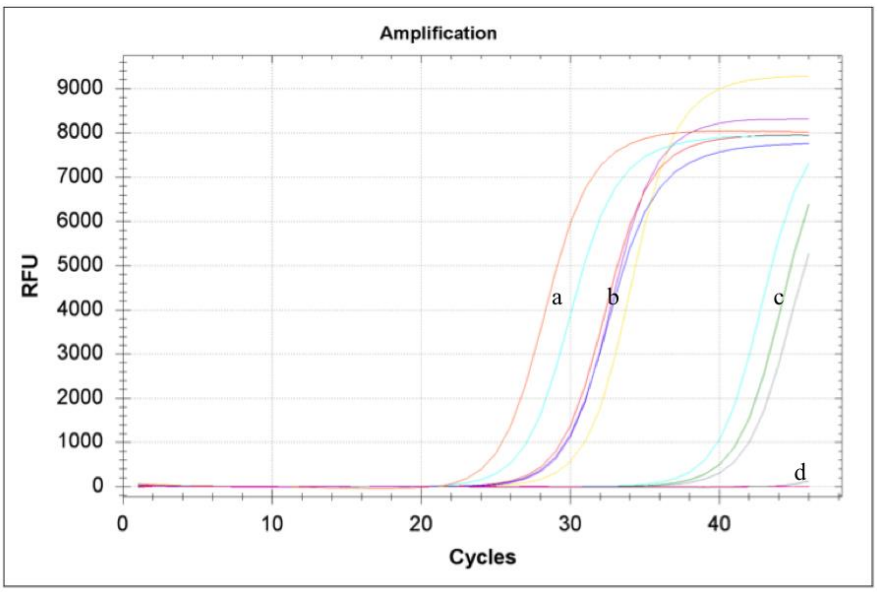


Fig. 3.1. *T. pallidum* organisms (n=5,000) were added to syphilis-negative blood stored in Tempus™ tubes and RNA was extracted both with and without an additional Proteinase K digestion step. Following reverse transcription, cDNA from a positive control (a); Tempus™-extracted samples without Proteinase K digestion (b) and Tempus™-extracted samples with additional Proteinase K digestion (c) were quantified, in triplicate, in a SYBR-green based qPCR. Aliquots of (a), (b) and (c) taken prior to reverse transcription were also subject to qPCR to control for gDNA carry-over. The mean cycle threshold for (a), (b) and (c) detection was 26.04 (std. dev. 0.91), 28.84 (std. dev. 0.91) and 40.37 (std. dev. 1.48), respectively. No amplification was observed from no-template or from gDNA carry-over controls (d).

Table 3.3 *T. pallidum* RNA extraction from blood

	Method	Mean cycle threshold (cycles)*	Standard deviation
a	**Tempus™ extraction alone**	31.70	0.59
b	**No oligonucleotide control**	32.31	0.49
c	**Human RNA depleted**	34.20	0.73
d	**PCR no template control (water)**	neg	n/a
e	**PCR positive control**	25.59	0.32

Table 3.3. *T. pallidum* negative blood (3ml) stored in Tempus™ tubes was spiked with 5000 *T. pallidum* organisms. RNA was extracted and divided into three aliquots: (a) Tempus™ extraction alone; (b) Tempus™ extraction with a negative bead control (where magnetic beads were added, but the capture oligonucleotide omitted) and (c) Tempus™ extracted RNA with bead-mediated depletion of human RNA. Together with a no-template control (d) and a positive control of *T. pallidum* RNA (e), the extracted RNA was reverse-transcribed and cDNA quantified in a SYBR-green based qPCR.

* cycle thresholds given are a mean of three replicas in a single experiment.

3.2.3 Designing a qPCR for *T. pallidum* DNA quantification

3.2.3.1 tpp047 quantification in a qPCR using SYBR green

Initially, the genome copy number of *T. pallidum* in blood and ulcer exudates was determined with qPCR using the Roche LightCycler 2.0 instrument in a SYBR green-based assay (Chapter 2.12.1.1). Enumeration of *T. pallidum* was achieved by amplifying and detecting a 177bp section of the *tpp047* gene, which appears once on the bacterial chromosome and encodes an integral membrane protein (28). The gene is specific to the species *pallidum*, thus the qPCR was expected to detect other *pallidum* subspecies (e.g. *pertenue*), but not other organisms of the genus *Treponema*.

DNA was extracted from $6x10^6$ DGM-quantified *T. pallidum* Nichols strain organisms and diluted in a 1:10 dilution series to create an absolute quantification standard against which samples containing an unknown quantity of *T. pallidum* could be compared. The qPCR quantified the amount of DNA present after each round of PCR by detecting the green fluorescence emitted by SYBR green I dye (λ_{max} = 520 nm) which had bound to double-stranded DNA. By calculating the rate of rise of fluorescence, the cycle of PCR where the fluorescence curve of the sample turned sharply upward represented the Crossing point (Cp) for the reaction. This Cp was therefore a maximum on the second derivative plot of the reaction. Figure 3.2.B demonstrates the sigmoidal amplification plots of the PCR, which depict a near-doubling of detectable DNA after each round of PCR until a plateau where declining substrate concentrations limit reaction efficiency. Amplification curves shifted to the right with each 10-fold dilution of the starting amount of DNA, which represents the additional number of PCR cycles required to achieve detectable fluorescence.

By plotting the known starting quantity for each standard curve dilution against its crossing point, the LightCycler software produced a regression line (the standard curve) which it could then use to calculate the starting amount of target DNA in unknown samples. The standard curve showed a linear relationship between starting quantity and crossing point, with an efficiency of 1.818 and an error of 0.005 (Figure 3.2.). Next, by diluting the standard curve 1:10 to reach a concentration 100-fold less than one genome copy and showing no detection beyond a single genome equivalent, the analytical sensitivity of this assay was determined to be a single *T. pallidum* organism.

In order to ensure the correct target was being amplified in the reaction, melting curve analysis was performed as described in Chapter 2.12.1.1. Figure 3.2. demonstrates the melting curve analysis for *T. pallidum pallidum, T. pallidum pertenue* and *T. denticola*. Fluorescence increased gradually until 82.5°C, the temperature at which the amplicon DNA strands dissociated (melting temperature, Tm) and no longer bound SYBR green, resulting in a sharp fall in fluorescence. The Tm for the two *T. pallidum* subspecies is seen to be the same. Neither *T. denticola* nor the reaction negative control showed any detectable fluorescence. In addition to melt curve analysis, gene sequence analysis was performed on the *tpp047* amplicon and showed complete homology with published Nichols strain sequences. Verification of assay specificity had been previously sought by confirming that the assay did not amplify DNA from healthy human volunteers in addition to sexually transmitted bacteria and viruses and common skin bacteria and fungi (Full list of organisms given in Appendix 1) (231).

The inter-assay variance was determined by quantifying three dilutions of the Nichols strain standard curve ($1x10^4$ to $1x10^2$) in five separate experiments. The coefficient of variation (Table 3.4) was less than 1.2% for all three dilutions.

Figure 3.2 Development of a qPCR for *T. pallidum* quantification

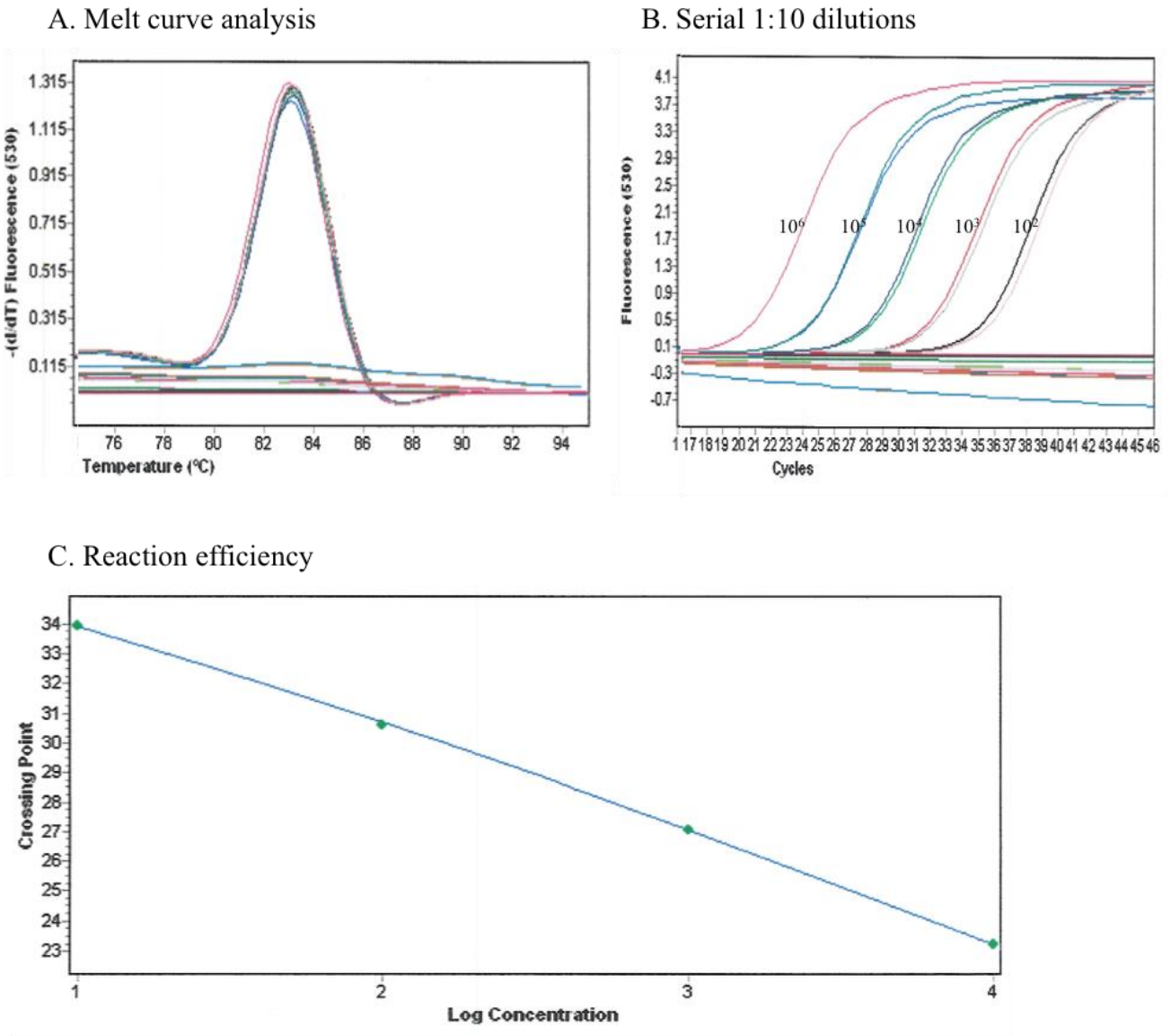


Fig. 3.2. *T. pallidum tpp047* detection and quantification was peformed with the Roche Lightcycler. Two technical replicates of each dilution were assayed. **Panel A**. Melt curve analysis of the product of *tpp047* PCR amplification from *T. pallidum pallidum* and *T. pallidum pertenue*. Neither the negative (water) control nor *T. denticola* show detectable PCR product. **Panel B**. Sigmoidal amplification curves resulting from PCR amplification of a pTp47 dilution series from 10^6 to 100 copies per reaction. **Panel C**. Plot of log pTp47 starting concentration (1:10 dilution series from 10^4 to 10) against crossing point for each concentration. The calculated reaction efficiency from this plot was 1.818.

Table 3.4 Inter-assay variance of *tpp047* qPCR

Dilution	Mean crossing point	Standard deviation	Coefficient of variation (%)
1 (10^4)	23.21	0.23	0.97
2 (10^3)	26.96	0.31	1.13
3 (10^2)	30.61	0.36	1.18

Table 3.4. Three dilutions of DNA extracted from *T. pallidum* Nichols strain organisms were quantified by qPCR using SYBR green I and a LightCycler 2.0 instrument. The mean crossing point for the detection of each dilution was calculated from five separate experiments.

3.2.3.2 tpp047 qPCR using a plasmid quantification standard in a TaqMan assay

qPCR based on the detection of SYBR green I was an efficient method of *T. pallidum* quantification. Another goal, however, was to design a multiplex qPCR capable of quantifying *tpp047* and an internal control sequence concurrently. The TaqMan assay has the capability to detect two distinct fluorescent probes and thus to measure amplification of both *tpp047* and a control sequence in the same reaction tube. As described in Chapter 2.12.1.2 a FAM-TAMRA TaqMan probe was designed for the quantification of *tpp047* in a qPCR using the same amplification primers as for the SYBR-green assay (TP1 and TP2, Appendix 2). The assay detected Nichols strain *T. pallidum* DNA, but amplified neither human DNA nor the test organisms listed in Appendix 1. Two further modifications were made to the assay. The first was replacing the Roche LightCycler 2.0 instrument with a Biorad CFX 1000, which allowed quantification of an additional 96 samples in a single experiment. Secondly, the standard curve DNA extracted from *T. pallidum* organisms was replaced with a recombinant plasmid, pTp47. Construction of this plasmid is described in Chapter 2.9. Biorad CFX manager 2.0 software was used to analyse qPCR data. In a similar way to the Roche light-cycler instrument software, fluorescence is measured at the end of each round of PCR, although this time the fluorescent source is hydrolysed Taqman probe molecules. The rate of rise in fluorescence is then subject to regression analysis and used to determine the cycle threshold - the cycle of PCR where fluorescence begins to rise exponentially and which is closely associated with the starting quantity of DNA in the reaction. Importantly, the determination of cycle threshold is calculated by the software and is not, therefore, subject to observer error.

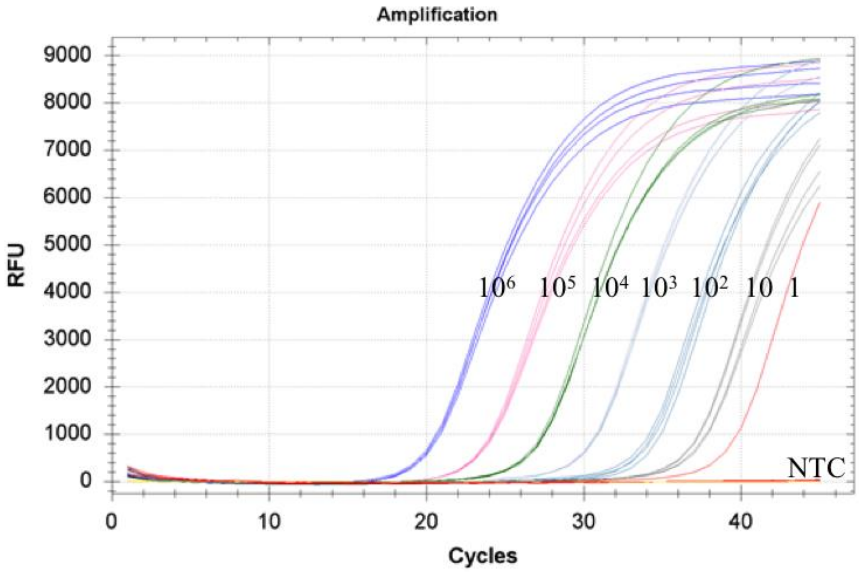
Before use as a quantification standard, the plasmid pTp47 was linearised and determined to have a concentration of 6.06×10^9 *tpp047* copies/μl using spectrophotometry. The plasmid was then resuspended to obtain a concentration of 1×10^9 copies/μl and subject to serial 1:10 dilution down to a single copy per microlitre. qPCR of dilutions five (10^5 copies/μl) to nine (one copy/μl) was performed in quadruplicate and calculated to be 103% efficient with an average of 3.17 cycles between each 10-fold dilution of pTp47 (Figure 3.3). The lowest dilution at which *tpp047* was amplified in all four replicas was 10 copies/μl – the analytical sensitivity of the assay. The probability of a negative reaction at one copy per microliter (the volume of DNA added to each PCR) was 0.75. According to the equation in Chapter 2.9 this

probability at a dilution factor of 10^9 equated to an original concentration of $4.7x10^9$ copies/µl, which is comparable to the $6.06x10^9$ copies/µl calculated by spectrophotometry.

Inter-assay variability was determined by quantifying pTp47 (*tpp047*) in dilutions of the newly constructed standard curve containing 10^4, 10^3 and 10^2 copies/µl. Four separate experiments were averaged (Table 3.5) and the coefficient of variation was found to be below 2.4%. Intra-assay variability was also low, at less than 1.05% for all three concentrations assessed (Table 3.6).

Figure 3.3 Construction of a standard curve for *tpp047* quantification

A. Amplification curves



B. Standard Curve Efficiency

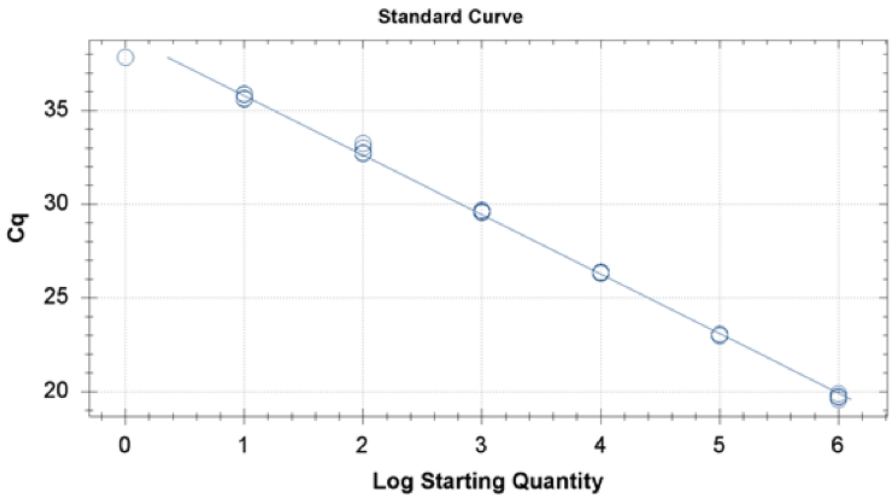


Figure 3.3. The plasmid pTp47 was serially diluted in water from $1x10^5$ copies/µL to 1 copy/µL and quantified in a TaqMan qPCR using the primers TP1&TP2 and the probe TPP (**Panel A**). Four (technical) replicates of each sample were quantified using the Biorad CFX real-time PCR instrument. A standard curve was produced where each 10-fold dilution corresponded to an increase in cycle threshold of 3.17 cycles (**Panel B**).

Table 3.5 Inter-assay variation of the *tpp047* qPCR

Dilution	Mean cycle threshold	Standard deviation	Coefficient of variation (%)
1 (10^4)	26.98	0.63	2.35
2 (10^3)	30.41	0.70	2.32
3 (10^2)	34.56	0.67	1.95

Table 3.5. Three dilutions of the plasmid pTp47 were quantified by qPCR using a TaqMan assay with the primers TP1 and TP2 and the probe TPP (Appendix 2). The mean cycle threshold for the detection of each dilution was calculated from four separate experiments.

Table 3.6 Intra-assay variation of the *tpp047* qPCR

Dilution	Mean cycle threshold	Standard deviation	Coefficient of variation (%)
1 (10^4)	26.08	0.172	0.66
2 (10^3)	29.40	0.308	1.04
3 (10^2)	32.57	0.308	0.95

Table 3.6. Three dilutions of the plasmid pTp47 were quantified by qPCR using a TaqMan assay with the primers TP1 and TP2 and the probe TPP (Appendix 2). The mean cycle threshold for a given plasmid concentration was calculated from four technical replicas quantified in a single experiment.

3.2.4 Designing a qPCR for *T. pallidum* RNA quantification

3.2.4.1 16S rRNA was selected as the target gene for T. pallidum RT-qPCR

Three candidate *T. pallidum* RNA targets (*tpp047* mRNA, *flaB2* mRNA and *16S rRNA* were investigated for use in an RT-qPCR assay. The targets *tpp047* and *flaB2* were selected as both have demonstrated a high level of expression in a rabbit model of syphilis, moreover they have sequences specific to *T. pallidum.* (141) The third target was bacterial 16S rRNA, previously shown to a sensitive target for RT-PCR, and for which *T. pallidum* specific primers had been identified, despite considerable homology of 16S rRNA among bacteria. (153)

During an initial assessment, *T. pallidum* Nichols strain DNA was amplified in two novel TaqMan assays (*flaB2* and *16S rRNA* detection). The primer and probe sequences for these assays are given in Appendix 2 and all three assays were able to detect *T. pallidum* DNA. Given the potential for 16S rRNA to be present in large amounts in actively replicating organisms, this gene was selected as the target of choice, despite the later cycle threshold seen for DNA detection when initially compared with *tpp047* and *flaB2*.

3.2.4.2 RT-qPCR for 16S rRNA quantification was optimised

Despite showing no cross-reactivity with a panel of bacterial DNA (Table 3.7), the initial 16S rRNA assay primers (16S_F1, 16S_R1) amplified a human gene when tested against reverse-transcribed *T. pallidum* RNA extracted from spiked Tempus™ blood samples (Figure 3.4). This non-specific amplification was confirmed by agarose gel electrophoresis, where the expected 366bp band was seen together with a second band at 450bp and a third at 840bp. In order to identify the bands, these were gel-extracted and sequenced bidirectionally using the 16S_1F and 16S_1R primers, according to the methods described in Chapters 2.4 and 2.6. The 366bp band was confirmed to be *T. pallidum 16S rRNA*. The 840bp band had significant homology with homosapien non-muscle myosin heavy chain 9 (*MYH9*) mRNA on a NCBI Nucleotide database BLAST search. Gene sequence analysis of the 450bp band failed, suggesting it was amplified non-specifically.

In order to improve the efficiency and specificity of the 16S rRNA qPCR a new TaqMan probe (16S_P3) and five new antisense primers were designed (Table 3.8). The sense primer (16S_F(seattle)) selected had been shown previously to amplify a section of the *T. pallidum 16S rRNA* gene that has little homology with other bacteria including members of the genus *Treponema (153)*. These four new primer pairs were compared with each-other and *tpp047* RNA quantification in a one-step RT-qPCR (Chapter 2.12.2) quantifying RNA extracted from 10^4 *T. pallidum* organisms. Table 3.8 shows the mean cycle thresholds for each of the new primer pairs. The cycle threshold for *tpp047* was significantly higher (34.75 cycles, p=<0.005) than any of the *16S rRNA* assays and when used to quantify RNA from human blood spiked with *T. pallidum* all four assays amplified only a single product (Figure 3.4). Finally, an attempt was made to quantify bacterial DNA from the test panel with the anti-sense primers 16S_R4 and 16S_R7 in combination with the sense primer 16S_F(Seattle) and the probe 16S_P3 (Table 3.7). The 16S_R4 primer proved to be non-specific and amplified 10 other bacterial species, however 16S_R7 proved specific and following confirmation that the correct target was amplified using gene sequence analysis this assay was selected for *T. pallidum* RNA quantification by RT-qPCR.

Figure 3.4 Specificity of *T. pallidum* RNA detection

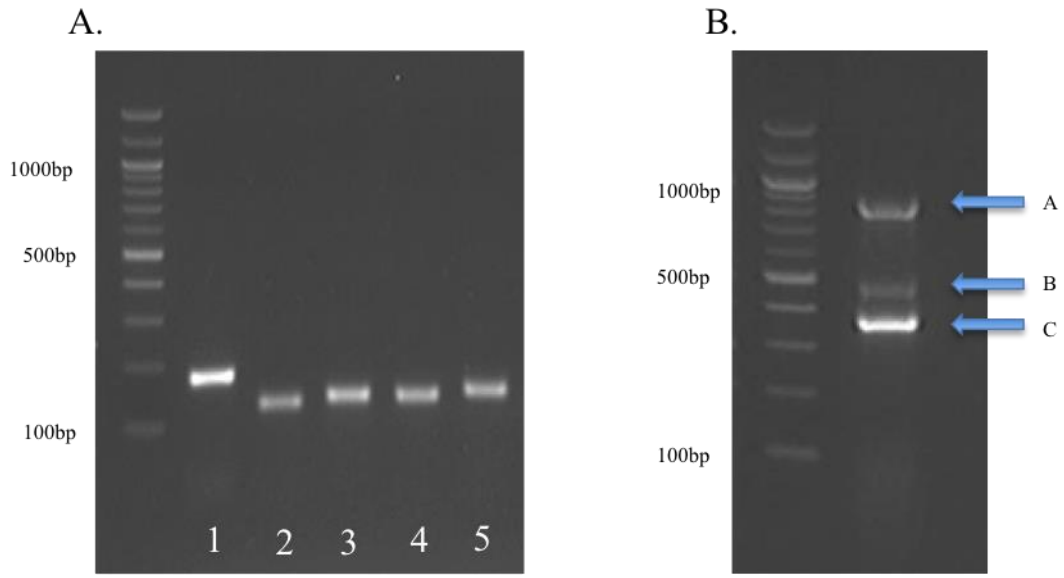


Fig. 3.4. Human whole blood collected into Tempus™ RNA preservation tubes was spiked with 5,000 *T. pallidum* organisms and RNA was extracted. **Panel A.** *T. pallidum* RNA was amplified in a one-step TaqMan assay with the sense primer 16S_F(Seattle), the probe 16S_P3 and one of four different anti-sense primers (sequences are given in Table 3.8). Amplicons were subsequently visualised with 2% agarose gel electrophoresis to ensure amplification of a single product. Lane 1) positive control (*tpp047* amplification); lane 2) 16_R4; lane 3)16_R5; lane 4) 16S_R6; lane 5) 16S_R7. **Panel B.** Following reverse transcription, cDNA was amplified with the primers 16S_F1 and 16S_R1. Three products were identified with 2% agarose gel electrophoresis: Band A, an 840bp product (subsequently identified as human *MYH9* mRNA); band B, an unidentifiable 450bp product and band C, the expected 366bp *16S rRNA* product.

Table 3.7 Specificity of three *T. pallidum* qPCRs

Organism	**Assay and Primers**		
	TaqMan qPCR (I) 16S_F1 16S_R1 16S_P1	TaqMan qPCR (II) 16S_F(seattle) 16S_R4 16S_P3	TaqMan qPCR (III) 16S_F(Seattle) 16S_R7 16S_P3
Methicillin Resistant Staphylococcus aureus	-	+	-
Staphylococci spp.	-	-	-
Coliforms	-	+	-
Streptococcus pyogenes	-	-	-
Group B *Streptococci*	-	-	-
E. coli	-	+	-
Campylobacter	-	-	-
Staphylococcus aureus	-	+	-
Corynebacterium	-	-	-
Haemophilus influenzae	-	+	-
Neisseria gonorrhoeae	-	+	-
Enterococcus spp	-	-	-
alpha-haemolytic Streptococci	-	+	-
Streptococcus agalactiae	-	-	-
Lactobacillus spp	-	+	-
Neisseria meningitidis	-	-	-
Haemophilus parainfluenzae	-	+	-
Streptococcus pneumoniae	-	+	-
T. pallidum subspecies pertenue (Gauthier strain)	not tested	+	+
T. denticola	not tested	-	-
T. pallidum (Nichols strain)	+	+	+
No Template Control (water).	-	-	-

Table 3.7. DNA from 22 different bacterial cultures/propagations was subjected to qPCR with three different primer combinations (I, II and III) in order to assess species specificity of *T. pallidum 16S rRNA* gene amplification. Positive (+) and Negative (-) reactions are shown. All samples were run in duplicate.

Table 3.8 Determination of the optimal assay for *T. pallidum* RT-qPCR

	Sequence 5'- 3'	Tm (°C)	Length (bp)	Product size (bp)	Mean* cycle threshold (cycles)
Probe 16S_P3	6-FAM-CGGCCTCGCTCAGAGATGAGC-TAMRA	69.1	21	-	-
Sense Primer				-	-
16S_F(seattle)	CTCTTTTGGACGTAGGTCTTTGAG	60.7	24		
Anti-sense Primers					
16S_R4	TCTCAGGTCGGATACCCATC	59.9	125	125	23.51
16S_R5	ACCCTCTCAGGTCGGATACC	60.3	129	129	23.46
16S_R6	CACCCTCTCAGGTCGGATAC	59.5	130	130	23.95
16S_R7	TCACCCTCTCAGGTCGGATA	60.6	131	131	24.15

Table 3.8. Four anti-sense primers were designed for use in a one-step RT-qPCR with the sense primer 16S_F(Seattle) and the TaqMan probe 16S_P3. RNA. Tempus™ tubes were filled with 3ml of syphilis seronegative blood and spiked with 5000 *T. pallidum* organisms. RNA was extracted and the mean cycle threshold for *T. pallidum* 16S rRNA detection was then determined with each of the four reaction conditions.

*Mean of three replicas from the same experiment

3.2.4.3 Constructing an RNA standard curve for 16S rRNA quantification

Chapter 2.9 describes the production of a recombinant plasmid containing 131bp of the *T. pallidum 16S rRNA* gene for use as a template to produce an RNA quantification standard. This 4086bp plasmid (p16SRNA) had a spectrophotometric concentration 210.9ng/μl, which equated to an initial concentration of $4.75x10^{10}$ copies/μl. The parent PCR-4 TOPO plasmid vector contained a T7 promoter site to allow T7 polymerase binding and RNA transcription of the plasmid. Prior to T7 polymerase treatment, a linear plasmid structure was needed to enable the production of separate copies of the plasmid sequence and not a single large transcript containing multiple copies. The restriction enzyme *ncoI* sucessfully cut the plasmid at a single site to produce a linear structure (Figure 3.5). The linearised plasmid was then quantified with spectrophotometry and diluted in water to give a starting concentration of $1x10^{10}$ copies/μl and a 1:10 dilution series made down to a single copy/μl. Dilutions containing 10^6 to one copy/μl were then quantified in quadruplicate in the *16SrRNA* TaqMan qPCR (Table 3.9). At a concentration of one copy/reaction, there was a 50% chance of a negative reaction, which equated to $6.93x10^9$ copies/μl and comparable to the starting concentration of $1x10^{10}$ copies/μl.

An RNA quantification standard was constructed by using the plasmid p16SRNA as a template for RNA transcription. Initially, the Invitrogen T7 polymerase was assessed and found to produce RNA detectable at an average of 20 RT-qPCR cycles when 50 units were added to 10^6 copies of p16SRNA for one hour (Chapter 2.13.2). When less than 10^4 copies of plasmid were added to the reaction, no detectable RNA was produced and when 10^7 or more copies were added, unacceptable amounts of plasmid DNA were still detectable following DNAse treatment (Figure 3.6). Three method modifications to increase the amount of RNA produced by this enzyme were subsequently assessed (Table 3.10). The first was extending the one-hour polymerase incubation time to 20 hours, which was found to lower the cycle threshold for RNA detection by 1.97 cycles compared with one hour, but this difference did not reach significance (p=0.123). The second modification was to incubate the reaction at 30°C and 40°C in addition to the usual 37°C. Over two experiments, the mean cycle threshold for RNA detection at 30°C was not significantly different from 37°C, however incubation at 40°C (19.16 cycles) compared with 37°C showed a 2.3 cycle improvement (p=0.050). Finally, the concentration of rNTPs was increased from 1mM to 2mM, which resulted in a

significant ($p=<0.05$) 1.86 cycle improvement in RNA detection. The maximum difference in cycle threshold observed for any of the modifications was less than three cycles, and the amount of RNA produced too low for the construction of a standard curve with an adequate quantification range.

A final attempt to improve the amount of RNA produced compared two new T7 polymerases with the original Invitrogen polymerase. Table 3.11 shows the detection of *16S rRNA* and residual plasmid DNA following T7 polymerase transcription of 10^6 copies of the plasmid p16SRNA with three different polymerases (each used in accordance with its manufacturer's protocol). RNA produced with all three methods was subject to the same DNAse treatment with a Roche highpure column-based method (Chapter 2.1.2). The Ambion T7 RiboMAX express kit produced RNA detectable at 17 cycles of RT-qPCR, significantly more ($p<0.001$) than either of the other enzymes. Minimal residual plasmid DNA was detected in the RiboMAX-produced RNA after DNAse treatment (29.17 cycles).

Carrier RNA (poly-adenylated RNA) was added to the RNA transcribed from p16SRNA to reduce the potential for RNAse degradation during handling and storage. In order to assess the potential impact of carrier RNA on target detection during RT-qPCR, eight different amounts of carrier RNA (0.1µg, 0.15µg, 0.2µg, 0.25µg, 0.3µg, 0.35µg, 0.4µg and 0.45µg) were added to $2.3x10^5$ copies of p16SRNA-produced RNA in a total volume of 100µl water and quantified in a one-step TaqMan RT-qPCR with the primers 16S_F(Seattle) and 16S_R7 and the probe 16S_P3. Compared with the addition of no carrier RNA (26.73 cycles), the average number of PCR cycles for detection was 0.32 higher with a carrier RNA concentration of 0.01µg/µl and 1.14 higher with a concentration 0.045µg/µl carrier RNA. To make the final standard for *16S rRNA* RT-qPCR, 80µl of RNAse-free water were added to 10µl of a 0.1µg/µl solution of carrier RNA (final concentration 0.01µg/µl) and 10µl of water containing $2.3x10^7$ copies/µl of 16S rRNA (final concentration $2.3x10^6$ copies/µl). This RNA was stored at -80°C in 10 aliquots of 10µl.

Figure 3.5 Linearisation of the plasmid p16SRNA

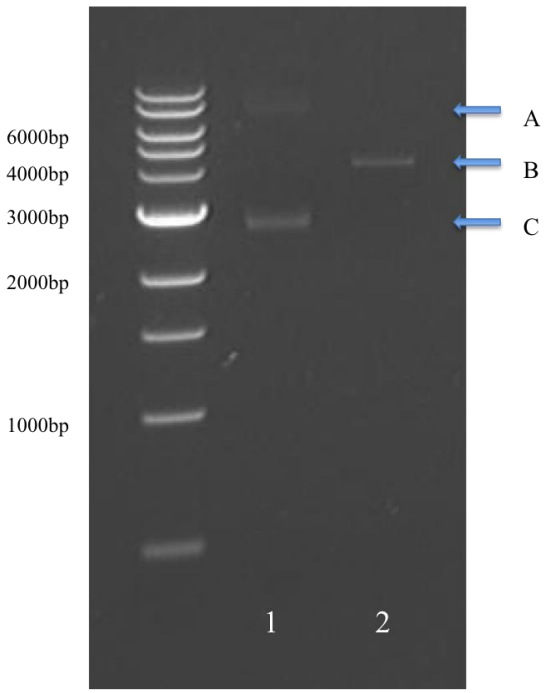


Figure 3.5. The plasmid p16SRNA (4086bp) was subject to restriction digestion with the enzyme *ncoI* and visualised following electrophoresis in a 1% agarose gel. Lane 1 shows the open (A) and supercoiled (C) species of an uncut control. In lane 2 there is only band B (linearised plasmid).

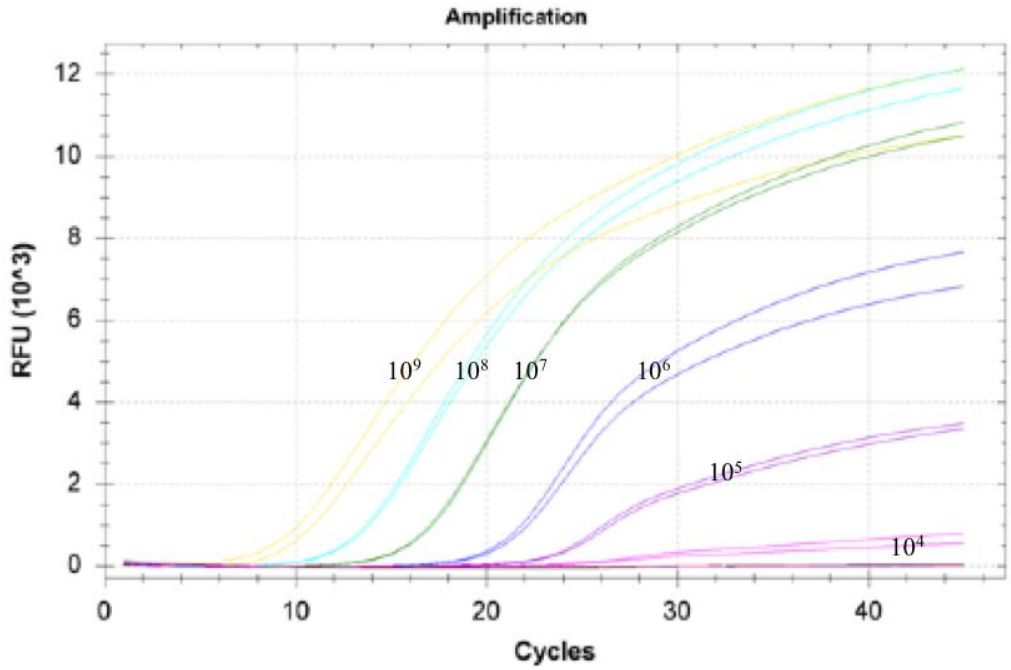
Table 3.9 Limiting dilution assay to quantify p16SRNA

Starting concentration (copies/reaction)	Number of replicas positive	Mean (std. dev) cycle threshold
$1x10^6$	+ + + +	19.27 (0.07)
$1x10^5$	+ + + +	22.70 (0.04)
$1x10^4$	+ + + +	25.85 (0.09)
$1x10^3$	+ + + +	28.98 (0.04)
$1x10^2$	+ + + +	32.19 (0.23)
10	+ + + +	35.80 (0.78)
1	+ + - -	38.84 (2.96)

Table 3.9. The primers 16S_F(Seattle) and 16S_R7 were used to PCR-amplify of a 131bp section of the *T. pallidum 16S rRNA* gene from a 1:10 dilution series of the recombinant plasmid p16SRNA. The hydrolysis (TaqMan) probe 16S_P3 was used to detect amplification in real-time and four technical replicates of each dilution were analysed.

Figure 3.6 Assessment of T7 polymerase activity

A. 16S rRNA quantification



B. gDNA carry-over

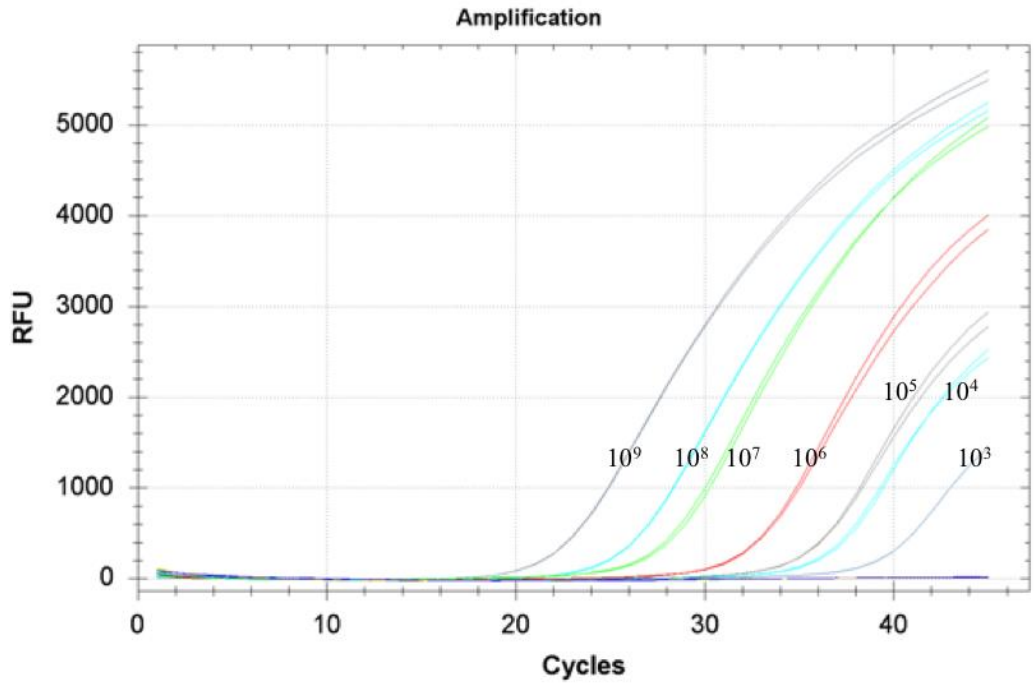


Fig. 3.6. Panel A. Amplification of *T. pallidum* 16SrRNA in a TaqMan RT-qPCR with the primers 16S_F(Seattle) and 16S_R7 and the probe 16S_P3. Template RNA was transcribed using Invitrogen T7 polymerase from a 1:10 dilution series of plasmid p16SRNA ranging from 10^9 to 10 copies/reaction and subject to DNAse treatment prior to quantification. **Panel B**. Quantification of p16SRNA DNA in a TaqMan qPCR with the primers 16S_F(Seattle) and 16S_R7 and the probe 16S_P3 to assess gDNA 'carry-over' into the RT-qPCR.
Two technical replicates of each sample were included for both 16S rRNA and gDNA.

Table 3.10 Optimisation of T7 polymerase reaction conditions

Variable	Mean** cycle threshold RNA (std. dev.)	Mean cycle threshold DNA (std. dev)	t-test (2-sided)
T7 polymerase incubation temperature			
30°C	20.87(2.20)	30.83 (0.26)	p=0.050
37°C*	21.49 (0.57)	30.75 (0.03)	-
40°C	19.16 (1.58)	30.25 (0.31)	p=0.62
Incubation time			
1 hour*	25.25 (2.46)	31.7 (3.62)	p=0.272
20 hours	23.29 (1.65)	33.45 (0.73)	
rNTP concentration			
1mM*	26.87 (1.05)	29.94 (0.30)	p=0.046
2mM	24.70 (0.41)	32.39 (1.06)	

Table 3.10. A total of 10^6 copies of the plasmid p16SRNA were added to 50 units of Invitrogen T7 polymerase and, following DNAse treatment, resultant RNA quantified by RT-qPCR with the primers 16S_F(Seattle) and 16S_R7 and the probe 16S_P3. Residual plasmid DNA was then quantified by qPCR with the same primers and probe. Changes to reaction conditions were assessed individually. Incubation at 40°C and doubling the rNTP concentration improved polymerase activity, whereas increasing the incubation time decreased it.

*Standard reaction conditions with which modified conditions were compared.

** Mean of three replicas in two separate experiments.

Table 3.11 Comparison of T7 RNA polymerases

Method and Reaction Conditions		Mean† (std. dev.) cycle threshold of DNA detection	Mean† (std. dev.) cycle threshold RNA detection
Reagent	**[final]**		
I. Promega T7 Polymerase			
rNTP mix (Invitrogen, UK)	2mM	33.39 (0.63)	27.62 (0.99)
Reaction buffer	1x		
DTT (Promega, UK)	10mM		
RNAse inhibitor	50units		
T7 polymerase	40units		
II. Invitrogen T7 Polymerase			
rNTP mix (Invitrogen, UK)	2mM	37.28 (1.03)	26.53 (0.18)
Reaction buffer	1x		
DTT (Promega, UK)	5mM		
RNAse inhibitor	50units		
T7 polymerase	50units		
III. Promega T7 RiboMAX Express			
Ribomax buffer (contains rNTPs)	1x	29.71 (0.15)	17.11* (0.06)
T7 express enzyme (contains RNAsin)	Not given. 2μl used in 20μl reaction		

Table 3.11. A total of 10^6 copies of *ncoI*-linearised p16SRNA were used as a template for the assessment of three T7 polymerases (I, II and III). Reaction components are listed beneath the name of each enzyme. DNA was quantified in a TaqMan qPCR with the primers 16S_F(Seattle) and 16S_R7 and the probe 16S_P3. RNA was quantified in a one-step RT-qPCR with the same primers and probe. *p=<0.001 (two-sided t-test) compared with Invitrogen T7 polymerase. † The mean of three technical replicates in a single experiment.

3.2.4.4 Sensitivity of 16S rRNA detection

RNA for standard determination was diluted in a 1:10 dilution series from a spectrophotometry-determined concentration of 10^9 copies/ 5µl to 10^2 copies/µl. A total of 10 replicas of each of the dilutions 10^4, 10^3 and 10^2 were then quantified by RT-qPCR according to the method in Chapter 2.12.2. None of the dilutions at 10^3 copies/reaction showed any *16S rRNA* amplification and nine of 10 replicas at the 10^4 copies/reaction concentration were positive with a mean cycle threshold of 40.82 cycles (std. dev. 1.365). At a dilution factor of 10^5 (a starting concentration of 10^4 copies/reaction by spectrophotometry), the natural log of a 10% probability of a negative reaction equated to a starting concentration of $2.3x10^6$ copies/µl. This suggested that the original 10^4 dilution contained approximately 23 copies which, given an average cycle threshold of 40.82 cycles for this dilution, was deemed accurate. A quantification standard curve was then constructed comprising a 1:10 dilution series from $2.3x10^5$ to 23 copies/reaction. In order to confirm its analytical sensitivity, efficiency of amplification and to ensure acceptable inter and intra-assay variability, three separate dilution series were quantified (each dilution was quantified in triplicate). A fourth dilution series, containing both *16S rRNA* and carrier RNA was also quantified for comparison. The results of these four experiments are shown in Figure 3.7 and Table 3.12 summarises the variability of the assay. With perfect amplification efficiency, the amount of DNA in a PCR should have doubled with each cycle – every copy of DNA becoming the template for a new copy. In reality, an amplification efficiency of 90-105% is ideal and was shown here to vary from 85.2% to 92.9% (mean 87.77%). The slope of the plot cycle threshold vs starting copy number was also a measure of efficiency. In the perfect reaction, where the quantity of DNA doubles after each round of PCR, there would have been a 3.32 cycle difference (or a slope of -3.32 on a plot of cycle threshold vs DNA quantity) between each 1:10 dilution of the standard according to the equation:

$$Efficiency = 10^{\left(\frac{-1}{slope}\right)} - 1$$

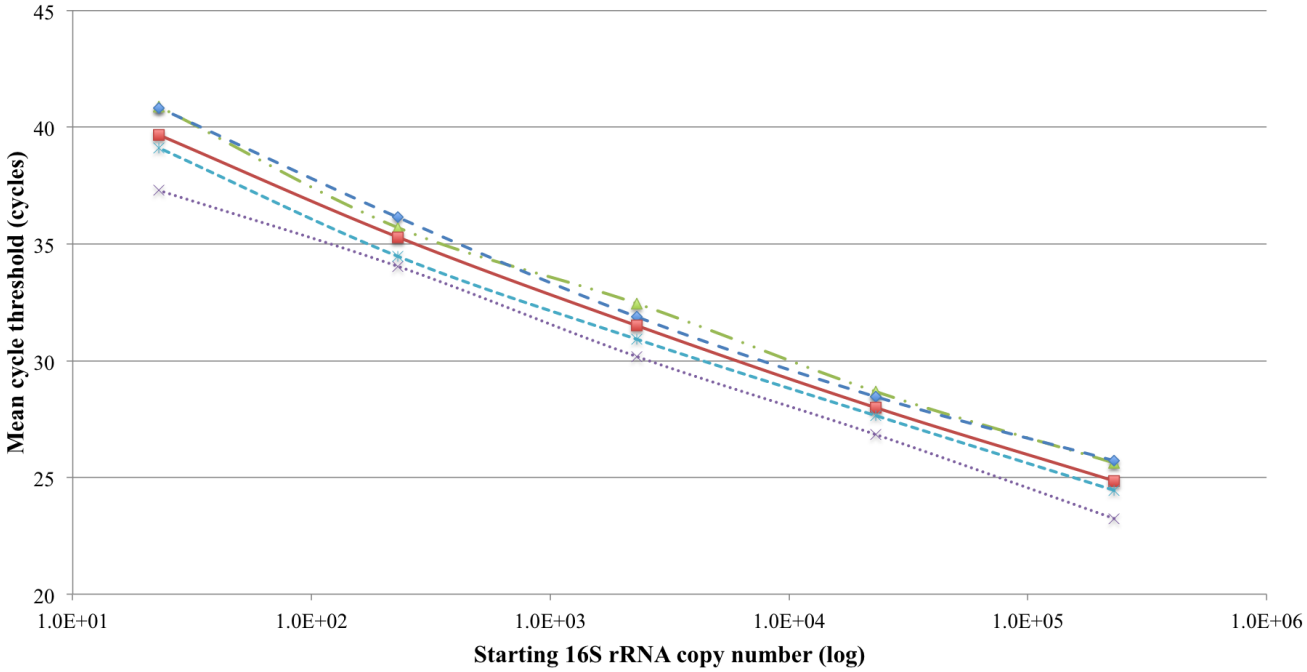
For example,

The mean slope calculated for this standard curve was -3.66 (std. dev. 0.10). Finally, the 'linearity' of the plot was calculated by the real-time instrument software and indicated variability across replicates by measuring how well the experimental data fitted the regression line. Also referred to as the coefficient of variation, the target for linearity is ≥ 0.98 and was

shown here to be 0.97 (std. dev 0.04). Intra-assay variability was calculated to be low and inter-assay variability, dependent on RNA stability, reverse transcriptase and PCR efficiency was around three percent.

Figure 3.7 Construction of a standard curve for 16S rRNA quantification

A. Standard curves



B. Absolute quantification

Calculated copies	17/05/12	18/05/12	22/05/12	25/5/12*	Mean**	std. dev.
$2.3x10^5$	25.71	25.61	23.24	24.45	24.85	1.40
$2.3x10^4$	28.46	28.68	26.84	27.65	27.99	1.00
$2.3x10^3$	31.88	32.45	30.18	30.92	31.50	1.18
$2.3x10^2$	36.14	35.7	34.04	34.47	35.29	1.11
$2.3x10^1$	40.83	40.88	37.29	39.11	39.67	2.06

C. Standard curve parameters

	Amplification efficiency (%)	Linearity (R^2)	Slope (cycles)
18.5.12	86.40	0.98	-3.70
22.5.12	91.70	0.99	-3.53
17.5.12	85.20	0.92	-3.74
25.5.12*	92.90	0.99	-3.50
mean**	87.77	0.97	-3.66
st dev	3.46	0.04	0.10

Fig. 3.7. Four dilution series of RNA to be used as a quantification standard for *T. pallidum 16S rRNA* were amplified in a TaqMan RT-qPCR. **Panels A and B** compare the mean cycle threshold (red squares) during the PCR for target detection with the starting concentration of RNA for each of the four experiments conducted. All four were identical, except for the addition of carrier RNA to the dilution series 25/5/12 (blue crosses). **Panel C** demonstrates the amplification efficiency (a perfect doubling of DNA after each cycle would have an efficiency of 100%) together with a measure of linearity (coefficient of determination) and the slope of the line (where doubling each cycle would give a gradient 3.32 cycles).

* This replica contained carrier RNA ** Excluding dilution series containing carrier RNA

Table 3.12 Intra-assay and inter-assay variability of 16S rRNA quantification

Copies/reaction	Mean Cycle Threshold	Standard Deviation	Coefficient of variation (%)
	*Intra-assay variability**		
$2.3x10^4$	26.80	0.09	0.34
$2.3x10^3$	30.18	0.23	0.75
$2.3x10^2$	34.04	0.38	1.12
	*Inter-assay variability***		
$2.3x10^4$	28.64	0.87	3.05
$2.3x10^3$	32.09	0.93	2.90
$2.3x10^2$	35.75	0.94	2.63

Table 3.12 Intra-assay variation of *T. pallidum* 16S rRNA quantification was assessed by comparing the cycle thresholds of three technical replicates at three RNA starting concentrations. Inter-assay variation was determined by repeating the assay on three separate days.

*Mean of three replicas in a single experiment

**Mean of four separate experiments

On the basis of the results in this section methods described for use with clinical samples were finalised. To summarise:

- DNA and RNA extraction from snostrip samples were to proceed in the same extraction process using the Qiagen all-prep kit (Chapter 2.3.4.2).
- Blood for RNA extraction was to be collected into Tempus™ tubes and extracted according to the method in Chapter 2.3.3.1). No additional steps for bacterial lysis or human mRNA capture were required.
- *T. pallidum* DNA was to be quantified in a TaqMan qPCR assay with the primers TP1 and TP2 and the FAM-labelled probe TPP. A calibration curve was to be included in each run constructed from *ncoI* linearised plasmid pTp47 (Chapter 2.12.1.2).
- *T. pallidum* RNA was to be quantified in a one-step RT-qPCR. The RNA target was 16S rRNA and the quantification assay used the primers 16S_R7 and 16S_F(seattle) and the probe 16S_P3 (Chapter 2.12.2). An in-run RNA absolute quantification standard was to be used which was made by the activity of T7 polymerase on the recombinant plasmid p16SRNA (2.13.1).

3.2.5 Clinical validation of *T. pallidum* quantification in blood and ulcer samples

Samples for *T. pallidum* quantification were collected from patients during three separate studies, which are described in Chapter 2.14. The first study was a clinical validation of the SYBR-green based qPCR described in Chapter 3.2.3.1 which recruited patients with suspected syphilis, and was conducted in London between July 2006 and January 2008. The study was designed to recruit both those with syphilis and non-syphilis diagnoses in order to compare cases (those in whom syphilis was confirmed) and controls (with non-syphilis diagnoses). The results of this validation described here were then followed by both further assay development and patient recruitment for two subsequent studies.

3.2.5.1 Patient recruitment and demographic data

A total of 99 patients were recruited between July 2006 and January 2008. Figure 3.8 demonstrates that 44 patients had a microbiologically confirmed diagnosis of syphilis comprising 14 primary cases; 19 secondary cases; eight latent cases and three patients with neurosyphilis (all during the secondary stage). Of the syphilis-uninfected patients, 19 were confirmed to have infection with *Herpes simplex* virus and 19 patients, for whom no evidence of an STI could be found, were deemed to have non-specific genital ulceration (which is consistent with other studies examining the causes of genital ulceration) (232, 233). The final 17 patients were asymptomatic sexual contacts of infectious syphilis who received treatment and subsequently showed no clinical or microbiological sign of infection during follow-up. Case definitions for both syphilis and non-syphilis diagnoses are given in Appendix 5. The study population was representative of the current UK syphilis epidemic being predominantly male (97/99, 98%), of which 80.9% were homosexual; Caucasian (88/99, 88.7%) and around half (52/99, 52.5%) were HIV-1 infected. Neither of the two women recruited were HIV-1 infected. (1)

Figure 3.8 Patient recruitment to qPCR validation study

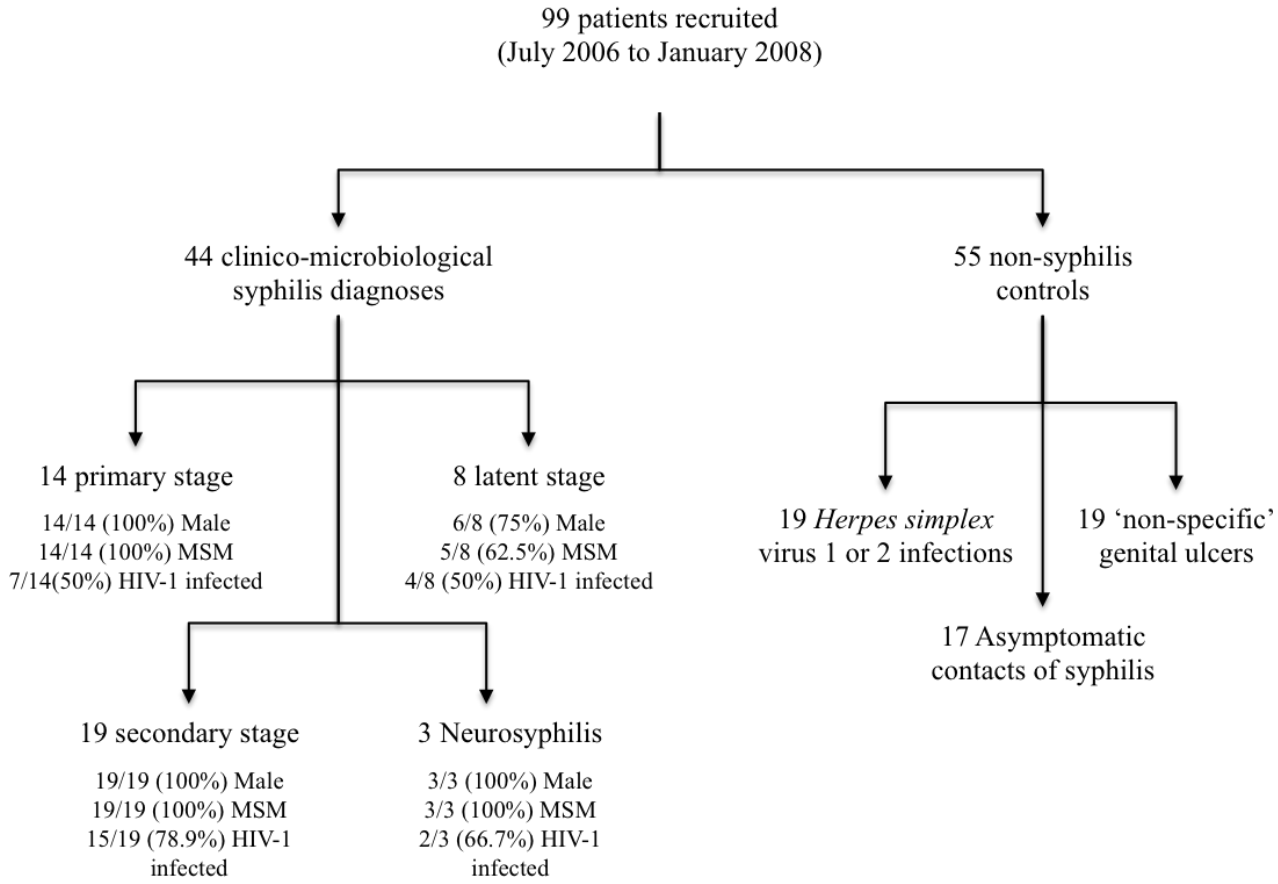


Fig. 3.8. Summary of patient recruitment to a study between July 2006 and January 2008 which aimed to validate a method for *T. pallidum* DNA and RNA quantification in blood and ulcer samples

3.2.5.2 T. pallidum detection in three blood components

An aliquot of each EDTA whole blood sample collected during the current study was separated into plasma and PBMCs using Histopaque, as described in Chapter 2.2.1. Table 3.13 shows qPCR results for 16 samples where at least one of the three blood components (whole blood, plasma, PBMCs) contained detectable *T. pallidum tpp047* DNA. Plasma and whole blood results were available for all 16 samples and PBMC results for eight. In total, *T. pallidum* was detected in 94% (15/16) of the whole blood samples, in the plasma of 75% (12/16) and PBMC fraction of 75% (6/8). The mean *T. pallidum* load in whole blood (269.19 copies/ml) exceeded that of both plasma (154.88 copies/ml) and PBMC (143.63 copies/ml) fractions. These differences (whole blood vs plasma and whole blood vs PMBCs) were not statistically significant, but whole blood appeared to be the better sample for *T. pallidum* detection, both qualitatively and quantitatively.

3.2.5.3 Qualitative PCR: Sensitivity and Specificity analysis of tpp047 detection in ulcer and whole blood samples

Patients' final clinico-microbiological diagnoses (Appendix 5) were compared with *T. pallidum tpp047* qPCR results. Table 3.14 illustrates the proportion of whole blood and ulcer samples that contained *tpp047* DNA for each stage of disease. All ulcer samples from patients with primary or secondary disease and one sample from a patient with a non-specific ulcer contained *T. pallidum* DNA. This equates to a sensitivity of 100% and a specificity of 97.14%, respectively. The potentially false positive ulcer PCR result was from a patient with a one-week history of multiple non-tender genital ulcers thought to be atypical genital herpes and treated with Acyclovir. The DGM was negative, as was the contemporaneous and convalescent syphilis serology and PCR for both *herpes simplex virus* types. It is not known whether this patient received coincidental treponemocidal antibiotics that could have prevented seroconversion (111). All 14 patients with primary syphilis had PCR-positive ulcers and 10 were DGM-positive. All four patients with DGM-negative ulcers had positive serology at enrolment (and thus microbiological confirmation of their diagnosis). Three of the five patients with ulcers in secondary disease had DGM performed and two were positive. Compared with the final diagnosis, the sensitivity and specificity of DGM for syphilis ulcers (primary and secondary combined) were 70.59% and 87.88%, respectively, with no

significant difference observed for primary or secondary cases analysed individually. The sensitivity of PCR for the diagnosis of primary disease was 100%.

Whole blood samples from all 44 patients with a syphilis diagnosis (primary, secondary, latent or neurosyphilis) were subjected to *T. pallidum tpp047* PCR and a total of 14/44 (31.80%) were positive (Table 3.14). No patient with a non-syphilis diagnosis had detectable *T. pallidum* DNA in their blood, including the 17 who were contacts of infectious syphilis. When all 99 whole blood PCR results were compared with the donor patients' final clinical diagnosis, the sensitivity for *T. pallidum* detection was 31.80%, with a specificity of 100%. When calculated according to disease stage, the sensitivity of whole blood PCR to identify secondary disease was 74.0% (specificity 100%); 21.0% for primary disease (specificity 100%) and 12.5% for latent disease. Sensitivity and specificity results, including 95% confidence intervals are given in Table 3.15.

Table 3.13 Comparison *T. pallidum* detection in three blood components

		tpp047 quantification		
	Patient Study Number	**Plasma (copies/ml)**	**Whole Blood (copies/ml)**	**PBMC (copies/ml equivalent)**
Primary Syphilis				
	9	195	127	120
	32	614	168	0
	90	122	127	†
Secondary Syphilis				
	16	673	886	119
	17	203	1026	309
	22	121	518	0
	27	0	632	185
	29	10	58	242
	34	0	495	†
	45	0	6	†
	49	16	10	†
	53	284	28	†
	61	144	91	†
	81	95	65	†
Latent Syphilis				
	28	0	70	174
Neurosyphilis				
	56	1	0	†
Mean		154.88	269.19	143.63
Std. dev		210.66	333.99	108.29
*t-test (2 tailed)**		p=0.26		p=0.32**

Table 3.13. An aliquot of EDTA whole blood samples from 16 patients was separated into plasma and PBMC fractions. DNA was extracted and *T. pallidum tpp047* DNA was quantified by qPCR. Whole blood contained the largest amounts of detectable *T. pallidum,* although the difference between detection in whole blood compared with PBMCs or plasma was not statistically significant.

† Result was unavailable for analysis

*statistical significance versus quantification in whole blood.

**Missing were results excluded from the analysis.

Table 3.14 *T. pallidum* quantification according to clinico-microbiological diagnosis

Diagnosis	HIV-1 co-infected n (%)	Mean CD4 (cells/μl) (std. dev.)	Median RPR (min-max)	Mean *tpp047* copies/ml of Whole Blood (std. dev.)	Mean *tpp047* copies/snostrip (std. dev.)
Primary (n=14)	7 (50.0)	501 (234)	n=12 6 (0-64)	3/14 +ve 140.67 (23.67)	14/14 +ve 2480.96 (3419.15)
Secondary (n=19)	15 (78.9)	469 (152)	n=19 128 (8-512)	10/19 +ve 346.82 (380.02)	5/5 +ve 303.10 (236.15)
Latent (n=8)	4 (50.0)	450 (178)	n=8 0.5 (0-256)	1/8 +ve 70	n/a
Neurosyphilis (n=3)	2 (67.7)	475 (261)	n=3 32 (18-64)	0/3 +ve	n/a
Contact (n=17)	9 (52.9)	515 (143)	n/a	0/17 +ve	0/2 +ve
HSV (n=19)	4 (21.1)	595 (177)	n/a	0/19 +ve	0/18 +ve
Non-specific genital ulcer (n=19)	8 (42.1)	433 (210)	n/a	0/19 +ve	1/15 +ve 46 copies/strip

Table 3.14. *T. pallidum* DNA was quantified in blood and ulcer samples (collected with Snostrips) taken from 99 patients with suspected syphilis. Following examination and investigation, including a Rapid Plasma Reagin (RPR) syphilis test, patients were given a clinic-microbiological diagnosis. Detection was most reliable in the blood of patients with secondary disease, and was found at higher load that those with primary or latent infection. A single sample from a patient with a non-syphilis diagnosis was found to contain *T. pallidum* DNA.

n/a = not applicable

Table 3.15 Sensitivity and Specificity of *T. pallidum* detection in whole blood

			95% Confidence Interval	
		%*	Upper Limit (%)	Lower Limit (%)
All Stages (n=44)	**Sensitivity**	31.8	19.0	48.0
	Specificity	100.0	91.0	100.0
Secondary Disease (n=19)	**Sensitivity**	74.0	49.0	90.0
	Specificity	100.0	92.0	100.0
Primary Disease (n=14)	**Sensitivity**	21.0	8.0	51.0
	Specificity	100.0	91.0	100.0
Latent Disease (n=8)	**Sensitivity**	12.5	1.0	53.0
	Specificity	100.0	92.0	1.0

Table 3.15. The clinical sensitivity and specificity of *T. pallidum tpp047* DNA detection in whole blood samples from patients with syphilis was assessed according to stage of disease. Samples from patients given a non-syphilis final diagnosis were used as negative controls.

* Sensitivity and specificity (expressed as a percentage) compared with 55 syphilis-uninfected controls.

3.2.5.4 Quantitative PCR results according to stage of disease

Whole blood and ulcer specimens were collected and handled as described in Chapter 2.2. The first 39 samples collected were processed immediately, DNA extracted (according to the methods in Chapter 2.3) and *T. pallidum* DNA quantified by qPCR (Chapter 2.12.1). The remaining 60 samples were stored at -80°C for up to two years prior to DNA extraction and quantification. DNA from 400µl whole blood was extracted and resuspended in 120µl of water. Ulcer exudate was absorbed onto a single snostrip and extracted DNA was similarly resuspended in 120µl water. The *tpp047* gene was quantified in three separate 10µl replicates and a mean starting copy number calculated by the instrument software, according to cycle threshold results obtained from an in-run quantification standard. The starting concentration of *tpp047* copies/ml a whole blood was calculated as follows:

$$tpp047\ copies/ml\ =\ (mean\ copies\ in\ 10\mu l\ reaction\ \times \left(\frac{volume\ of\ eluate}{10}\right) \times 2.5\ \dagger$$

† DNA was extracted from 400µl of blood.

For snostrips, the equation was as follows:

$$tpp047\ copies/strip =\ mean\ copies\ in\ a\ 10\mu l\ reaction\ \times \left(\frac{volume\ of\ eluate}{10}\right)$$

Quantification data for all patients are presented in Appendix 6.

Primary syphilis

A total of 14 patients were diagnosed with primary syphilis, of whom half were HIV-1 infected, 11 had a positive *T. pallidum* EIA result and 13 had a positive RPR (mode 4, range 64). *T. pallidum tpp047* DNA was detected in ulcer exudates of all 14 patients diagnosed with primary disease. Figure 3.9 A shows the mean copy number of *tpp047* DNA from ulcers to be 2480 copies/strip (std. dev. 3419). When sub-divided, a statistically significant difference was not observed between samples processed immediately (mean 3233 copies/strip, std. dev. 4116) and those processed after storage (mean 1126 copies/strip, std. dev. 734) (p=0.29). A

DGM result was available for all patients and 10 were positive. A mean of 562 (std. dev. 464) copies/strip were detected in DGM negative ulcers vs 3248 (std. dev. 3811) in those which were DGM positive (p=0.196) (Figure 3.10). Three patients had *tpp047* DNA detectable in their blood with a copy number one to two logs lower (mean 30.14 copies/mL, std. dev. 60.61) than for ulcers.

Secondary syphilis

Secondary syphilis was diagnosed in 19 patients of whom 15 (78.9%) were HIV-1 infected. All patients had both a positive *T. pallidum* EIA and RPR (modal RPR was 128 with a range of 504). As seen in Table 3.14, DNA was detected in 10/19 whole blood samples with an average (excluding negative samples) of 346 *tpp047* copies/ml (std. dev. 380). Moreover, Figure 3.9 B demonstrates that immediate DNA extraction improved *T. pallidum* detection and was associated with higher (602 copies/ml vs 40 copies/ml) bacterial load (p=0.005).

A total of five ulcer samples were available from five patients, two of which were taken from muco-cutaneous oral lesions and three from healing primary chancres. DGM was performed on samples from the healing chancres and spirochetes were seen in two. All five samples contained *T. pallidum* DNA with a mean bacterial load of 303 (std. dev 206) copies/strip. When compared, the mean *tpp047* DNA copy number was lower in these three healing primary chancres (351 copies/strip) than primary chancres (2480 copies/strip) (p=0.219).

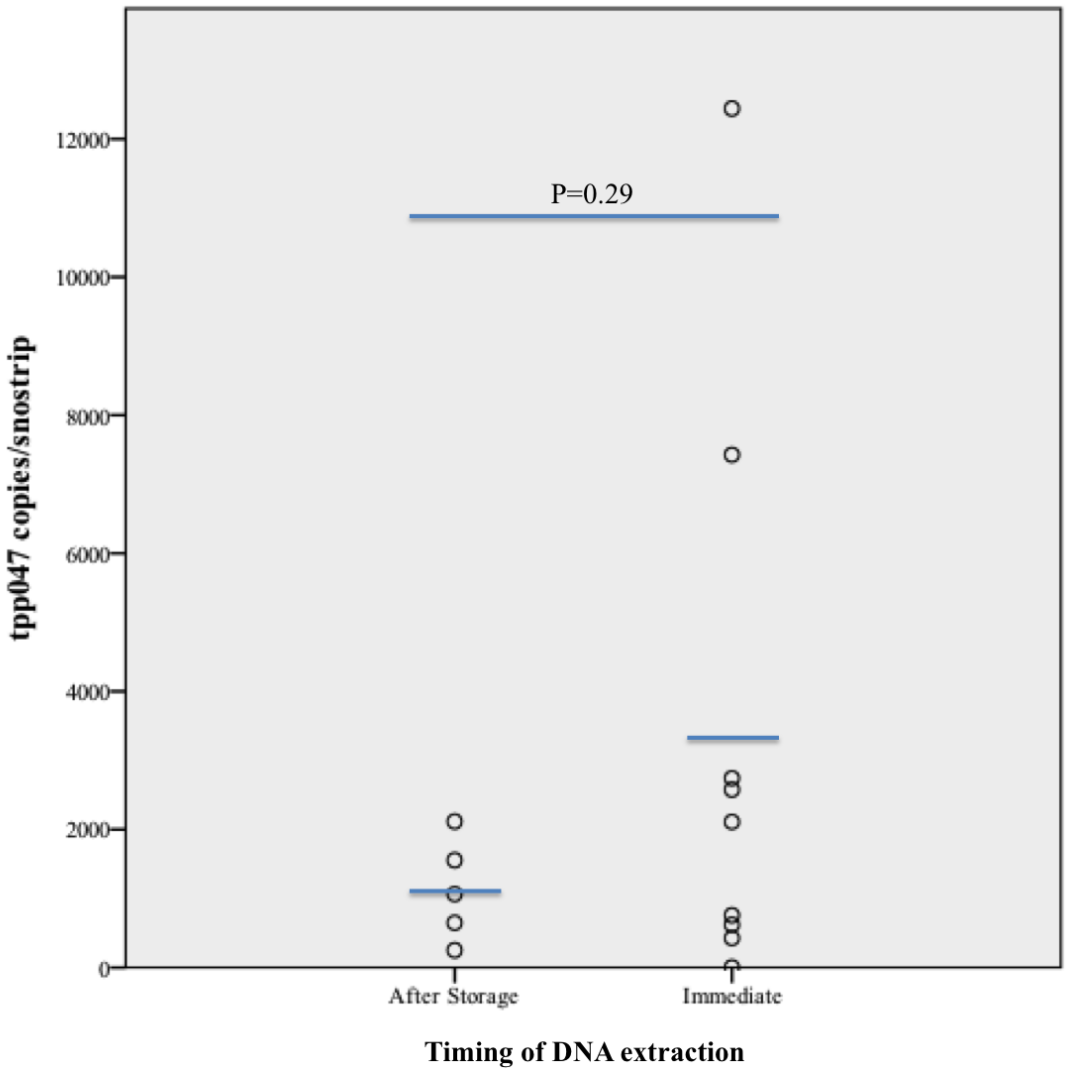
Other cases

Latent disease (early or late) was diagnosed in a total of eight patients of whom four were HIV-1 infected. All had a positive *T. pallidum* EIA and four had a positive RPR (two were neat, one was 1:16 and one was 1:256). *tpp047* DNA was detected in the EDTA whole blood of a single patient who was HIV-1 infected, had an RPR titre of 1:256 and had not received prior syphilis treatment. The patient had been initially diagnosed with syphilis two years prior to study entry (with an RPR of 1:131,072), but had been lost to follow-up. Three patients recruited to the study were diagnosed with neurosyphilis. All three had serological results consistent with early disease (EIA positive and an RPR ≥8) together with neurological

symptoms and CSF examination typical of central nervous system involvement. None had detectable *T. pallidum* DNA in either their blood or CSF.

Figure 3.9 Assessment of the impact of sample storage on *T. pallidum* bacterial load

A. Primary disease Snostrip samples



B. Secondary disease whole blood samples

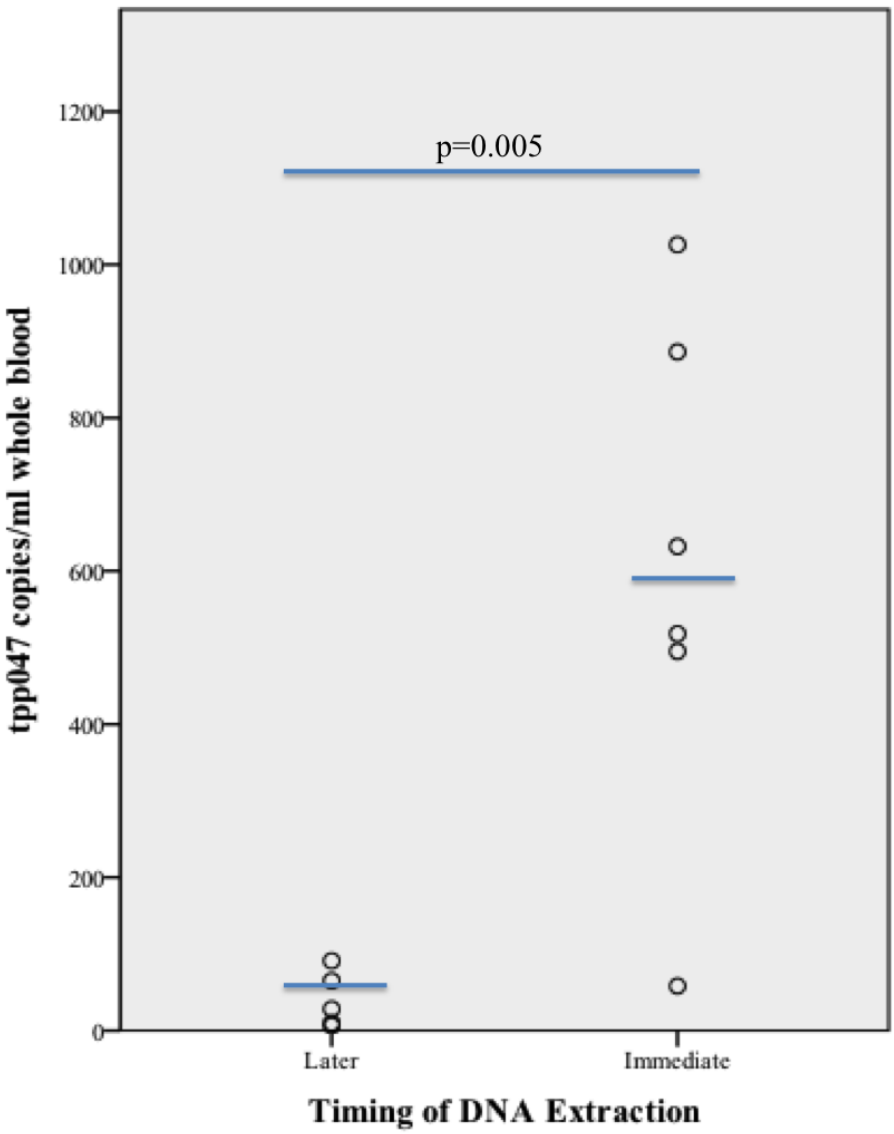


Fig. 3.9. A. *T. pallidum* DNA (*tpp047*) from primary ulcer snostrip samples was either extracted and quantified (using qPCR) immediately following collection, or after sample storage at -80°C for up to 5 years. **B.** Whole blood samples collected into EDTA were available from patients with secondary syphilis. DNA was extracted and quantified (*tpp047* qPCR) immediately or following long-term storage at -80°C. A student's t-test was used to compare mean bacterial loads.

Figure 3.10 Comparison of *T. pallidum* ulcer bacterial load with DGM detection

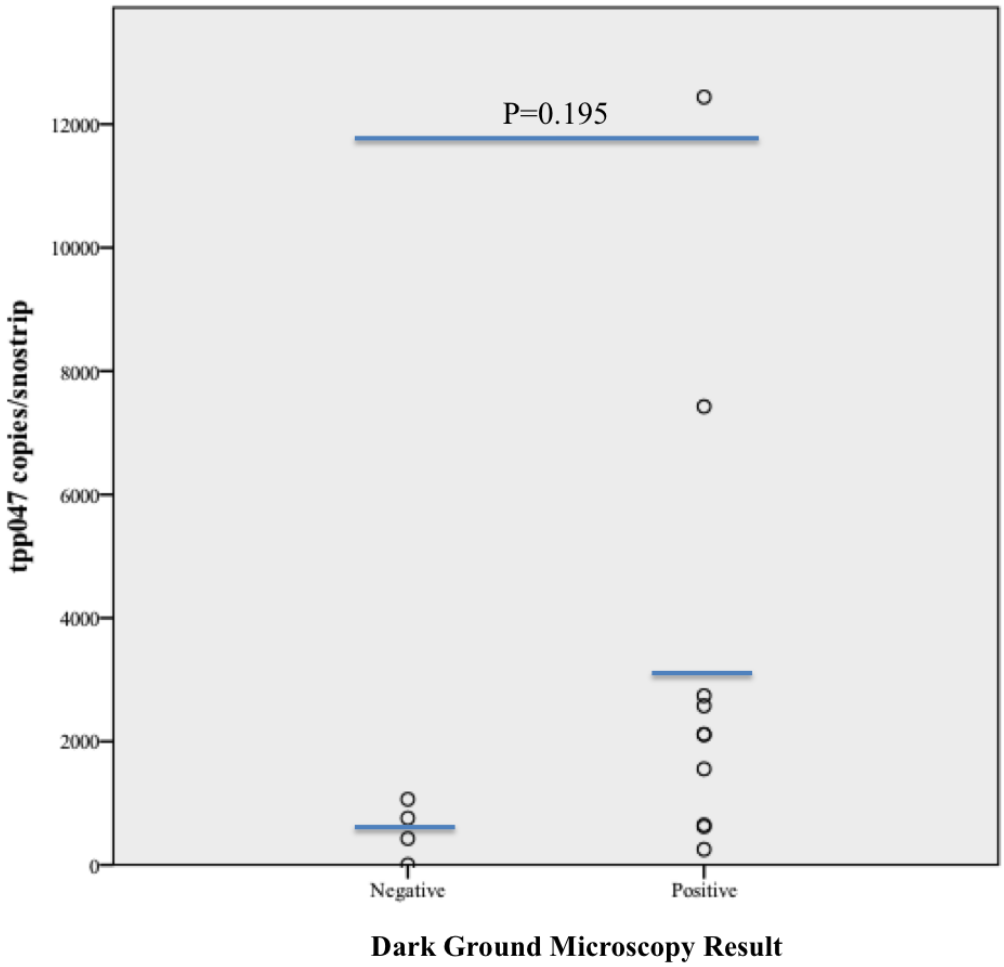


Fig. 3.10. During a cross-sectional study of early syphilis, ulcers suspected to be primary syphilis chancres were examined by DGM prior to sampling for *T. pallidum* DNA. Following *T. pallidum* qPCR (*tpp047* target), the bacterial load in DGM positive and negative ulcers was compared. A student's t-test was used to compare mean ulcer load.

3.2.5.5 Impact of HIV infection of T. pallidum load

Of the 14 patients recruited to the study suffering from primary disease, half were HIV-1 infected and, as demonstrated previously, all 14 ulcer samples collected contained *T. pallidum* DNA. Figure 3.11 demonstrates the difference in *T. pallidum* bacterial load observed in primary disease ulcers according to HIV seropositivity. In HIV-1 infected patients, a mean of 3772 (std. dev. 4513) *tpp047* copies/strip were detected compared with a mean of 1190 (std. dev. 1035) for HIV-1 uninfected patients (p= 0.166). The mean CD4 count of the seven HIV-1 infected patients was 501 (std. dev. 234) and CD4 count did not correlate with *T. pallidum* load.

T. pallidum DNA was detectable in the blood of 11 patients with secondary disease, of whom eight (73%) were HIV-1 infected. No statistically significant difference was observed in whole blood bacterial load between HIV-1 infected (276 *tpp047* copies/ml, std. dev. 344) and uninfected (534 *tpp047* copies/ml, std. dev. 484) subjects (p=0.343), although it must be noted that the standard deviations for both measurements were wide. Across all stages of disease, the mean CD4 count of bacteraemic HIV-1 infected patients was 446 (std. dev. 166) and not significantly different from those patients without detectable bacteraemia (494, std. dev. 183) (p=0.497).

Figure 3.11 *T. pallidum* load in ulcer samples according to HIV-1 infection status

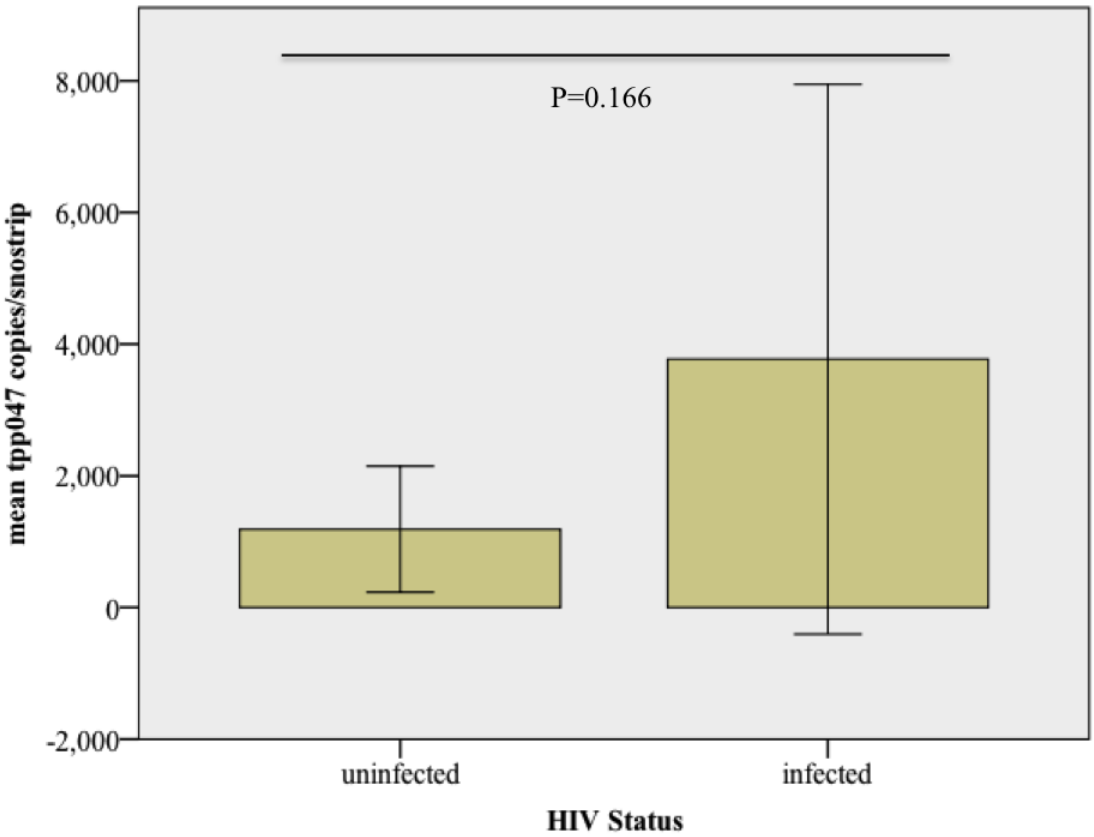


Fig. 3.11. *T. pallidum* DNA *(tpp047*) was quantified by qPCR in 14 ulcer samples collected from patients suffering from primary disease, of whom seven were HIV-1 co-infected. Error bars show 95% confidence intervals.

3.2.5.6 RPR correlates with bacterial load (tpp047 DNA)

Figure 3.12 demonstrates a moderate correlation (Spearman's correlation coefficient = 0.354) between *T. pallidum* load detectable in whole blood and baseline RPR results (taken on day of recruitment to the study) (p=0.021). This association was not seen when stages of disease were analysed separately.

3.2.5.7 Initial RNA quantification

Tempus™ RNA blood samples were collected from a total of 70 patients. Tubes were stored at -80°C for an average of 5 years prior to RNA extraction and quantification as described in Chapters 2.3.3.1 and 2.12.2. During extraction, 16 samples were noted to be clotted and were discarded, leaving 54 samples for analysis by RT-qPCR. A total of 22 samples were from patients with syphilis, of which 11 (5 primary, 5 secondary and 1 latent) were found to contain *T. pallidum* 16S rRNA. None of the 32 samples from patients with non-syphilis diagnoses contained detectable *T. pallidum* 16S rRNA.

Table 3.16 summarises RT-qPCR data for patients with syphilis diagnoses. RNA detection was more sensitive than DNA for the correct identification of primary (62.5% vs 25%), secondary (55.5% vs 55%) and latent (16.4% vs 0%) syphilis and was equally specific (100%). Moreover, the amount of RNA quantified exceeded that of DNA for both primary (p=0.086) and secondary (p=0.023) disease.

Figure 3.12 Correlation of blood *T. pallidum* load and baseline RPR titre

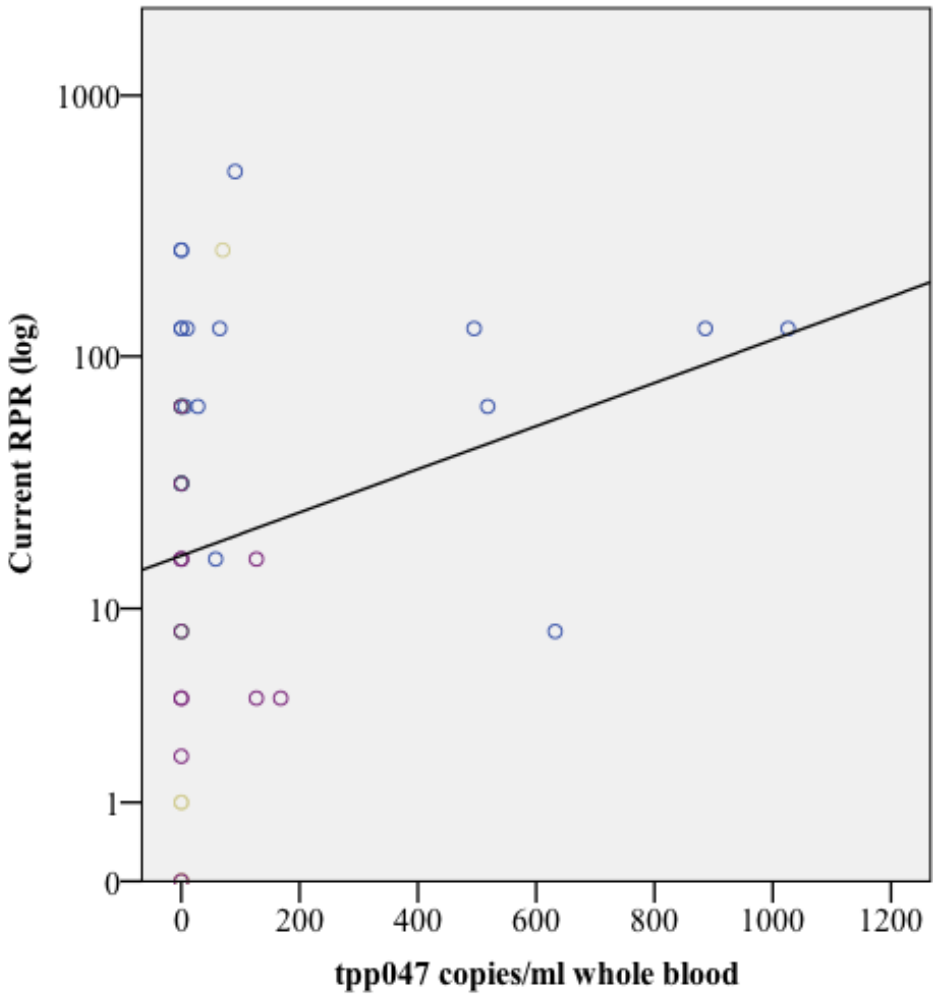


Fig 3.12. Baseline (pre-treatment) RPR titres (log^{10}) of 19 patients with secondary stage syphilis (blue), eight with latent stage (light green) and 12 with primary stage (purple) were compared with the results of *T. pallidum tpp047* quantification in whole blood samples. Spearman's correlation coefficient was 0.354 (p=0.021, 2-tailed))

Table 3.16 Validation of *T. pallidum* RNA and DNA qPCRs

Disease Stage	Patient study number	*16S rRNA* copies/ml blood	*tpp047* copies/ml blood	
Primary	33	3841	0	
	69	0	0	
	90	6892	127	
	93	0	0	
	26	0	0	
	32	3015	168	
	52	48	0	
	89	60	0	
	mean (std. dev.)	1732.00 (2600)	36.88 (69)	p=0.086*
	sensitivity (%)	62.5	25	
	specificity (%)	100	100	
Secondary	29	2437	58	
	27	2061	632	
	34	1776	495	
	45	4376	6	
	61	2433	91	
	74	0	0	
	75	0	0	
	77	0	0	
	83	0	0	
	mean (std. dev.)	1453.67 (1556)	142.44 (243)	p=0.023*
	sensitivity (%)	55.5	55	
	specificity (%)	100	100	
Latent	44	0	0	
	67	82	0	
	95	0	0	
	59	0	0	
	84	0	0	
	mean (std. dev.)	16.40 (16.40)	n/a	
	sensitivity (%)	20	0	
	specificity (%)	100	100	

Table 3.16. *T. pallidum* DNA (*tpp047*) and 16S rRNA detection and quantification were compared in 22 whole blood samples (only samples where both RNA and DNA results were available were included). Sensitivity and specificity were calculated by comparing *T. pallidum* detection with patients' clinic-microbiological diagnosis. Patients included in the study with non-syphilis diagnoses (not shown) became negative controls.

*Statistical significance was calculated with a two-tailed student's t-test.

3.2.6 CDTB samples

Following clinical validation of the *tpp047* qPCR, blood and ulcer samples from a a further cross-section of patients (described in Chapter 2.14.2) were used to assess modifications made to the *tpp047* qPCR (Chapter 3.2.3.2); to continue to monitor the prevalence of *T. pallidum* macrolide resistance (Chapter 5), to quantify *T. pallidum* 16S rRNA in freshly drawn samples and to assess *T. pallidum* load in blood from more patients with latent disease.

3.2.6.1 Demographics

As shown in Figure 3.13, 62 patients were recruited. A total of 58 syphilis diagnoses (nine primary, 21 secondary and 28 latent) three cases of genital herpes and one patient with a non-specific genital ulcer. In contrast to the validation study, patients with neurological signs or symptoms were considered to be secondary or latent cases, and sub-divided into those with and without neurological symptoms. In this cross-section, seven patients had neurological involvement and all were in the secondary stage. Patients with syphilis comprised two women (3.4%) and 58 men (96.6%) of whom 52 (89.7%) were MSM and 36 (62.1%) were HIV-1 infected (with a mean CD4 count of 603.5 (std. dev. 160). Overall, patients had a mean age of 37.45 (range 43) and the majority (62.1%) were white.

Figure 3.13 Summary of CDTB sample collection

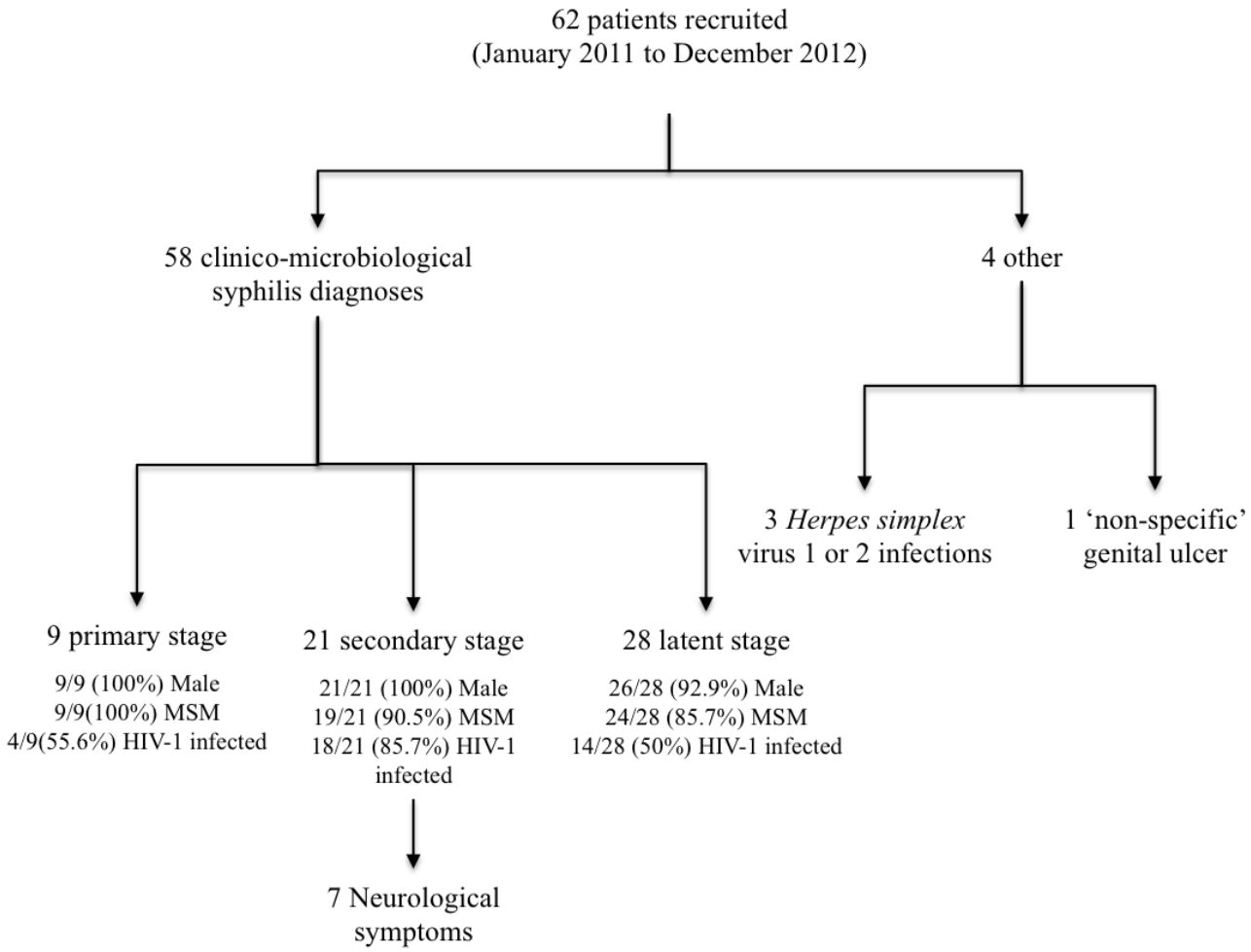


Fig. 3.13. A total of 62 patients donated samples to the CDTB between January 2011 and December 2012 of which 58 patients were diagnosed with syphilis, and four were given an alternative diagnosis.

3.2.6.2 T. pallidum bacteraemia and disease stage

None of the samples collected from patients without syphilis contained detectable *T. pallidum* nucleic acids. DNA quantification results were available from all 58 patients with syphilis, and RNA results from 44 (Tempus™ samples were not collected from the first 10 patients recruited and four samples were spoiled). Table 3.17 and Figure 3.14 show quantification data from all samples and compare DNA and RNA quantification in the sub-group of patients where both samples were available for qPCR. The stage of disease during which both DNA and RNA were most frequently detected was secondary (81% and 61.75%), followed by primary (55.5% and 40%) and finally latent (25% and 39%). Bacterial load, measured by both qPCR and RT-qPCR, was also greatest during the secondary stage and was significantly higher than detected in latent disease (p=0.0001).

The standard deviation of mean RNA levels at any stage of disease were seen to exceed those of DNA, suggesting that DNA level may be a better predictor of disease stage as it is less variable. Figure 3.14 demonstrates that the nine highest *tpp047*-measured bacterial loads (1303 copies/ml and above) were all from patients with secondary disease. When this level of 1303 copies/ml was taken as a cut-off (i.e. levels above this value indicated secondary disease), *T. pallidum* DNA detection in whole blood had a sensitivity of 0.58 (95% CI 0.33 - 0.80) and specificity of 1 (95% CI 0.69 - 1) to identify correctly secondary disease. When a level of <1303 was used to predict latent disease, the sensitivity was 100% (95% CI 0.84 – 1), however specificity was low (33%. 95% CI 0.17-0.52). Binary logistic regression was then employed to calculate the predicted probability of both secondary and latent disease at each observed level of whole blood *T. pallidum* DNA. Next, these probabilities were used to plot receiver operator characteristics (ROC) curves for both stages (Figure 3.15) comparing the true positive rate (sensitivity) on the ‘y’ axis with the false positive rate (1-specificity) on the ‘x’ axis. The area under the curve for secondary disease was 0.784 (asymptotic significance = 0.010) signifying that a whole blood sample containing *T. pallidum* DNA has a 78.4% chance of being from a patient with secondary disease. Latent disease, however, had an area under the curve (AUC) of 0.569 (asymptotic significance = 0.543) confirming that presence *T. pallidum* bacteraemia cannot be used to predict latent infection. Too few quantification results were available from patients diagnosed with primary disease to make detailed analysis of this group meaningful.

Table 3.17 Comparison of the quantification of *T. pallidum* DNA and RNA in whole blood.

Disease Stage	All samples		Paired DNA and RNA results*		
	DNA *(tpp047)*	**16S rRNA**	**DNA (*tpp047*)**	**16S rRNA**	**t-test (2-sided)**
Primary					
no of positive samples (%)	5/9 (55.5)	2/5 (40)	3/5 (60)	2/5 (40)	
mean copies/ml (std. dev.)	205 (388)	486 (850)	248 (518)	486 (850)	p=0.608
Secondary					
no of positive samples (%)	17/21 (80.9)	11/16 (68.75)	12/16 (75)	11/16 (68.75)	
mean copies/ml (std. dev.)	4776 (1396)	2608 (2969)	1458** (1408)	2608** (2969)	p=0.171
Latent					
no of positive samples (%)	7/28 (25)	9/23 (39)	7/23 (30)	9/23 (39)	
mean copies/ml (std. dev.)	124 (315)	796 (1698)	151** (342)	796** (1698)	p=0.081

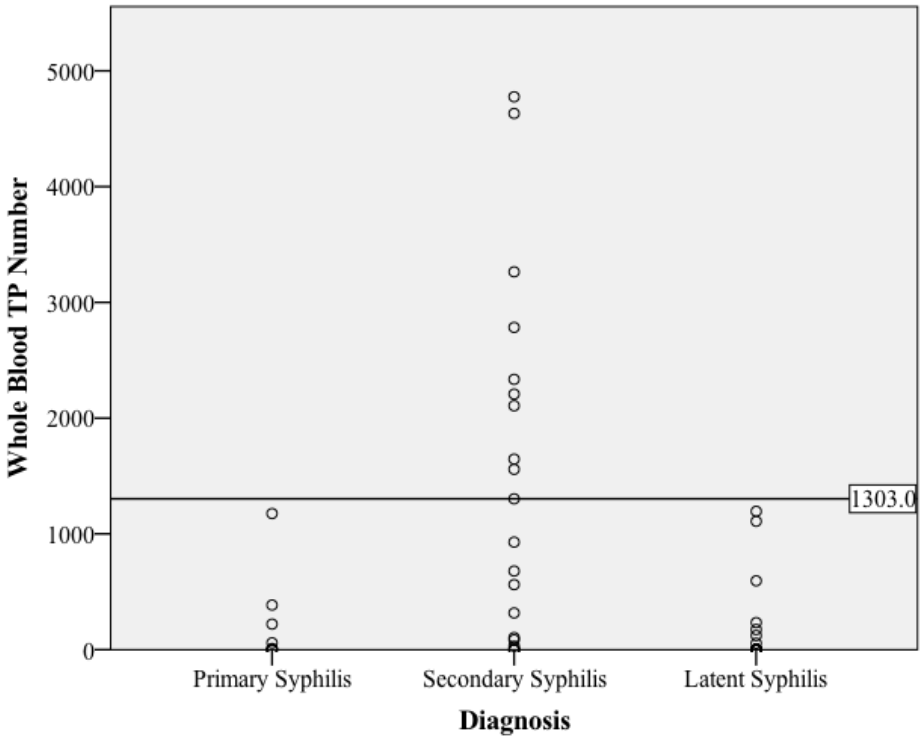
Table 3.17. *T. pallidum tpp047* DNA and 16S rRNA were quantified in 58 whole blood samples collected into EDTA and 44 Tempus™-preserved samples from patients with primary, secondary and latent syphilis.

*Results are only displayed where both DNA and RNA quantification data were available for the same patient.

** RNA and DNA quantification in samples from patients with secondary disease compared with samples from latently infected patients. Secondary vs latent (RNA) p=0.02; secondary vs latent (DNA) (p=0.0001).

Figure 3.14 Comparison of bacterial load in whole blood samples taken from patients with different stages of disease

A. *tpp047* quantification



B. 16S rRNA quantification

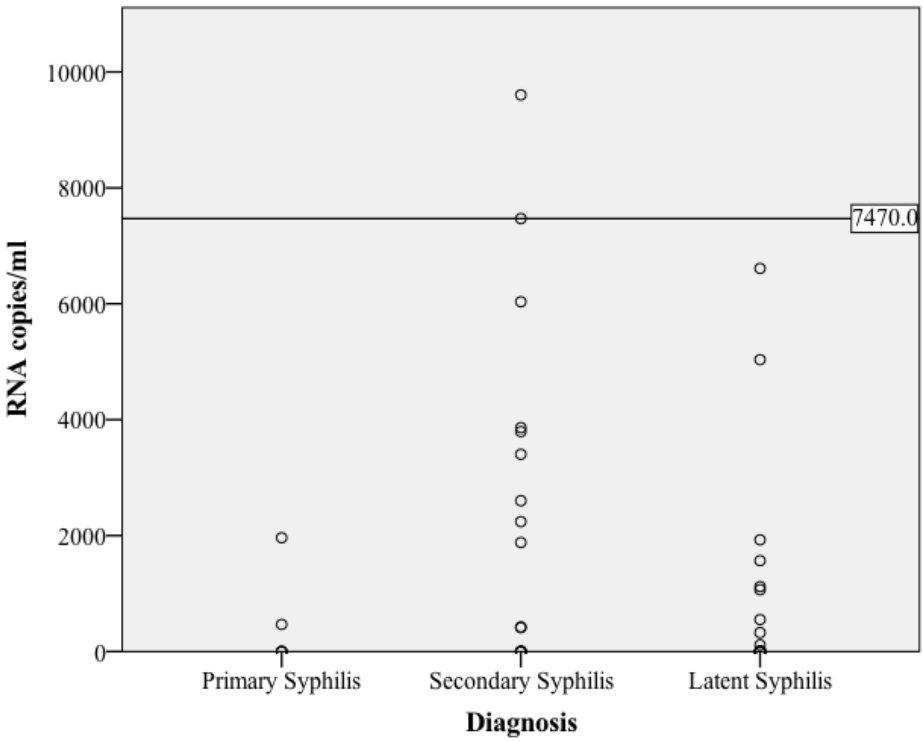
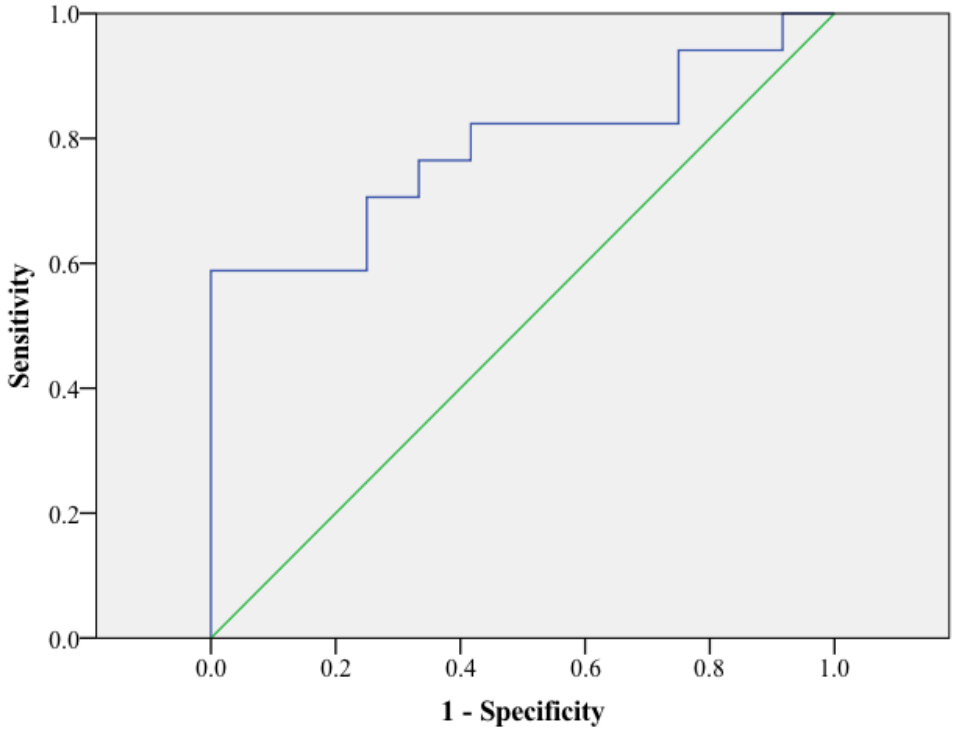


Fig 3.14. A. Whole blood from patients with primary (n=5), secondary (n=16) or latent (n=23) syphilis was collected into EDTA and *T. pallidum* DNA (*tpp047*) quantified by qPCR. **B**. Quantification of *16S rRNA* in Tempus™ blood samples collected from the same patients.

Figure 3.15 Receiver operator characteristics (ROC) of *T. pallidum* detection in blood for the diagnosis of syphilis.

A. Secondary Disease



B. Latent Disease

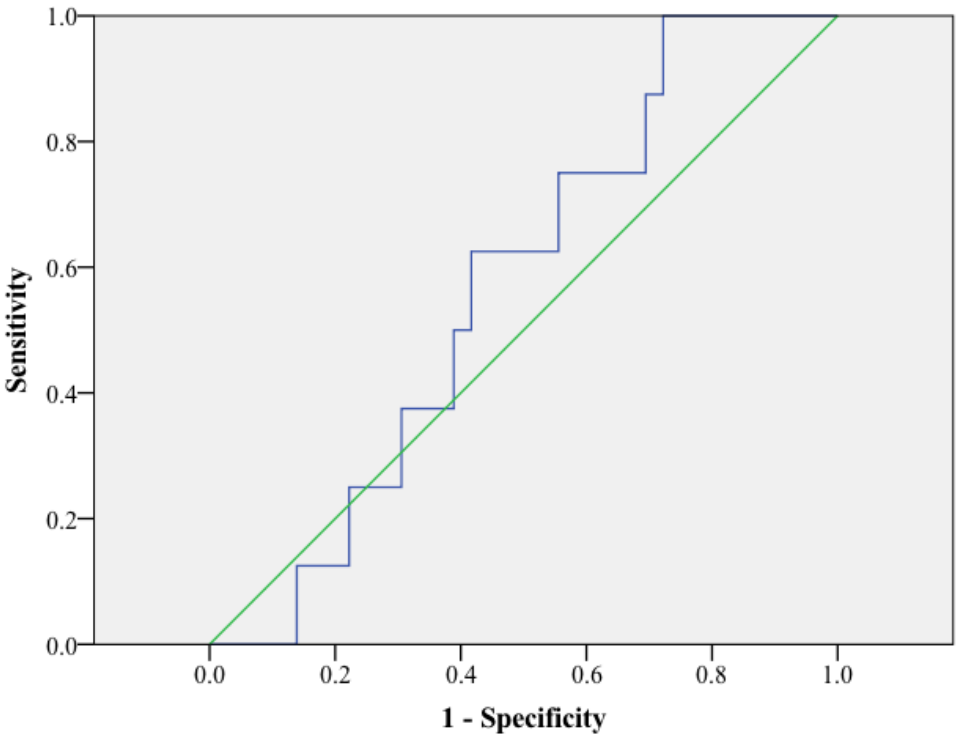


Fig. 3.15. A binary logistic regression was used to calculate the probability of secondary and latent syphilis being correctly identified across a range of whole blood bacterial loads. These ROC curves (in blue) compare the sensitivity and 1-specificity for each probability. An area under the curve (AUC) of 0.5 (green line) represents the null hypothesis that any association between disease and detectable bacteremia is due to chance alone. **A.** demonstrates the AUC for secondary disease to be 0.784 with an asymptotic significance value of 0.010 (95% CI 0.617-0.952). **B.** shows the AUC for latent disease is 0.569 with an asymptotic significance value of 0.543 (95% CI 0.385-0.754).

3.2.6.3 HIV-1 infection is not associated with higher levels of T. pallidum bacteraemia at any stage of disease

In total, 62% of patients included in the current study were HIV-1 infected with a mean CD4 count of 603.5 cells/μl. This percentage varied according to stage of syphilis, as demonstrated in Figure 3.13. Of note, the percentage of HIV-1 infected patients who presented with secondary disease was 85.7% compared with 55.6% and 50% for those diagnosed with primary or latent disease, respectively.

Figure 3.16 demonstrates that no significant differences in *T. pallidum* bacterial load (measured with both *tpp047* and *16S rRNA*) were identified when patients with and without HIV-1 infection were compared.

3.2.6.4 Concordance of RNA and DNA detection in samples from both cross-sectional studies.

The detection of *tpp047* DNA and 16S rRNA in samples collected during the two UK cross-sectional studies were compared. Only those Tempus™ blood samples which were intact (not clotted) were included in this analysis. For patients with secondary disease, there was 100% concordance between DNA and RNA detection, with 15/23 (65%) of samples positive for both nucleic acids and 8/23 (35%) of samples negative for both. Samples from patients with primary disease were 70% concordant and 89% concordance was observed for patients with latent disease. RNA detection was the more reliable in primary and latent disease with 3/13 (23%) and 3/28 (11%) of samples being RNA positive and DNA negative, respectively. A single primary sample (7%) and no latent samples were DNA positive, RNA negative.

Figure 3.16 *T. pallidum* quantification in blood according to HIV-1 status

A. 16S rRNA quantification

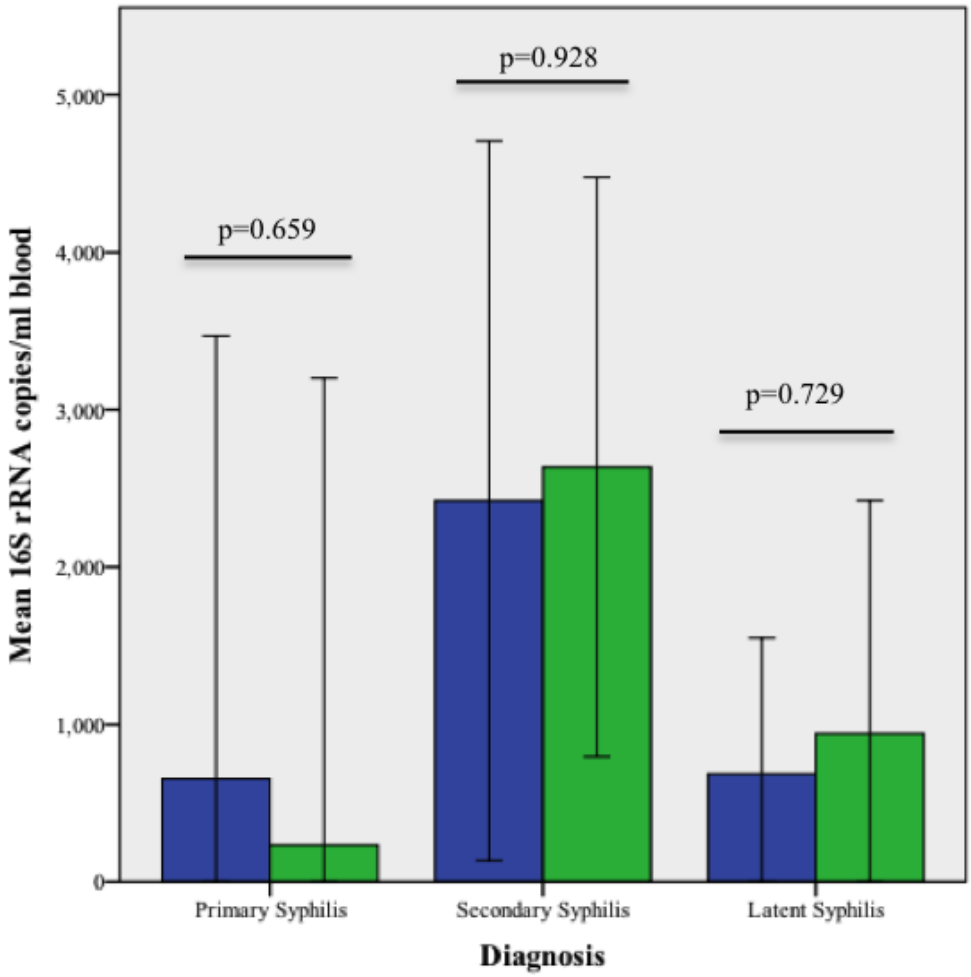


Fig 3.16. B. *tpp047* DNA quantification

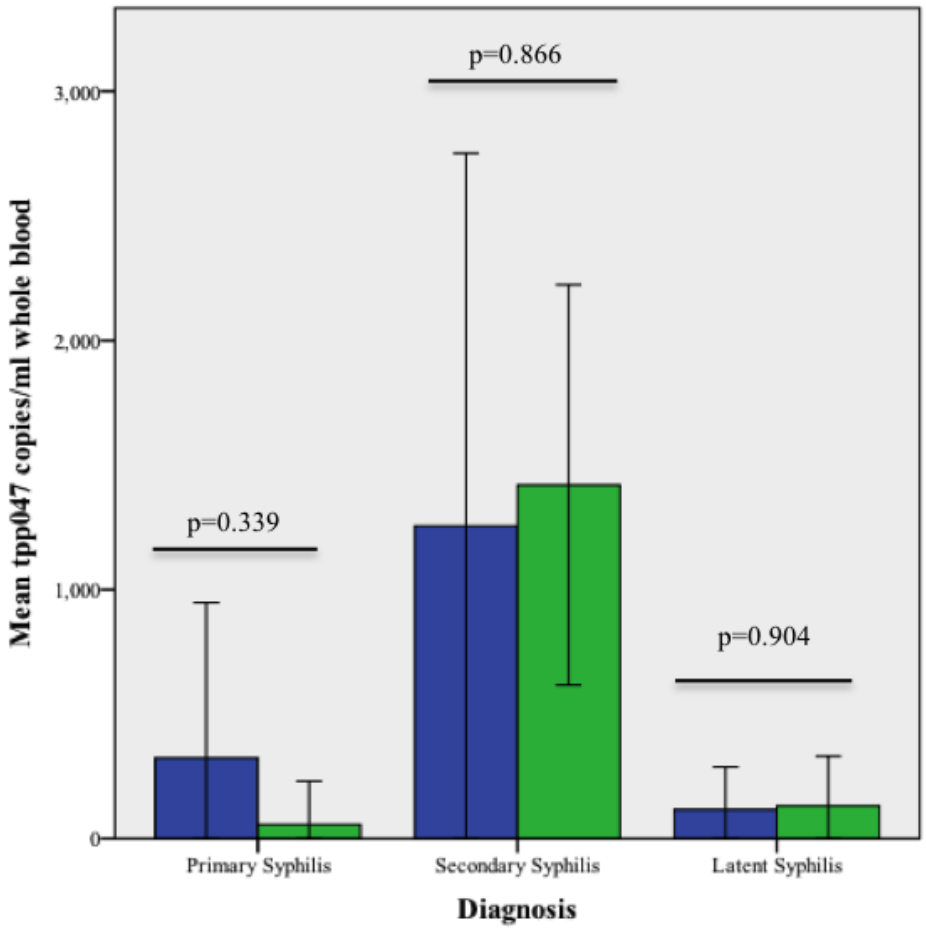


Fig 3.16 *T. pallidum* 16S rRNA (**A**) and *tpp047* DNA (**B**) and were quantified in Tempus™ samples and whole blood collected into EDTA from 58 patients with (green bars) and without (blue bars) HIV-1 co-infection. Error bars define the 95% confidence intervals.

3.2.7 Cross-sectional study of *T. pallidum* in Sri Lanka

A total of 25 exudate samples were collected from 25 patients with microbiologically confirmed syphilis between January and August 2012. All subjects except one were male of whom 11 (45.8%) were MSM. The mean age of subjects (excluding a neonate with congenital disease) was 29.5 (std. dev. 9.7) years and all except two patients were Sinhalese.
T. pallidum DNA and RNA were detected in 24/25 samples (96% sensitive, 95% CI 76-99%). The single negative sample was taken from a female patient who presented with a DGM-negative labial ulcer and positive syphilis serology (RPR 1:64). Excluding the negative sample, 14 samples were from primary chancres, 9 were from secondary lesions and one was a skin swab (neonatal sample). A summary of recruitment and samples is given in Figure 3.17.

3.2.7.1 T. pallidum quantification according to stage of disease

T. pallidum DNA and RNA were quantifiable in 14 primary chancres. These chancres had a mean area of 57.15mm^2 (std. dev. 36.42) and contained an average of 3.84x10^4 (std. dev. 4.78x10^4) *tpp047* copies/strip and significantly more 16S rRNA (6.79x10^6 (std. dev. 7.52x10^6) copies/strip) (p=0.002). Secondary lesions had a mean area of 32.71mm^2 (std. dev. 20.50) and contained 8.42x10^3 (std. dev. 1.08x10^4) *tpp047* copies/strip and 1.41x10^6 *16S rRNA* copies/strip (std. dev. 2.26x10^6) (p=0.08). When the quantity of DNA and RNA present in primary and secondary lesions were compared (Figure 3.18), it was noted that primary chancres contain more of both nucleic acids than secondary lesions, but that only the difference in RNA reached significance (p=0.05).

3.2.7.2 Ulcer size correlates with bacterial load in secondary disease

Ulcer area data were available for seven secondary lesions and found to correlate with *tpp047* (DNA) measured in those lesions. Figure 3.19 highlights this association and Pearson's correlation coefficient (for normally distributed data) was calculated to be 0.831 (p=0.021). A similar correlation was not observed with primary lesions using Spearman's rank (non-parametric data) (0.006, p=0.986). Ulcer RNA levels showed no correlation with ulcer size

for either stage of disease, although there was on overall weak correlation between ulcer load measured by DNA and RNA ($R^2 = 0.367$) (Figure 3.20)

Figure 3.17 Recruitment of patients from Colombo, Sri Lanka

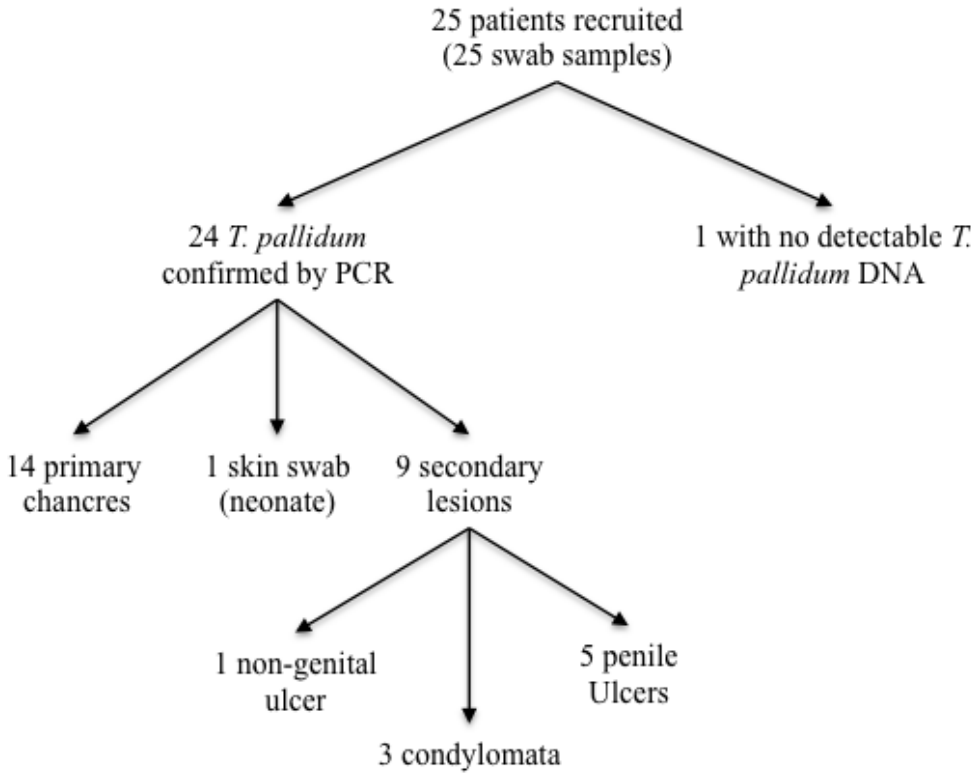


Fig 3.17. Summary of the recruitment of patients diagnosed with early syphilis to a cross-sectional study between January and August 2012 at the National Centre for STD control in Colombo, Sri Lanka.

Figure 3.18 *T. pallidum* bacterial load in primary and secondary lesions

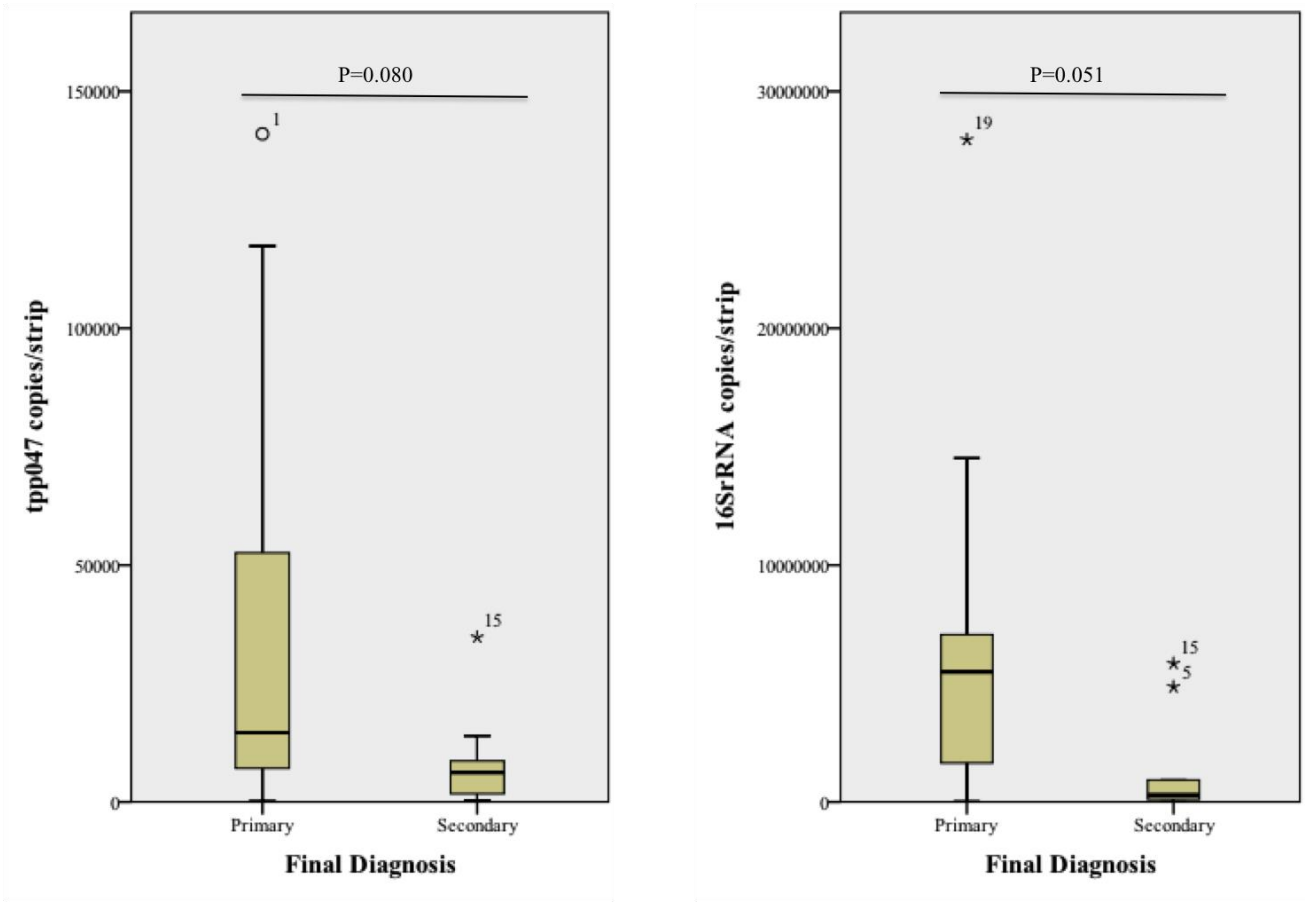


Fig. 3.18. *T. pallidum* DNA (*tpp047*) and 16S rRNA were quantified in 14 primary chancres and nine moist secondary lesions by qPCR and RT-qPCR, respectively. The symbols ○[1], ★[5], ★[15], ★[19] represent outlying values (the numbers refer to the study number assigned to each patient).

Figure 3.19 Correlation between ulcer size and ulcer bacterial load

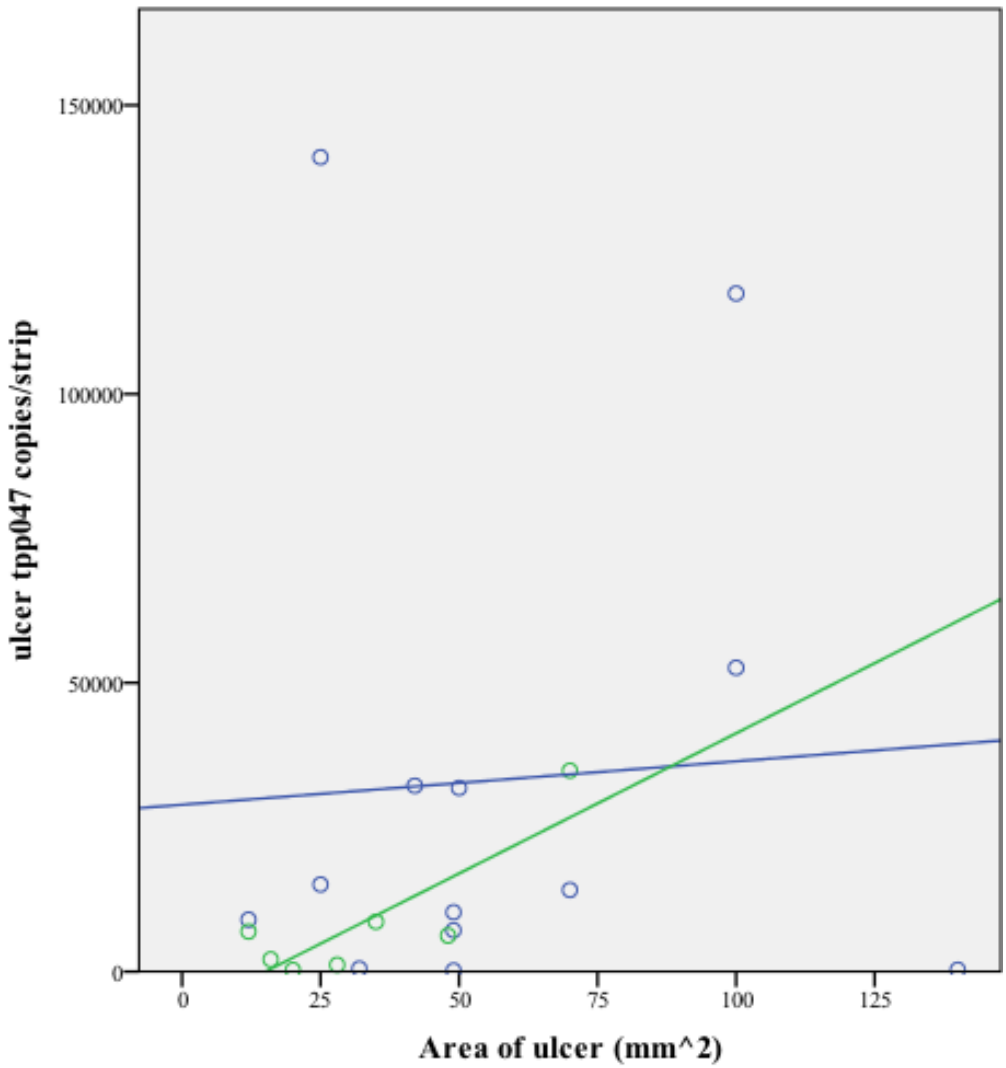


Fig 3.19. The area of primary chancres (blue) and moist secondary lesions (green) was measured prior to ulcer exudate collection. The quantity of *T. pallidum* DNA (*tpp047*) in the lesions was then determined by qPCR and compared with ulcer area.

Figure 3.20 Correlation between *T. pallidum* DNA and RNA levels in ulcers

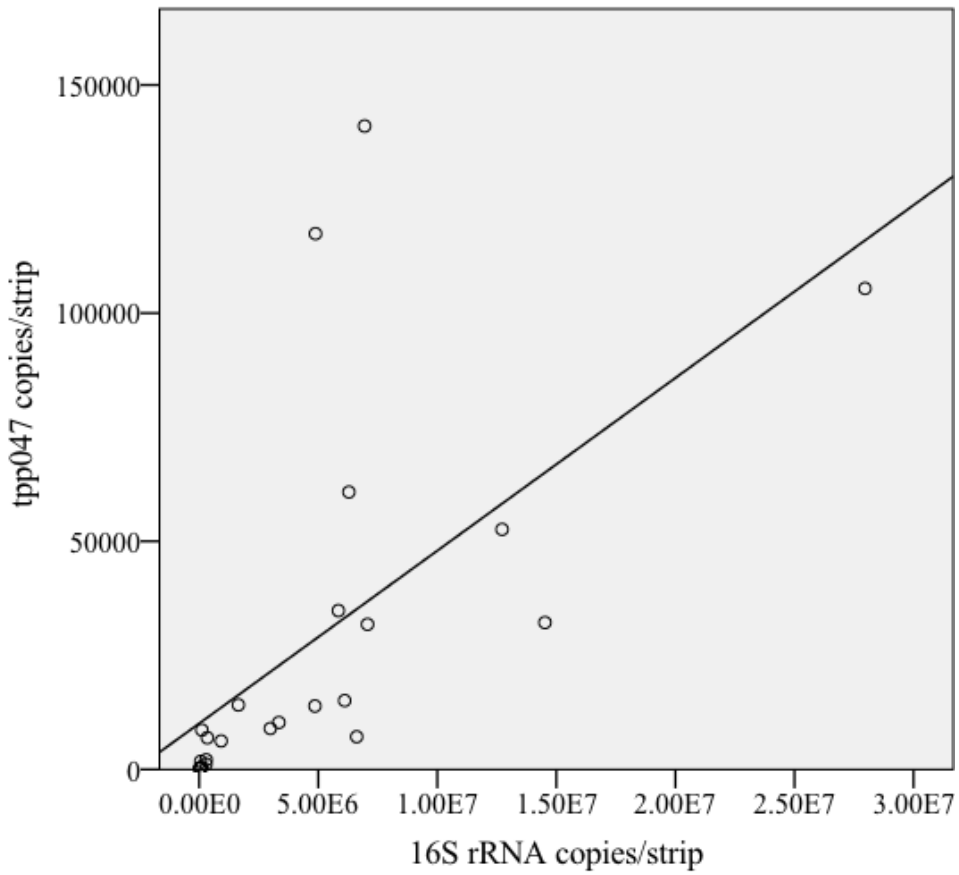


Fig 3.20. *T. pallidum* DNA (*tpp047*) and 16S rRNA were quantified in 23 samples (14 primary chancres and nine secondary lesions). When compared, a positive correlation was observed between DNA and RNA-measured bacterial loads ($R^2 = 0.367$).

3.2.7.3 Association between duration of symptoms and T. pallidum load in ulcer exudates

Data concerning the duration of symptoms prior to presentation were available for 13 patients with primary and nine patients with secondary disease. The mean quantity of *T. pallidum* DNA present in a primary chancre of 4-7 days' duration was $2.14x10^4$ (std. dev. $1.83x10^4$) *tpp047* copies/strip, which increased to $3.18x10^4$ (std. dev. $4.94x10^4$) at 8-14 days and $5.61x10^4$ (std. dev. $5.98x10^4$) if present for longer than two weeks. This trend, shown in Figure 3.21, was not statistically significant. RNA levels in primary chancres are seen to rise between 4-7 days ($7.21x10^6$) and 8-14 days ($9.08x10^6$), but then fall after 14 days ($5.68x10^6$). Only a single sample was available from a secondary lesion within 7 days of onset and both DNA and RNA levels in secondary lesions at 8-14 days and over 14 days were similar.

Figure 3.21 Duration of symptoms and *T. pallidum* bacterial load

A. DNA quantification

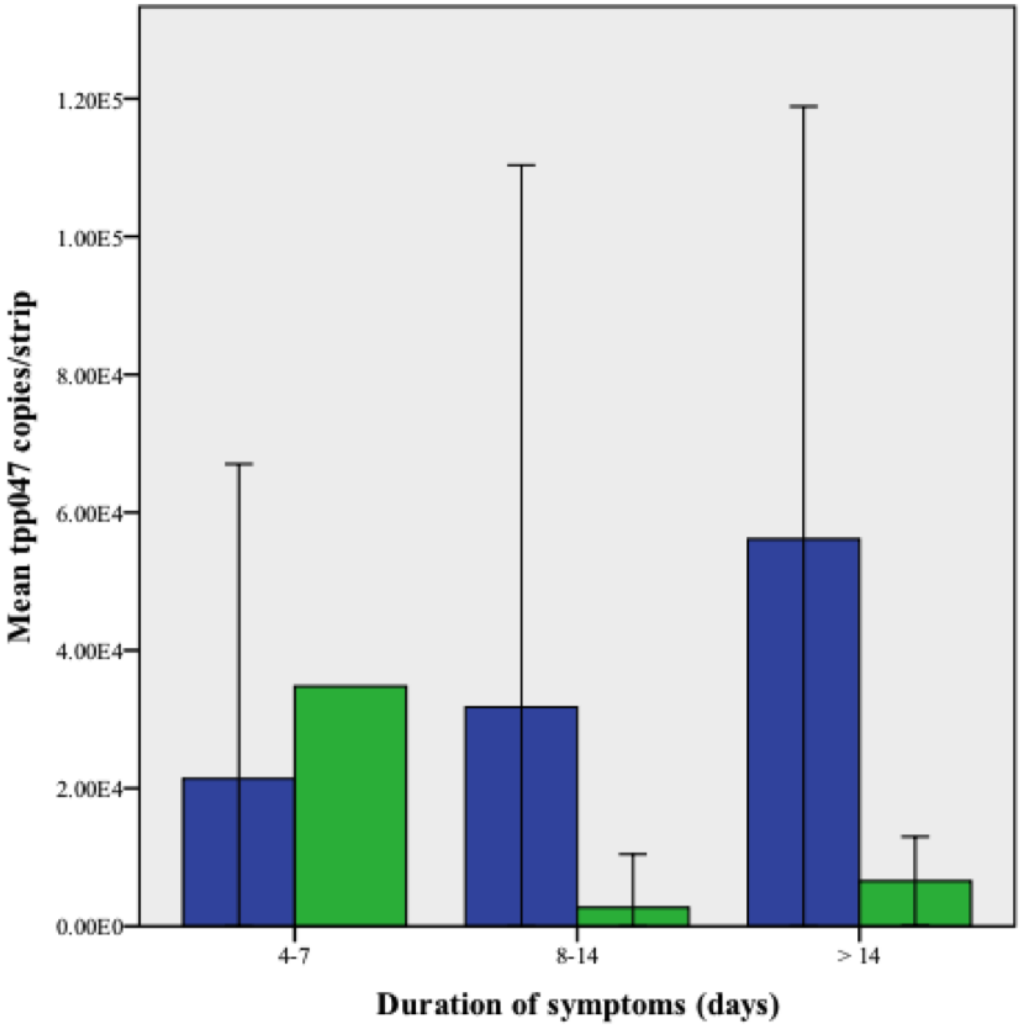


Fig 3.21. B. RNA quantification

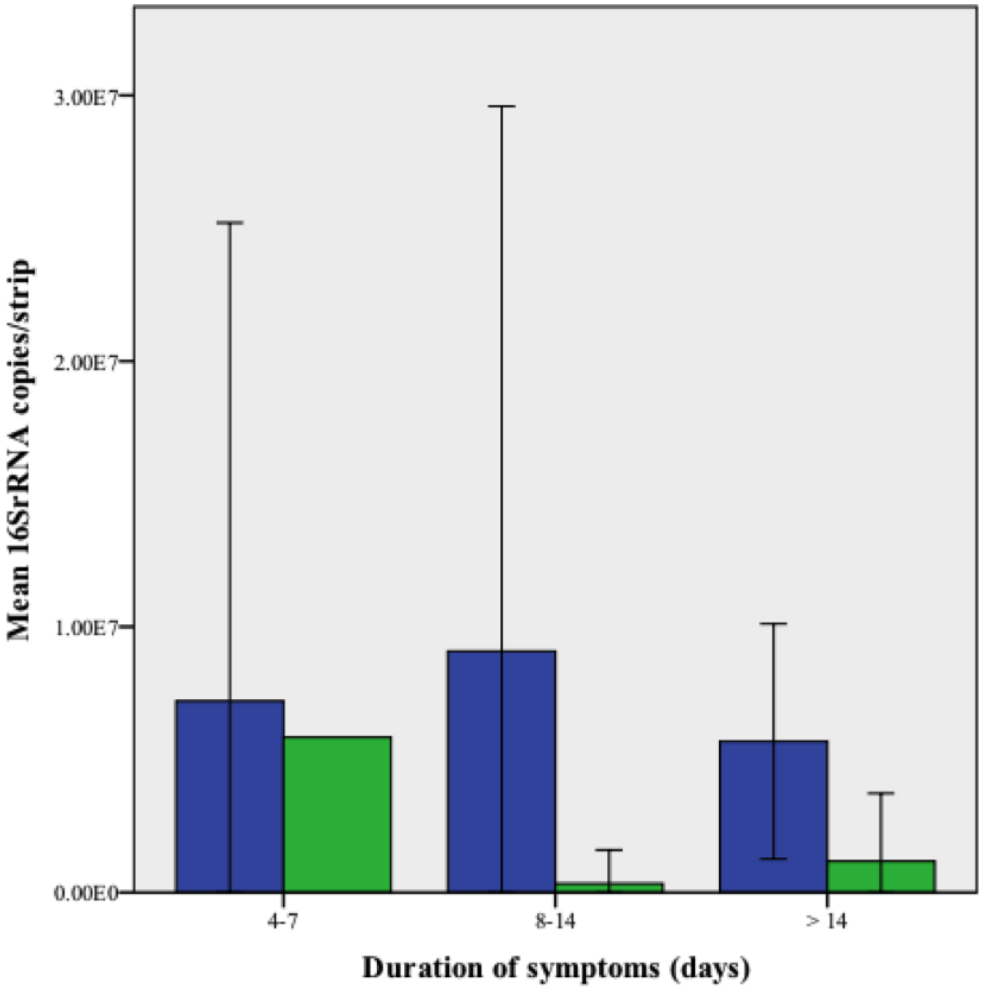


Fig 3.21. Duration of symptoms (days since reported onset) was compared with ulcer *T. pallidum* load for 13 patients with primary (blue) and nine patients with secondary disease (green). **A.** shows load according to DNA (*tpp047* gene) quantification and **B**. shows 16S rRNA quantification. Error bars depict 95% confidence intervals.

3.2.8 Impact of long-term sample storage on *T. pallidum* nucleic acid detection and quantification

The blood and ulcer samples collected for each of the three current studies were the same, but the methods for storage, DNA extraction and *T. pallidum* quantification were not. Differences in sample handling are described fully in Chapter 2 and are summarised in Table 3.18. For this reason, quantification results for each study have been considered separately. It is, however, possible to compare quantification data between studies to assess further the impact of long-term sample storage.

Table 3.19 A compares DNA and RNA quantification data from blood samples collected in the periods 2006-8 and 2011-12. Combining all stages of disease, *T. pallidum* DNA was detected in 28% of whole blood samples collected in 2006-8, which had been stored long-term at -80°C, compared with 50% of samples from 2011-12. Moreover, a significantly higher quantity of DNA was detected in samples collected during the later period. A similar pattern was seen when each stage of disease was analysed separately (data not shown). In contrast, RNA detection and quantification appeared to be unaffected, with roughly 50% of samples from both periods containing detectable RNA and at similar levels.

T. pallidum load in primary disease ulcers and secondary disease lesion samples from all three studies are compared in Table 3.19 B. A similar pattern is seen, where samples collected in 2006-8 and stored at -80°C until 2011 contained smaller amounts of DNA. RNA quantification data are presented for completeness, but were only available for samples collected in Colombo.

Table 3.18 Comparison of sampling handling and qPCR methods

	2006-8 (qPCR validation)	2011-12 (CDTB)	Cross-sectional study in Colombo
Snostrips	Stored dry at -80°C.	Extracted immediately or stored for 1-2 months in Lysis buffer at -80°C	Stored in RNAlater® at -80°C.
Whole Blood	Histopaque separation and storage at -80°C or divided into 400µl aliquots and stored at -80°C	500µl added to 500µl Lysis Buffer and extracted immediately or stored for 1-2 months at -80°C	n/a
Tempus™	Stored at -20°C until 2011, then -80°C until extracted in late 2012.	Stored at -80°C for 1-2 months until extracted.	n/a
qPCR method	SYBR-green based qPCR (Roche light-cycler instrument). Standard curve made from whole organism DNA. RT-qPCR with 16SrRNA TaqMan assay (Biorad instrument). Plasmid standard (p16SRNA)	TaqMan q-PCR and RT-qPCR with plasmid standards (pTp47 and p16SRNA) (Biorad instrument)	TaqMan q-PCR and RT-qPCR (Biorad instrument) with plasmid standards (pTp47 and p16SRNA)

Table 3.18. Summary of sample handling techniques and *T. pallidum* quantification methods in three cross-sectional studies.

Table 3.19 Detection and quantification of *T. pallidum* nucleic acids

A. Quantification in blood samples

Study		DNA (whole blood)	RNA (Tempus™)
2006-8 (qPCR validation)	n (% pos)	21 (28%)	21 (48%)
	Mean copies/ml	51	1202
	std. dev.	141	1952
2011-12 (CDTB)	n (% pos)	58 (50%)	44 (50%)
	Mean copies/ml	597	1420
	std. dev.	1108	2335
	t-test (2-sided)	p=0.027	p=0.713

B. Quantification in ulcer samples

Study		DNA	RNA
2006-8 (qPCR validation)	n (% positive)	14 (100%)	
	Mean copies/strip	2480	
	std. dev.	3419	
2011-12 (CDTB)	n (% positive)	4 (75%)	
	Mean copies/strip	18042	
	std. dev.	19668	
Cross-sectional study in Colombo	n (% positive)	23 (100%)	23 (100%)
	Mean copies/strip	26646	4.68×10^6
	std. dev.	40168	6.51×10^6

Table 3.19 A. Comparison of *T. pallidum* DNA (*tpp047* gene) and 16S rRNA quantification in Tempus™ samples and whole blood collected into EDTA from patients in two separate studies. **B**. Quantification of *T. pallidum* DNA (*tpp047*) and 16S rRNA in ulcer samples collected during three separate cross-sectional studies.

3.3 Discussion

For 20 years, PCR has been used on a variety of clinical and rabbit specimens to detect, type and quantify *T. pallidum* (51, 132, 150, 151, 153). Now PCR reliably detects *T. pallidum* in ulcer specimens, often as part of a multiplex assay, for the diagnosis of primary syphilis (132). The more modest success of *T. pallidum DNA* PCR in other clinical samples has been attributed to their lower treponemal load and the presence of PCR inhibitors, especially in blood samples (133, 134). There are few data on the quantification of *T. pallidum* by PCR in clinical samples thus the timing of bacteraemia and duration post-treatment, as well as the impact of HIV-1 infection on the course of *T. pallidum* infection, is uncertain. Although cross-sectional, the data presented here provide insight into the quantity of *T. pallidum* organisms, measured by qPCR, that are present in blood and ulcer samples taken from patients at different stages of syphilis, both with and without HIV-1 infection.

The initial focus of the current study was to establish an efficient method to extract *T. pallidum* nucleic acids from clinical samples. A number of extraction methods exist, which are generally solution or column-based, and many have been developed into commercial kits. Chomczynski and Sacchi described perhaps the best known solution-based method for DNA/RNA/Protein isolation using guanidium thiocyanate-phenol-chloroform (234). Column-based methods, however, are quicker, do not rely on toxic and irritant chemicals and are not subject to problems with phase separation (235). The Roche highpure, Qiagen QiAamp (both mini and midi) and Qiagen All-prep kits employed for *T. pallidum* nucleic acid extraction were all solid phase extraction methods using silica-containing micro-centrifuge columns. In order to avoid phenol:cholorform-based co-extraction of DNA and RNA from ulcer exudate samples, it was initially decided to extract DNA and RNA separately, by dividing the sample into two aliquots. When compared, however, the Qiagen all-prep method for DNA/RNA co-extraction was more efficient at recovering *T. pallidum* DNA from snotrips stored in RNAlater® than a combination of QiAamp and Roche highpure methods. Moreover, all-prep was quicker (important to avoid RNA degradation) and had the advantage of not requiring the limited material of the ulcer samples to be divided. All extraction processes begin with cell disruption/lysis to release nucleic acids, followed by inactivation of nucleases and separation of cell debris and salts from the target molecule (DNA, RNA, proteins). Tempus™ tubes achieve cell lysis and RNA preservation immediately after blood is drawn and mixed

vigorously with the detergent and guanidium thiocyanate inside the tube (236). Moreover, this method of blood collection and RNA extraction better preserves RNA (237, 238). When compared with alternative methods of *T. pallidum* RNA extraction from human blood, the Tempus™ system was equally effective and required less processing time. Moreover, concerns that adequate bacterial lysis would not be achieved with the Tempus™ detergent alone were unfounded and an additional proteinase K digestion step resulted in decreased RNA yield. Another concern was that the overwhelming majority of human RNA species and the presence of other PCR inhibitors (e.g. haem, calcium) in Tempus™ blood samples would inhibit the detection of bacterial RNA or result in high levels of non-specific target amplification (239). A single attempt was made to deplete human poly-adenylated RNA from Tempus™-extracted samples using magnetic bead capture, which was counterproductive in improving *T. pallidum* detection. While it is likely that a proportion of human RNA was removed from the samples tested, however, it is equally likely that additional sample handling resulted in *T. pallidum* RNA degradation and negated any benefit thus further exploration of the technique was abandoned.

The *T. pallidum* gene *tpp047* is the most commonly used PCR target for *T. pallidum* amplification and had an excellent likelihood ratio for the correct identification of primary syphilis in a recent meta-analysis (52, 132, 133, 143, 240). New primers were designed to amplify this target in two real-time PCR assays. The first was based on the detection of SYBR-green fluorescence and performed in a Roche LightCycler 2.0 instrument. This assay was specific for *T. pallidum* detection with no cross-reactivity to *T. denticola* or any of the common commensal skin flora and sexually transmitted organisms tested. Analytical sensitivity was maintained when quantifying *T. pallidum* DNA extracted from human whole blood samples (that have been previously spiked with *T. pallidum*) with no evidence of inhibition from co-extracted human DNA. Absolute quantification of *T. pallidum* was achieved by comparing cycle thresholds of unknown samples to a standard of DNA extracted from whole organisms. The limit of detection of this assay was determined to be a single organism (genome equivalent) per 10µl reaction and inter-assay variance (reproducibility) was acceptable. It must be noted, however, that assuming a Poisson distribution and a 95% chance of including at least one copy in a given reaction, that the theoretical limit of detection for any qPCR is three copies/reaction (241). It is possible that the DNA from a single *T. pallidum* contained more than a single copy of the gene *tpp047*, although this has no foundation in published whole gene sequences of Nichols strain *T. pallidum* (28). An

alternative explanation is inaccuracy of the initial quantification by DGM of spirochetes from which DNA for the quantification standard were extracted.

A further goal was to design an internal control that could be co-extracted with *T. pallidum* DNA then co-amplified in a single qPCR. It is not possible using SYBR-green to quantify two or more targets in the same qPCR, thus an alternative assay was designed and validated, this time with a Biorad real-time PCR instrument. This assay paired the existing *tpp047* amplification primers with a novel hydrolysis (TaqMan) probe and absolute quantification was achieved by comparison to an in-run plasmid (pTp47) DNA quantification calibrator. Inter-assay variance remained low at <2.35%, intra-assay variance (repeatability) was equally low <1.05% and the analytical sensitivity of the new qPCR was determined to be 10 *tpp047* copies/reaction. Multiplex detection of *tpp047* and internal control sequences is described in Chapter 4.

A recent meta-analysis compared 46 studies aiming to diagnose syphilis using PCR (52) and showed the sensitivity of *T. pallidum* detection in whole blood across all stages of disease to be 41.2% (95% CI 26.5 – 57.6%). *T. pallidum* was most reliably detected in blood from neonates with congenital disease (sensitivity 83%) and in adults with secondary disease (sensitivity 54.2%). A previous attempt to improve this sensitivity was made by targeting bacterial RNA in place of DNA. A 100-fold greater sensitivity of the detection of diluted whole *T. pallidum* organisms was shown using a 16S rRNA target which, based on data from other bacterial species, was predicted to be present at 5,000 to 10,000 copies per organism (153). Of interest are data from an animal model demonstrating that DNA from heat-killed *T. pallidum* could persist in skin lesions for 15-30 days (155), whereas RNA was predicted to be more quickly cleared due to its inherent instability. The potential for more sensitive *T. pallidum* detection with a novel 5' nuclease (TaqMan) RT-qPCR was, therefore, assessed in the current study. The prediction of a more rapid clearance of this RNA target following syphilis treatment is addressed in Chapter 5. Following the trial of three probes and four primer pairs, assay design settled on the previously published sense primer sequence (153) together with a novel anti-sense primer and hydrolysis probe. This combination was specific for *T. pallidum* species detection and did not result in amplification of any member of a panel of skin and sexually transmitted bacteria nor of human DNA or RNA. A 'one-step' assay method, where reverse transcription and PCR take place in a single tube, was selected to reduce to potential for contamination with RNAse and cross-contamination between samples

(242). Moreover, given that a one-step RT-qPCR uses the same reaction primers for both reverse transcription and for qPCR, the potential for non-specific target amplification is reduced. An RNA quantification standard was required for the absolute quantification of *T. pallidum 16S rRNA* as despite being more stable and therefore reproducible, a DNA standard would not control for reverse transcription (243, 244). RNA for this standard was produced from the recombinant *nco-1* linearised plasmid (p16SRNA) by the action of T7 polymerase and template plasmid DNA subsequently removed with DNAse treatment. The RNA transcript was then purified, subject to serial dilution and used to test rigorously the amplification characteristics of the assay. The average efficiency of the quantification standard curve was established as 87.7%. Analytical sensitivity was 23 copies/reaction with an intra-assay variance of 1.12% or less and a maximum inter-assay variance of 3.05%, thus demonstrating adequate sensitivity, repeatability and reproducibility. It is likely that the slight (1%) increase in inter-assay variability of the RT-qPCR compared with the *tpp047* qPCR was caused by inherent variation in the efficiency of reverse transcription (245).

With the assays established, attention turned to the validation of *T. pallidum* quantification in clinical samples. SYBR-green-based qPCR was used to quantify *T. pallidum* DNA in ulcer samples collected from 19 patients with syphilis (15 primary, five secondary) and 35 controls in the qPCR validation study and from a further 25 patients (nine secondary, 15 primary, one congenital) in the Sri Lankan cross-sectional study. The sensitivities of *T. pallidum* detection in primary chancres were 100% and 96%, respectively and compare favourably to a meta-analysis comparing 15 studies (mean sensitivity 78.4%, 95% CI 68.2 – 86.0) (52). In the qPCR validation study, the mean bacterial load (DNA) in primary chancres was measured at 2480 organisms/strip. *T. pallidum* DNA was also detected in samples from secondary syphilis lesions with three quarters of genital and two thirds of oral lesions being positive. Fewer organisms were detected in these lesions compared with primary chancres (303 organisms/strip). Although a difference in extraction and PCR techniques does not permit direct comparison of quantification data between studies, the same pattern was seen in Sri Lankan samples where significantly more 16S rRNA was detected in primary ulcers compared with secondary lesions. This observed reduction in *T. pallidum* load as disease progressed from early to latent is supported by studies from the pre-PCR era, which demonstrated using DGM and bacterial staining that exudates from primary disease were rich in *T. pallidum*, but treponemes were scant in histological samples from latently infected patients (224). Of further note is an assessment of *T. pallidum* dissemination during

experimental rabbit infection when bacterial load in testicular tissue (measured by DGM) rose from $1.4x10^6$ treponemes/ml exudate at day 3 post-innoculation to $8.2x10^7$ at day 11 then decreased thereafter (246). This parallels the observation in primary chancres from Sri Lanka that *T. pallidum* load increased with the duration of symptoms. However, this study was cross-sectional, not longitudinal, and the observed difference was not statistically significant. Analysis of DGM results from both studies provides further support for the validity of qPCR data. In the hands of a skilled operator, DGM has a sensitivity of approximately 80% for the identification of *T. pallidum* in lesions of both primary and secondary disease (114, 247). In the validation study, 71% of primary and 40% of secondary lesions were DGM positive, compared with only 53.8% of primary and 57% of secondary Sri Lankan ulcer samples. If used as a diagnostic test, qPCR would have identified an additional 16 cases of syphilis in these studies. It was further observed that for primary chancre samples collected in London, DGM positivity was associated with bacterial load, such that those ulcers with high load were more likely to have been DGM positive. To summarise, ulcer PCR is a sensitive and specific test for the diagnosis of primary and secondary syphilis. It out-performed traditional DGM, although it is not a point-of-care diagnosis. Moreover, the concept of lesion bacterial load, measured by qPCR, has been validated and results shown to be biologically plausible.

In the current study, *T. pallidum* DNA was more frequently detected and at a higher load in whole blood than in plasma or PMBC fractions. This finding is similar to a previous description of *T. pallidum* detection in rabbit blood where significantly more organisms were recovered from whole blood and plasma compared with serum and PBMCs (51). *T. pallidum* bacteraemia is thought to occur from the outset of the primary infection and to persist at all stages of disease thereafter, albeit at markedly different levels (146), and a number of studies have attempted to detect *T. pallidum* in whole blood samples from patients. The mean sensitivity of detection in primary disease was previously calculated to be 36.1%, rising to 54.2% in secondary disease and falling to 20.2% in the latent stage (52). While the stages of disease are clinically well defined and the likelihood of detectable bacteraemia is measurable, the rate of bacterial dissemination and clearance (both immunological and following treatment) in humans is not known. In rabbits inoculated intra-testicularly, a low load of *T pallidum flaA* DNA could be detected in blood after 24 hours and a much higher load was detected after 7-10 days. This increased detection after one week also corresponded with rising RPR titres and widespread treponemal dissemination into organs (51). In both the qPCR validation and CDTB samples, detection rates observed in whole blood samples from

patients with any stage of syphilis were lower than for ulcers, but were comparable with those previously reported (151, 240). This could be partly explained by the presence of PCR inhibitors in blood-extracted samples, however the ulcer is the site of bacterial inoculation, multiplication and invasion and would be expected to contain more *T. pallidum*. Moreover, the current data demonstrate that improved detection of bacteraemia during secondary syphilis corresponded to higher bacterial load, a similar picture to that reported in the rabbit model 7-10 days post-innoculation. Were the difference in ulcer and blood load due to PCR inhibition alone, such dynamic change according to stage would perhaps not be expected.

Combining qualitative and quantitative findings a pattern of bacterial dissemination emerges. Primary disease is characterized by high numbers of organisms in ulcers, but low-level or undetectable *tpp047* DNA in blood. In secondary disease, healing ulcers and mucosal lesions contain *T. pallidum* at lower load, but bacteraemia is more frequently detected and at higher load than in primary disease. By the latent stage, bacteraemia is less commonly detected and when present is at low load.

The greater number of blood samples collected for the CDTB from patients with secondary and latent disease enabled more detailed analysis of *T. pallidum* load during those stages of disease. Bacteraemia was associated with a 78.4% probability of being from a patient with secondary disease, and only patients in this stage had more than 1303 *tpp047* copies/ml whole blood. Low bacterial load, however, was equally associated with primary or latent disease and could not be used to predict disease stage, although a load over 1303 *tpp047* copies/ml effectively excluded latent disease.

An assay to quantify 16S rRNA was developed to improve the sensitivity of *T. pallidum* detection and to exploit the inherent instability of RNA in order to follow changes in bacterial load more quickly. RNA captured in Tempus™ blood samples from both the qPCR validation and CDTB was extracted and quantified with the same quantification standard and the assay was 100% specific in both studies. Using the qPCR validation samples, qualitative RNA detection identified three additional primary cases and one additional latent case when compared with DNA detection. Mean 16S rRNA levels for both these stages were higher than those of DNA, although not significantly so. The same number of secondary cases were positive by both qPCR and RT-qPCR, but significantly more RNA was detected at this disease stage than DNA.

When CDTB samples were analysed a similar, although not identical, picture was seen as 16S rRNA detection identified two additional latent cases, but one less secondary and one less primary case. Again, RNA levels were higher than DNA at all stages, although not significantly so for any stage this time. It was predicted that *T. pallidum* RNA levels in clinical samples would far exceed those of DNA, given that each bacterium is expected to contain a single copy of the *tpp047* gene but many hundreds of 16S ribosomes. While RNA levels were undoubtedly higher than DNA, however, this did not translate to a large improvement in *T. pallidum* detection.

The near-parity of DNA and RNA detection for CDTB samples could be explained by superior *tpp047* detection with a hydrolysis probe compared with SYBR-green, although a direct comparison of these assays was not possible (due to insufficient remaining sample volumes). The key, however, is likely to be the much longer storage time of samples collected for the qPCR validation. Across all stages, *T. pallidum* DNA was detected in 50% of contemporaneously-extracted CDTB whole blood samples compared with 28% of those stored for the qPCR validation study. A similar observation was made in a study of secondary syphilis where detection of *T. pallidum* DNA in stored whole blood was 30% compared with 63% for samples processed within a few hours of collection (151). The maximum interval between blood sampling and DNA extraction, which is likely to be short, remains to be established.

In contrast to DNA, Tempus™-preserved RNA in the current studies was detected in 50% of intact samples and at similar levels, regardless of length of storage, indicating that Tempus™ tubes can preserve *T. pallidum* RNA effectively. RNAlater® solution was also effective at preserving both DNA and RNA in ulcer samples that were shipped to the UK at ambient temperature (around 48 hours in total) and then stored at -80°C. On average, around 100-fold more RNA than DNA was detected in ulcer exudates compared with up to 10-fold more RNA than DNA in whole blood samples. Differences in assays and potential PCR inhibition aside, the most likely explanation for this difference is the metabolic activity of the treponemes sampled. During primary infection, *T. pallidum* are more virulent and have been shown to divide more quickly than those observed in latent disease (4). Moreover, studies of the disease trachoma have shown that active disease (with clinical signs) is associated with 100-fold greater detection of 16S rRNA from its causative organism, *C. trachomatis* (248). It follows, therefore, that once *T. pallidum* have invaded and are detectable in the blood they become less

metabolically active and the difference between quantifiable DNA and RNA falls to between two-fold and 10-fold.

Another observation was made possible by comparing RNA quantification in CDTB samples from patients with secondary and latent disease. Although the difference measured according to stage was significant, the spread of RNA levels measured far exceeded that of DNA and, in some ways, made RNA quantification a less useful measure of bacterial load. Although this may be explained by variation in metabolic activity, it likely reflects the greater inter-assay variance determined for 16S rRNA quantification as a result of differences in reverse transcription.

A final observation was made by comparing the qualitative detection of *tpp047* DNA with that of 16S rRNA from samples collected during both UK cross-sectional studies. In this analysis RNA appeared the better sample type identifying an additional two cases of primary disease and three cases of latent disease compared with DNA detection alone. Again, this is likely to be the result of more effective nucleic acid preservation within RNA Tempus™ tubes, especially for those samples collected during the validation study and stored for a number of years prior to extraction. It may also indicate that the best method for qualitative *T. pallidum* detection in blood should be based on 16S rRNA detection.

Syphilis and HIV-1 are known to interact in a number of ways. Both HIV-1 transmission and acquisition are facilitated by the presence of syphilitic ulcers and there is evidence that those ulcers and the clinical course of syphilis can be affected by HIV-1 induced immunosuppression (197). In a previous study, 214 patients with genital ulcers were recruited and HIV-1 positive men were found to be 11% more likely to have secondary syphilis with a concomitant primary chancre (209). Similarly, in the current study HIV-1 infected patients with secondary syphilis were more likely to present with unhealed primary ulcers. Moreover, HIV-1 positive patients had a higher number of organisms in primary ulcers than patients known to be HIV-1 uninfected. Although this association did not reach statistical significance, the higher bacterial load suggests HIV-1-positive patients with primary syphilis may be more infectious, helping to explain high rates of syphilis-HIV-1 co-infection.

It should also be noted that the cohort of HIV-1 infected patients included in the current studies were not significantly immunosuppressed, but had good median CD4 counts. The

clinical course of syphilis has also been demonstrated to change in the context of HIV-1 infection. Patients are more likely to present with secondary disease, bypassing the primary stage and some suggest that enhanced syphilis treatment is required in HIV-1 infected patients to reduce the risk of treatment failure (249, 250). It may be significant, therefore, that despite no observed difference in whole blood bacterial load between HIV-1 infected and uninfected patients, the majority of patients (85%) with secondary disease who donated samples to the CDTB were HIV-1 infected compared with 50% for both primary and latent.

The studies presented in this chapter demonstrate that *T. pallidum* qPCR has the potential to determine the kinetics of syphilis infection and the impact of a number of factors, including HIV-1 co-infection, on those kinetics. They demonstrate that both ulcer and blood samples contain quantifiable *T. pallidum* nucleic acids and that bacterial load varies according to stage of disease in a way that is biologically plausible. It is noted that some of the data presented are derived from small samples, which limited the statistical significance of associations. It is also appreciated that longitudinal sampling of the same patients during the course of their infection, as opposed to these cross-sectional data, would more accurately reflect infection kinetics but this would not be ethical. Importantly, these data indicate the potential for qPCR as a tool for monitoring treatment response in patients with early syphilis – the focus of Chapter 5.

Chapter 4

The kinetics of early Syphilis treatment *in vivo*

4.1 Introduction

4.1.1 Background

As long as 50 years ago, researchers were interested in *T. pallidum* persistence following standard syphilis treatment. In one study from that period, lymph nodes taken from two prisoners who had been previously treated for syphilis with penicillin were disturbingly found to contain viable *T. pallidum* by rabbit inoculation testing (RIT) (221). It is noteworthy that both of these patients had VDRL results that would be consistent with cure by today's standards. More recently, by combining PCR and RIT in an animal model it was demonstrated that viable treponemes could persist in untreated animals for more that 120 days (155). Of course, RIT and surgical lymph node removal are impractical for the routine diagnosis of cure, thus, current clinical practice relies on serological testing, however inefficient, as a marker of treatment efficacy (104). PCR has given us the capability to detect micro-organisms in clinical specimens without the need for culture (251). However, the process required for this detection is complex and dependent on, not only the efficiency of PCR, but also that of nucleic acid extraction and, in the case of RNA quantification, reverse transcription. Internal controls for these processes, often added to samples at the start of nucleic acid extraction, have become commonplace and are now commercially available (228).

This Chapter describes the progress of four patients with early syphilis following treatment of their disease with benzathine penicillin. Each patient is monitored with both standard serological tests and by qPCR detection of bacterial load in their blood and for one patient, ulcer exudate. Additionally, it presents the construction of an internal control for the *T. pallidum* bacterial load assay and an initial clinical validation of its utility.

Results

4.2.1 Assay Development

Assays for the quantification of *T. pallidum* DNA (targeting *tpp047*) were described in Chapter three. In order to control for differences in bacterial load as a result of inefficiencies in extraction, an internal control was designed. Initially, splice-overlap PCR was used to produce a 178bp section of *tpp047* containing an alternative probing sequence complementary to a new TaqMan probe that was labelled with a different fluorescent dye to the *T. pallidum* probe TPP. Both the mutated and original probe-site sequences were flanked with the primer-binding sites TP1 and TP2 so that both would be amplified in a multiplex qPCR. An alternative internal control was also assessed. This time, a section of a gene was taken from a tomato plant virus and amplified in a multiplex PCR containing two primer pairs and two differently labelled TaqMan probes (one set for the virus and the other for *T. pallidum*). Finally, control sequences were cloned into recombinant plasmids and, together with *T. pallidum* bacteria, were spiked into whole blood samples in order to assess the utility of this approach as an internal control *in vitro*.

4.2.1.1 Splice-overlap PCR to produce an internal control of tpp047 amplification

The method for site directed mutagenesis of the *T. pallidum tpp047* gene is described in Chapter 2.8. It proceeded with three separate rounds of PCR using six different primers to produce a mutated 178bp section of the gene that contained a 19bp sequence complementary to the TaqMan probe TPP-QC. As pictured in Figure 4.1, the first step was to PCR-amplify the *tpp047* gene with the primers TP3 and TP4. Two separate reactions produced amplicons that were 5'-tailed with the sense and anti-sense sequence of the TPP-QC binding site. During the final step, these two tails were allowed to anneal and the outer primers (TP3 and TP4) employed a second time to fill in the sequence either side. The region of interest, containing the new probe site, was then amplified with the *tpp047* qPCR primers TP1 and TP2, gel extracted and quantified ($4.44x10^{10}$ copies/µl) with spectrophotometry. Gene sequence analysis was then performed to confirm both the sequence and position of the TPP-QC probe site within *tpp047*.

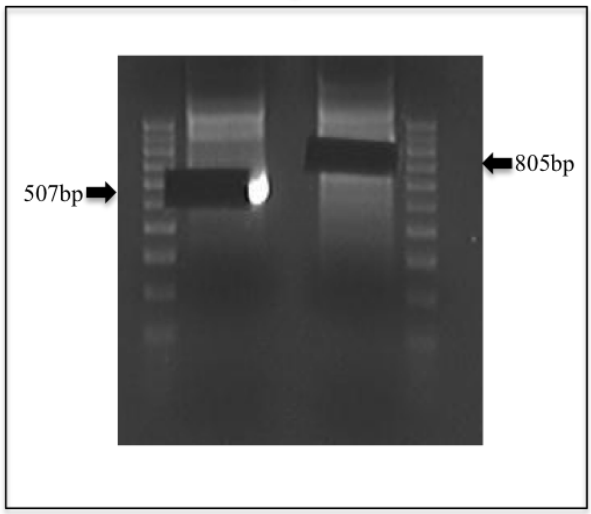
Once the sequences were confirmed, the specificity of each of the probes TPP and TPP-QC was assessed by qPCR. Two 1:10 dilution series from 10^6 to 10 copies of the TPP-QC amplicon and Nichols strain DNA (extracted from whole organisms) were constructed. All four series were then subject to qPCR with the amplification primers TP1 and TP2 and either the probe TPP or TPP-QC (Figure 4.2). Across the range of dilutions, only Nichols strain DNA amplification was detected by the probe TPP and TPP-QC DNA with the probe TPP-QC.

Figure 4.1 Production of an internal control for *T. pallidum* extraction and qPCR

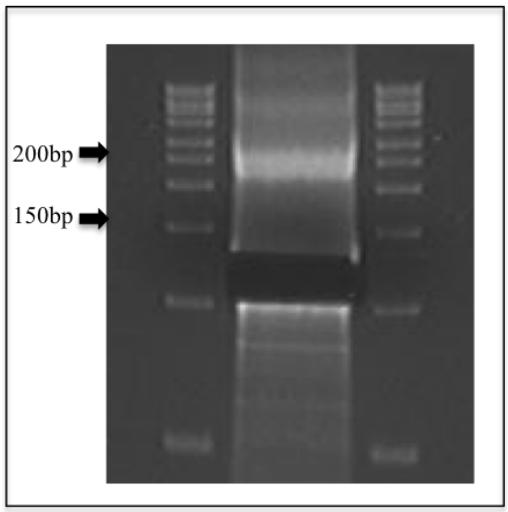
A. Primer Design

CTGCAGGGGGGGGGGGGGGG**CCCTCTATACGGAGGTGTAAT**CGTGAAAGTGAAATACGCACTACTTTCTGCCGGAGCGCTGC
AGTTGTTGGTTGTAGGCTGTGGCTCGTCTCATCATGAGACGCACTATGGCTATGCGACGCTAAGCTATGCGGACTACTGGGCCGGGGA
GTTGGGGCAGAGTAGGGACGTGCTTTTGGCGGGTAATGCCGAGGCGGACCGCGCGGGGGATCTCGACGCAGGCATGTTCGATGCAG
TTTCTCGCGCAACCCACGGGCATGGCGCGTTCCGTCAGCAATTTCAGTACGCGGTTGAGGTATTGGGCGAAAAGGTTCTCTCGAAGC
AGGAGACCGAAGACAGCAGGGGAAGAAAAAAGTGGGAGTACGAGACTGACCCAAGCGTTACTAAGATGGTGCGTGCCTCTGCGTC
ATTTCAGGATTTGGGAGAGGACGGGGAGATTAAGTTTGAAGCAGTCGAGGGTGCAGTAGCGTTGGCGGATCGCGCGAGTTCCTTCAT
GGTTGACAGCGAGGAATACAAGATTACGAACGTAAAGGTTCACGGTATGAAGTTTGTCCCAGTTGCGGTTCCTCATGAATTAAAAGG
GATTGCAAAGGAGAAGTTTCACTTCGTGGAAGACTCCCGCGTTACGGAGAATACCAACGGCCTTAAGACAATGCTCACTGAGGATAG
TTTTTCTGCACGTAAGGTAAGCAGCATGGAGAGCCCGCACGACCTTGTGGTAGACACGGTGGGTACCGTCTACCACAGCCGTTTTGG
TTCGGACGCAGAGGCTTCTGTGATGCTGAAAAGGGCTGATGGCTCTGAGCTGTCGCACCGTGAGTTCATCGACTATGTGATGAACTT
CAACACGGTCCGCTACGACTACTACGGTGATGACGCGAGCTACACCAATCTGATGGCGAGTTATGGCACCAAGCACTCTGCTGACTC
CTGGTGGAAGACAGGAAGAGTGCCCCGCATTTCGTGTGGTATCAACTATGGGTTCGATCGGTTTAAAGGTTCAGGGCCGGGATACTA
CAGGCTGACTTTGATTGCGAACGGGTATAGGGACGTAGTTGCTGATGTGCGCTTCCTTCCCAAGTACGAGGGGAACATCGATATTGGG
TTGAAGGGGAAGGTGCTGACCATAGGGGGCGCGGACGCGGAGACTCTGATGGATGCTGCAGTTGACGTGTTTGCCGATGGACAGCC
TAAGCTTGTCAGCGATCAAGCGGTGAGCTTGGGGCAGAATGTCCTCTCTGCGGATTTCACTCCCGGCACTGAGTACACGGTTGAGGT
TAGGTTCAAGGAATTCGGTTCTGTGCGTGCGAAGGTAGTGGCCCAGTAGAAGAGGGGTGTCCTATCCCGTGTGTCTTAA

B. PCR



C. Product extension



D. Sequence Analysis

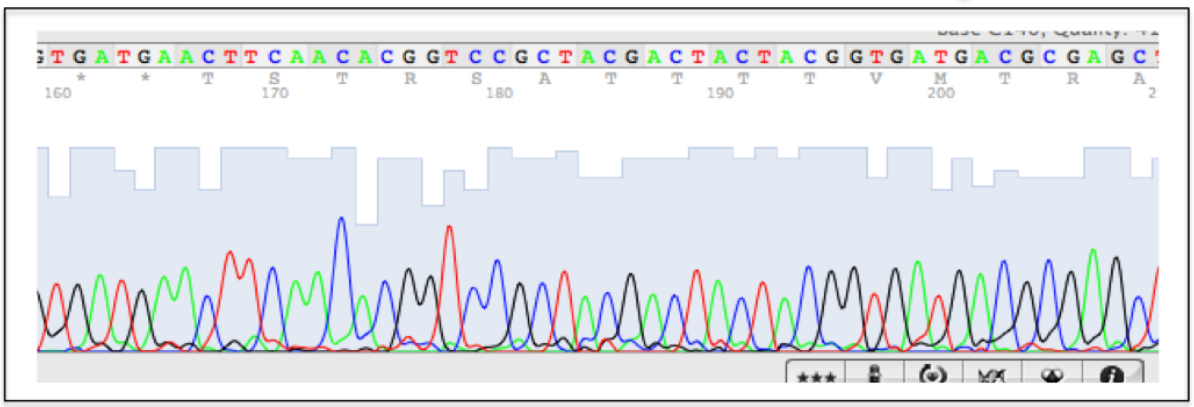
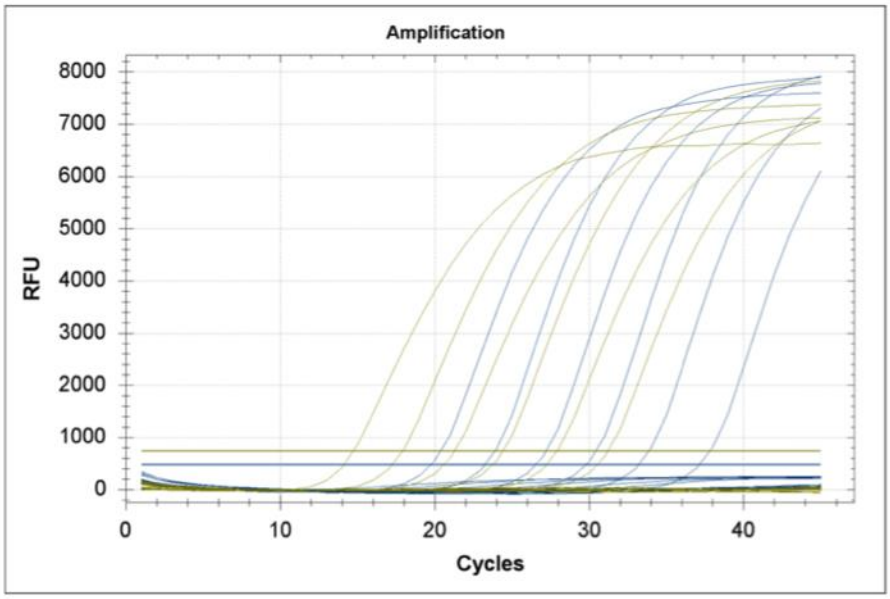


Fig 4.1. Panel A. DNA was extracted from Nichols strain *T. pallidum* organisms and primers **TP3** and TP4 were used to amplify a 1357bp section of the *tpp047* gene. **Panel B**. The purified product of step A was amplified in two separate PCRs with the primer pairs TP3 (outer sense) & TP5 (inner anti-sense) and TP4 (outer antisense) & TP6 (inner sense). One amplicon was 5'-tailed with the 19bp internal control sense sequence and the other with the anti-sense internal control sequence. **Panel C**. Equimolar concentrations of the two products from 'B' were annealed, then extended with the primers TP3&4 to produce a 1357 bp section of tpp047 in which the TPP probing site was replaced with TPQC. Next, the diagnostic primers TP1 and TP2, which flank the probing site, amplified a 178bp section which was gel-extracted and purified. **Panel D**. Gene sequence analysis of the product from step C using the primers TP1&2 confirmed the location and sequence of the alternative probing site. NB. In panels B and C, the gels were pictured after the desired DNA had been cut from them.

Figure 4.2 Assessment of TaqMan probe specificity

A. Amplification Curves



B. Detection cycle thresholds

	FAM (TPP) detection (PCR cycles)	**VIC (TPP-QC) detection (PCR cycles)**
***T. pallidum* (Nichols) DNA genome equivalents/reaction**		
$1.0x10^6$	19.72	-ve
$1.0x10^5$	23.24	-ve
$1.0x10^4$	26.53	-ve
$1.0x10^3$	29.79	-ve
$1.0x10^2$	33.19	-ve
10.0	37.15	-ve
NTC (water)	-ve	-ve
Purified TPP-QC amplicon copies/reaction		
$4.4x10^6$	-ve	14.72
$4.4x10^5$	-ve	17.92
$4.4x10^4$	-ve	21.26
$4.4x10^3$	-ve	24.39
$4.4x10^2$	-ve	27.91
44.0	-ve	31.20
NTC (water)	-ve	-ve

Fig. 4.2. Two 1:10 dilution series from 10^6 to 10 copies of both a TPP-QC amplicon and Nichols strain DNA (extracted from whole organisms) were constructed. **Panel A**. All four series were subject to qPCR with the amplification primers TP1 and TP2 and either the probe TPP for *T. pallidum* DNA detection (blue) or TPP-QC for internal control detection (green). **Panel B**. Cycle thresholds for PCR product detection.

4.2.1.2 Construction of the recombinant plasmid pTpQC and its concurrent detection with tpp047 DNA

Following the successful production and qPCR detection of the TPP-QC oligonucleotide, it was cloned into the vector pCR4-TOPO to produce the plasmid pTpQC and amplified in *E. coli* according to the method in Chapter 2.9. Following extraction, the plasmid was linearised with the restriction enzyme *ncoI* (pictured in Figure 4.3 A) for use as a *T. pallidum* DNA extraction control.

In order for pTpQC to be a reliable internal control it needed to be detected consistently in a multiplex qPCR without impacting on the detection of *tpp047*. Initially, 70 copies of pTpQC were spiked into a 1:10 dilution series of *T. pallidum* DNA extracted from whole organisms (Figure 4.3 panel B and Table 4.1). Cycle thresholds for the detection of *tpp047* at starting concentrations from 10^5 to 10 *T. pallidum* genome equivalents were similar both with and without the addition of pTpQC. However, whereas a single genome equivalent was detectable in the single-target qPCR, it was not detected in the multiplex. The detection of pTpQC showed a different pattern with no amplification when *T. pallidum* DNA concentrations were above 10^3 genome equivalents/reaction. Then, as the Nichols DNA concentration decreased, so did the cycle threshold for pTpQC detection, which fell from 33.64 to 32.08 cycles. Subsequently, two further concentrations of pTpQC (14 and 35 copies per reaction) were assessed and the results are compared in Table 4.1. Both concentrations were seen to impair the detection of *tpp047* in reactions containing DNA from between one and 10 *T. pallidum* genome equivalents. Moreover, pTpQC could not be detected at higher starting quantities of *T. pallidum* DNA. In summary, detection of *tpp047* was impaired at <10 *T. pallidum* genome equivalents at all concentrations of pTpQC and pTpQC detection was unreliable when DNA from more than 1000 *T. pallidum* organisms was present. The functional range of this assay, assuming that amplification of both targets was required, was thus between 10 and 1000 organisms.

The possibility that co-detection of the fluorescent dyes FAM and VIC were impairing the detection of one-another was explored by assessing an alternative TaqMan probe for pTpQC detection that was labeled with the dye ROX (TPP-QCR) (sequence and structure are given in Appendix 2). A dilution series of DNA extracted from 10^5 to one *T. pallidum* Nichols strain

organisms was quantified by qPCR using the probe TPP (FAM detection) and the primers TP1 and TP2 (Table 4.2). This was compared with the detection of either TPP (FAM) or TP-QC in a multiplex qPCR containing the same dilution series of *T. pallidum* plus 35 copies of the plasmid pTpQC using the primers TP1 and TP2 and the probes TPP (FAM) and TP-QCR. pTpQC could not be detected in reactions containing more than 100 genome equivalents of *T. pallidum* although *T. pallidum* detection appeared unimpaired by the co-amplification of pTpQC.

Figure 4.3 Construction and quantification of a recombinant plasmid

A. Plasmid linearisation **B.** Multiplex qPCR

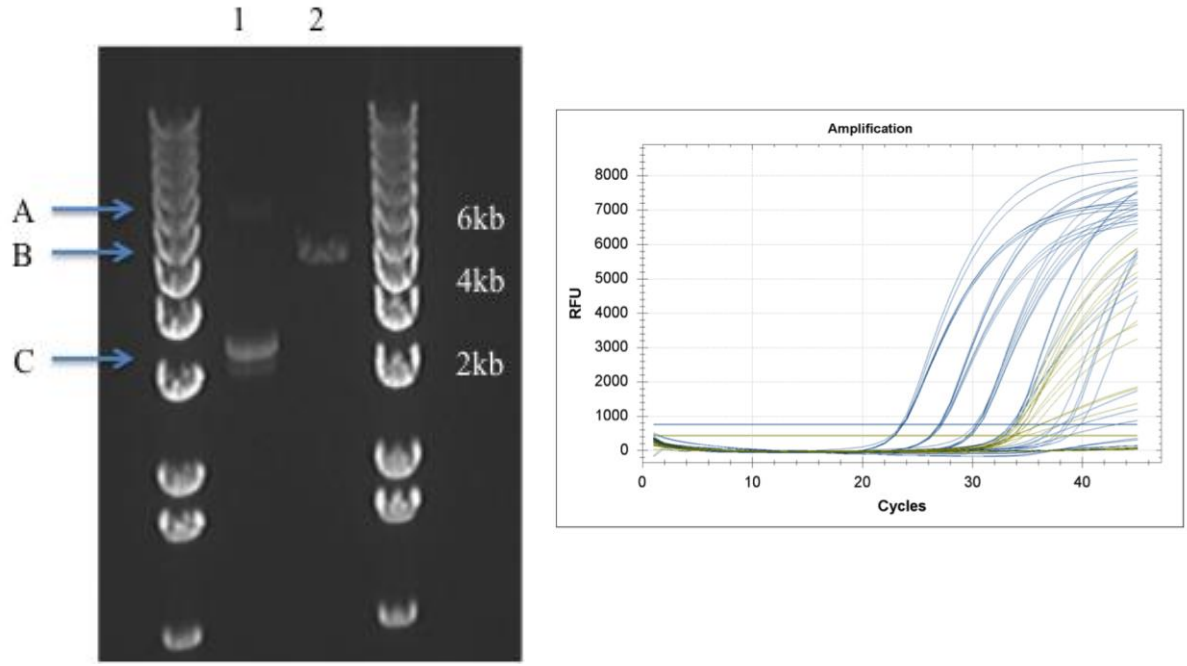


Fig 4.3. Panel A. *ncoI*-linearisation of the recombinant plasmid pTpQC. Lane 1: undigested plasmid separated into open conformation (A) and super-coiled (C) species. Lane 2: Linearised plasmid (B) (5314bp). **Panel B**. Nichols strain *T. pallidum* DNA spiked with 70 copies of internal control plasmid (pTp47) was amplified in a TaqMan qPCR with the primers TP1&TP2 and the probes TPP (*tpp047* detection) and TPPQC (pTpQC detection). Samples were assayed in quadruplicate and amplification is shown in blue (*tpp047*) and green (pTpQC).

Table 4.1 Multiplex qPCR of *T. pallidum* and an internal control

Reaction Conditions	14 pTpQC copies/reaction				35 pTpQC copies/reaction				70 pTpQC copies/reaction			
Copies of pTpQC/reaction	0	14	14	0	0	35	35	0	0	70	70	0
Probes added to reaction	TPP	TPP& TPP-QC	TPP& TPP-QC	TPP	TPP	TPP& TPP-QC	TPP& TPP-QC	TPP	TPP	TPP& TPP-QC	TPP& TPP-QC	TPP
Fluorescence detection	FAM	FAM	VIC	VIC	FAM	FAM	VIC	VIC	FAM	FAM	VIC	VIC
Nichols concentration (copies/reaction)												
$1x10^5$	23.89	24.07	neg	neg	23.49	23.82	neg	neg	23.49	23.36	neg	neg
$1x10^4$	27.40	27.34	neg	neg	26.68	27.16	neg	neg	26.69	26.52	neg	neg
1000	30.89	30.87	39.03	neg	30.19	30.72	neg	neg	30.23	30.32	33.64	neg
100	33.95	34.10	35.70	neg	33.84	33.98	37.43	neg	33.78	33.64	32.82	neg
10	37.72	36.95	35.13	neg	38.30	37.63	36.58	neg	37.25	37.14	32.52	neg
1	39.61	neg	35.31	neg	39.34	neg	37.26	neg	39.52	neg	32.08	neg

Table 4.1. Plasmid (pTpQC) was added to Nichols strain *T. pallidum* DNA in one of three concentrations: 14 copies (columns 3,4), 35 copies (columns 7,8) and 70 copies (columns 11,12). Cycle thresholds shown are the mean of two experiments and four technical replicates of each PCR were performed.

Table 4.2 Detection of an internal control with an alternative TaqMan probe

***T. pallidum* (Nichols) DNA/reaction (genome equivalents)**	**Mean* cycle threshold of detection**		
	Condition one	Condition two	Condition three
10^5	23.93	neg	23.92
10^4	27.12	neg	27.08
10^3	30.50	neg	30.57
10^2	33.78	43.34	34.27
10	37.03	37.03	36.96
1	41.33	37.46	39.80

Table 4.2. A 1:10 dilution series from 10^5 to a single genome equivalent of DNA extracted from Nichols strain *T. pallidum* were added to 35 copies of the internal control plasmid pTpQC and quantified in a multiplex qPCR. All reactions included the *T. pallidum* detection probe, TPP. In addition, reactions in condition one included the FAM-labelled TaqMan probe (TPP-QC) for pTpQC detection and those in condition two included an alternative ROX-labelled TaqMan probe (TP-QCR). Condition three was a control including only the probe TPP and to which pTpQC was not added.

* Denotes the average of three replicas in a single experiment.

4.2.1.3 Assessment of an alternative internal control of T. pallidum amplification

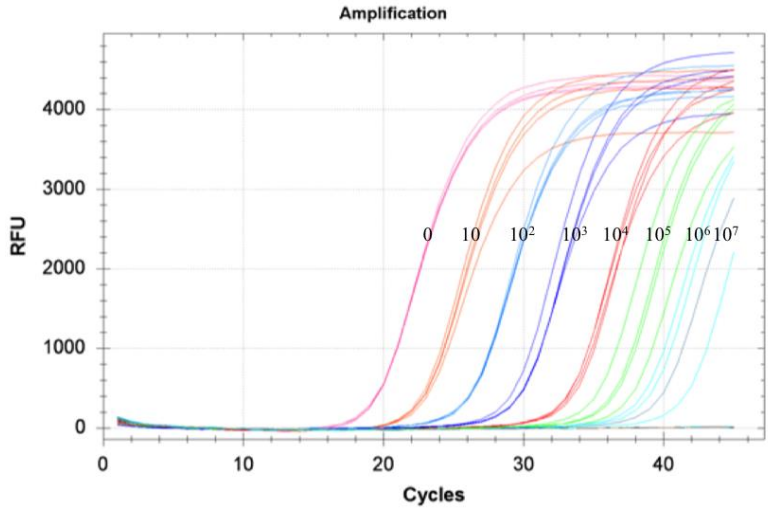
Whilst an analytical sensitivity of 10 *T. pallidum* organisms in the multiplex assay was considered acceptable, failure to detect the internal control pTpQC when over 1000 organisms were present was unacceptable. A new assay was required with separate amplification primers for the internal control sequence and *tpp047* to avoid competition in the reaction. In order to avoid cross-reactivity with any human or *T. pallidum* gene, a sequence from the plant germinivirus TYLCV gene was selected as the alternative internal control (Chapter 2.1.4). The design of the new assay settled on the previously described amplification primers TYLCV_F3 and TYLCV_R3 (226) and a new Yakima Yellow-labelled TaqMan probe (P_TYLCV) (Appendix 2). A 100bp section of the TYLCV genome was amplified with the primers TYLCV_F3 and TYLVC_R3 and cloned into the pCR-4TOPO plasmid vector (Chapter 2.9). Following amplification in transformed *E. coli*, the new plasmid (pTYLCV) was linearised with the restriction enzyme *ncoI* and quantified, firstly by spectrophotometry and then by limiting dilution qPCR (Figure 4.4). Next, the multiplex detection of *tpp047* and pTYLCV was assessed taking note of the following observations:

- the maximum bacterial load seen in whole blood during the cross-sectional studies (Chapter 3) was 4770 *tpp047* copies/ml.
- low-level *tpp047* detection was impaired by the co-amplification of pTpQC.
- high starting *tpp047* concentrations impaired the detection of pTpQC.

DNA extracted from 100 Nichols strain *T. pallidum* was spiked with 10, 50, 100, 500 or 1000 copies of pTYLCV and both targets detected simultaneously in the multiplex qPCR (Table 4.3 A). It was established that a minimum of 100 pTYLCV copies were needed in order to be detected reliably and that *tpp047* detection of 100 *T. pallidum* organisms was not affected by the addition of up to 1000 copies of pTYLCV. This observation was confirmed in a second experiment where the detection of both 100 and 10 *T. pallidum* genome equivalents was not impacted by the addition of 500 pTYLCV copies (Table 4.3 B).

Figure 4.4 Limiting dilution and quantification of an alternative internal control plasmid

A. Constructing a standard curve



B. Limiting dilution series

Starting concentration (pTYLCV copies/reaction)	Number of replicas positive	Mean* (std. dev) cycle threshold
10^6	+ + + +	18.62 (0.024)
10^5	+ + + +	21.92 (0.098)
10^4	+ + + +	25.38 (0.079)
10^3	+ + + +	28.72 (0.247)
10^2	+ + + +	32.36 (0.189)
10	+ + + +	35.26 (0.965)
1	+ + + -	38.66 (2.308)
0.1	+ - - -	**38.19

Fig 4.4. A. Following linearisation with the restriction enzyme *ncoI*, the plasmid pTYLCV was quantified by spectrophotometry and diluted to a concentration of $1x10^6$ copies/µl. A 1:10 dilution series was then made (dilution factors displayed) and quantified by qPCR with the primers TYLCV_F3 and TYLCV_R3 and the probe P_TYLCV. **B**. Mean cycle thresholds and the number of PCR-positive replicas at each dilution.

* The mean of four replicas in a single experiment ** Only a single replica was positive at this dilution.

Table 4.3 Multiplex PCR detection of *T. pallidum* and pTYLCV

A.

	pTYLCV copies added to each reaction					
	1,000	**500**	**100**	**50**	**10**	**(0) control**
***T. pallidum* detection**						
mean cycle threshold	29.87	29.96	30.05	30.20	30.09	30.04
std. dev.	0.142	0.160	0.031	0.057	0.102	0.030
TYLCV detection						
mean cycle threshold	33.51	33.22	35.77	35.79*	neg	neg
std. dev.	1.993	0.920	0.101	n/a		

B.

Nichols strain				
***T. pallidum* DNA (copies/reaction)**	**100**	**100**	**10**	**10**
pTYLCV (copies/reaction)	**500**	**0**	**500**	**0**
***T. pallidum* detection**				
mean cycle threshold*	30.05	29.90	33.08	33.59
std. dev.	0.137	0.231	0.412	0.332
TYLCV detection				
mean cycle threshold*	33.51	neg	**34.74	neg
std. dev.	1.993	n/a	n/a	n/a

Table 4.3. *T. pallidum* and TYLCV DNA were co-quantified in a multiplex qPCR.

A. Nichols strain *T. pallidum* DNA (100 genome equivalents) were added to either 1,000, 500, 100, 50, 10 or 0 copies of pTYLCV. Cycle thresholds are displayed for both the detection of *T. pallidum* and TYLCV.

*only a single technical replica was positive ** two of three of replicas were positive.

B. pTYLCV (500 copies) was added to either 100 or 10 *T. pallidum* genome equivalents and both targets quantified. Additionally, the potential impact of pTYLCV amplification on the ability to detect *T. pallidum* was assessed by including two no-pTYLCV controls.

* The mean of two technical PCR replicates each for two separate samples ** Only a single replica PCR-positive.

It was predicted that the process of DNA extraction from blood was not 100% efficient. It was, therefore, necessary to assess the amount of pTYLCV that was to be added to whole blood samples in order to be reliably detected by qPCR. Aliquots of 500µl of whole blood were spiked with 5000 *T. pallidum* genome equivalents and either 500, 1000 or 5000 copies of pTYLCV followed by DNA extraction (Chapter 2.3.1) and multiplex qPCR. Table 4.4 A demonstrates that starting concentrations of pTYLCV of less than 5000 copies/reaction were undetectable, and that while cycle thresholds for *tpp047* detection were variable, differences were not significant. A second experiment was conducted with the same conditions to confirm that the recovery and detection 5000 pTYLCV copies was reproducible. Four separate blood samples were spiked with 5000 *T. pallidum* organisms, two were then spiked again with 5000 copies of pTYLCV. The mean cycle threshold for pTYLCV detection was 38.18 cycles, which was significantly different from the previous experiment (Table 4.4 B). However, the cycle thresholds for the two replicas in this experiment were similar (std. dev. 0.432).

The internal control selected for the quantification of *tpp047* DNA was 5000 copies of pTYLCV. The final method is described in Chapter 2.12.1.3.

Table 4.4 Co-extraction of *T. pallidum* and an internal control

A.

Nichols strain *T. pallidum* (genome equivalents/reaction)	**5,000**	**5,000**	**5,000**	**5,000**	**0**
pTYLCV (copies/reaction)	**5,000**	**1,000**	**500**	**0**	**0**
***T. pallidum* detection**					
mean cycle threshold*	36.25	37.03	37.87	35.75	neg
std. dev	0.344	1.017	0.342	0.354	/a
TYLCV detection					
mean cycle threshold*	40.08	neg	neg	neg	neg
std. dev	1.159	n/a	n/a	n/a	n/a

B.

Nichols strain *T. pallidum* (genome equivalents/reaction)	**5000**	**5000**	**0**	**LB only**
pTYLCV (copies/reaction)	**5000**	**0**	**0**	**LB only**
***T. pallidum* detection**				
*mean** cycle threshold*	36.80	35.75	neg	neg
std. dev	1.19	0.354	n/a	n/a
TYLCV detection				
*mean** cycle threshold*	38.18	neg	neg	neg
std. dev	0.432	n/a	n/a	n/a

Table 4.4. A. Whole blood collected into EDTA (500µl) was spiked with 5,000 *T. pallidum* bacteria and either 5,000, 1,000, 500 or 0 copies of pTYLCV. Following DNA extraction, both *T. pallidum* (*tpp047*) and TYLCV were quantified in a multiplex qPCR (three technical replicates of each sample). A negative control of unspiked blood was also included.

* The mean of three PCR technical replicas for a single spiked sample. **B**. Whole blood (500µl) was spiked with 5,000 *T. pallidum* genome equivalents and with 5,000 or 0 copies of pTYLCV. DNA was extracted and both targets amplified in a multiplex qPCR (three technical replicates of each sample). Two negative controls were included, the first was un-spiked blood and the second was lysis buffer.

* The mean of three PCR technical replicas for a single spiked sample. **DNA was extracted from two separate spiked whole blood samples.

4.2.2 Longitudinal clinical study of *T. pallidum* clearance in four patients with early syphilis

As described in Chapter 2.14.4, regulatory and ethical approvals were obtained to recruit four patients with early syphilis to a longitudinal study that aimed to monitor the clearance of *T. pallidum* from blood and, where present, ulcer exudates following standard treatment. Briefly, patients were required to have a microbiologically confirmed (DGM or serology) diagnosis of primary or secondary syphilis; to have been prescribed benzathine penicillin as treatment and to have not received anti-treponemal antibiotics in the three months prior to study recruitment.

4.2.2.1 Patient recruitment and characteristics

A total of four patients, three with secondary and one with primary disease, were recruited to the study. All four patients were homosexual men and three were HIV-1 co-infected. The presentations and clinical course of each of the subjects are as follows and standard serological results for each patient are presented in Table 4.5.

Case one (STS1)

A 48 year-old white British and HIV-1 infected homosexual man presented with a two week history of rash. His HIV-1 infection was well controlled on combination anti-retroviral therapy (Lopinavir/Ritonavir/Tenofovir) and his latest CD4 count (one month prior to syphilis diagnosis) was 640 cells/mm^3. On examination, the rash comprised discrete erythematous macules over his trunk, and more diffusely involved his face, arms and legs. The palms of his hands were affected, but the soles of his feet were not. An area of dry healing skin was noted on glans penis, which he described as being an ulcer when it had first appeared four weeks previously. Examination was otherwise normal, with no apparent signs of central nervous system involvement. Syphilis serology had been negative one month prior to presentation, but was now positive (EIA positive, TPPA positive, RPR 1:64). He was diagnosed with secondary syphilis and his physician selected three once-weekly injections of intra-muscularly (IM) administered benzathine penicillin (2.4 megaunits) as treatment. The patient described feverish symptoms during the night after his treatment (at around eight hours post-dose) and his rash was noticed to deepen at 24 hours, before starting to improve at day 5.

Case two (STS2)

A 45 year-old homosexual white Brazilian man presented with a two day history of rash. He was HIV-1 infected, but well controlled on combination anti-retroviral therapy (Tenofovir/Emtricitabine/Efavirenz) with a CD4 count of 776 cells/mm^3 (measured two weeks prior to study enrolment). On examination, the patient had a feint, erythematous, macular rash over his trunk, arms and legs that spared his face, palms and soles (Figure 4.5 A). Examination was otherwise unremarkable and he gave no history of recent mucosal (oral, genital or peri-anal) ulcer. The patient described the rash as identical to one present during a previous diagnosis of syphilis three years before. Serological testing revealed a rise in RPR titre from negative to 1:32 and the patient was diagnosed with secondary syphilis. His physician selected treatment with a single dose of IM benzathine penicillin (2.4 megaunits). Approximately 16 hours after treatment the patient experienced feverish symptoms, which he did not report at the time and, thus, temperature was not recorded. His rash was noticed to deepen over 48-72 hours (Figure 4.5 B) before beginning to improve at day 5.

Case three (STS3)

A 26 year-old black British homosexual man presented with a four-week history of a painless penile ulcer. He had initially presented to a different genito-urinary medicine service, where blood for syphilis and HIV serology had been taken. He had been contacted and informed of a negative HIV-1 antibody/antigen test, but positive syphilis serology (EIA positive, RPR 1:32) and attended for treatment. He had no previous history of syphilis. On examination he had a firm, painless penile ulcer on the coronal sulcus anteriorly measuring 14x8mm at its widest points, which was associated with bilateral inguinal lymphadenopathy. Physical examination was otherwise unremarkable with no signs of secondary syphilis. A diagnosis of primary syphilis was confirmed with DGM and treatment with a single dose of IM benzathine penicillin (2.4 megaunits) was selected. Six hours following treatment, the patient felt unwell and was found to be pyrexial at 38.9°C. Physical examination revealed a tachycardia at 110 beats/minute and blood pressure of 145/85. No further physical signs of infection, other than the unchanged penile chancre were elicited. The patient continued to spike fevers over the following 24 hours. Blood chemistry revealed a C-reactive protein of 37 IU/ml at 6 hours,

which rose to 87 IU/ml at 30 hours and two sets of blood cultures were negative. At 48 hours, the patient was fully recovered.

Figure 4.5 Clinical presentations.

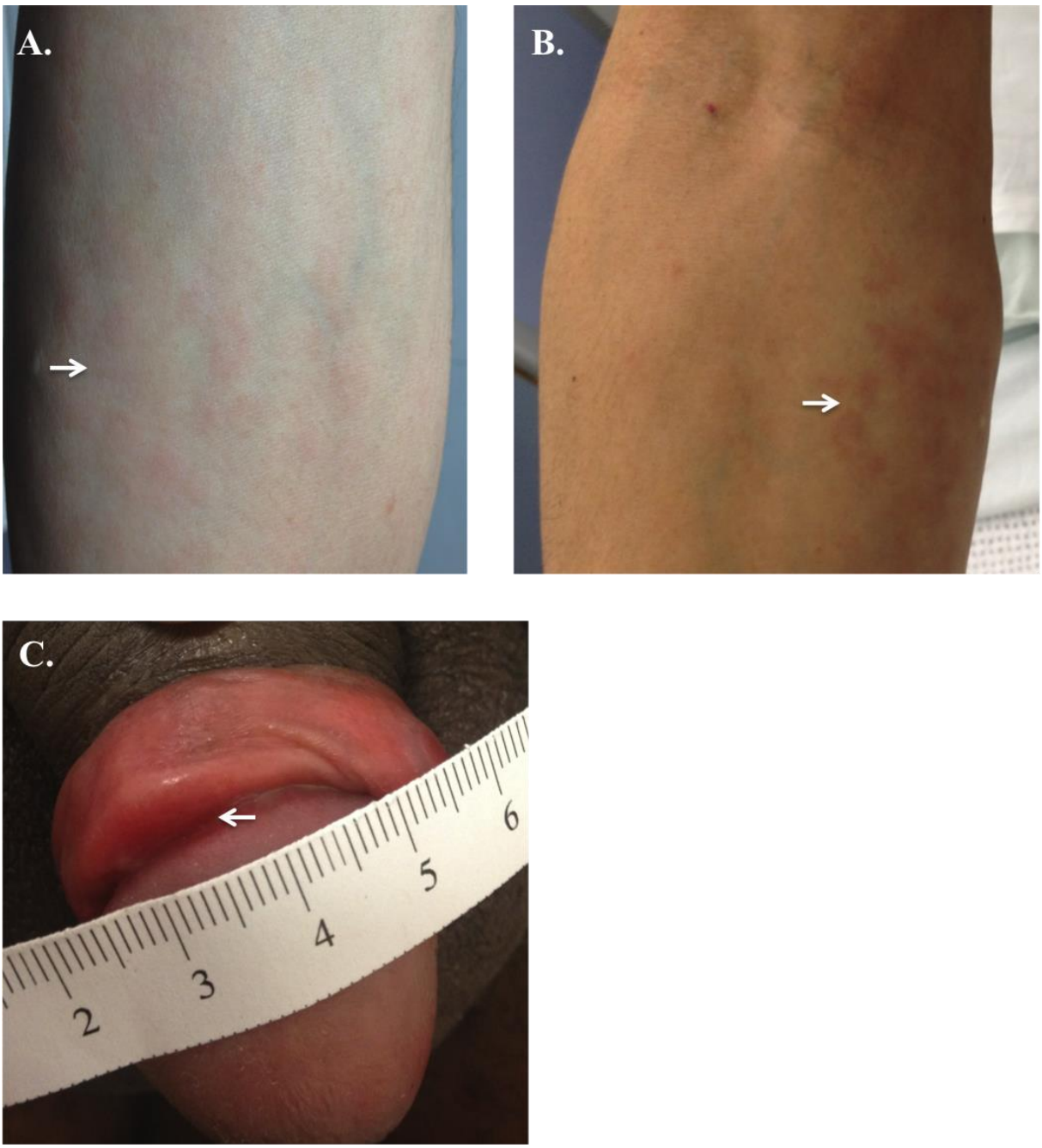


Fig 4.5. Panel A. Right forearm of patient STS2 at the time of presentation. **Panel B**. Right forearm of patient STS2 taken 48 hours after administration of benzathine penicillin. **Panel C**. Penile chancre of patient STS3 photographed at the time of presentation.

Case four (STS4)

A 35-year-old white Brazilian homosexual man was recalled to the HIV clinic for the investigation of an RPR of 1:64 and a rise in the liver enzyme alanine aminotransferase (ALT) to 89IU/L (upper limit of normal 40 IU/L). Syphilis serology six months prior had revealed a positive EIA, but a negative RPR. The patient was known to be HIV-1 infected and was stable on combination therapy (Darunavir/Ritonavir/Tenofovir/Emtricitabine) with an undetectable HIV-1 viral load (<50 copies/ml) and a CD4 count of 820 cells/mm^3. The patient described no rash or mucosal ulcers over the last few weeks, but mentioned that his regular sexual partner had developed a strange rash one month previously. He did complain of a headache on and off for the previous 'few' weeks, which was like a migraine, but no other neurological symptoms. On examination, he was noted to have a diffuse erythematous macular rash over his back and chest (pictured in Figure 4.6 A&B). The patient was uncertain how long this had been present. He was also noted to have a 5mm diameter red macule on glans penis (Figure 4.6 D), which had the appearance of a healing ulcer. The patient had noticed 'something' but had felt it to be traumatic in origin. Neurological and cardiovascular examinations were normal. It was felt that the patient had secondary syphilis with potential neurological involvement. He underwent a CT brain scan, which was normal, and a lumbar puncture was performed. CSF microscopy revealed 5 white blood cells per mm^3 (no differential given) and no red blood cells together with a protein of 0.47g/L (range 0.01-0.40) and a glucose of 3.1mmol/L. Bacterial culture and a CSF syphilis RPR were negative. A single dose of IM benzathine penicillin (2.4 megaunits) was selected as treatment. The patient was then transferred to the inpatient study unit for monitoring and sample collection. The patient described no symptoms of a JH reaction and no pyrexia was recorded. Any change in his rash (Figure 4.6 C) following treatment was modest.

Figure 4.6 Clinical presentation

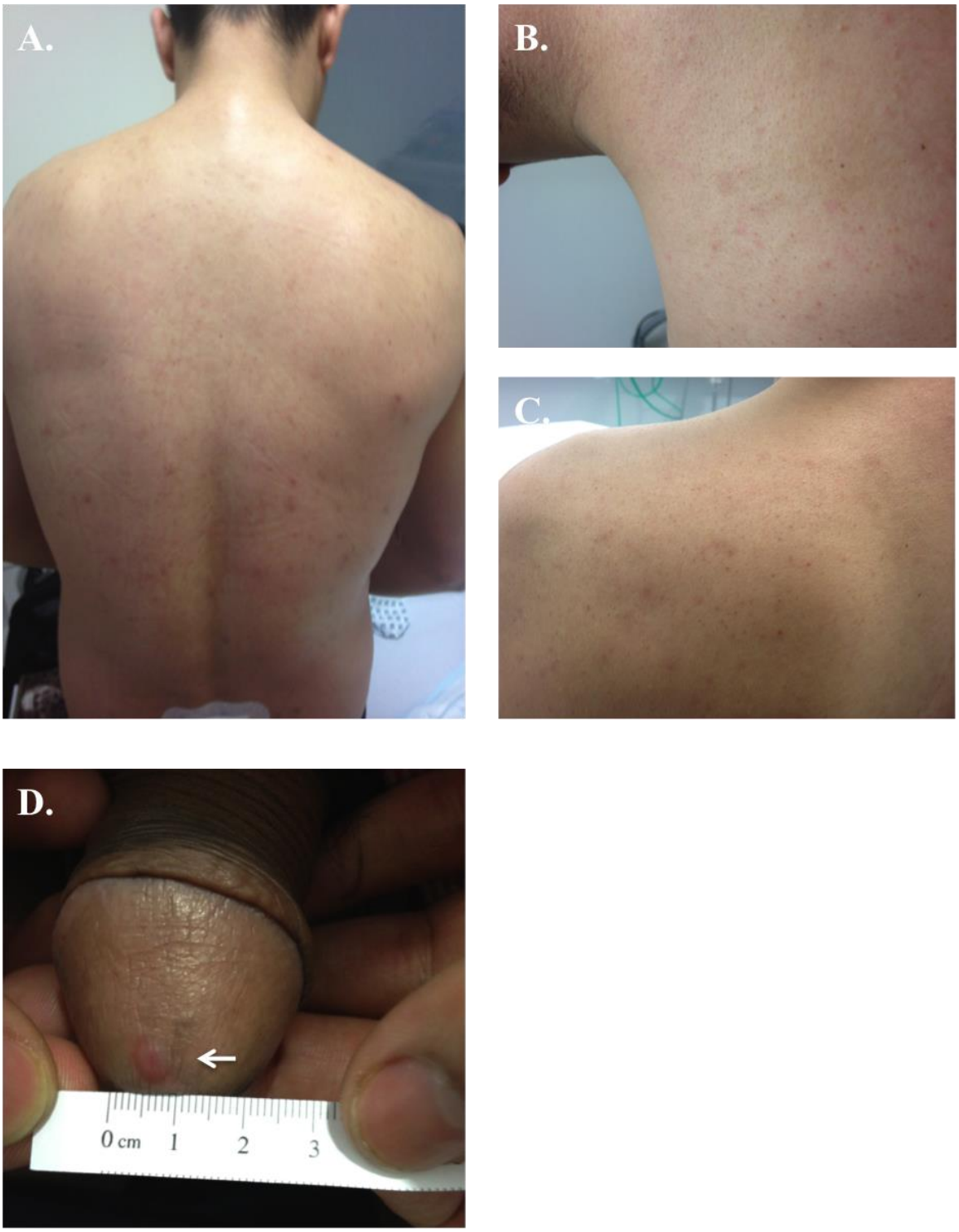


Fig 4.6. Clinical presentation of patient STS4. **Panel A**. shows the rash on patients back at time of presentation and a close-up is show in **Panel B.**. **Panel C.** depicts a close-up of same shoulder at 48 hours after treatment. **Panel D**. shows a healed ulcer present on glans penis at the time of presentation.

Table 4.5 Serological follow-up

	Rapid Plasma Reagin titre			
	Baseline	**One month**	**Three months**	**Six months**
Case 1	1:64	1:4	negative	negative
Case 2	1:32	1:4	1:1	1:4
Case 3	1:32	1:2	DNA*	negative
Case 4	1:64	unavailable	1:4	1:2

Table 4.5. Following the end of the study, patients were followed up with syphilis serological tests according to local guidelines. RPR titres, a measure of disease activity, were determined at one, three, six and 12 months following treatment. A two-dilution (four-fold) decrease in RPR titre confirmed treatment success.

*Result is missing as the patient did not attend for phlebotomy.

4.2.2.2 Patient management and sample handling

Following informed written consent, initial blood (and ulcer) samples and medical photographs were taken and patients then received a single dose of 2.4 mega-units of benzathine penicillin by deep intramuscular injection. Following transfer to an inpatient study facility, an intravenous cannula was inserted and blood samples (1x Tempus™ blood tube and 1ml EDTA whole blood) were collected according to the sample schedule in Appendix 7. Samples were immediately placed on dry ice prior to transport to the laboratory. Patient three (STS3) had a sample of penile ulcer exudate absorbed onto a snostrip at each sampling time point. These were placed into cryovials containing 1.3ml RNAlater solution and then stored on dry ice. *T. pallidum* DNA extraction and quantification proceeded in batches of up to eight samples. A positive (5000 Nichols organism-spiked blood or snostrip) and negative (2x lysis buffer) control were included with each batch. DNA was extracted from 500µl whole blood samples without the addition of 5000 pTYLCV copies. Following the end of sample collection from the fourth and final patient, a second 500µl aliquot of EDTA whole blood from each patient time-point was spiked with 5000 pTYLCV copies and DNA extracted according to the same method as unspiked samples.

Samples were taken from patient one (STS1) for the entire sampling schedule (a total of 22 samples over 150 hours). It was noted that once undetectable in two consecutive whole blood samples, *T. pallidum* qPCR remained negative in all subsequent samples. Allowing for sample batching, blood collection from patients two, three and four ended once a minimum of two consecutive blood samples were *tpp047* negative.

RNA from Tempus™ tube-collected blood samples was also extracted in batches of up to eight samples with a positive (syphilis seronegative blood in a Tempus™ tube spiked with 5000 *T. pallidum*) and a negative (PBS) control. Tempus™ tube-collected sample extraction, and likewise DNA/RNA co-extraction from ulcer samples was performed after sample collection from all patients was complete. All qPCR (*tpp047* quantification) and RT-qPCR (*16S rRNA* quantification) assays included three technical replicates of each sample. Insufficient whole blood and ulcer exudate were available to perform more than one biological replicate for each assay at each time point.

4.2.2.3 Kinetics of T. pallidum clearance from whole blood

All four of the patients recruited to the study had detectable *T. pallidum* bacteremia and it was possible to document a fall in the bacterial load, measured by both DNA and RNA, over time. The fourth patient recruited, STS4, had only low-level traces of *T. pallidum* nucleic acids in his blood. Indeed the first two time-points (pre-treatment and 2 hours post-treatment) were negative for both. *tpp047* was subsequently detected at hours 10, 18, 26 and 30 (maximum was 49 copies/ml at 18 hours) and 16S rRNA at hours 18 and 30 only (maximum was 237 copies/ml at 30 hours) (Figures 4.7 D and 4.8 D). It must be noted that Tempus™ samples from times one to six (0-14 hours) were spoiled, thus 16S rRNA may have been detectable prior to 18 hours. Both DNA and RNA were undetectable at hours 34 and 38, after which sample collection ended.

Samples from patients STS1, STS2 and STS3 showed a different pattern (Figures 4.7 A,B&C and 4.8 A,B&C). The average peak *tpp047* level was 1626 copies/ml (std. dev. 652), which did not occur pre-treatment, rather two to 10 hours post-treatment. Peak 16S rRNA levels (8879 copies/ml, std. dev. 11109) were detected at similar time-points (Table 4.6) and while the peak of *tpp047* DNA was similar for all three patients, that of 16S rRNA for patient STS3 was a log higher than for STS1 and STS2. Following peak bacteraemia, the level of both RNA and DNA was then seen to fall quickly with neither detectable in any patient after 56 hours. The decline in DNA and RNA levels for patients STS1 and STS2 is seen to oscillate rather than to fall smoothly as is the case for STS3. The magnitude and frequency of these oscillations appears greater for DNA than for RNA.

The time elapsed from the administration of treatment to halfway between the first of the consecutively negative samples and last detectable sample was referred to as ‘time to clearance’ (Table 4.6). For samples STS1, STS2 and STS3, the mean time to *tpp047* DNA clearance was 34 hours (std. dev. 8) and for 16S rRNA is was 29 hours (std. dev. 23.86). The greater variability of time to 16S rRNA clearance is explained by the 56-hour clearance time measured for STS3, which was the result of low-level (35 copies/ml) 16S rRNA detection at 56 hours following two negative samples at 42 and 46 hours. This 16S rRNA amplification at 56 hours was observed in two of three technical replicates during two separate experiments. Furthermore, both the negative Tempus™ extraction control and the two no-template PCR

negative controls were negative. In order to account for both the difference in time of peak bacteraemia and for outliers, clearance half-lives were calculated. Half-life is defined as the time it takes for a substance undergoing decay to decrease to half its initial quantity and usually describes exponential decay. Taken from the peak in bacterial load onwards, the plots of mean 16S rRNA and *tpp047* levels in blood and ulcers (Panel E in Figures 4.7 and 4.8) appeared to fall exponentially over time. In order to determine whether this relationship was truly exponential, regression lines for linear, quadratic, logarithmic and exponential relationships were fitted to the data (Figure 4.9). By measuring the coefficient of variation (R^2) for each regression line (where values of zero and one indicate no fit and perfect fit respectively) the decay of RNA and DNA in both ulcer and blood samples were deemed to be exponential, thus half-life was an appropriate measure of clearance. The calculation of half-life used was based on the exponential decay equation:

$$N(t) = N_0 \left(\frac{1}{2}\right)^{t/t_{1/2}}$$

where:

N(t) = the quantity that still remains and has not yet decayed after a time t

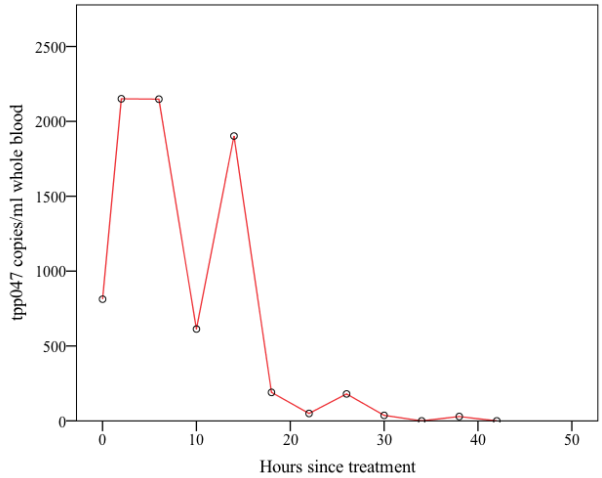
N_0 = the initial quantity of the substance that will decay

$t^{1/2}$ = the half-life of the decaying quantity

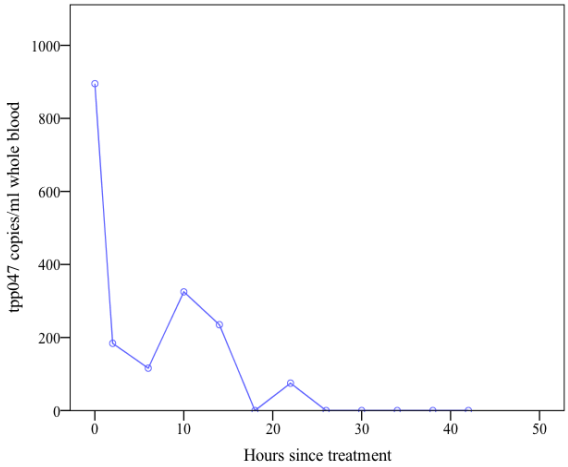
The mean half-life of *T. pallidum tpp047* DNA clearance from blood was determined to be 5.68 hours (std. dev 0.53) and was significantly longer than the 3.89 hours (std. dev. 0.84) calculated for 16S rRNA (p=0.035).

Figure 4.7 Clearance of *T. pallidum* DNA from blood

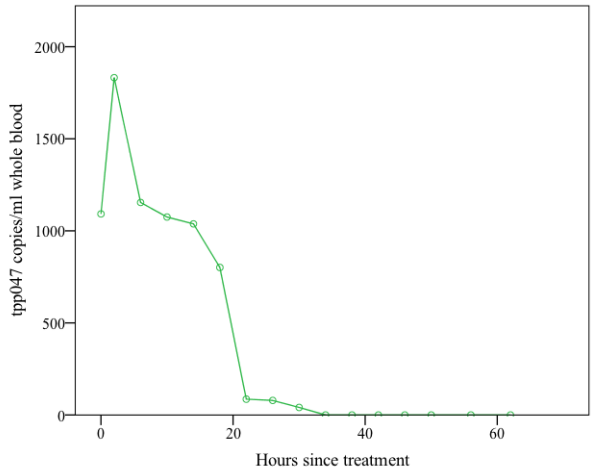
A. STS1



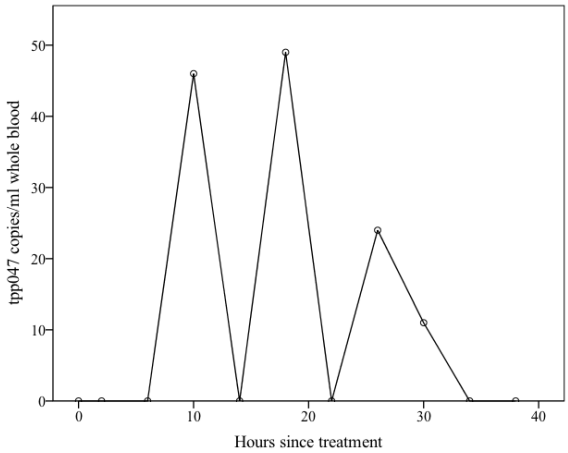
B. STS2



C. STS3



D. STS 4



E.

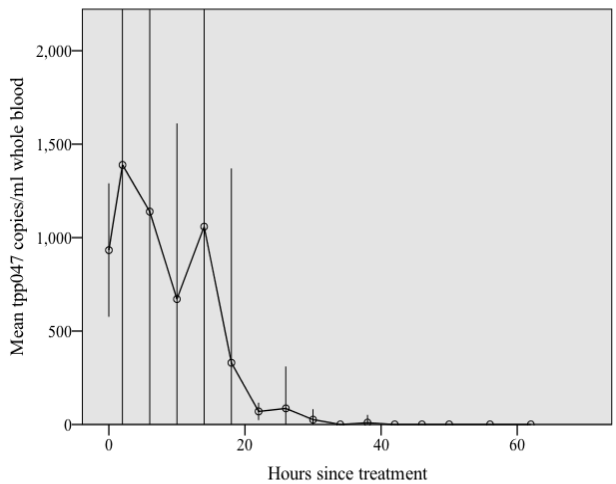
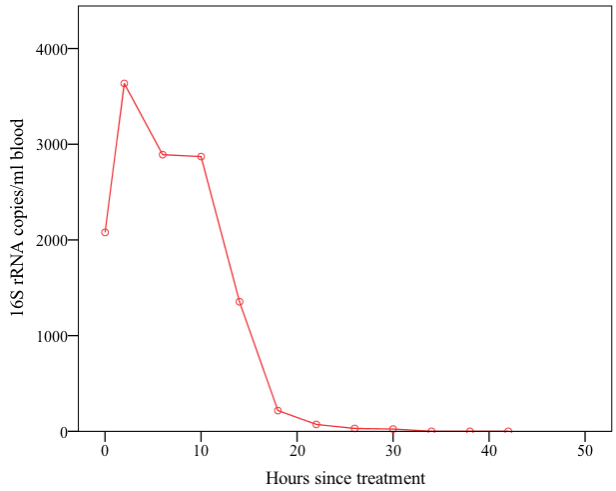


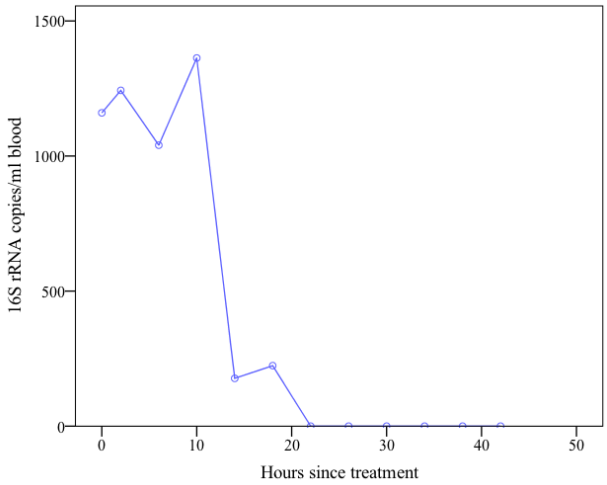
Fig 4.7. Four patients with microbiologically confirmed early syphilis were treated with a 2.4 megaunit dose of benzathine penicillin at time zero. Blood samples were collected into EDTA immediately pre-treatment, at two hours post-treatment and then four hourly for up to 50 hours. *T. pallidum* DNA was extracted from whole blood and quantified by qPCR (*tpp047* gene). Each of the **Panels A** to **D** represents a different patient and **Panel E** shows the average of *tpp047* quantification for patients STS1, STS2 and STS3. Error bars depict 95% confidence intervals

Figure 4.8 Clearance of *T. pallidum* RNA from blood

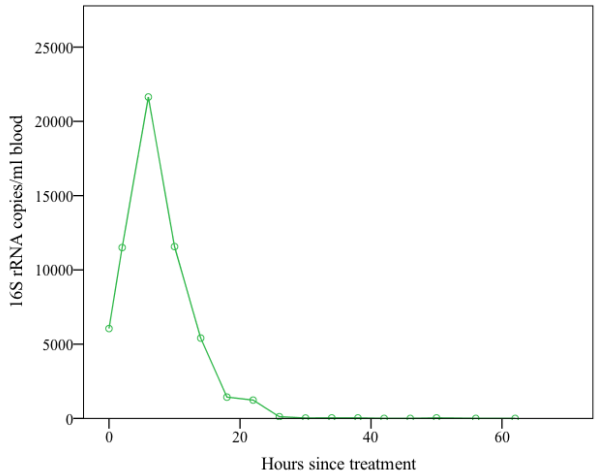
A. STS1



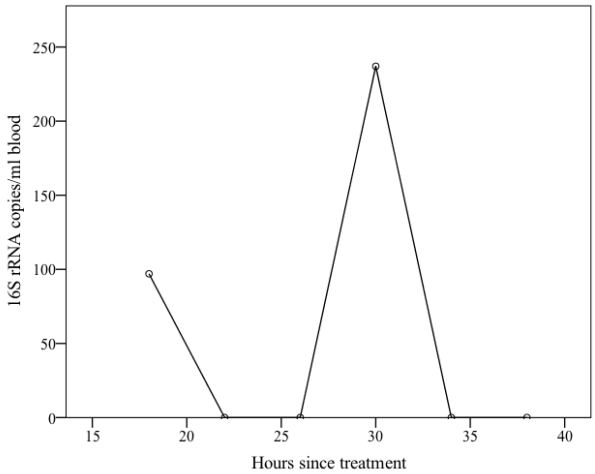
B. STS2



C. STS3



D. STS4



E.

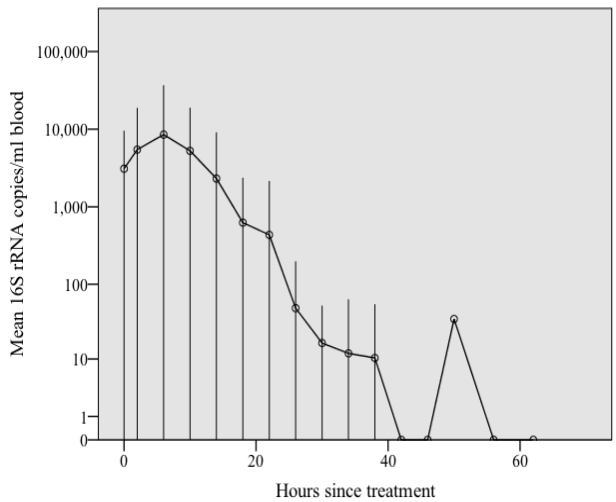
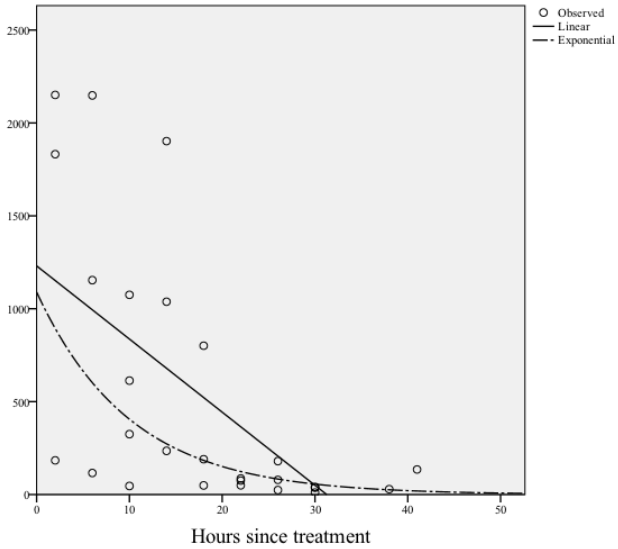


Fig. 4.8. Four patients with microbiologically confirmed early syphilis were treated with a 2.4 megaunit dose of benzathine penicillin at time zero. Blood samples were collected into RNA preservation tubes immediately pre-treatment, at two hours post-treatment and then four hourly for up to 50 hours. *T. pallidum* RNA was extracted from whole blood and quantified by RT-qPCR (*16S rRNA* target). Each of the **Panels A to D** represents a different patient and **Panel E** shows the average of *16S rRNA* quantification for patients STS1, STS2 and STS3. The first six time-points for STS4 are missing as the Tempus™ samples were spoiled during RNA extraction. Error bars represent 95% confidence intervals.

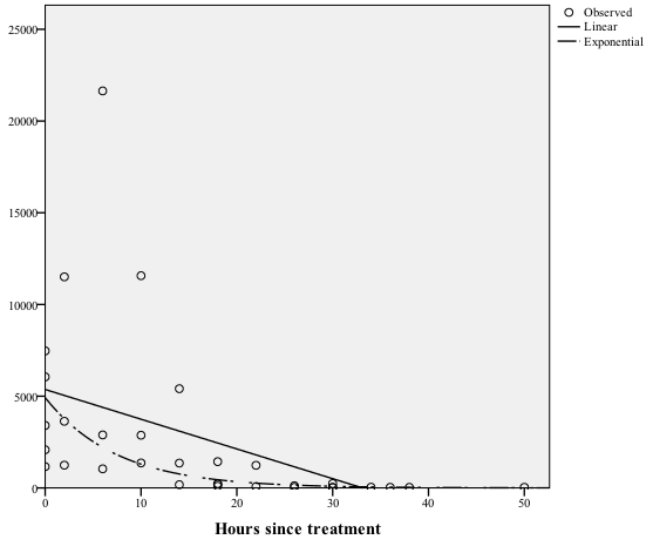
Figure 4.9 Decay of *T. pallidum* DNA and RNA following treatment

A. Blood DNA



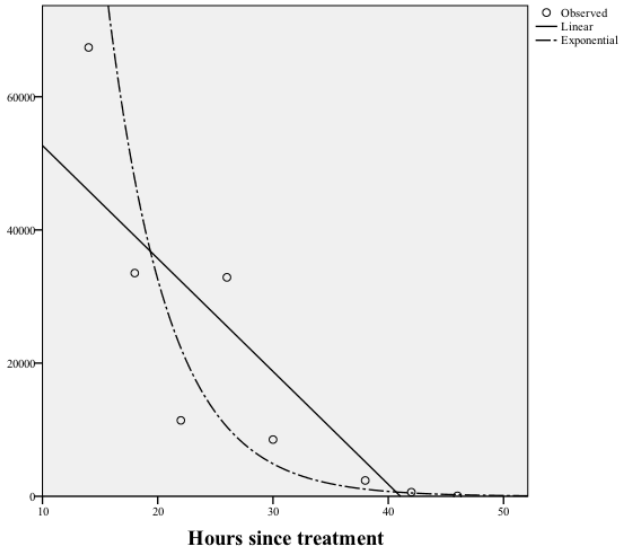
	R^2	sig
Linear	0.363	0.001
Exponential	0.470	0.001

B. Blood RNA



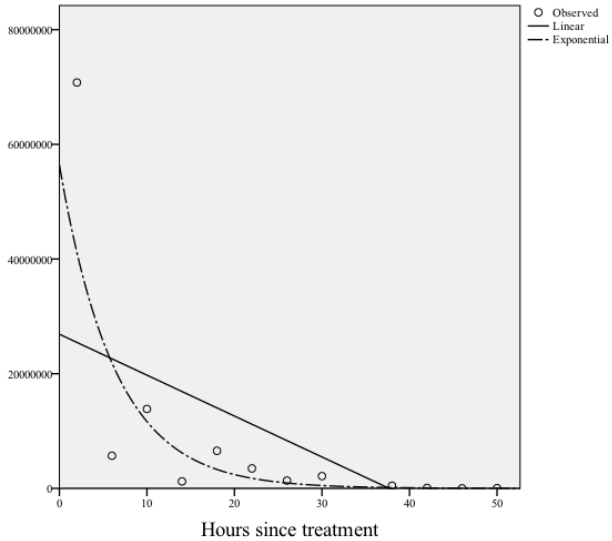
	R^2	sig
Linear	0.218	0.007
Exponential	0.70	0.001

C. Ulcer DNA



	R^2	sig
Linear	0.698	0.010
Exponential	0.878	0.001

D. Ulcer RNA



	R^2	sig
Linear	0.332	0.050
Exponential	0.839	0.001

Fig 4.9. The quantity of *T. pallidum* DNA (**Panel A**) and RNA (**Panel B**) detected in blood samples from four separate patients with early syphilis were plotted against time elapsed since treatment. Linear and exponential regression lines were then fitted to the data to assess the nature of the observed decay in bacterial load. The correlation coefficient, R^2 provides a measure of 'fit' of the data with the regression line where a value of one indicates a perfect fit and zero represents no fit. The significance of the R^2 values are also given. **Panels C** and **D** show the same analysis for DNA and RNA quantification in a penile ulcer from a single patient.

Table 4.6 Measurement of *T. pallidum* clearance following treatment.

	STS1	STS2	STS3	Mean	Std. dev.
Blood DNA					
$t^{1/2}$ (hours)	5.79	6.15	5.11	5.68*	0.53
Peak bacteraemia (tpp047 copies/ml)	2151	895	1832	1626.00	652.85
Peak Bacteraemia (hours)	2	0	2	1.33	1.15
Time to clearance (hours)	40	24	32	32.00	8.00
Blood RNA					
$t^{1/2}$ (hours)	3.87	3.07	4.75	3.89*	0.84
Peak bacteraemia (16S rRNA copies/ml)	3635	1363	21640	8879.33	11109.30
Peak bacteraemia (hours)	2	10	6	6.00	4.00
Time to clearance (hours)	32	20	53	35.00	16.70
Ulcer DNA					
$t^{1/2}$ (hours)	x	x	3.19	x	x
Peak ulcer load (tpp047 copies/strip)	x	x	67400	x	x
Time of peak load (hours)	x	x	14	x	x
Time to clearance (hours)	x	x	50	x	x
Ulcer RNA					
$t^{1/2}$ (hours)	x	x	4.08	x	x
Peak ulcer load (16S rRNA copies/strip)	x	x	7.08e7	x	x
Time of peak load (hours)	x	x	2	x	x
Time to clearance (hours)	x	x	56**	x	x

Table 4.6. Three patients were treated for early syphilis with benzathine penicillin. Following injection of the drug, *T. pallidum* DNA and RNA clearance from the blood of all three patients and the penile ulcer of a single patient was measured and used to calculate time to clearance and the half-life for clearance.

*p=0.035 (2-sided t-test of $t^{1/2}$(RNA clearance) vs $t^{1/2}$(DNA clearance)

** The last sample collected was the first negative sample.

4.2.2.4 Relationship between DNA and RNA clearance

A direct comparison of DNA-measured and RNA-measured whole blood bacterial load revealed a positive correlation between these levels for each of the patients STS1 (R^2 = 0.69), STS2 (R^2 = 0.44) and STS3 (R^2 = 0.64) (Figure 4.10 A). It is also noteworthy that for a given level of DNA the amount of RNA detected in the blood of patient STS3 was higher than that of patients STS 2 or STS 1. When expressed as a ratio of RNA:DNA and plotted against time since treatment (Figure 4.10 B) the decay of bacterial load was similar to that observed for DNA or RNA considered individually. The mean RNA:DNA ratio over all time periods was 7.19 (std. dev. 6.22) for STS3 compared with 1.44 (std. dev. 1.36) for STS1 and 3.66 (std. dev. 3.62) for STS2. Only the difference between STS1 and STS3 ratios was significant (p=0.002).

Figure 4.10 *T. pallidum* nucleic acid levels in blood during early syphilis

A. RNA vs DNA

B. DNA:RNA ratio vs time

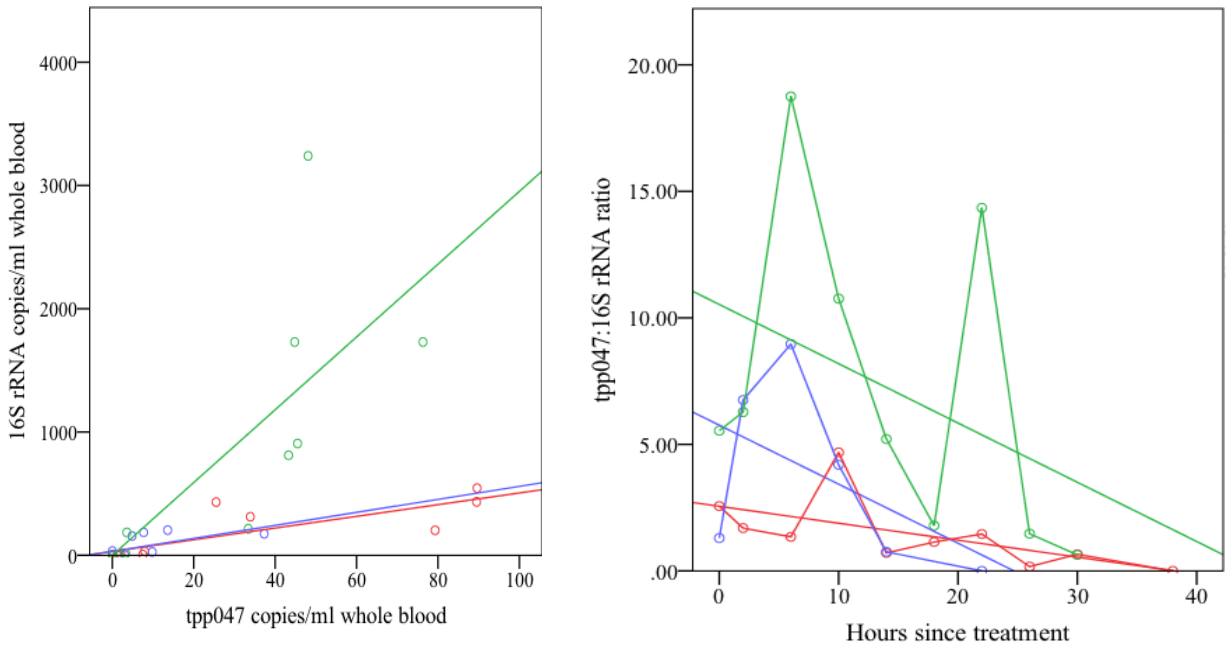


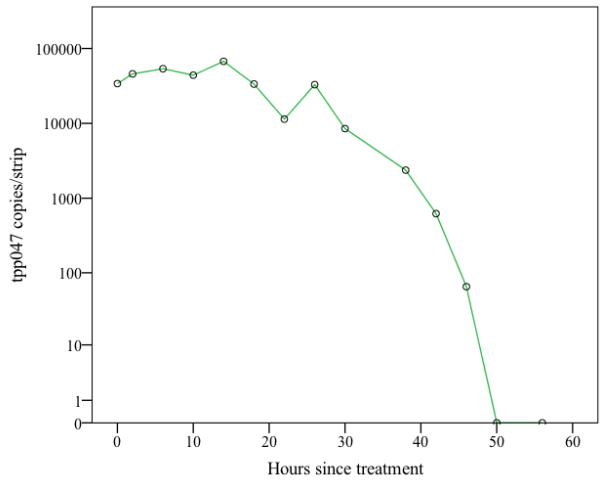
Fig 4.10. Panel A. Comparison of *tpp047* DNA and 16S rRNA levels detected in the blood of three patients (STS1 in red, STS2 in blue and STS 3 in green) with early syphilis measured over a 40-hour period following treatment. Regression lines are shown and were used to calculate the following correlation coefficients: STS1 R^2=0.686; STS 2 R^2=0.435 and STS 3 R^2=0.637 **Panel B**. The ratio of 16S rRNA:*tpp047* DNA quantity measured in the blood of three patients over 40 hours following treatment for early syphilis. Trendlines are shown with colour coding as for panel A.

4.2.2.5 Kinetics of T. pallidum clearance from a penile chancre

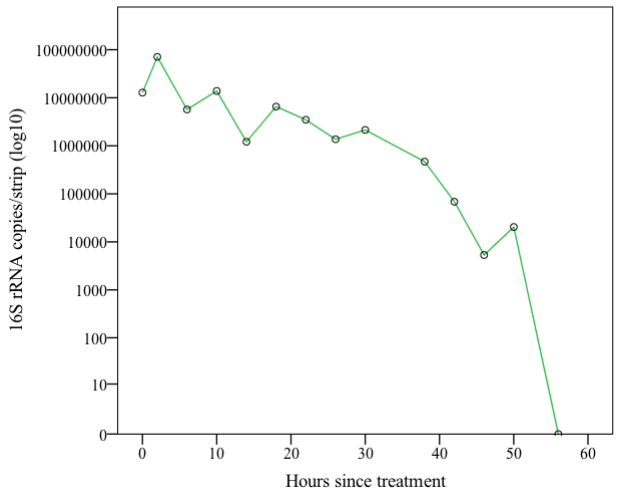
A single subject presented with a DGM positive genital ulcer that was suitable for sampling. Samples were taken at the same time-points as for blood samples, however, nucleic acid extraction and quantification was not performed in parallel with blood samples. Figure 4.11 demonstrates an initial ulcer *tpp047* load of 34000 copies/strip, which peaked at 67400 copies/strip at 14 hours and remained high until 26 hours when it began to fall rapidly. At 50 hours following treatment, *tpp047* was no longer detectable. The decay of *tpp047* DNA fitted an exponential regression line (R^2 = 0.878) (Figure 4.11) and the $t^{1/2}$ (clearance) was calculated to be 3.19 hours (Table 4.6). Quantification of 16S rRNA followed a similar pattern, although started from a level 1000 times greater than that seen for DNA ($1.28x10^7$ copies/strip) and peaked earlier at two hours post-treatment ($7.08x10^7$ copies/strip). After 40 hours, the level was seen to decrease 10-fold every four hours until becoming undetectable at 56 hours. It must be noted that no ulcer samples were taken beyond the 56 hour time-point as the end of sample collection was informed by whole blood DNA results (which were negative after 34 hours). Again, the decay of 16S rRNA was deemed to be exponential (R^2 = 0.839) and $t^{1/2}$ (clearance) was 4.08 hours.

Figure 4.11 Ulcer bacterial clearance

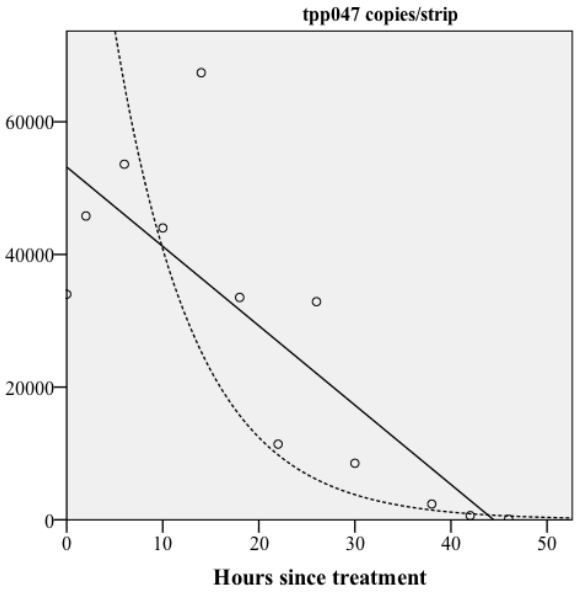
A. *tpp047* quantification



B. *16S rRNA* quantification



C. *tpp047* curve estimation



D. 16S rRNA curve estimation

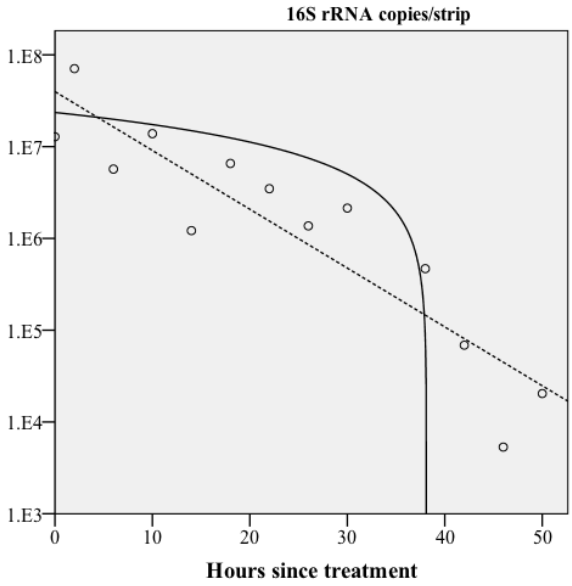


Fig 4.11. A patient presenting with a DGM positive penile ulcer (primary syphilis) was administered a 2.4 megaunit dose of benzathine penicillin at time zero. Ulcer samples for *T. pallidum* DNA and RNA quantification were collected pre-treatment and then four hourly for 56 hours. **Panel A**. presents DNA (*tpp047*) decay and **Panel B**. shows 16S rRNA decay. **Panel C.** depicts the fit of *tpp047* clearance data to linear (R^2 =0.66) and exponential (R^2 = 0.78) regression lines. **Panel D**. shows the fit of 16S rRNA data to exponential (dotted line) (R^2 =0.83) and linear (solid line) (R^2 =0.29) regression lines.

4.2.2.6 Further observations of T. pallidum quantification at two or more time points from samples collected in a cross-sectional study

During the cross-sectional study described in Chapter 2.14.2 four patients from whom pre-treatment samples were taken were subsequently admitted to hospital for the investigation and treatment of suspected neurosyphilis. All four patients were HIV-1 infected MSM with a mean age of 38.25 years (range 9) and a mean CD4 count of 689 cells/ml (range 712). Moreover, all received treatment, at least initially, with 6-hourly intravenous infusions of aqueous benzylpenicillin. Patient TP124 had two blood draws (at baseline and 36 hours post-treatment) expressly for *T. pallidum* quantification, which included Tempus™ samples. Pre-treatment samples for the remaining three patients were also taken for *T. pallidum* qPCR, however subsequent EDTA whole blood samples were left over full blood count (EDTA whole blood) samples which had been stored in the haematology department at 4°C. Three of the four patients attended for syphilis follow-up and all met the serological criteria for cure (Table 4.5).

Table 4.7 summarises the *T. pallidum* quantification data obtained for each of these four patients. It is noteworthy that none of the samples taken from patient TP103 at 7, 10, and 13 days post-treatment or the samples donated by patient TP105 at 12 and 14 days following treatment contained detectable *T. pallidum* DNA. A $t^{1/2}$ (clearance) of 9.97 hours has been calculated for patient TP140 as four samples over 86 hours were available for *tpp047* quantification. However, this calculation is based on the peak DNA being that which was recorded at baseline as a second sample was not available until 41 hours into treatment. Finally, patient TP 124 was shown to have cleared *T. pallidum* DNA by 36 hours, but not 16S rRNA, which remained detectable at 36 copies/ml.

Table 4.7 Clinical characteristics and *T. pallidum* qPCR results

Patient Study Number	**TP 103**		**TP 105**		**TP124**		**TP140**	
Age (years)	44		44		35		36	
CD4 count (cells/mm^3)	460		540		880		590	
Anti-retroviral treatment	naive		naive		naive		Darunavir Ritonavir Tenofovir Emtricitabine	
Syphilis treatment received	notes unavailable		2.4g IV benzylpenicillin 4-hourly for 14 days		2.4 megaunits IM benzathine, *and* 11 x 2.4g IV benzylpenicillin, *and* 3 once-weekly injections of benzathine 2.4 megaunits IM		2.4g IV benzylpenicillin 4-hourly for 14 days	
Steroids therapy administered	notes unavailable		none		none		3x 40mg prednisolone on 1st 3 days of treatment	
HIV-1 viral load (copies/ml)	79092		75		496936		<50	
CSF results								
WCC (cells/mm^3)	3		<1		6		4	
RBC (cells/mm^3)	4		<1		18		<1	
protein (g/L)	no result		<0.4		0.27		0.52	
TPPA	positive		positive		no result		positive	
RPR	no result		negative		no result		not done	
***tpp047* DNA (copies/ml blood)**	*Days*	*Load*	*Days*	*Load*	*Hours*	*Load*	*Hours*	*Load*
Sample 1	0	679	0	30	0	105	0	2335.2
Sample 2	7	0	12	0	36	0	41	135.36
Sample 3	10	0	14	0			62	0
Sample 4	13	0					86	0
t $^{1/2}$ (*tpp047* clearance)	n/a		n/a		n/a		9.97 hours	
16S rRNA (copies/ml blood)	n/a		n/a					
Sample 1					0	3404	0	7470.4
Sample 2					36	34		
RPR titre								
Pre-treatment	128		512		32		64	
1 month	64		128		Did not attend follow-up		16	
3 months	2		16				2	
6 months	4		4				-	

Table 4.7. Characteristics of four HIV-1 infected patients admitted to hospital for the treatment of suspected neurosyphilis. At least two whole blood (EDTA) samples were obtained from each patient for *T. pallidum* quantification and were stored in the CDTB at -80°C prior to DNA extraction. One or two Tempus™ blood samples were also available for *T. pallidum* 16S rRNA quantification from patients TP124 and TP140.

4.2.2.7 Multiplex results

The design of an internal control is described in Section 4.2.1 and was intended to measure the efficiency of DNA recovery from clinical samples and its subsequent amplification by PCR. It was to be spiked into whole blood samples and co-'extracted' with *T. pallidum* DNA, then co-amplified in a multiplex qPCR. In order to assess the feasibility of this process, 30 whole blood samples from patients STS1, STS2 and STS3 were selected and 500µl aliquots were spiked with 5000 copies of *ncoI*-linearised pTYLCV DNA. The samples chosen were those in which *tpp047* DNA was previously detected and the first two *T. pallidum* negative samples from each of the patients.

DNA extraction proceeded in batches of eight samples (Chapter 2.3.1) and was followed by *tpp047* and pTYLCV quantification in a multiplex qPCR (Chapter 2.12.1.3), which included three technical replicates of each sample, a no-template control and positive *tpp047* and pTYLCV controls. The cycle thresholds for detection of both targets were determined and compared with the two absolute quantification standards included in each run.

Panel A of Figure 4.12 and Table 4.8 describe the clearance of *tpp047* DNA, measured by the multiplex assay, from the blood of all three patients. On average, peak bacteraemia was detected at 0.7 hours (std. dev 1.15) and bacterial clearance was achieved by 28 hours (std. dev. 6). Panel B of Figure 4.12 and Table 4.9 summarise the quantity of pTYLCV DNA 'recovered' from whole blood samples. The average, across all patients and time points was 1949 copies (std. dev 3108). The efficiency of DNA extraction (E) was calculated by dividing the number of pTYLCV copies 'recovered' from a given sample (R) by the number spiked (5000) and expressing the result as a percentage, i.e.

$$E = \left(\frac{R}{5000}\right) \times 100$$

Next, in order to adjust the quantity of *tpp047* DNA detected in a given sample (Q) for the DNA extraction efficiency of that sample (E), the copy number was multiplied by E^{-1}, i.e

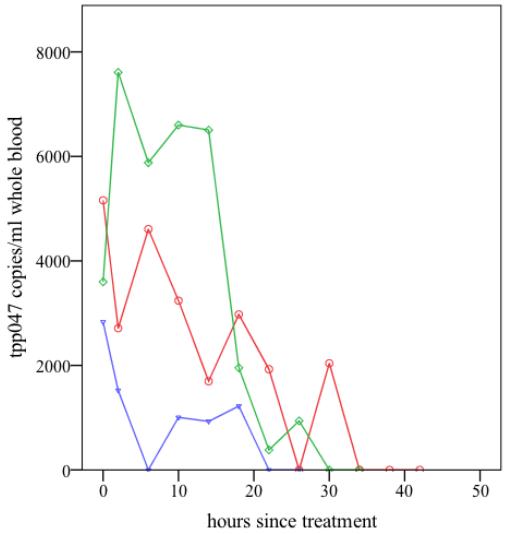
$$adjusted\ tpp047\ copies = \left(\frac{1}{R/5000}\right) \times Q$$

Table 4.9 demonstrates that the efficiency of DNA extraction from the whole blood samples included in this assessment was, on average, 39% (min seven, max 338%). Analysed separately, the extraction efficiency of STS 3 samples (extracted as a single batch) was significantly lower than that for STS1 (p=0.033) or STS 2 (0.056). Of course, the observed extraction efficiency was also subject to the performance of the multiplex qPCR used to quantify pTYLCV. To minimise variation, DNA from all sample time-points for a given patient was quantified during the same experiment using the same 'master' dilution as the starting concentration for each standard curve. The performance characteristics of each of these absolute quantification standard curves are described in Table 4.10 and are broadly similar.

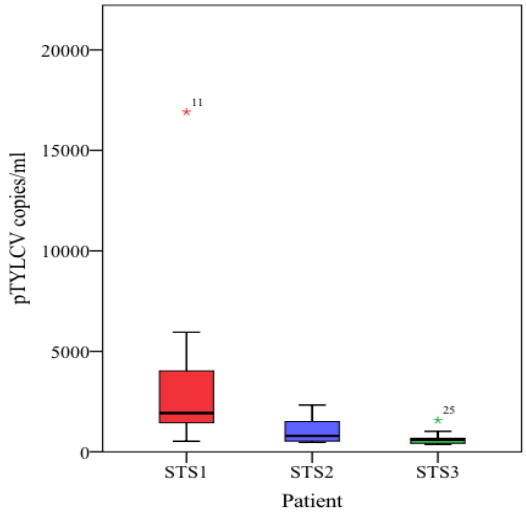
Figure 4.12 and Table 4.8 also provide a comparison of *tpp047* quantification with and without adjustment for extraction efficiency. Moreover, they demonstrate the impact of multiplex detection on *tpp047* quantification by providing a comparison with the single-target qPCR. Panel C demonstrates that the patterns of clearance for each of these three measurement methods (single-target assay, multiplex, and adjusted multiplex) are similar although the overall levels of *tpp047* measured are different. Table 4.8 allows closer inspection of the differences and demonstrates that while there is a trend towards higher *tpp047* levels following adjustment for extraction efficiency and lower levels in the single target assay these differences are not significant. $t^{1/2}$(clearance) measured by the 'multiplex' method is seen to be 9.38 and 7.74 hours longer than 'single-target' and 'adjusted' methods of *tpp047* quantification. Again, this difference did not reach statistical significance. Finally, when analysed separately, extraction efficiency-adjusted levels of *tpp047* for patient STS3 are markedly higher than for patients STS2 and STS1

Figure 4.12 Efficiency of *T. pallidum* DNA extraction

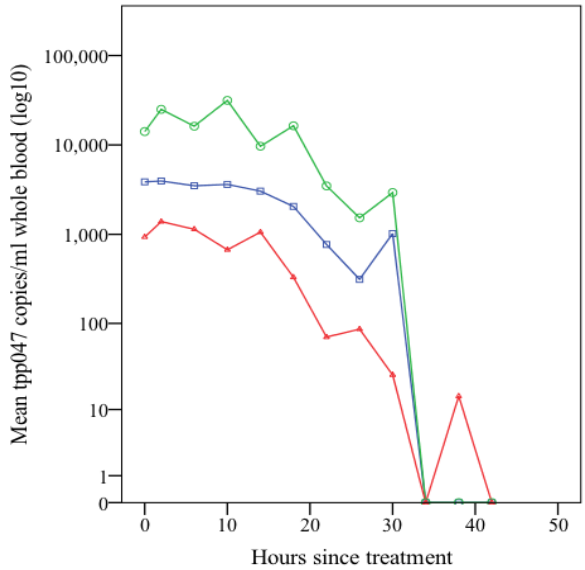
A. *tpp047* quantification



B. pTYLCV quantification



C. Singleplex vs multiplex *tpp047* quantification



D. Adjusted *tpp047* quantification

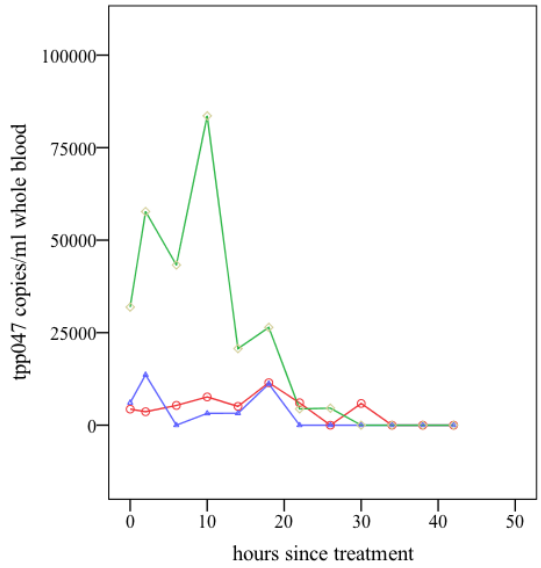


Fig 4.12. Multiplex qPCR detection of *tpp047* and pTYLCV DNA in the whole blood of three patients treated for early syphilis with single-dose benzathine penicillin at time zero. Each 500μl whole blood sample was spiked with 5000 pTYLCV copies prior to DNA extraction and qPCR. **Panels A** and **B** show *tpp047* and pTYLCV quantification for patients STS1 (red), STS2 (blue) and STS3 (green), respectively. **Panel C** compares the mean *tpp047* detection for all three patients at each time-point in both single-target (*tpp047*) (red) and multiplex (*tpp047* and pTYLCV) (blue) qPCRs. Additionally, quantification of *tpp047* was adjusted for DNA extraction efficiency (green) **Panel D** compares *tpp047* quantification adjusted for DNA extraction efficiency for each of the three patients (STS1 in red; STS2 in blue and STS3 in green) separately.

Table 4.8 *T. pallidum* clearance with and without adjustment for DNA extraction efficiency

	Patient					
	STS1	**STS2**	**STS3**	**Mean**	**Std. dev.**	**t-test***
A. Single-target *tpp047* detection						
$t^{1/2}$ (hours)	5.79	6.15	5.11	5.68	0.53	
Peak bacteraemia (tpp047 copies/ml)	2151	895	1832	1626	653	
Peak Bacteraemia (hours)	2.00	0	2.00	1.33	1.15	
Time to clearance (hours)	42	26	34	34	8	
B. Multiplex detection of *tpp047*						
$t^{1/2}$ (hours)	22.4	14.84	7.94	15.06	7.23	0.089
Peak bacteraemia (tpp047 copies/ml)	5160	2832	7608	5200	2388	0.067
Peak Bacteraemia (hours)	0	0	2.00	0.66	1.15	0.002
Time to clearance (hours)	34	22	30	28	6	0.357
C. *tpp047* detection adjusted for DNA extraction efficiency						
$t^{1/2}$ (hours)	12.37	5.77	3.82	7.32	4.48	0.56
Peak bacteraemia (tpp047 copies/ml)	11481	13605	83587	36224.33	41031.02	0.22
Peak Bacteraemia (hours)	18	2	10	10.00	8.00	0.14
Time to clearance (hours)	34	22	30	28.67	6.11	0.41

Table 4.8 Whole blood was collected into EDTA from three patients with early syphilis before administration of 2.4 megaunits of benzathine penicillin. Samples were then taken four-hourly post-treatment and *T. pallidum* DNA (*tpp047* gene) quantified at each time point. The level and time of peak bacteraeamia were determined in addition to the total time and half-life required for its clearance (**A**). A second aliquot of blood from each time-point was spiked with 5000 copies of pTYLCV and both *tpp047* and pTYLCV were quantified in a multiplex qPCR (**B**). DNA extraction efficiency were determined by calculating the fraction of spiked pTYLCV recovered and *tpp047*-measured *T. pallidum* load adjusted accordingly (**C**) *comparison of single-target and either multiplex or adjusted *tpp047* results.

Table 4.9 Efficiency of pTYLCV extraction

Patient Study Number (number of samples)	pTYLCV copies 'recovered' (copies/reaction)		Extraction efficiency (%)		
	mean	*std. dev*	*mean*	*min*	*max*
STS 1 (n=12)	3602	4476	72	11	338
STS 2 (n=8)	1059	671	21	10	47
STS 3 (n=10)	678	370	13	7	31
All patients (n=30)	1949	3108	39	7	338

Table 4.9. Linearised internal control plasmid (pTYLCV) (n=5,000 copies) were spiked into 500µl whole blood collected during a longitudinal study of *T. pallidum* clearance following treatment. The quantity of pTYLCV recovered following DNA extraction was determined in a multiplex qPCR. All samples for a given patient were extracted in the same batch and quantified in a single qPCR experiment (comprising three technical replicates of each sample). pTYLCV copy numbers and extraction efficiencies presented are the mean of all samples collected from each patient.

Table 4.10 Multiplex qPCR efficiency for *T. pallidum* and pTYLCV amplification

Patient	PCR Target	Efficiency	R^2 (linearity)	Slope (cycles)
STS1	pTYLCV	1.98	0.98	-3.46
	tpp047	1.43	0.95	-4.27
STS2	pTYLCV	1.63	1.00	-3.85
	tpp047	1.64	1.00	-3.83
STS3	pTYLCV	2.18*	1.00	-3.45
	tpp047	1.90	1.00	-3.45

Table 4.10. A multiplex qPCR was used to quantify pTYLCV and *tpp047* by comparing the cycle thresholds for detection in samples with that of two absolute quantification standard curves. The PCR instrument software calculated reaction efficiency and plotted cycle threshold vs cycle number to calculate the number of PCR cycles between each 10-fold dilution of standard curve DNA. Finally, R^2 was calculated which denotes how well the experimental standard curve line fits a theoretical regression line.

*It is noted that this result is above the theoretical limit of PCR efficiency.

Discussion

qPCR is used routinely in clinical practice to monitor patient responses to HIV-1 and hepatitis C therapy and for both diseases an undetectable 'viral load' has become synonymous with treatment success (252, 253). In both cases, RNA is extracted from patients' blood and subject to reverse transcription and quantification in one of a number of commercially available assays. The principal aim of the current study was to replicate this approach for early syphilis by measuring bacterial load following treatment. It has previously been shown that qPCR can quantify *T. pallidum* in blood and ulcer samples taken from patients with early disease (131, 151). Moreover, in Chapter three it was demonstrated that *T. pallidum* bacterial load in both blood and ulcers varies according to stage of disease in a way that is both biologically plausible and potentially clinically useful.

4.3.1 Controlling for the efficiency of DNA recovery with an internal control

Both the *T. pallidum* qPCR for *tpp047* DNA quantification and RT-qPCR for 16S rRNA quantification described in Chapter two include internal calibration with an in-run absolute quantification standard. By observing for changes in the amplification characteristics of these standards, the overall efficiency of the PCR was controlled. It was also important, especially in a quantitative assay, to control for the efficiency of DNA recovery, which can be variable (254). This is most often achieved by adding an internal control to samples prior to processing which is then co-extracted with the same efficiency as target DNA. Both targets can then be quantified in the resultant DNA mixture with separate or multiplex qPCR (228, 255, 256). Of course, it is not possible to control for the efficiency of bacterial lysis using this method as the control plasmid is not contained within intact *T. pallidum* organisms. The first of two internal control plasmids assessed in the current study (pTpp047) was produced using site-directed mutagenesis of the *tpp047* gene. The rationale for this method was that both *tpp047* and control DNA could be amplified with a single primer pair, thus reducing the potential for one or other target to dominate the reaction. Three concentrations of pTp47 were assessed and 70 copies/reaction appeared to be optimal. However, the utility of this multiplex qPCR proved limited as competition in the qPCR prejudiced against the detection of low bacterial loads (<10 *T. pallidum* genome equivalents), which had previously been detectable in the singleplex *tpp047* qPCR. Moreover, at higher (>1000 genome equivalents/reaction) starting quantities of

T. pallidum DNA, pTpp047 was not detected. This effect is thought to result from earlier amplification of the higher concentration target exhausting PCR reagents such that the reaction has plateaued before the lower concentration target has been amplified beyond the threshold of its detection. For example, a similar pattern was described for the Roche Cobas AMPLICOR multiplex assay, which detects *Neisseria gonorrhea, Chlamydia trachomatis* and an internal control. When one or both of the organisms were present at a high starting quantity they would be amplified in preference to the low-level internal control DNA and the assay reported as a failure despite correctly identifying the target organism (257). Another consideration was the potential for interference between the detection of the two fluorescent dyes in the reaction. This was easily assessed by replacing the VIC fluorophore on the probe TPP-QC with ROX and observing no demonstrable change in the detection of either target over a range of *T. pallidum* DNA starting concentrations. One approach to removing competition in multiplex qPCRs is to limit the concentration of one target's primer pair (258). In this way amplification of that target will reach plateau phase earlier and allow time (and reagents) for the second target, which is not primer concentration-limited, to amplify. To put this in context, at low starting amounts of *T. pallidum* DNA a reduction in the internal control primer concentration should allow both targets to amplify adequately before other reaction substrates (magnesium, taq polymerase, DNTPs) become limited. Of course, this approach is only possible for multiplex qPCRs using two distinct primer pairs thus a new internal control was constructed from a section of the TYLCV genome. By having two sets of primers the concentration of both could be varied to optimise the detection of both targets. During an initial assessment, the detection of neither 100 nor 10 *T. pallidum* genome equivalents was affected by the addition of 500 pTYLCV to the multiplex qPCR, even without limiting TYLCV reaction primer concentrations.

The next step was to determine the optimal amount of pTYLCV to be added to whole blood samples prior to DNA extraction. When samples were spiked with 500 copies, the amount of DNA recovered was below the threshold of detection in the qPCR and it was ultimately demonstrated that spiking with 5000 pTYLCV copies was required. Again, *tpp047* detection, this time from 1000 *T. pallidum* organisms, was unaffected by co-recovery and qPCR amplification of pTYLCV. The DNA extraction efficiency of Qiagen's QIAamp method has been shown previously to have variable efficiency. In one study, the recovery of marker DNA (a familiar 500 copies/ reaction of a linearised plasmid target) was 38%. However, another assessment of 12 DNA extraction methods favoured the QIAamp technique, not only for a

DNA recovery rate of 55.7%, but additionally for the higher purity and integrity of the DNA extracted (255, 256, 259). In the current study an *ex vivo* assessment of DNA recovery efficiency and the performance of the multiplex qPCR was made using blood collected during the longitudinal treatment study. The overall efficiency of DNA recovery, according to the average fraction of pTYLCV recovered, was 39%, with a range of over 300%. Potential explanations for the unexpectedly high variation of DNA recovery lie in the spiking and DNA extraction methods. Firstly, pTYLCV was diluted to a concentration of 1000 copies/μl and each sample was spiked separately with 5μl. The potential existed therefore, for variation in the amount of DNA spiked into each sample and adding pTYLCV to the lysis buffer used during the first step of the process may have improved consistency. Secondly, following DNA elution into 400μl of water, a DNA precipitation step was performed. The last step of this process required the DNA pellet to be dried (ethanol allowed to evaporate) and then re-suspended in 60μl of water. Any remaining ethanol would have likely affected the qPCR and over-drying led to difficulty in re-suspending the pellet. It may have been preferable, in the context of a quantitative and not qualitative PCR, to replace this step with an alternative method. Finally, the measurement of extraction efficiency also relied on the efficiency of the qPCR used to detect the extracted DNA. All samples from a given patient were quantified in triplicate during a single multiplex qPCR run and compared with the same absolute quantification standard. Variation in efficiency of the calibration curves between runs was noted (1.98, 1.63, 2.18) and would have contributed to the reported variation in extraction efficiency to some extent. This may also explain the higher DNA-measured bacterial loads detected with the multiplex assay compared with the single-target *tpp047* assay. Unfortunately, repeating the extractions to produce the average extraction efficiency for each time point was not possible as insufficient sample remained. Despite these limitations, a similar pattern of *T. pallidum* clearance was observed for *tpp047* single-target detection; *tpp047* mutliplex detection and *tpp047* detection adjusted for DNA recovery efficiency.

4.3.2 *T. pallidum* clearance as a marker of anti-microbial efficacy

T. pallidum persistence after syphilis treatment has been the subject of research for more than 50 years. Prior to the description of PCR, these studies relied on DGM examination and RIT of clinical and animal specimens (221, 224). In the first description of penicillin use for the treatment of syphilis, four men with DGM positive chancres were given four-hourly IM

penicillin injections and ulcers sampled with each injection. By the 16th hour, no spiral forms were observed in any of the ulcers (18). PCR has since provided a tool for sensitive detection of bacterial DNA and has been used in rabbits to demonstrate that *T. pallidum* injected intra-testicularly were cleared two weeks after the animals were treated with penicillin and probenecid (155). In a more recent clinical study, PCR detection revealed that *T. pallidum* in the CSF of patients with asymptomatic neurosyphilis could persist for up to 36 months after treatment (144). By serially sampling blood following treatment the current study demonstrates, in four patients, that it is possible to monitor the treatment of early syphilis with qPCR and that bacterial clearance (to below the level of detection) can be observed within three days of treatment. All three patients were also shown, by following their RPR titres, to be 'cured' following treatment according to the currently recognised definition of serological cure (104). Of course, a far greater number of patients would be needed to prove definitively that DNA clearance equates with cure and further assessment should include patients with both serological failure and those with serofast treatment outcomes in addition to serological successes. Moreover, it could be argued that RIT is the current gold standard for *T. pallidum* detection and should be used in future studies to diagnose cure, however impractical.

The ultimate goal of establishing the kinetics of *T. pallidum* clearance is to develop a speedy and robust way of predicting patient treatment outcomes without the need for months of serological follow-up. It was important for this pilot study to select patients who were likely to have detectable bacterial loads at baseline. Therefore, patients with early syphilis, specifically those with late primary or early secondary disease, were selected as they were previously shown in the cross-sectional data (Chapter 3.2) to have more frequently a detectable blood bacterial load. This strategy proved successful as all of the four patients recruited were found to be bacteraemic. It must be noted, however, that patient STS4 had a much lower blood bacterial load than the other three recruits, and while data from this patient were of interest; they were excluded from many of the analyses. Overall, the half-life for blood clearance of *T. pallidum* DNA following the administration of a single dose of benzathine penicillin was 5.68 hours and for RNA was 3.89 hours. All patients had an undetectable blood bacterial load (DNA and RNA) 56 hours after treatment. Importantly, this correlated with serological cure, defined as a two-dilution fall in RPR titre, measured at either one or three months following treatment. It is also significant that patient STS1, whose blood was sampled until 150 hours (6.25 days) post-treatment continued to have an undetectable bacterial load following clearance. Moreover, inpatients undergoing treatment for

neurosyphilis who had detectable bacteraemia pre-treatment had consistently negative bacterial loads following treatment up to 14 days after treatment. This time period more than adequately covers the 32-hour replication time of *T. pallidum* and the theoretical potential for a late recrudescence in bacterial load (4). It could be envisaged, therefore, that for patients with active PCR-positive syphilis at baseline, a clearance half-life of under six hours and/or a negative bacterial load at three days would be adequate to indicate cure and could potentially be a measure of treatment success in a trial comparing syphilis treatments. Of course, the results presented in the current study would need to be replicated in a much larger study in order to validate this strategy.

A further question arises when considering the use of bacterial load as a marker of *T. pallidum* clearance (and syphilis cure): Is DNA or RNA the superior measure of bacterial clearance? If time is of the essence, then RNA may be deemed superior as the clearance half-life for DNA was significantly longer than RNA (5.68 hours vs 3.89 hours). Of course, the number of samples taken in a study assessing the efficacy of a syphilis treatment will be limited, and sampling points could easily be tailored to the marker being used, which makes the five-hour difference between DNA and RNA clearance less important. If laboratory time and costs are also considered then RNA, which can be extracted in 50 minutes, is the clear favourite. The DNA extraction method requires two three-hour incubations and a DNA precipitation step that, as discussed previously, may contribute to a wide variability in extraction efficiency. It must be noted, however, that an internal control of the RNA extraction process has not been assessed. One final consideration is the effect that bacterial metabolic activity has on RNA levels detected, which has been predicted to be substantial. Again, if clearance half-life is to be the measure of anti-microbial efficacy then DNA may prove the superior measure as it is less variable. However, if a qualitative approach is favoured, then the greater sensitivity offered by RNA detection may be preferable.

All four subjects in the current study presented within one month of the onset of symptoms and had serological results consistent with early disease, however they were not an homogenous group. Most notable is patient STS4 who had only low-load *T. pallidum* bacteraemia, similar to levels seen for latent disease in Chapter three. Moreover, he was the only one of the four not to exhibit a marked worsening of his rash or other symptoms of a JHR following treatment. It is possible, therefore, that he had already progressed to the early latent stage by the time he was recalled for treatment and enrolled into the study, which may

explain the significantly lower level of *T. pallidum* bacteremia recorded. If this is indeed the case, it may indicate a role for testing the blood of patients with suspected latent disease on more than one occasion in order to increase the sensitivity of *T. pallidum* detection. Also of interest, is patient STS3 who presented with a primary chancre of 4 weeks' duration, which was still strongly DGM positive. Measured by DNA, his blood bacterial load at presentation was not significantly higher that that of patients STS1 or STS2. The RNA-measured load, however, far exceeded that of the other three patients. When expressed as a ratio of RNA:DNA and plotted against time since treatment a similar pattern of clearance was noted for patients STS1, STS2 and STS3. It is notable, however, that patient STS3's blood had a higher RNA:DNA ratio than that of patients STS1 or STS2. This could simply represent differences in extraction and PCR efficiency for samples collected for this patient, however, it may well represent increased metabolic activity of *T. pallidum* organisms actively invading the patients blood stream from his primary chancre (which contained 1000 times more RNA than DNA pre-treatment).

Three of the four patients in the study developed symptoms and signs consistent with a JHR in the hours following their treatment with benzathine penicillin. A reported 40-80% of patients experience a JHR following treatment and it has been associated with younger age (maximal at 27 years); early disease (primary or secondary) and with previous penicillin administration (190, 218, 260, 261). Furthermore, increased risk of JHR has been associated with RPR titres greater than 1:32 compared with less than 1:32 (82.1% vs 55.6%, p=0.001) (261). It is also of note that in a recent post-hoc multivariate analysis of a randomised syphilis treatment trial, younger patients who experienced a JHR following treatment had two-fold greater odds of cure than those who did not (218). In the current study, patient STS3, a 26 year-old with primary disease experienced the most severe JHR, which resulted in 24 hours of fever, malaise and headache. Of interest, this patient had the highest RNA-measured blood bacterial load; a baseline RPR titre of 1:32; was the youngest patient recruited and had achieved serological cure at one month. Patient STS4, a 35 year-old with secondary disease and who had the lowest blood bacterial load was the only patient not to develop the JHR despite a baseline RPR titre of 1:64. One might conclude, therefore, that higher bacterial loads are associated with both increased risk and greater severity of JHR although a much larger number of patients would be needed to prove this association. It is also of note that in a study measuring inflammation (CRP and complements C3 and C4) following early syphilis treatment the only patient not to experience a JHR was one with early latent disease (192). A

further interesting observation is that a peak in RNA bacterial load was noted between two and 10 hours after treatment in patients STS1, STS2 and STS3. A similar peak was noted in DNA levels at two hours post-treatment for patients STS1 and STS3 and a peak in neither DNA nor RNA was noted for patient STS4. This phenomenon was initially suspected to be the result of variations in extraction and PCR efficiency, but became harder to explain when it was noted in three patients' blood and for both nucleic acids. A possible explanation would be that the rapid killing of *T. pallidum* following the administration of penicillin results in the release of bacterial nucleic acids, proteins (including ribosomes) and other cellular components into the bloodstream. Indeed, the peak plasma concentration of penicillin does not occur until 13-24 hours after the administration of benzathine thus the drug and nucleic acids levels may peak simultaneously. This transient increase in the amount of detectable *T. pallidum* nucleic acids in the circulation also coincides with the JHR and may explain its aetiology. (100). Of course, at a higher pre-treatment bacterial load (patient STS3 had the highest pre-treatment levels of both DNA and RNA) this effect would be magnified and at lower loads (such as documented for patient STS4) it may be absent.

In summary, it has been demonstrated in a prospective longitudinal study of four patients that qPCR can be used to monitor decay in *T. pallidum* load following treatment and that bacterial clearance correlates with both resolution of clinical symptoms and serological cure. RNA decay is the quicker measure both in terms of laboratory processing time and time taken for blood clearance, and may be subject to less variation during the extraction process. A larger study is now proposed, using time points selected on the basis of these data, to show definitively that qPCR is superior to serology as a measure of *T. pallidum* clearance and thus to diagnose cure in early syphilis. These data will then be used to inform to design of new clinical trials comparing syphilis treatments in which cure or treatment failure can be quickly diagnosed.

Chapter 5

Molecular characterisation of *T. pallidum* strains

5.1 Introduction

5.1.1 Background

The British and Sri Lankan guidelines for the management of syphilis include macrolide antibiotics as alternative treatments for early disease. (104) In Britain, this advice comes with the caveat that intrinsic macrolide resistance is present in some strains of *T. pallidum*, although until now the prevalence of such strains in either country was unknown. Following clinical trials showing comparable efficacy to benzathine penicillin (161, 164, 219, 220, 262), there was enthusiasm for treating incubating and early disease with single-dose azithromycin (160). Unfortunately, this introduction was shortly followed by treatment failures in some patients (mostly MSM), which was attributed to a single point mutation in the 23srRNA gene (168, 169). Macrolide antibiotics are not widely used in Sri Lanka and the level of resistance in another bacterium, *Streptococcus pneumoniae*, is reported to be low. It is predicted that *T. pallidum* macrolide resistance, if present, will be equally low-level (263) .

Molecular studies of *T. pallidum* go beyond the identification of resistance and can also be used to reveal a number of strains based on analysis of the acidic repeat protein (*arp*); *T. pallidum* repeat (*tpr*) subfamily 1 genes and the *T. pallidum 548* gene (*tpp0548*) (82, 86). Population studies have shown that whilst there are common types, there is also considerable strain diversity within epidemics (87, 147). These data have led to theories regarding the origin of macrolide resistant strains. Previous studies have reported partial strain-typing data due to a low success rate of amplifying the *arp* gene by PCR (89, 264). There are also reported differences in the success of typing according to the clinical samples from which *T. pallidum* DNA is extracted. Finally, the original CDC method for *T. pallidum* strain-typing did not include *tpp0548* sequence analysis which may improve strain discrimination. (86)

In this study, macrolide resistance and strain-typing data from St Mary's Hospital, London, UK and the central clinics of the National STD/AIDS control programme in Colombo, Sri Lanka are presented. Additionally, an assessment is made of the addition of *tpp0548* gene sequence analysis to the standard CDC typing together with discussion of optimal PCR conditions for *arp* gene amplification and the most appropriate samples for strain-typing.

5.2 Results

5.2.1 Assessment of assay modifications to improve amplification of the *T. pallidum arp* gene

Successful amplification of DNA targets containing repeated sequences and a high proportion of guanine and cytosine bases can be troublesome. (265) In order to optimise the amplification of the *T. pallidum arp* gene, which contains a variable number of 60bp repeats and has a high GC content, four optimisation steps were assessed.

Initially, the concentration of the amplification primers ARP F and ARP R were increased from 0.2μM to 0.4μM and the amount of taq polymerase per reaction increased from one to two units. Figure 5.1 A demonstrates the superior PCR efficiency with the higher enzyme and primer concentrations. Moreover, that the limit of detection for this assay is approximately 10^3 genome equivalents of DNA. Secondly, the addition of betaine to the reaction (Figure 5.1 B) resulted in inferior *arp*-amplification at each of the three concentrations assessed. Likewise, the addition of GC-rich solution to the reaction (Figure 5.1 C) failed to enhance amplification, regardless of concentration. Finally, reaction efficiency was unaltered by the variation of magnesium concentration from 1-4 mM per reaction(data not shown).

On the basis of these results, the final reaction mixture selected for *arp* amplification contained: 0.4μM of each primer; 2 units of taq polymerase; 2mM magnesium (as supplied in PCR reaction buffer) and no betaine or GC-rich solution. Failed reactions were repeated with 10μL extracted *T. pallidum* DNA in place of the standard 5μL.

Figure 5.1 Optimisation of *T. pallidum* PCR amplification

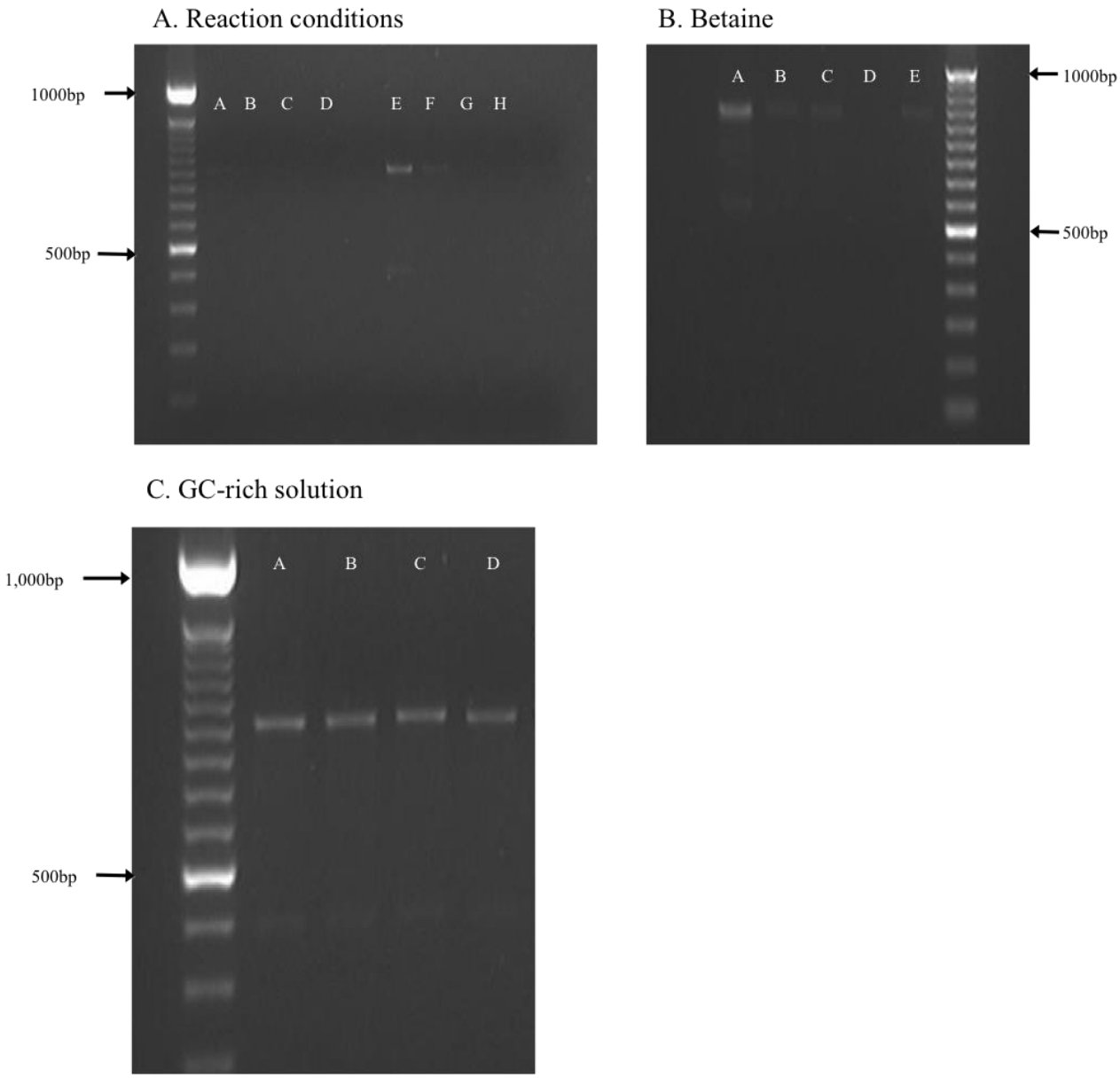


Fig. 5.1. *T. pallidum* strain typing required amplification of the gene *arp*. Different PCR conditions were assessed with the aim of improving *arp* amplification efficiency. **Panel A** shows PCR amplification in reactions with different concentrations of taq polymerase and reaction primers. Reactions in lanes A-D contained one unit taq polymerase and 0.2μM of ARP_F2 and ARP_R2 primers. Reactions in lanes E-H contained two units taq polymerase and 0.4μM of the primers ARP_F2 and ARP_R2. The impact of *T. pallidum* DNA starting concentration was also assessed for both conditions. Accordingly, 10^4 genome equivalents were added to reactions in lanes A&E; 10^3 in lanes B&F and 10^2 in lanes C&G. Reactions shown in lanes D&H were no template (water) controls. **Panel B** demonstrates the effect of adding betaine on the amplification of *arp* DNA extracted from 10^4 Nichols strain *T. pallidum*. The concentration of betaine added to reactions was either 40mM (lane E); 35mM (lane D); 30mM (lane C); 25mM (lane B) or none (lane A). Panel C assessed in impact of 'GC-rich solution' (Roche, Germany) at final concentrations of 0.4x (lane B); 0.6x (lane C) or 0.4x (lane B). Lane A contained no GC-rich solution.

5.2.2 Patient Recruitment and Demographics

London

Samples were collected from patients attending the sexual health department at St Mary's Hospital in London during two separate periods (described in sections 2.14.1 and 2.14.2). A total of 99 patients were recruited for the qPCR validation study between July 2006 and March 2008 and a further 62 patients donated samples between January 2011 and November 2012. In total, 102 patients had a clinic-microbiological diagnosis of syphilis and at least one sample (blood or ulcer) from 58 of these patients contained *T. pallidum* DNA. Following initial *tpp047* quantification by real-time PCR, any remaining DNA was subject first to mutation analysis, and then to strain-typing. Blood-extracted DNA was used only when ulcer-extracted DNA was unavailable.

Mutation analysis was achieved for 41 samples and a full or partial strain type subsequently identified in 23. All 41 samples were from men, of whom 38 (92.7%) were MSM and 26 (63.4%) were HIV-1 infected. The mean age of patients was 37.34 years (std. dev. 9.47) and the majority were of white British (22, 53.7%) or white 'other' (13, 31.7%) ethnic origin. A total of 16 (39%) patients were diagnosed with primary disease, 21 (48.8%) with secondary and 5 (12.2%) were latently infected. Patient recruitment and available samples are summarised in Figure 5.2.

Sri Lanka

Exudate was collected from the primary and secondary syphilis lesions of 25 patients attending the clinics of the centre for STD/AIDS control in Colombo, Sri Lanka between January and August 2012. All patients had a microbiological diagnosis of syphilis following dark-ground microscopy or serological testing. The presence of *T. pallidum* DNA was confirmed by PCR (*tpp047* gene) in 24 of the 25 samples of which 14 were from primary chancres, nine were from moist secondary lesions (or healing ulcers still present in secondary disease) and one was a skin swab. In total, 11 patients had self-referred to the service and 13, including a neonate with congenital syphilis, were referred from general practice or another hospital department. The majority of patients were Sinhalese (88%) and the remainder (12%) Sri Lankan Tamil. Excluding the neonate, patients had a mean age of 29.52 years (std. dev. 9.70). All patients were male and 11 (47.8%) identified as MSM (including seven identifying as bisexual). The majority (91.3%) reported no previous travel outside Sri Lanka.

Figure 5.2 Molecular characterisation of *T. pallidum* strains in London

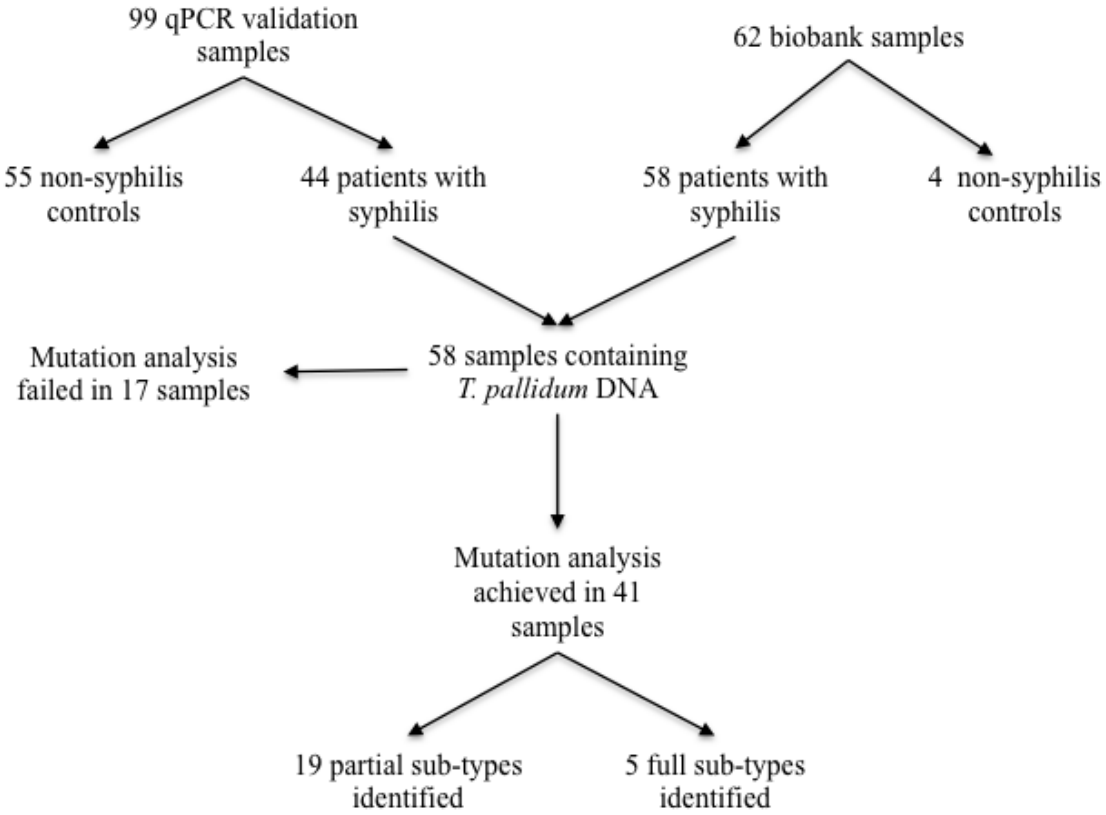


Fig. 5.2. *T. pallidum* strain-type and antibiotic resistance mutation analysis was performed on samples collected in London during two cross-sectional studies.

5.2.3 Sensitivity of *T. pallidum* strain-typing techniques

In London, 27 samples were subjected to strain type analysis resulting in the identification of five full and 19 partial subtypes. It was not possible to amplify any of the three typing genes in the remaining three samples. Figure 5.3 summarises these data and compares the success of strain-type identification from blood and ulcer specimens. The five full subtypes identified were all from ulcers with an average of 1798 *tpp047* copies/strip (std. dev. 1104) and of the 19 partial sub-types, 11 were from ulcers (mean 8789 copies/strip, std. dev. 14429) and eight from blood samples (mean 729 copies/strip, std. dev. 905). The difference in ulcer bacterial load between those yielding a full or a partial sub-type was not statistically significant (p=0.302). Ulcer samples taken in Sri Lanka allowed identification of 13 full and 11 partial strain types and a significant difference in bacterial load was observed between those samples yielding a partial (8482 *tpp047* copies/strip, std. dev. 17745) or a full subtype (44643 copies/strip, std. dev. 46255) (p=0.023).

tprEGJ MseI restriction digest patterns were identified in 5/19 (26.3%) samples from London and 13/24 (54.1%) of samples from Colombo. The only DNA samples from either country for which this technique was successful were ulcer-derived. A *tpp0548* gene sequence was identified in all 24 Sri Lankan samples analysed. As shown in Figure 5.4, *tpp0548* sequence analysis was successful in 20/24 (83%) samples from London. Of the four samples for which PCR amplification failed, two were low-load blood samples (six and 234 copies/ml) and a further sample was a low-load ulcer (113 copies/strip). The fourth sample was a high load (36720 copies/strip) ulcer sample. The size of the *arp* amplicon was established in 23/24 (95.8%) Sri Lankan strains and 13/24 (54.2%) of London strains. In London samples, *arp* amplification was achieved in 11 ulcer samples with a mean of 9444 *tpp047* copies/strip (st.dev. 14039) and failed in five ulcer samples (357 copies/strip, std .dev. 182) (p=0.178). Bacterial load in the failures did not exceed 648 copies/strip, whereas in all but two of the successes, bacterial load was in excess of 1024 copies/strip. Successful *arp* amplification was achieved in only A total of two blood samples containing an average of 1165 *tpp047* copies/ml (std. dev. 1475) yielded and *arp* type, and amplification failed in a further eight (mean 446 copies/ml, std. dev. 708) (p=0.312). Unlike ulcer samples, there was no clear cut-off above which successful *arp* amplification was more likely.

Finally, *23S rRNA* mutation analysis was achieved in all 24 *tpp047*-positive Colombo samples and 41/58 London samples (16 ulcer and 25 blood). The 41 samples on which the assay was successful included five blood-extracted and two ulcer-extracted samples for which all three strain-type analysis PCRs failed.

Figure 5.3 *T. pallidum* strain-type analysis from different sample types

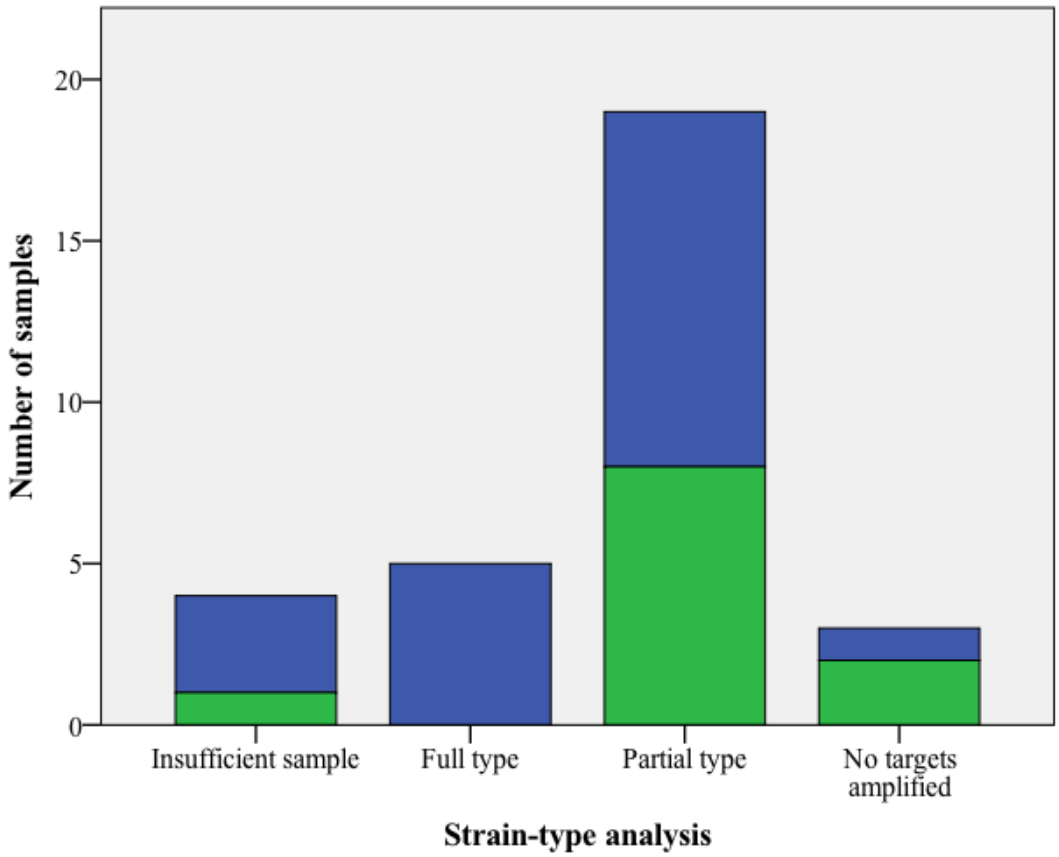


Fig. 5.3. *T. pallidum* strain type analysis requires analysis of three genes. Successful amplification of all targets yielded a 'full' subtype and identification of one or two targets, a 'partial' subtype. Partial or full *T. pallidum* strain-types were identified in 27 samples, comprising 16 ulcers (blue) and 11 blood samples (green). PCR amplification failed for all targets in three samples.

Figure 5.4 Assessment of an additional target for *T. pallidum* strain type analysis

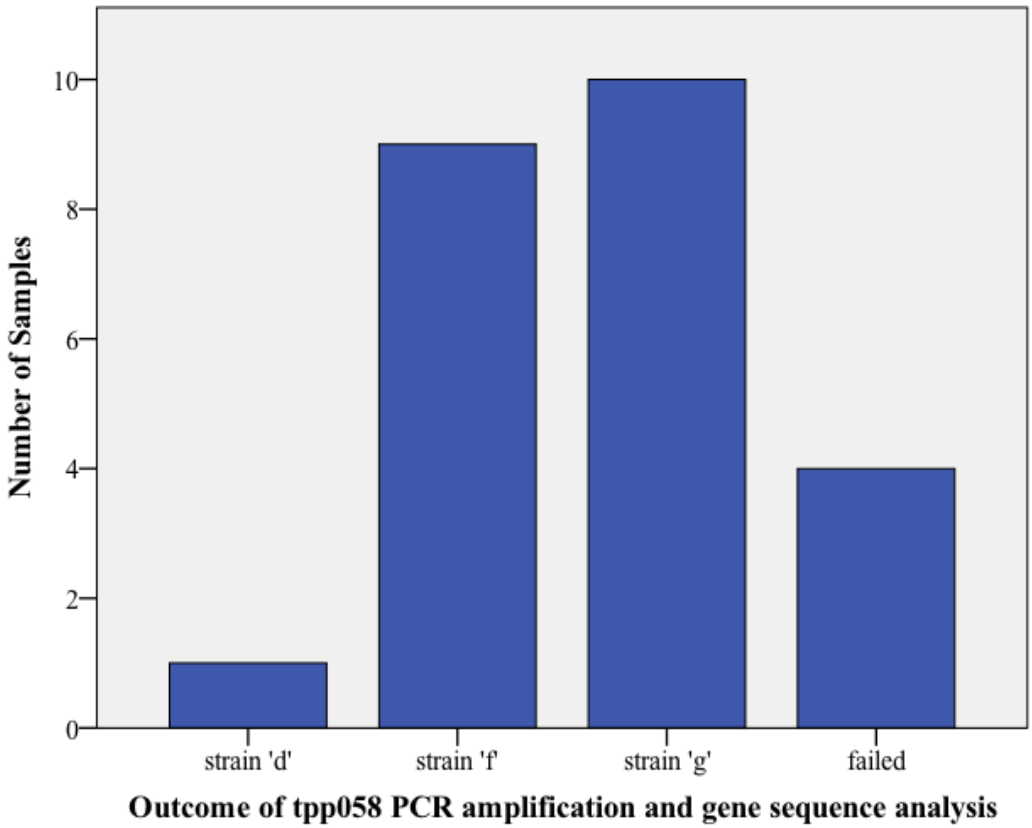


Fig. 5.4. The *T. pallidum* gene *tpp0548* was PCR-amplified and subject to gene sequence analysis to better discriminate between molecular sub-types. In total, *tpp0548* was successfully sequenced in 20 samples and three strains (‘d’. ‘f’ and ‘g’) were identified. Amplification failed in four samples.

5.2.4 Mutation analysis and strain-typing in London

5.2.4.1 T. pallidum macrolide resistance in London is of high and increasing prevalence

23S rRNA mutations conferring macrolide resistance were identified in 73% (30/41) of *T. pallidum* strains collected between July 2006 and November 2012. Whilst a single A2059G substitution was identified, the majority (29/30) of mutations were A2058G.
An increase in the prevalence of resistant *T. pallidum* strains from 66% (12/18) in 2006-8 to 78% (18/23) in 2011-2012 was observed but this was not statistically significant (p=0.50) (Figure 5.5).

All resistant strains were identified in samples from MSM of whom 62.1% (18/29) were HIV-1 infected and 86.6% (26/30) were of white British or white 'other' ethnic origin (Figure 5.6). The demographics of patients infected with resistant strains matched the overall study population. Therefore, while white ethnicity, sexual orientation and HIV status are seemingly associated with risk of *T. pallidum* infection, they cannot be analysed as risk factors associated with macrolide resistance.

Data concerning the use of antibiotics in the three months prior to syphilis diagnosis were available for 51.2% (21/41) patients. Of these, two had received macrolide antibiotics (clarithromycin or azithromycin) and both had genotypic evidence of resistance. One further patient with resistance had received the fluoroquinolone ofloxacin. None of the four patients infected with wild-type strains (and for whom antibiotic use data were available) had received antibiotics. However, mutant strains were identified in 14 patients who had not received antibiotics prior to their diagnosis.

Figure 5.5 Prevalence of macrolide resistant *T. pallidum* strains

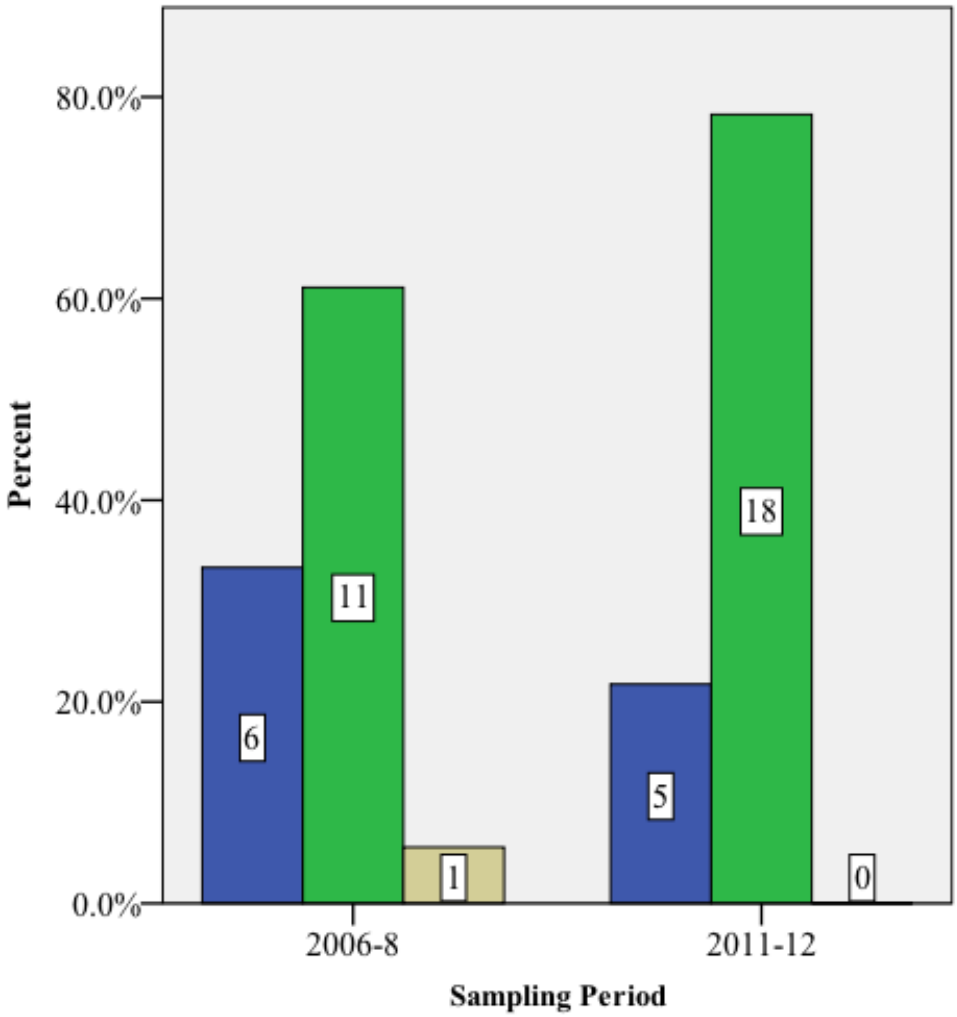
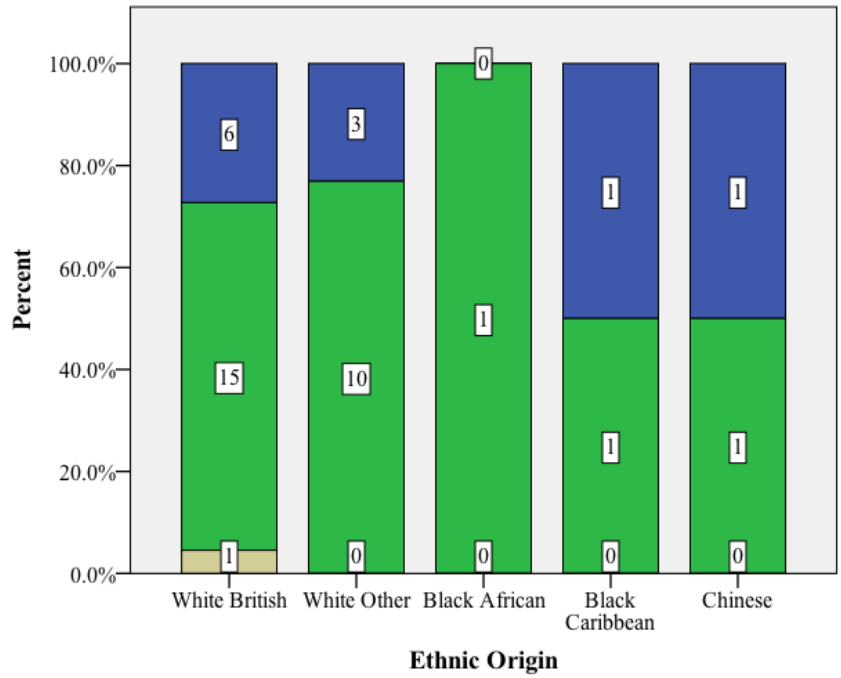


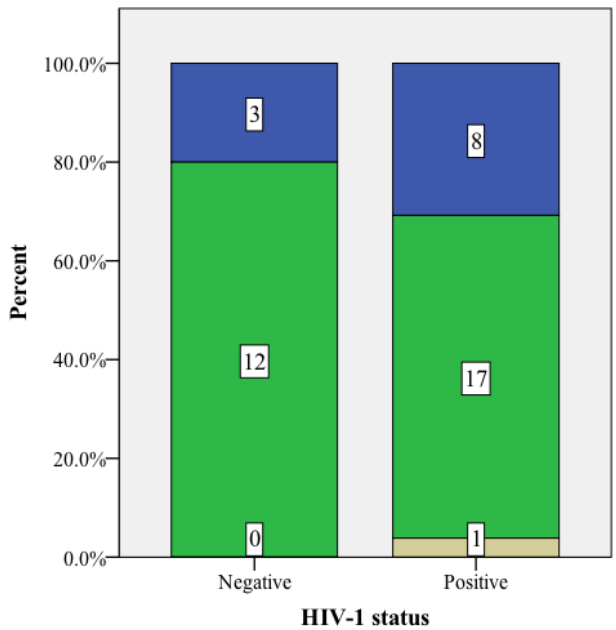
Fig 5.5. Blood and ulcer samples containing *T. pallidum* were subject to a point mutation assay for the identification of macrolide resistance conferred by A2058G and A2059G base substitutions in the *23S rRNA* gene. Samples were collected at St Mary's Hospital, London during two separate periods (July 2006 to March 2008 and January 2011 to December 2012) and the frequency of resistant and wild-type strains were determined. Blue bars depict wild-type strains; green bars show A2058G mutants and beige bars are wild-type strains. Sample numbers for each type are displayed within bars.

Figure 5.6 *T. pallidum* macrolide resistance according to demographic factors

A. Ethnic origin



B. HIV-1 co-infection



C. Sexual orientation

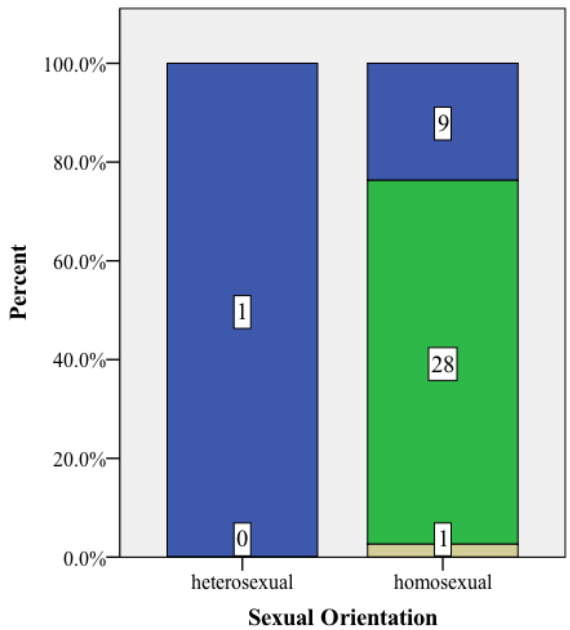


Fig 5.6. *23S rRNA* mutation analysis was performed on a total of 41 *T. pallidum* strains. Results are presented here according to patients' ethnic origin (**Panel A**); HIV-1 status (**Panel B)** and sexual orientation (**Panel C**). Blue bars depict wild-type strains; green bars show A2058G mutants and beige bars are wild-type strains. Sample numbers for each type are displayed within bars.

5.2.4.2 Analysis of T. pallidum strain types

All strain types identified in London, including 23S rRNA analysis are listed in Table 5.1 and summarized in Figure 5.7. A total of three unique complete (ie. successful amplification and analysis of the genes *arp, tpp0548 and trpEGJ*) strain types (two 14d/g, one 14d/f and two 14g/g) and a further seven partial strain types were identified. All 13 samples for which *arp* amplification was possible had 14 tandem repeats and of the five identifiable *tprEGJ* digest patterns three were 'd' and two were 'g'.

The *23S rRNA* A2058G mutation was found in each of the three unique full sub-types. No sub-type analysis was possible from the single A2059G mutant strain. The A2058G *23S rRNA* mutation was found in all 10 samples with the *tpp0548* 'g' sequence and in only one third of type 'f' strains with the remainder being wild-type.

All full and the majority of partial strain types identified include a *tpp0548* gene sequence. Figure 5.4 demonstrates sequences 'f' and 'g' to be the most prevalent at 45% and 50% respectively, and that a single strain was found to have sequence 'd' (5%). No demographic differences were noted between patients with *tpp0548* 'f' or 'g' gene sequences and HIV-1 status, ethnic origin, gender, sexual orientation.

Table 5.1: *T. pallidum* strain-typing and mutation analysis in London

Study ID	Disease Stage	Sample Type	Macrolide Type	Strain Type
Patients recruited from qPCR validation study (2006 to 2008)				
23	Primary	Ulcer Swab	A2058G	14d/g
24	Primary	Ulcer Swab	A2058G	14d/g
26	Primary	Ulcer Swab	wild-type	++/+
32	Primary	Ulcer Swab	A2058G	14g/g
33	Primary	Ulcer Swab	A2058G	xx/g
52	Primary	Ulcer Swab	A2058G	14d/f
62	Primary	Ulcer Swab	wild-type	14x/f
69	Primary	Ulcer Swab	A2058G	xx/g
89	Primary	Ulcer Swab	A2058G	14g/g
90	Primary	Ulcer Swab	wild-type	14x/f
93	Primary	Ulcer Swab	wild-type	xx/f
22	Secondary	Ulcer Swab	A2058G	xx/g
34	Secondary	Ulcer Swab	wild-type	xx/f
27	Secondary	Whole Blood	A2058G	++/+
45	Secondary	Whole Blood	A2058G	xx/x
53	Secondary	Whole Blood	wild-type	++/+
61	Secondary	Whole Blood	A2058G	++/+
81	Secondary	Whole Blood	A2059G	++/+
Samples from patients recruited to the CDTB (2011 to 2012)				
109	Primary	Ulcer Swab	A2058G	14x/g
108	Primary	Whole Blood	wild-type	xx/f
127	Primary	Whole Blood	A2058G	14-/-
129	Primary	Whole Blood	A2058G	14-/x
107	Secondary	Ulcer swab	A2058G	14x/d
100	Secondary	Whole Blood	A2058G	xx/f
103	Secondary	Whole Blood	A2058G	xx/g
105	Secondary	Whole Blood	A2058G	xx/g
106	Secondary	Whole Blood	wild-type	xx/f
110	Secondary	Whole Blood	A2058G	xx/f
115	Secondary	Whole Blood	A2058G	--/-
124	Secondary	Whole Blood	wild-type	--/-
134	Secondary	Whole Blood	A2058G	14-/-
138	Secondary	Whole Blood	A2058G	--/-
140	Secondary	Whole Blood	A2058G	--/-
151	Secondary	Whole Blood	A2058G	--/-
154	Secondary	Whole Blood	A2058G	--/-
155	Secondary	Whole Blood	A2058G	--/-
113	Latent	Whole Blood	wild-type	xx/x
123	Latent	Whole Blood	wild-type	--/x
128	Latent	Whole Blood	A2058G	14-/-
136	Latent	Whole Blood	A2058G	--/-
142	Latent	Whole Blood	A2058G	--/-

Table 5.1. Blood and ulcer samples collected during two cross-sectional studies in London were subject to molecular analysis to identify 23S rRNA mutations associated with macrolide resistance and T. pallidum strain types. Strain types are presented in the order: *arp* repeats, *tprEGJ* digest pattern (*tpp0548* sequence). E.g. 14d/d. The symbol 'x' denotes a typing target which failed to amplify; '+' indicates that insufficient sample was available for the analysis and '-' signifies that amplification was not attempted.

Figure 5.7 *T. pallidum* sub-types identified in London

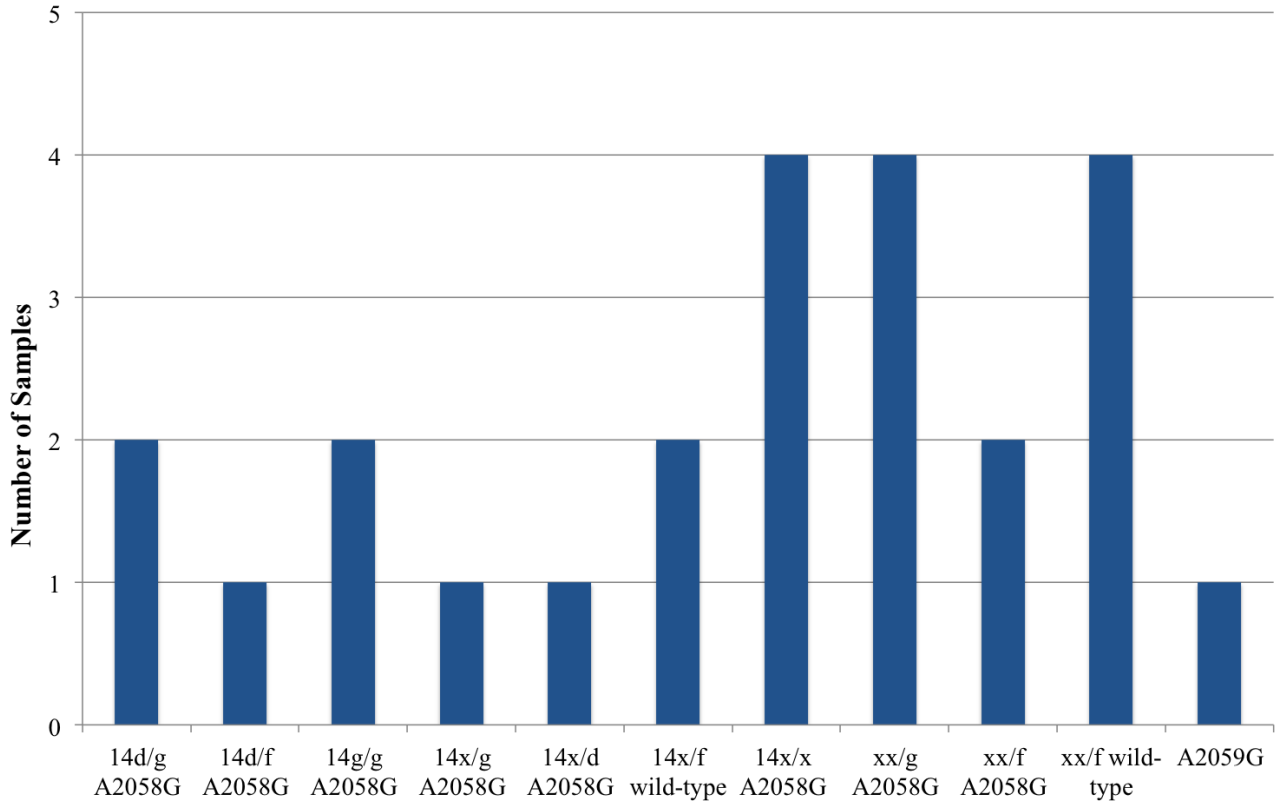


Fig 5.7. Molecular analysis of 24 T. *pallidum* DNA samples identified five full and 19 partial sub-types. The four components of the sub-type are presented in the order: *arp* repeats; *tprEGJ* digest pattern; *tpp0548* sequence and *23S rRNA* mutation analysis. Eg. 14d/d A2058G. Missing components replaced with 'x'

5.2.5 Strain-typing and mutation analysis in Sri Lanka

A total of 24/25 samples collected from the central STD clinics of Colombo, Sri Lanka contained *T. pallidum* DNA sufficient for mutation and strain type analysis. *T. pallidum* DNA was not detected in one sample, collected from a female patient with a DGM positive labial ulcer and serology consistent with a diagnosis of early syphilis. PCR amplification of *tpp047* and RT-PCR of *16S rRNA* both failed.

In total, 11 partial and 13 full strain-types were identified and are listed in Table 5.2. The full sub-types comprised a single 13b/f strain and twelve (92.3%) 14b/f strains. Similarly, of the partial sub-types characterised, the majority (10/11, 90.0%) were 14x/f and a single 13x/f strain was found. When analysed for 23S rRNA mutations, neither substitution (A2058G or A2059G) was identified in any of the 24 strains.

Two patients were infected with a strain containing 13 *arp* tandem repeats, whereas the majority of strains had 14 (Figure 5.8). These two strains were otherwise similar, both possessing a *tpp0548* 'f' sequence and neither having a mutated *23S rRNA* gene. The patients were both men of a similar age and Sinhalese (Table 5.3). Patient eight gave a history of travel to Saudia Arabia and Germany and of sex with other men. Patient 13 had never left Sri Lanka and was heterosexual.

Figure 5.8 Determination of the size of *T. pallidum arp* amplicons in Sri Lankan strains

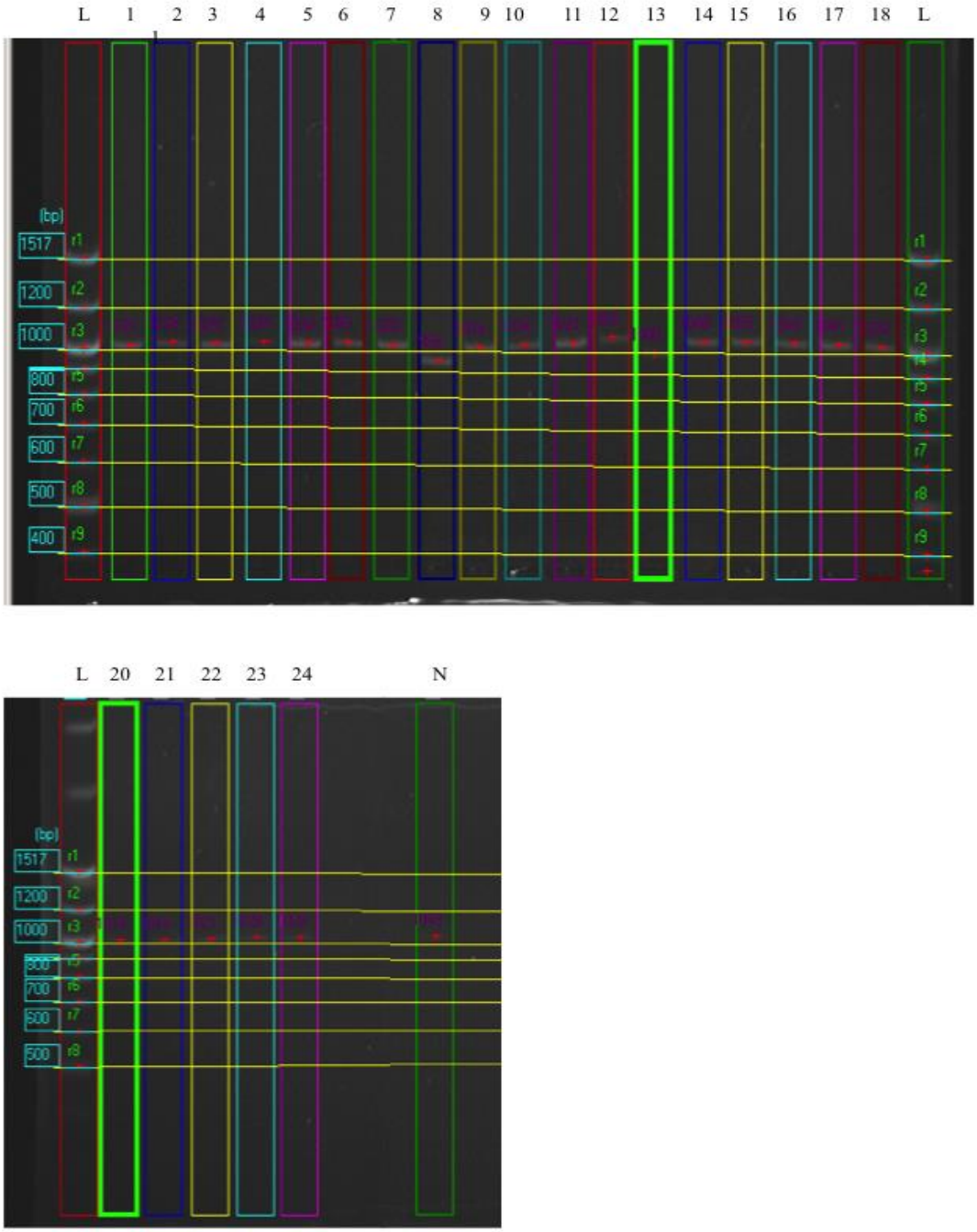


Fig 5.8. The *arp* gene from Sri Lankan *T. pallidum* strains was PCR-amplified and visualised with SYBRsafe stain and agarose gel electrophoresis. Using a molecular marker and LabWorks Image Acquisiton Software sample band sizes were determined and compared with that of the Nichols reference strain. L=Ladder, N=Nichols (reference) *T. pallidum* strain.

Table 5.2 Molecular analysis of *T. pallidum* strains in Sri Lanka

Study ID	Disease Stage	Macrolide Type	Strain Type
1	Primary	wild-type	14b/f
2	Primary	wild-type	14x/f
3	Primary	wild-type	14b/f
4	Secondary	wild-type	14x/f
5	Secondary	wild-type	14b/f
6	Secondary	wild-type	14x/f
7	Secondary	wild-type	14b/f
8	Primary	wild-type	13b/f
9	Secondary	wild-type	14b/f
10	Secondary	wild-type	14x/f
11	Primary	wild-type	14x/f
12	Secondary	wild-type	14x/f
13	Primary	wild-type	13x/f
14	Primary	wild-type	14b/f
15	Secondary	wild-type	14b/f
16	Primary	wild-type	14b/f
17	Primary	wild-type	14b/f
18	Primary	wild-type	14b/f
19	*No T. pallidum DNA was detected in sample*		
20	Primary	wild-type	14b/f
21	Primary	wild-type	14b/f
22	Primary	wild-type	14x/f
23	Secondary	wild-type	14x/f
24	Secondary	wild-type	14x/f
25	Primary	wild-type	xx/f

Table 5.2. Ulcer samples were collected from patients with early syphilis in Colombo, Sri Lanka. The presence of *T. pallidum* (*tpp047*) DNA was confirmed and samples were then subject to molecular analysis for strain-type identification and to determine the prevalence of 23S rRNA mutations which confer macrolide antibiotic resistance. Strain Types data is presented in the order: *arp* repeats, *tprEGJ* digest pattern and *tpp0548* sequence. Eg. 14d/d.
'X' denotes amplification failure of target to amplify.

Table 5.3 Clinical details of two Sri Lankan patients

	Patient 8	Patient 13
Age	27	25
Gender	Male	Male
Ethnic Origin	Sinhalese	Sinhalese
Disease Stage	Primary	Primary
Sexual Orientation	Bisexual	Heterosexual
Travel outside Sri Lanka (ever)	Saudi and Germany	None
Referral Source	GP referral	Self-Presentation

Table 5.3. Two patients were found to have been infected with *T. pallidum* strains containing 13 *arp* gene tandem repeats, whereas the remaining 22 strains identified contained 14 repeats. Demographic data collected for these two patients is presented including history of travel outside Sri Lanka.

5.3 Discussion

The data presented in this chapter are the first description of *T. pallidum* macrolide antibiotic resistance in both the United Kingdom and Sri Lanka and have implications for syphilis treatment guidelines. Associated strain-typing data both enhances understanding of the origins of macrolide resistance and helps to inform the choice of samples and techniques for *T. pallidum* molecular characterisation.

Several studies have presented *T. pallidum* macrolide resistance data since 2004 when a 23S rRNA point mutation (A to G at position 2058) was linked definitively with phenotypic macrolide resistance (168-172, 176, 266). In London, *T. pallidum* samples were collected during two separate studies. The first (qPCR validation) was conducted in 2006-8 and the second (CDTB sample collection) in 2011-12. In the three years between studies, the prevalence of macrolide resistance, which was already 66% in 2006, increased by 14% to near 80%. This increase was not statistically significant but similar prevalence data have been reported in San Francisco, where resistance rose from 4% to 56% percent between 2000-2002 and 2004, and British Columbia where a temporary fall in syphilis cases following an azithromycin intervention in 2001 was not sustained. (171, 176). The populations affected by syphilis on both sides of the Atlantic are also similar, being predominantly white MSM with high (30%) rates of HIV-1 co-infection (1, 171). Patients with syphilis sampled in Sri Lanka also bore a resemblance to their North American and British counterparts being mostly male and 50% MSM. In stark contrast however, no macrolide resistant strains were identified in any samples collected in Colombo during this study.

Two theories have been suggested to explain the rapid spread of macrolide resistance in *T. pallidum* (266). The first is that a strain with a pre-existing 23S rRNA mutation is introduced into a small and well-defined population and becomes endemic. The strain ‘street 14’, which carries the 23S rRNA mutation A2058G, was isolated from a patient in San Francisco in 1977 and despite multiple passages in rabbits (in the absence of antibiotic pressure) the mutation has remained stable (167, 266). It is conceivable that this strain, or another like it, spread within the city’s MSM population during the 80s and 90s and resulted in treatment failure among patients treated with azithromycin following inclusion of the drug in clinical guidelines in 2000 (168, 171). The prevalence of resistance then increased as selection

pressure singled out this clone. According to this theory a mutant strain would have been introduced into the London MSM population (8% of infections are thought to be acquired abroad) and would have come to predominate under selection pressure from wide-spread azithromycin use in this population for genital *Chlamydia trachomatis* infection and non-specific urethritis treatment (1). By the same analysis, the absence of a resistant strain in Colombo, and the relative seclusion of the Sri Lankan population (only two patients in our study had ever been outside Sri Lanka) could explain why no macrolide-resistant strains were identified. Further support for this theory is found in Madagascar, where none of 141 strains studied contained macrolide resistance mutations, and where sexual networks in which syphilis is transmitted are distinct from those of North America and Europe where MSM predominate (175).

Strain-typing data available from this and other studies support a second theory on the origins of macrolide resistance. If the result of endemic spread of a mutant clone in a given sexual network, then a single resistant strain-type would be expected within that population. It has been demonstrated widely, however, that epidemics comprise several strains both with and without resistant genotypes (96). More specifically, when 20 resistant strains collected in King County, Washington, USA were examined for the presence of an insertion in the *tpp0126-tpp0127* intergenic spacer it was only present in 80%, signifying at least two clones (266). The results of strain-type analysis in the current study showed three full sub-types (14d/g, 14g/g and 14d/f), which were all A2058G mutants. This indicates that resistance is not the result of a single endemic clone, rather a number of resistant strains circulating in London's MSM population, and that resistance may be developing *de novo* as a result of selection pressure. Regarding antibiotic use, two patients in this study had received a macrolide antibiotic in the three months prior to their syphilis diagnosis, and both were infected with a resistant strain. Moreover, of the four patients with wild-type strains for whom antibiotic use data were available, none had recently received a macrolide. This association is supported by a previous observation that the relative risk of harbouring a resistant strain was 2.2 in patients who had received a macrolide in the year prior to a diagnosis of syphilis. Given that the prevalence of resistant strains is now 80%, and knowing that the mutant 23S ribosome is stable enough to persist in the absence of antibiotic pressure, the majority of patients with resistance in this study were likely infected with an already resistant strain, of which at least three are currently circulating. In order to assess the true impact of resistance developing *de*

novo, samples prior to 2006 would be required to demonstrate much lower resistance prevalence in the context of similar strain diversity.

In addition to helping identify the origin of antibiotic resistance in a population, strain-typing data are useful to map epidemics and to identify populations at risk of disease. The data presented here are the first to describe *T. pallidum* subtypes in England. Data from Scottish samples collected in 2006-7 revealed six subtypes among 58 typeable specimens, the majority of which (80%) were collected from white MSM. The commonest subtype was 14d (76%), followed by 14e (12%), 14j (5%), 14b (3%), 14p and 14k (2% each) (87). Given that the population sampled was relatively homogenous, the authors felt the strain diversity surprising, although no other epidemiological data were available to explain it. The predominance of syphilis among MSM in Scotland is reflected more generally in the UK and in other high income countries. The London study population presented here is a relatively homogenous group of white MSM thus the majority of patients infected with resistant strains are also white MSM. UK national data demonstrate that 27% of syphilis diagnoses of any stage are found in men with HIV-1 co-infection (1). Moreover, a phenomenon known as, 'sero-sorting' has been described where people choose sexual partners of the same HIV serostatus in order to engage in unprotected sex resulting in a 61% increase in risk of STI transmission (267). The potential exists, therefore, to identify different *T. pallidum* strains from HIV-1 infected and un-infected men as their sexual networks are relatively discrete. However, no evidence was found in this study for such a separation as strains identified from both groups were similar across all three strain-typing targets and were equally likely to harbour *23S rRNA* mutations.

In contrast to London, *T. pallidum* strain-diversity in Colombo was limited. Just two strain types were identified (14b/f and 13b/f) neither being macrolide resistant. This limited strain diversity may be a reflection of the limited population sampled, which was similarly-aged Sinhalese men. Moreover, whilst half of the patients identified as MSM the majority were bisexual, not homosexual, making sexual networks less distinct. It is also of note that only two patients gave a history of any foreign travel, so the potential for the introduction of new strains is limited. By contrast, the population sampled in London was ethnically diverse with around half of patients describing non-British ethnicity, which could explain the increased diversity of *T. pallidum* strains observed. A recent meta-analysis compared strain-types identified in eight countries and found between four and fourteen circulating strains. The majority of strains worldwide carry 14 *arp* tandem repeats and strain type 14d, thought to be the original circulating subtype in many parts of the world, predominates (83). Whilst this is

true of the London strains identified, all containing 14 *arp* tandem repeats and being either tprEGJ ‘d’ or ‘g’, strains from Colombo were more unique. The most common of the two types identified was 14b and while less frequent world-wide it has been reported previously in China, South Africa, Madagascar and Scotland (82, 87, 89). Type 13b, however, was described in this study for the first time. Further examination of the two infections caused by type 13b and 13x strains in Colombo revealed a history of foreign travel (to Germany and Saudi Arabia) in one patient but not the other and different sexual networks (one homosexual and the other heterosexual). No typing data are available from Germany or Saudi to demonstrate that this strain was imported from either of those countries, but this remains a possibility. It may also be the case that strain 13b is more prevalent in Colombo than the small number of patients sampled for this study can demonstrate.

The discussion of *T. pallidum* strain-types has focussed mainly on the CDC typing method which characterises two typing targets (*arp* and *tprEGJ*) (82). Addition of *tpp0548* gene sequence analysis to the typing algorithm was recently shown to improve discrimination between sub-types, turning 14 CDC types into 25 ‘enhanced’ sub-types (86) and this study presents the first enhanced typing data in Europe. Whilst *tpp0548* sequence analysis did not identify any additional strains in Colombo, it sub-classified two 14d types into 14d/f and 14d/g in London and identified a unique partial strain-type which possesses a *tpp0548* ‘d’ sequence, adding further to the correlation between population and strain diversity observed in both countries.

An additional typing target, especially one requiring gene sequence analysis, adds both time and expense to strain identification. This study demonstrates, however that the success of *tpp0548* PCR-amplification and sequencing far exceeds that of the original two CDC typing targets. For example, if *arp* and *tpp0548* analysis alone were used to identify strain types, an additional 10 full types would have been described in Colombo. Moreover, the success of *tpp0548* amplification from whole blood samples matched that from ulcer samples making it useful in latent and secondary disease. Of the four *tpp0548* amplification failures across all studies, two were from low-load blood samples and one from a low-load ulcer sample. The final amplification failure was from a high-load ulcer sample, perhaps inhibited by the high starting DNA concentration (143). *tpp0548* sequence analysis is therefore validated in these studies as a useful and effective *T. pallidum* strain-typing target.

Far less successful was the amplification by PCR of *arp* and particularly of *tprEGJ* genes. Several previous studies have reported partial strain type results as a result of *arp* amplification failure, which is unsurprising given its high GC content and the presence of a number of tandem repeats (264, 265, 268). Efforts made here to improve the amplification of *arp* were generally unsuccessful. The addition of betaine and Roche GC rich solution which aim to reduce secondary structure in the *T. pallidum* template DNA failed to improve reaction efficiency, and likely inhibited it further. Increasing the concentration of the reaction primers and of taq polymerase, however, was found to increase PCR efficiency. *MseI* restriction digest analysis of *tprEGJ* was found to be the least reliable typing target and was successful in 26.3% and 54.1% of samples from London and Colombo, respectively. All of these strains were derived from ulcer samples, which were also found to be the best sample type for *arp* amplification. An observation made previously and confirmed here is that the improved detection of *T. pallidum* typing targets in ulcer samples is likely the result of higher bacterial load (83). *tpp047* copies in ulcer samples were 26 times more numerous for *arp* amplification successes compared with failures. It was also demonstrated that *arp* amplification from blood-extracted samples was more successful at higher bacterial loads. Overall, full strain-types were only identified from ulcer samples, and the majority of partial strain types from blood samples tended to be only a *tpp0548* sequence. The identification of macrolide resistance from blood samples was more successful than strain-typing and results are available for five patients with latent syphilis. This is perhaps unsurprising as amplification of the two *23S rRNA* alleles proceeds with two rounds of PCR in a 'nested' reaction which can theoretically detect a single gene copy.

There are two main avenues of future research to go down in this field. The first is to recognise the power of next-generation sequencing techniques (e.g Illumina/Solex and ABISOLiD paltforms) for the identification of *T. pallidum* strain types. Whole genome sequences of strains would allow sequence-based analysis of bacterial strains which would enhance both strain-discrimination and the ability to correlate strain types with clinical phenotypes, such as the association made between subtype 14d/f and neurosyphilis (86). The second, is to be vigilant for the development of *T. pallidum* resistance to other antibiotics, specifically tetracyclines. *T. pallidum* has a small genome which most likely does not include plasmid DNA, thus the only mechanism available to the organism for developing resistance is to mutate genes encoding targets antibiotic targets (28). Resistance to macrolides was

conferred in this way, and *16S rRNA* mutations have been shown to cause tetracycline resistance in other bacterial species (160).

In summary, 23srRNA mutations conferring macrolide resistance in *T. pallidum* have been discovered to be common in MSM in London and caution is recommended in the use of macrolides for the treatment of incubating or early syphilis in the MSM population as well as for other populations. Although limited, typing data suggests that the epidemic in London is not a single clone of resistant *T. pallidum*, raising the possibility that resistance mutations are developing *de novo*, secondary to the selective pressure of macrolide use for other indications. There is currently no evidence of *T. pallidum* macrolide resistance strains in Colombo, Sri Lanka and strain-diversity there is limited. Given the rapid spread of macrolide resistance in other populations and the small sample size of this study, caution in using single dose azithromycin for the treatment of early syphilis in Sri Lanka would seem prudent. The identification of macrolide resistance is possible in both blood and ulcer samples in early disease, but also in latently infected patients with detectable bacteremia. Ulcer samples, which generally contain more DNA, are better for *T. pallidum* strain type analysis than blood and *tpp0548* gene sequence has been shown to be a reliable typing assay, which enhances strain discrimination.

Chapter 6

Discussion

General conclusions

In 1943 penicillin was discovered to be an effective treatment for syphilis and its widespread use over the following decade led to near-eradication of the disease (18). However, 70 years on syphilis is resurgent in the UK in an epidemic characterised by male-to-male transmission in urban centres that is paralleled in a number of high-income countries (1). In the post-penicillin years, the field of syphilis research made slow but steady progress despite continuous culture of *T. pallidum* remaining elusive. Specific antibody tests for clinical use have been developed; the enigmatic bacterial surface is beginning to yield its secrets and provide insight into the organism's success in evading the immune response and the genome of several *T. pallidum* strains are now available and providing the answers to key questions about the organism's metabolism and pathogenicity (14, 72, 247). Molecular strain type analysis has also shown the potential to identify neuro-invasive phenotypes and is providing useful information to the syphilis epidemiologist (95). Other antibiotic classes, namely macrolides, tetracyclines and cephlasporins have also been demonstrated to have activity against *T. pallidum* and are used as syphilis treatments (160). A recent shortage of benzathine penicillin, however, highlighted the fragility of current syphilis treatment strategies given that the evidence for using cephalosporins is weak; macrolide resistance is now widespread and doxycycline cannot be used in pregnancy or for treating children with congenital disease (104, 269). Moreover, the new syphilis epidemic is closely related to that of HIV-1, each disease impacting on the clinical course of the other and 30% of those with syphilis being HIV-1 co-infected (199). Two randomised controlled trials have compared standard benzathine penicillin syphilis treatment with oral azithromycin and both demonstrated equivalent efficacy. However, in the absence of a speedier alternative both trials were reliant on standard serological outcomes to diagnose cure, one at six months and the other at nine months following treatment (161, 219). This thesis sought to describe an alternative method for the assessment of treatment success in patients with early syphilis.

The first step was to design assays capable of detecting and quantifying *T. pallidum* DNA and RNA and to validate these assays using clinical samples. This required the selection of robust methods for nucleic acid extraction; identification of the best sample types and sensitive assays for their detection. Early attempts at *T. pallidum* PCR recognised the need for efficient DNA extraction, given the potential for PCR inhibition by haem, proteins and other biological

molecules contained in clinical samples (130). DNA extraction from blood was found to be efficient with a proprietary DNA extraction kit modified by adding a proteinase K digestion step to ensure sufficient bacterial lysis. For blood RNA extraction the Tempus™ system was selected, on account of ease, speed and extraction efficiency. There is currently no consensus on which blood fraction is the most reliable for *T. pallidum* detection (52). The bacterium is known to bind to many nucleated cell types and is extracellular, thus could potentially be detected equally well in cellular and non-cellular blood fractions (59). The current validation study identified whole blood as the best sample type and was supported by data from both RIT testing (reliant on the presence of viable organisms), prior to the discovery of PCR, in which whole blood has been found more likely to harbor *T. pallidum* than plasma and an ex vivo experimental model where viable *T. pallidum* were spiked into blood and their distribution among fractions assessed (51, 126). The current study describes the development of four qPCRs: two for *tpp047* DNA detection, a multiplex assay targeting TYLCV and *tpp047* simultaneously, and an RT-qPCR for 16S rRNA quantification. Each assay was deemed to have an acceptable analytical sensitivity, all were specific for *Treponema pallidum* subspecies and quantified target DNA by comparison with in-run absolute quantification standards with acceptable efficiency. Furthermore, the assays were assessed according to published guidelines for the minimum information for publication of quantitative real-time PCR experiments (MIQE) and all except two of the essential criteria for publication were met (254). The two unmet criteria resulted from non-measurement of nucleic acid concentration and purity following extraction and prior to qPCR. The MIQE guidelines also recommend the use of internal controls of DNA recovery and PCR efficiency thus an attempt was made to use a recombinant plasmid containing TYLCV sequence in this role. While results were encouraging, the technique would benefit from further optimisation to improve inter-assay reproducibility. The use of PCR has also been described for the diagnosis of a number of other bacterial infections and is now used widely for the rapid identification of *Mycobacterium tuberculosis* (270, 271). The particular advantage of PCR is that it can detect very small amounts of bacterial DNA without the need for weeks of bacterial culture.

T. pallidum DNA detection in samples from two cross-sectional studies was found to have a sensitivity and specificity similar to previously described assays. In ulcer samples, collected both in the UK and Sri Lanka, detection sensitivity was 100% and 96%, respectively. Compared with DGM, PCR identified an additional 16 cases of syphilis (using final clinical diagnosis as the gold standard). It is important to note, however, that in the hands a skilled

operator DGM will correctly identify *T. pallidum* in 80% of primary chancres and provides a point-of-care diagnosis in minutes whereas PCR takes at least four hours, and given laboratory logistics, potentially days (247). PCR may have a role in the diagnosis of early syphilis when DGM is not available, or there is concern of contamination by commensal treponemes, but should not, be considered a replacement for DGM. The quantity of DNA detected in lesions correlated with both ulcer area and duration, thus larger and earlier ulcers should potentially be considered more infectious to sexual partners. It was also notable that ulcers containing greater amounts *T. pallidum* by PCR were more likely to be DGM positive, further validating the quantitative capability of the technique. The level of RNA in ulcer samples far exceeded that of DNA, thought to be the result of increased metabolic activity of organisms in the primary chancre thus a greater number of 16S ribosomes. A trend was also noticed towards patients with HIV-1 co-infection having higher *T. pallidum* bacterial loads (DNA) in primary chancres. This fits well with previous observations that HIV-1 infected patients are more likely to present with multiple and deeper chancres. Moreover, that chancres take longer to heal in the context of HIV-1 infection (209).

The sensitivity of *tpp047* DNA detection in blood samples during primary, secondary and latent syphilis was also similar to previously published data (52, 83). Sensitivity was greatest in secondary disease and a PCR-positive blood sample in the current study had a 78.4% chance of coming from a patient in the secondary stage. It was also notable that only patients diagnosed with secondary syphilis had blood bacterial loads in excess of 1303 copies/ml. Of course, a number of patients with secondary disease had an undetectable blood bacterial load, thus the assay could not function as a diagnostic test. It is likely, although difficult to prove, that these patients had a longer duration of symptoms and were progressing towards *T. pallidum* immune clearance and disease latency. It was also noteworthy that a quarter of asymptomatic patients diagnosed with latent disease had low-level bacteraemia. This may represent patients with early latent disease who are either poised to clear the bacteraemia, or are heading towards disease relapse, which is common within a year of latency (103). The current rationale behind latent disease treatment is that those patients in late latency receive a three-week course of penicillin as a longer duration of infection results in attenuated, slow-growing *T. pallidum* which take longer to kill. Those in early latency receive a single dose of penicillin. The cut-off between early and late latent disease is an arbitrary year in the USA and two years in the UK (104, 272). A potential role for *T. pallidum* qPCR in latent disease could therefore be in defining which patients require single-dose or multi-dose treatment,

although this strategy would need confirmation in a larger clinical study. *T. pallidum* qPCR cannot help, however, to differentiate between patients with latent non-bacteraemic infection and those with previously treated disease as both patient groups will have serological tests comprising a positive treponemal antibody and a negative or 'serofast' non-treponemal test result. In current clinical practice, this question is often addressed by differentiating between IgM and IgG antibodies with EIA assays. Whereas IgG persists into late disease, IgM is often negative thus differentiating between these two antibody types can help to determine disease stage (104). Another potential strategy to address this problem is detailed in the future work section below.

Analysis of blood bacterial load in patients with and without HIV-1 co-infection revealed no trend towards HIV-1 infected patients having higher-load disease. It is noteworthy that HIV-1 infected patients in this cross-section had a mean CD4 count of 603 cells/mm^3 (std. dev. 160) thus could not be considered to be highly immunosuppressed. Previous studies of HIV-1 and syphilis co-infection have noted that CSF abnormalities are more frequent at $CD4^+$ T-lymphocyte counts under 350 cells/mm^3 and that while changes to the presentation of early disease were more common, there were no significant differences in syphilis treatment outcomes between HIV-1 infected and HIV-1 uninfected patients (212, 262). Moreover, the majority of case reports describing more severe syphilis in HIV-1 infected individuals appeared in the pre-anti-retroviral era in the context of (sometimes profound) immunosuppression. It is still the case, however, that syphilis results in a transient fall in $CD4^+$ T-cell count and rise in HIV-1 plasma viral load, even in patients with suppressed virus on anti-retroviral treatment and genital ulcer disease remains a risk factor for both HIV-1 transmission and acquisition thus the interaction between the two disease cannot be considered insignificant (214, 273).

The quantification of *T. pallidum* RNA in both blood and ulcers was described in the current study for the first time. When compared, RNA levels were greater at all stages of disease than DNA. This is not surprising considering a great many more copies of the 16S rRNA target are present in each organism that the single copy of *tpp047* DNA (153). The greater abundance of the RNA target also improved the sensitivity of *T. pallidum* detection in blood for primary and secondary disease where three samples from each stage were positive for 16S rRNA, but negative for *tpp047*. In only one case (a patient with primary disease) was a blood sample RNA negative but DNA positive. Detection during secondary disease, where bacterial load is

higher, was equal for both DNA and RNA. The variability of RNA levels was high, explained partly by differences in the efficiency of reverse transcription, but also in the metabolic activity of bacteria. This effect was most marked in primary disease ulcers from the Sri Lankan cross-sectional study where RNA levels greatly exceeded those of DNA.

These observations raise the question of how disease stage should be classified in the absence of symptoms but, in the presence of increased bacteraemia, especially given that RPR results did not correlate with blood bacterial load. Current dogma dictates that latent disease is a less active stage, characterised by slowly dividing bacteria and requiring prolonged treatment (three doses of benzathine penicillin over three weeks). By contrast, secondary disease requires only a single dose of benzathine penicillin (104). It could be argued, therefore, that asymptomatic, yet bacteraemic, disease is a different disease entity, with greater similarity to the secondary than the latent stage. It remains to be proven, however, that this new stage would respond adequately to a single dose of penicillin. It is also important to note that 25% of patients with early disease in the Oslo study were noted to develop relapses of secondary disease and that a further quarter of relapsers experienced more than one relapse (103). An alternative hypothesis, therefore, is that those patients with latent disease and detectable bacteraemia would, if left untreated, relapse. In these cross-sectional studies, patients with latent disease were, by definition, asymptomatic and all were treated on or around the time of sample collection.

It was also demonstrated that levels in primary ulcers were much higher than those in secondary lesions reflecting a smaller number of bacteria, diminished metabolic activity, or, in the latter, both. It must be noted, however, that ulcer sampling with the SnoStrip technique is somewhat operator dependent, despite attempts to standardise collection. Overall, it seems that the qualitative detection of *T. pallidum* is improved using the 16S rRNA target. Moreover, the Tempus™ system makes sample collection and RNA extraction quick and simple. For quantification, however, the greater variability of RNA levels must be weighed against improved sensitivity of detection.

Following validation of *T. pallidum* quantification in blood and ulcer samples, the techniques were applied to a longitudinal study to assess the clearance of *T. pallidum* following treatment. Patients with symptomatic early disease were selected, and the rate of bacterial clearance from blood and ulcer samples monitored following the injection of benzathine

penicillin. After a mean of 35 hours, none of the four patients studied still had detectable *T. pallidum* bacteraemia, measured by either qPCR or RT-qPCR. The increased variability of RT-qPCR was reflected in a standard deviation of RNA levels twice that observed for DNA, however, by calculating the half-life of clearance RNA was shown to clear significantly more quickly. This confirmed the hypothesis that the instability of RNA would make it the more dynamic measure of clearance despite much larger starting quantities. Ulcer *T. pallidum* clearance occurred over a similar period although the difference between DNA and RNA-measured clearance was not significant. Two other important observations were made. Firstly, all four patients had serological results consistent with cure, supporting the hypothesis that bacterial clearance can predict cure. Secondly, through longitudinal sampling of an additional four patients for up to 14 days it was evident that once cleared, there was no recrudescence in detectable bacteremia. Finally, the ability to observe patients during treatment also made observation on the JHR possible. Anecdotally, the patient with the lowest blood bacterial load was the only one not to experience a JHR and the strongest reaction was in a patient with a high blood bacterial load (and a very high ulcer load). This observation, correlating strength of JHR with bacterial load remains to be proven in a larger observational study of bacterial loads, although JHR is known to occur with greater frequency and severity in early disease compared with late (190).

The identification of genetic differences between *T. pallidum* strains, including those that confer macrolide resistance allowed comparison of syphilis cases in London and Colombo with those described elsewhere in the world. It was of interest that the predominant strain types in London were identical to the predominant circulating strains in much of the developed world, whereas those in Sri Lanka were unique (83). This was thought to represent relative isolation of the Colombo population compared with the melting pot of London. It was also of interest that macrolide resistance is of an increasingly high level in London, similar to San Francsico or Seattle, but that it was not detected in Sri Lanka, much as has been described in two other small island populations: Taiwan and Madagascar (84, 91, 175). The absence of resistance in Tawian is of particular interest given its proximity to China where resistant strains represent over 90% of circulating strains (90). Strain-typing in London was enhanced by the addition of a third typing target, helping to discriminate the predominant 14d subtype in two separate types. Other *T. pallidum genes* (t*pp0138*, *tpp0326*, *tpp0488*) also have variability that may make them suitable for strain-typing targets (14). Moreover, given the recent association of strain-type 14d/f with neuro-invasion it will be of great interest to

identify additional typing targets, which may predict other clinical phenotypes (86). Finally, previous studies have stated that ulcer samples are superior to blood for strain type identification (83). In the current study, those samples with highest bacterial load were the most suitable and were predominantly ulcer samples, thus it appears that DNA quantity is key, not necessarily the sample type.

Syphilis is a complex disease caused by an organism that does not yield its secrets easily. While clinicians deliberate on how best to diagnose and treat patients at each stage of disease, scientists work to establish how such a genetically simple bacterium is so remarkably invasive and evasive. The lack of a system for continuous culture is undoubtedly one of the greatest impediments to *T. pallidum* research, but modern molecular microbiological techniques are now answering previously unanswerable questions and informing clinicians on appropriate management strategies. With time, it is hoped that the technique for monitoring treatment described in this thesis will be a useful tool to assess new treatment strategies for this ancient, resurgent and fascinating disease.

Plan for future work

This work has demonstrated that qPCR and RT-qPCR can be used to monitor the treatment response of patients with early syphilis to treatment with benzathine penicillin injection. However, prior to concluding that qPCR is a suitable tool to monitor treatment efficacy in new clinical trials the following questions must be answered:

1. Are the results presented here for four patients reproducible in a larger study?
2. Do patients who are subsequently identified as treatment failures, either serological or clinical, exhibit the same qPCR treatment response as treatment successes?
3. Is there a difference in the rate of clearance according to the antibiotic used. For example, does clearance following orally administered doxycycline follow the pattern seen for injected penicillin?
4. Can treatment monitoring by qPCR be applied to the rabbit model of syphilis such that candidate antibiotics can be assessed using the tool in animals before human clinical trials?

5. Can ear lobe samping be used as a quick and easy method for sample collection and does capiliary blood contain enough T. pallidum DNA to be detected reliably?

In addition to developing the qPCR method for measuring treatment efficacy, two parallel avenues of research also warrant exploration. The first is to better characterise the JHR both by correlating propensity to develop the reaction with bacterial load and by characterising the inflammatory response in patients who experience it. To date, no studies have associated *T. pallidum* bacterial load with risk of developing a JHR nor shown whether the JHR accelerates bacterial clearance despite an association between serological syphilis cure at 6 months and developing a JHR (218). Secondly, the ability to discriminate untreated latent and treated syphilis, which often have identical serological findings, would significantly improve disease management. Acquired immunity to *T. pallidum* appears to increase as disease progresses towards latency, despite the resolution of clinical signs. Whilst the partial immunity to re-infection induced during early disease wanes after treatment. (100) T-cells derived from experimental models of syphilis react strongly to several *T. pallidum* lipoproteins in the outer cytoplasmic membrane leaflet (TpN47, TpN17, TpN15) and the flagellar sheath (TpN37, TpN35, TpN33, TpN30) (274). There is also evidence from both humans and rabbits that IFN-gamma produced by both $CD4^+$ and $CD8^+$ T-cells may be integral to bacterial clearance from primary ulcers, despite *T. pallidum* being an extra-cellular pathogen (187). An ELISpot assay is one means of quantifying the *ex vivo* response by enumerating T-cells or B-cells in peripheral blood from patients which recognise and respond to specific pathogen antigens by secreting antibody or, for example, IFN-gamma (124, 275). It is envisaged, therefore, to stimulate either B-cells or T-cells taken from patients with both treated and latent syphilis with *T. pallidum* peptides in order to measure a clinically relevant difference in cytokine or antibody secretion. Prior to ELISpot assay development it is also envisaged that the most appropriate effector cytokine will be determined in a series of flow cytometry experiments measuring the acitivity of cells seperated from the blood of patients with both early and latent disease.

Experimental Design and Methods

Animal study experimental plan

It is impossible to control the moment of infection in humans, moreover to assess differences in bacterial load and immune response between subjects and untreated controls is unethical. Rabbits will be used, therefore, to map both bacterial load and immunological response during infection and treatment. These results will establish an alternative animal model for the assessment of future candidate antibiotics using qPCR at the same time as characterizing the inflammatory response to *T. pallidum* infection and treatment, and correlating this with bacterial load. Taken together, these findings will inform further the choice of immune markers to be measured in the proposed human study.

It is envisaged that *T. pallidum* clearance and cytokine response will be measured in three groups of animals, which have been infected with *T. pallidum* by intradermal injection. One receiving single-dose benzathine penicillin, a second being given multi-dose oral doxycycline and a third group of both uninfected and infected but untreated controls. Blood and ulcer biopsy samples will be taken for measurement of IFN-gamma, IL-2, IL-4, IL-10, IL1-beta, IL-22, IL-6 and TNF-alpha and both *16S rRNA* and *tpp047* DNA quantification.

Human pilot study

A three-site clinical study will recruit 100 patients from sexual health clinics in Colombo (Sri Lanka), Kampala (Uganda) and London (UK). Participants will be given standard syphilis treatment, as selected by their doctor, and their blood monitored with both the novel RT-qPCR and standard serology. Time points at 0 (pre-treatment), 2 (peak bacteraemia), 24 (approaching clearance) and 38-48 hours (cleared) have been selected based on previous pilot data. A further sample will be taken at 72 hours to cover adequately the 33-hour generation time of *T. pallidum* and potential for a late recrudescence in DNA. Initial sample processing will be performed locally prior to shipping samples to the Jefferiss Trust Laboratories, Imperial College London for cytokine and *T. pallidum* nucleic acid quantification. The long-term goal is to use these data to inform the design of clinical trials comparing syphilis treatment efficacy with current and novel antibiotics in patients both with and without HIV-1 infection.Development of an ELISpot assay

Samples

Samples will be needed for the initial development of an ELISpot assay to differentiate patients with latent infection (requiring treatment) and those with previously treated disease.
The following patients will be approached and asked to donate 30ml EDTA blood (We aim to recruit 25 patients per group): Those with latent disease (no symptoms or signs and a positive *T. pallidum* EIA); previously treated patient with no symptoms or signs (positive *T. pallidum* EIA, negative or low and stable anti-cardiolipin antibody titre) and finally syphilis negative controls (asymptomatic; negative *T. pallidum* EIA; no sexual contact within the last 3 month).
At each time point, patients will also have a capillary blood sample collected from their ear lobe for the assessment of *T. pallidum* quantification in capillary blood. This sample type has been shown previously to be rich in *T. pallidum* DNA and may be a quick and easy method for serial blood sampling (133).

Assay development

Over-lapping peptides spanning the length of the *TpN47* protein will be purchased (to ensure that T-cell responses are not restricted to or biased by patient HLA haplotypes). The best peptides will be selected for future use. ELISpot plates pre-coated with monoclonal antibodies will be purchased from Mabtech and standard ELISpot assay conditions will be employed. The following will be determined:

Background number of spot-forming cells (SFC) per million PBMC for buffer-only negative controls (if any).
Background SFC/million PBMC for syphilis negative controls.
SFC/million PBMC for latently infected patients (clinically asymptomatic)
SFC/million PBMC for previously treated patients.

Once the working parameters of the assay have been established, a larger clinical validation study will be designed.

References

1. Health Protection Agency U. Syphilis and Lymphogranuloma Venereum: Resurgent Sexually Transmitted Infections in the UK: 2009 report2009. Available from: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1245581513523.

2. Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol. 1998;180(18):4765-74.

3. Paster BJ, Dewhirst FE. Phylogenetic foundation of spirochetes. J Mol Microbiol Biotechnol. 2000;2(4):341-4.

4. Lafond RE, Lukehart SA. Biological basis for syphilis. Clin Microbiol Rev. 2006;19(1):29-49.

5. Norris SJ, Cox DL, Weinstock GM. Biology of Treponema pallidum: correlation of functional activities with genome sequence data. J Mol Microbiol Biotechnol. 2001;3(1):37-62.

6. Antal GM, Lukehart SA, Meheus AZ. The endemic treponematoses. Microbes Infect. 2002;4(1):83-94.

7. A C. Syphilis and the Columbian Exchange: Was the Disease Really New? In: MG M, J C, editors. The Great Maritime Discoveries and World Health. Lisbon: Ecola Nacional de Saude Publica. p. 187-200.

8. Schaudinn F, Hoffmann E. Bericht Über das Vorkommen von Spirochaeten in syphilitischen Krankheitsprodukten und bei Papilloten. Arb Kaiserl Gesurndheisamte. 1905;22:527-32.

9. Ricord P. A practical treatise on venereal diseases. Rouvier et Le Bouvier, Paris, France.

10. Rothschild BM. History of syphilis. Clin Infect Dis. 2005;40(10):1454-63.

11. Wood CS. Syphilis in anthropological perspective. Soc Sci Med. 1978;12(1B):47-55.

12. LJ M, E E. Historical Evidence of Syphilis and Other Treponemes. In: JD R, SA L, editors. Pathogenic Treponema: Molecular and Cellular Biology. Caister: Caister Academic Press; 2006. p. 183-95.

13. FB L. On the origin of syphilis: an alternative hypothesis. Curr Anthropol. 1991;32:587-90.

14. Smajs D, Norris SJ, Weinstock GM. Genetic diversity in Treponema pallidum: implications for pathogenesis, evolution and molecular diagnostics of syphilis and yaws. Infect Genet Evol. 2012;12(2):191-202.

15. Stokes, JH. Modern Clinical Syphilology. Philadephia: WB Saunders; 1944.

16. Wasserman A, Neisser F, Bruck C. Eine serodiagnostische Reaktion bei Syphilis. Deutsche medicinische Wochenschrift. 1906;32:745-6.

17. The British Venereal Diseases Service, 1916-66. Br J Vener Dis. 1966;42(4):223-4.

18. Mahoney JF, Arnold RC, Harris A. Penicillin Treatment of Early Syphilis-A Preliminary Report. Am J Public Health Nations Health. 1943;33(12):1387-91.

19. Danehower WF, Schamberg IL. Effect of coincidental penicillin administration on the spread of syphilis: An attempt at quantitation. Arch Dermatol. 1963;88:930-4.

20. Lloyd NC, Morgan HW, Nicholson BK, Ronimus RS. The composition of Ehrlich's salvarsan: resolution of a century-old debate. Angew Chem Int Ed Engl. 2005;44(6):941-4.

21. HPA. Total number of STI diagnoses in genitourinary medicine clinics & community settings in England 2008 - 20102011. Available from: http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1215589015024.

22. Simms I, Ward H. Congenital syphilis in the United Kingdom. Sex Transm Infect. 2006;82(1):1.

23. WHO. Prevalence and Incidence of selected sexually transmitted infections: methods and result used by WHO to generate 2005 estimates. 2005.

24. Personal Communication with Weerasinge G. 2013.

25. Maity S, Bhunia SC, Biswas S, Saha MK. Syphilis seroprevalence among patients attending a sexually transmitted disease clinic in West Bengal, India. Jpn J Infect Dis. 2011;64(6):506-8.

26. Organisation WH. Guidelines for the management of sexually transmitted infections2003. Available from: http://www.who.int/hiv/pub/sti/pub6/en/.

27. Walker EM, Arnett JK, Heath JD, Norris SJ. Treponema pallidum subsp. pallidum has a single, circular chromosome with a size of approximately 900 kilobase pairs. Infect Immun. 1991;59(7):2476-9.

28. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, et al. Complete genome sequence of Treponema pallidum, the syphilis spirochete. Science. 1998;281(5375):375-88.

29. Krawiec S, Riley M. Organization of the bacterial chromosome. Microbiol Rev. 1990;54(4):502-39.

30. Miao RM, Fieldsteel AH. Genetic relationship between Treponema pallidum and Treponema pertenue, two noncultivable human pathogens. J Bacteriol. 1980;141(1):427-9.

31. Centurion-Lara A, Molini BJ, Godornes C, Sun E, Hevner K, Van Voorhis WC, et al. Molecular differentiation of Treponema pallidum subspecies. J Clin Microbiol. 2006;44(9):3377-80.

32. Cejková D, Zobaníková M, Chen L, Pospíšilová P, Strouhal M, Qin X, et al. Whole genome sequences of three Treponema pallidum ssp. pertenue strains: yaws and syphilis treponemes differ in less than 0.2% of the genome sequence. PLoS Negl Trop Dis. 2012;6(1):e1471.

33. Schmidt H, Hensel M. Pathogenicity islands in bacterial pathogenesis. Clin Microbiol Rev. 2004;17(1):14-56.

34. Achtman M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol. 2008;62:53-70.

35. Fieldsteel AH, Cox DL, Moeckli RA. Cultivation of virulent Treponema pallidum in tissue culture. Infect Immun. 1981;32(2):908-15.

36. Magnuson HJ, Eagle H, Fleischman R. The minimal infectious inoculum of Spirochaeta pallida (Nichols strain) and a consideration of its rate of multiplication in vivo. Am J Syph Gonorrhea Vener Dis. 1948;32(1):1-18.

37. Norris SJ. In vitro cultivation of Treponema pallidum: independent confirmation. Infect Immun. 1982;36(1):437-9.

38. Lombard M, Touati D, Fontecave M, Nivière V. Superoxide reductase as a unique defense system against superoxide stress in the microaerophile Treponema pallidum. J Biol Chem. 2000;275(35):27021-6.

39. Jovanović T, Ascenso C, Hazlett KR, Sikkink R, Krebs C, Litwiller R, et al. Neelaredoxin, an iron-binding protein from the syphilis spirochete, Treponema pallidum, is a superoxide reductase. J Biol Chem. 2000;275(37):28439-48.

40. Turner TB, Hollander DH. Studies on the mechanism of action of cortisone in experimental syphilis. Am J Syph Gonorrhea Vener Dis. 1954;38(5):371-87.

41. Hollander DH. Treponematosis from pinta to venereal syphilis revisited: hypothesis for temperature determination of disease patterns. Sex Transm Dis. 1981;8(1):34-7.

42. Stamm LV, Gherardini FC, Parrish EA, Moomaw CR. Heat shock response of spirochetes. Infect Immun. 1991;59(4):1572-5.

43. Cox DL. Culture of Treponema pallidum. Methods Enzymol. 1994;236:390-405.

44. Turner TB, Hollander DH. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monogr Ser World Health Organ. 1957(35):3-266.

45. Klein JR, Monjan AA, Hardy PH, Cole GA. Abrogation of genetically controlled resistance of mice to Treponema pallidum by irradiation. Nature. 1980;283(5747):572-4.

46. Nichols HJ. Demonstration of the Spirochaeta pallida in the cerebrospinal fluid. From a patient with nervous relapse following the use of salvarsan. JAMA. 1913;40:108-10.

47. Chacko CW. Accidental human infection in the laboratory with Nichols rabbit-adapted virulent strain of Treponema pallidum. Bull World Health Organ. 1966;35(5):809-10.

48. Magnuson HJ, Thomas EW, Olansky S, Kaplan BI, De Mello L, Cutler JC. Inoculation syphilis in human volunteers. Medicine (Baltimore). 1956;35(1):33-82.

49. Fitzgerald TJ, Repesh LA, Oakes SG. Morphological destruction of cultured cells by the attachment of Treponema pallidum. Br J Vener Dis. 1982;58(1):1-11.

50. Quist EE, Repesh LA, Zeleznikar R, Fitzgerald TJ. Interaction of Treponema pallidum with isolated rabbit capillary tissues. Br J Vener Dis. 1983;59(1):11-20.

51. Salazar J, Rathi A, Michael N, Radolf J, Jagodzinski L. Assessment of the kinetics of Treponema pallidum dissemination into blood and tissues in experimental syphilis by real-time quantitative PCR. Infect Immun. 2007;75(6):2954-8.

52. Gayet-Ageron A, Lautenschlager S, Ninet B, Perneger TV, Combescure C. Sensitivity, specificity and likelihood ratios of PCR in the diagnosis of syphilis: a systematic review and meta-analysis. Sex Transm Infect. 2012.

53. Thomas DD, Navab M, Haake DA, Fogelman AM, Miller JN, Lovett MA. Treponema pallidum invades intercellular junctions of endothelial cell monolayers. Proc Natl Acad Sci U S A. 1988;85(10):3608-12.

54. Thomas DD, Fogelman AM, Miller JN, Lovett MA. Interactions of Treponema pallidum with endothelial cell monolayers. Eur J Epidemiol. 1989;5(1):15-21.

55. Sell S, Gamboa D, Baker-Zander S, Lukehart S, Miller J. Host response to Treponema pallidum in intradermally-infected rabbits: evidence for persistence of infection at local and distant sites. J Invest Dermatol. 1980;75(6):470-5.

56. Dorfman DH, Glaser JH. Congenital syphilis presenting in infants after the newborn period. N Engl J Med. 1990;323(19):1299-302.

57. Orton S, Liu H, Dodd R, Williams A, Group AE. Prevalence of circulating Treponema pallidum DNA and RNA in blood donors with confirmed-positive syphilis tests. Transfusion. 2002;42(1):94-9.

58. Fitzgerald TJ, Repesh LA, Blanco DR, Miller JN. Attachment of Treponema pallidum to fibronectin, laminin, collagen IV, and collagen I, and blockage of attachment by immune rabbit IgG. Br J Vener Dis. 1984;60(6):357-63.

59. Fitzgerald TJ, Johnson RC, Miller JN, Sykes JA. Characterization of the attachment of Treponema pallidum (Nichols strain) to cultured mammalian cells and the potential relationship of attachment to pathogenicity. Infect Immun. 1977;18(2):467-78.

60. Timpl R, Dziadek M. Structure, development, and molecular pathology of basement membranes. Int Rev Exp Pathol. 1986;29:1-112.

61. Cameron CE, Brouwer NL, Tisch LM, Kuroiwa JM. Defining the interaction of the Treponema pallidum adhesin Tp0751 with laminin. Infect Immun. 2005;73(11):7485-94.

62. Cameron CE, Brown EL, Kuroiwa JM, Schnapp LM, Brouwer NL. Treponema pallidum fibronectin-binding proteins. J Bacteriol. 2004;186(20):7019-22.

63. Konishi H, Yoshii Z, Cox DL. Electron microscopy of Treponema pallidum (Nichols) cultivated in tissue cultures of Sf1Ep cells. Infect Immun. 1986;53(1):32-7.

64. Hougen KH. Further observations on the ultrastructure of Treponema pallidum nichols. Acta Pathol Microbiol Scand B Microbiol Immunol. 1972;80(2):297-304.

65. Holt SC. Anatomy and chemistry of spirochetes. Microbiol Rev. 1978;42(1):114-60.

66. Isaacs RD, Hanke JH, Guzman-Verduzco LM, Newport G, Agabian N, Norgard MV, et al. Molecular cloning and DNA sequence analysis of the 37-kilodalton endoflagellar sheath protein gene of Treponema pallidum. Infect Immun. 1989;57(11):3403-11.

67. Limberger RJ, Slivienski LL, Izard J, Samsonoff WA. Insertional inactivation of Treponema denticola tap1 results in a nonmotile mutant with elongated flagellar hooks. J Bacteriol. 1999;181(12):3743-50.

68. Walker EM, Zampighi GA, Blanco DR, Miller JN, Lovett MA. Demonstration of rare protein in the outer membrane of Treponema pallidum subsp. pallidum by freeze-fracture analysis. J Bacteriol. 1989;171(9):5005-11.

69. Penn CW, Cockayne A, Bailey MJ. The outer membrane of Treponema pallidum: biological significance and biochemical properties. J Gen Microbiol. 1985;131(9):2349-57.

70. Hardy PH, Nell EE. Study of the antigenic structure of Treponema pallidum by specific agglutination. Am J Hyg. 1957;66(2):160-72.

71. Radolf JD, Robinson EJ, Bourell KW, Akins DR, Porcella SF, Weigel LM, et al. Characterization of outer membranes isolated from Treponema pallidum, the syphilis spirochete. Infect Immun. 1995;63(11):4244-52.

72. Cox DL, Chang P, McDowall AW, Radolf JD. The outer membrane, not a coat of host proteins, limits antigenicity of virulent Treponema pallidum. Infect Immun. 1992;60(3):1076-83.

73. Radolf JD, Moomaw C, Slaughter CA, Norgard MV. Penicillin-binding proteins and peptidoglycan of Treponema pallidum subsp. pallidum. Infect Immun. 1989;57(4):1248-54.

74. Weigel LM, Brandt ME, Norgard MV. Analysis of the N-terminal region of the 47-kilodalton integral membrane lipoprotein of Treponema pallidum. Infect Immun. 1992;60(4):1568-76.

75. Baker-Zander SA, Hook EW, Bonin P, Handsfield HH, Lukehart SA. Antigens of Treponema pallidum recognized by IgG and IgM antibodies during syphilis in humans. J Infect Dis. 1985;151(2):264-72.

76. Hsu PL, Chamberlain NR, Orth K, Moomaw CR, Zhang LQ, Slaughter CA, et al. Sequence analysis of the 47-kilodalton major integral membrane immunogen of Treponema pallidum. Infect Immun. 1989;57(1):196-203.

77. Baker-Zander SA, Lukehart SA. Antigenic cross-reactivity between Treponema pallidum and other pathogenic members of the family Spirochaetaceae. Infect Immun. 1984;46(1):116-21.

78. Lukehart SA, Baker-Zander SA, Gubish ER. Identification of Treponema pallidum antigens: comparison with a nonpathogenic treponeme. J Immunol. 1982;129(2):833-8.

79. Leader BT, Hevner K, Molini BJ, Barrett LK, Van Voorhis WC, Lukehart SA. Antibody responses elicited against the Treponema pallidum repeat proteins differ during infection with different isolates of Treponema pallidum subsp. pallidum. Infect Immun. 2003;71(10):6054-7.

80. Stamm LV, Bergen HL. The sequence-variable, single-copy tprK gene of Treponema pallidum Nichols strain UNC and Street strain 14 encodes heterogeneous TprK proteins. Infect Immun. 2000;68(11):6482-6.

81. Giacani L, Brandt SL, Puray-Chavez M, Reid TB, Godornes C, Molini BJ, et al. Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J Bacteriol. 2012;194(16):4208-25.

82. Pillay A, Liu H, Chen CY, Holloway B, Sturm AW, Steiner B, et al. Molecular subtyping of Treponema pallidum subspecies pallidum. Sex Transm Dis. 1998;25(8):408-14.

83. Peng RR, Wang AL, Li J, Tucker JD, Yin YP, Chen XS. Molecular typing of Treponema pallidum: a systematic review and meta-analysis. PLoS Negl Trop Dis. 2011;5(11):e1273.

84. Katz KA, Pillay A, Ahrens K, Kohn RP, Hermanstyne K, Bernstein KT, et al. Molecular epidemiology of syphilis--San Francisco, 2004-2007. Sex Transm Dis. 2010;37(10):660-3.

85. Flasarová M, Pospíšilová P, Mikalová L, Vališová Z, Dastychová E, Strnadel R, et al. Sequencing-based molecular typing of treponema pallidum strains in the Czech Republic: all identified genotypes are related to the sequence of the SS14 strain. Acta Derm Venereol. 2012;92(6):669-74.

86. Marra CM, Sahi SK, Tantalo LC, Godornes C, Reid T, Behets F, et al. Enhanced molecular typing of treponema pallidum: geographical distribution of strain types and association with neurosyphilis. J Infect Dis. 2010;202(9):1380-8.

87. Cole M, Chisholm S, Palmer H, Wallace L, Ison C. Molecular epidemiology of syphilis in Scotland. Sex Transm Infect. 2009;85(6):447-51.

88. Molepo J, Pillay A, Weber B, Morse S, Hoosen A. Molecular typing of Treponema pallidum strains from patients with neurosyphilis in Pretoria, South Africa. Sex Transm Infect. 2007;83(3):189-92.

89. Pillay A, Liu H, Ebrahim S, Chen C, Lai W, Fehler G, et al. Molecular typing of Treponema pallidum in South Africa: cross-sectional studies. J Clin Microbiol. 2002;40(1):256-8.

90. Peng RR, Yin YP, Wei WH, Wang HC, Zhu BY, Liu QZ, et al. Molecular typing of Treponema pallidum causing early syphilis in China: a cross-sectional study. Sex Transm Dis. 2012;39(1):42-5.

91. Wu H, Chang SY, Lee NY, Huang WC, Wu BR, Yang CJ, et al. Evaluation of macrolide resistance and enhanced molecular typing of Treponema pallidum in patients with syphilis in Taiwan: a prospective multicenter study. J Clin Microbiol. 2012;50(7):2299-304.

92. Azzato F, Ryan N, Fyfe J, Leslie DE. Molecular subtyping of Treponema pallidum during a local syphilis epidemic in men who have sex with men in Melbourne, Australia. J Clin Microbiol. 2012;50(6):1895-9.

93. Martin IE, Tsang RS, Sutherland K, Anderson B, Read R, Roy C, et al. Molecular typing of Treponema pallidum strains in western Canada: predominance of 14d subtypes. Sex Transm Dis. 2010;37(9):544-8.

94. Sutton MY, Liu H, Steiner B, Pillay A, Mickey T, Finelli L, et al. Molecular subtyping of Treponema pallidum in an Arizona County with increasing syphilis morbidity: use of specimens from ulcers and blood. J Infect Dis. 2001;183(11):1601-6.

95. Marra C SS, Tantalo L, Reid T, Lukehart S, and Centurion A. *Treponema pallidum* Strain Type Is Associated with Neurosyphilis. CROI 2010; San Francisco2010.

96. Ho EL, Lukehart SA. Syphilis: using modern approaches to understand an old disease. J Clin Invest. 2011;121(12):4584-92.

97. Schober PC, Gabriel G, White P, Felton WF, Thin RN. How infectious is syphilis? Br J Vener Dis. 1983;59(4):217-9.

98. Garnett GP, Aral SO, Hoyle DV, Cates W, Anderson RM. The natural history of syphilis. Implications for the transmission dynamics and control of infection. Sex Transm Dis. 1997;24(4):185-200.

99. Schulz KF, Cates W, O'Mara PR. Pregnancy loss, infant death, and suffering: legacy of syphilis and gonorrhoea in Africa. Genitourin Med. 1987;63(5):320-5.

100. Radolf J, Lukehart S. Pathogenic Treponema: Molecular and Cellular Biology. Great Britain: Caister Academic Press; 2006.

101. Azar HA, Pham TD, Kurban AK. An electron microscopic study of a syphilitic chancre. Engulfment of Treponema pallidum by plasma cells. Arch Pathol. 1970;90(2):143-50.

102. Baughn RE, Musher DM. Secondary syphilitic lesions. Clin Microbiol Rev. 2005;18(1):205-16.

103. Gjestland T. The Oslo study of untreated syphilis; an epidemiologic investigation of the natural course of the syphilitic infection based upon a re-study of the Boeck-Bruusgaard material. Acta Derm Venereol Suppl (Stockh). 1955;35(Suppl 34):3-368; Annex I-LVI.

104. Kingston M, French P, Goh B, Goold P, Higgins S, Sukthankar A, et al. UK National Guidelines on the Management of Syphilis 2008. Int J STD AIDS. 2008;19(11):729-40.

105. Kampmeier RH. The Tuskegee study of untreated syphilis. South Med J. 1972;65(10):1247-51.

106. Roberts W, Ko J, Vowels T. Natural history of syphilitic aortitis. Am J Cardiol. 2009;104(11):1578-87.

107. O'donnell JA, Emery CL. Neurosyphilis: A Current Review. Curr Infect Dis Rep. 2005;7(4):277-84.

108. Anderson J, Mindel A, Tovey SJ, Williams P. Primary and secondary syphilis, 20 years' experience. 3: Diagnosis, treatment, and follow up. Genitourin Med. 1989;65(4):239-43.

109. Chapel TA. The signs and symptoms of secondary syphilis. Sex Transm Dis. 1980;7(4):161-4.

110. Creighton ET. Darkfield microscopy for the detection and identification of Treponema pallidum. In: Krause SJ, editor. A manual of tests for syphilis. 8th ed. Washington DC: American Public Health Association; 1990. p. 49-62.

111. Larsen S, Steiner B, Rudolph A. Laboratory diagnosis and interpretation of tests for syphilis. Clin Microbiol Rev. 1995;8(1):1-21.

112. Pangborn MC. Cardiolipin and its application in a chemically purified antigen for the serodiagnosis of syphilis. Proc N Y State Assoc Public Health Lab. 1946;26(1):26-9.

113. Harris A, Rosenberg AA, Riedel LM. A microflocculation test for syphilis using cardiolipin antigen. J Vener Dis Inf. 1946;27:169-74.

114. Anderson J, Mindel A, Tovey SJ, Williams P. Primary and secondary syphilis, 20 years' experience. 3: Diagnosis, treatment, and follow up. Genitourin Med. 1989;65(4):239-43.

115. Larsen SA, Steiner BM, Rudolph AH. Laboratory diagnosis and interpretation of tests for syphilis. Clin Microbiol Rev. 1995;8(1):1-21.

116. McGrew BE, DuCros MJ, Stout GW, Falcone VH. Automation of a flocculation test for syphilis. Am J Clin Pathol. 1968;50(1):52-9.

117. Fiumara NJ. Treatment of primary and secondary syphilis. Serological response. JAMA. 1980;243(24):2500-2.

118. Fiumara NJ. Serologic responses to treatment of 128 patients with late latent syphilis. Sex Transm Dis. 1979;6(4):243-6.

119. Nelson RA, Mayer MM. Immobilization of Treponema pallidum in vitro by antibody produced in syphilitic infection. J Exp Med. 1949;89(4):369-93.

120. Hunter EF, McKinney RM, Maddison SE, Cruce DD. Double-staining procedure for the fluorescent treponemal antibody absorption (FTA-ABS) test. Br J Vener Dis. 1979;55(2):105-8.

121. Wentworth BB, Thompson MA, Peter CR, Bawdon RE, Wilson DL. Comparison of a hemagglutination treponemal test for syphilis (HATTS) with other serologic methods for the diagnosis of syphilis. Sex Transm Dis. 1978;5(3):103-14.

122. Hoover KW, Radolf JD. Serodiagnosis of syphilis in the recombinant era: reversal of fortune. J Infect Dis. 2011;204(9):1295-6.

123. Marra CM, Maxwell CL, Smith SL, Lukehart SA, Rompalo AM, Eaton M, et al. Cerebrospinal fluid abnormalities in patients with syphilis: association with clinical and laboratory features. J Infect Dis. 2004;189(3):369-76.

124. Tabidze IL, Lee FK, Tambe P, Rocha E, Larsen SA, Stoll BJ, et al. Enzyme-linked immunospot assay for the diagnosis of active Treponema pallidum infection during the various stages of syphilis. Sex Transm Dis. 1999;26(8):426-30.

125. Turner TB, Hollander DH. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monogr Ser World Health Organ. 1957(35):3-266.

126. Turner TB, Hardy PH, Newman B. Infectivity tests in syphilis. Br J Vener Dis. 1969;45(3):183-95.

127. Hay P, Clarke J, Strugnell R, Taylor-Robinson D, Goldmeier D. Use of the polymerase chain reaction to detect DNA sequences specific to pathogenic treponemes in cerebrospinal fluid. FEMS Microbiol Lett. 1990;56(3):233-8.

128. Lukehart SA, Hook EW, Baker-Zander SA, Collier AC, Critchlow CW, Handsfield HH. Invasion of the central nervous system by Treponema pallidum: implications for diagnosis and treatment. Ann Intern Med. 1988;109(11):855-62.

129. Noordhoek GT, Wolters EC, de Jonge ME, van Embden JD. Detection by polymerase chain reaction of Treponema pallidum DNA in cerebrospinal fluid from neurosyphilis patients before and after antibiotic treatment. J Clin Microbiol. 1991;29(9):1976-84.

130. Grimprel E, Sanchez P, Wendel G, Burstain J, McCracken GJ, Radolf J, et al. Use of polymerase chain reaction and rabbit infectivity testing to detect Treponema pallidum in amniotic fluid, fetal and neonatal sera, and cerebrospinal fluid. J Clin Microbiol. 1991;29(8):1711-8.

131. Tipple C, Hanna M, Hills S, Daniels J, Goldmeier D, McClure MO, et al. Getting the measure of syphilis: qPCR to better understand early infection. Sexually Transmitted Infections [Internet]. 2011.

132. Orle K, Gates C, Martin D, Body B, Weiss J. Simultaneous PCR detection of Haemophilus ducreyi, Treponema pallidum, and herpes simplex virus types 1 and 2 from genital ulcers. J Clin Microbiol. 1996;34(1):49-54.

133. Castro R, Prieto E, Aguas M, Manata M, Botas J, Santo I, et al. Detection of Treponema pallidum sp pallidum DNA in latent syphilis. Int J STD AIDS. 2007;18(12):842-5.

134. Gayet-Ageron A, Ninet B, Toutous-Trellu L, Lautenschlager S, Furrer H, Piguet V, et al. Assessment of a real-time PCR test to diagnose syphilis from diverse biological samples. Sex Transm Infect. 2009;85(4):264-9.

135. Case definitions for infectious conditions under public health surveillance. Centers for Disease Control and Prevention. MMWR Recomm Rep. 1997;46(RR-10):1-55.

136. Kouznetsov A, Weisenseel P, Trommler P, Multhaup S, Prinz J. Detection of the 47-kilodalton membrane immunogen gene of Treponema pallidum in various tissue sources of patients with syphilis. Diagn Microbiol Infect Dis. 2005;51(2):143-5.

137. Genest DR, Choi-Hong SR, Tate JE, Qureshi F, Jacques SM, Crum C. Diagnosis of congenital syphilis from placental examination: comparison of histopathology, Steiner stain, and polymerase chain reaction for Treponema pallidum DNA. Hum Pathol. 1996;27(4):366-72.

138. Inagaki H, Kawai T, Miyata M, Nagaya S, Tateyama H, Eimoto T. Gastric syphilis: polymerase chain reaction detection of treponemal DNA in pseudolymphomatous lesions. Hum Pathol. 1996;27(8):761-5.

139. Horowitz HW, Valsamis MP, Wicher V, Abbruscato F, Larsen SA, Wormser GP, et al. Brief report: cerebral syphilitic gumma confirmed by the polymerase chain reaction in a man with human immunodeficiency virus infection. N Engl J Med. 1994;331(22):1488-91.

140. O'Regan A, Castro C, Lukehart S, Kasznica J, Rice P, Joyce-Brady M. Barking up the wrong tree? Use of polymerase chain reaction to diagnose syphilitic aortitis. Thorax. 2002;57(10):917-8.

141. Smajs D, McKevitt M, Howell JK, Norris SJ, Cai WW, Palzkill T, et al. Transcriptome of Treponema pallidum: gene expression profile during experimental rabbit infection. J Bacteriol. 2005;187(5):1866-74.

142. Jones SA, Marchitto KS, Miller JN, Norgard MV. Monoclonal antibody with hemagglutination, immobilization, and neutralization activities defines an immunodominant, 47,000 mol wt, surface-exposed immunogen of Treponema pallidum (Nichols). J Exp Med. 1984;160(5):1404-20.

143. Burstain J, Grimprel E, Lukehart S, Norgard M, Radolf J. Sensitive detection of Treponema pallidum by using the polymerase chain reaction. J Clin Microbiol. 1991;29(1):62-9.

144. Noordhoek GT, Wolters EC, de Jonge ME, van Embden JD. Detection by polymerase chain reaction of Treponema pallidum DNA in cerebrospinal fluid from neurosyphilis patients before and after antibiotic treatment. J Clin Microbiol. 1991;29(9):1976-84.

145. Casal CA, Silva MO, Costa IB, Araújo EaC, Corvelo TC. Molecular detection of Treponema pallidum sp. pallidum in blood samples of VDRL-seroreactive women with lethal

pregnancy outcomes: a retrospective observational study in northern Brazil. Rev Soc Bras Med Trop. 2011;44(4):451-6.

146. Marfin A, Liu H, Sutton M, Steiner B, Pillay A, Markowitz L. Amplification of the DNA polymerase I gene of Treponema pallidum from whole blood of persons with syphilis. Diagn Microbiol Infect Dis. 2001;40(4):163-6.

147. Martin I, Tsang R, Sutherland K, Tilley P, Read R, Anderson B, et al. Molecular characterization of syphilis in patients in Canada: azithromycin resistance and detection of Treponema pallidum DNA in whole-blood samples versus ulcerative swabs. J Clin Microbiol. 2009;47(6):1668-73.

148. Koek A, Bruisten S, Dierdorp M, van Dam A, Templeton K. Specific and sensitive diagnosis of syphilis using a real-time PCR for Treponema pallidum. Clin Microbiol Infect. 2006;12(12):1233-6.

149. Leslie D, Azzato F, Karapanagiotidis T, Leydon J, Fyfe J. Development of a real-time PCR assay to detect Treponema pallidum in clinical specimens and assessment of the assay's performance by comparison with serological testing. J Clin Microbiol. 2007;45(1):93-6.

150. Heymans R, van der Helm J, de Vries H, Fennema H, Coutinho R, Bruisten S. Clinical value of Treponema pallidum real-time PCR for diagnosis of syphilis. J Clin Microbiol. 2010;48(2):497-502.

151. Cruz A, Pillay A, Zuluaga A, Ramirez L, Duque J, Aristizabal G, et al. Secondary syphilis in cali, Colombia: new concepts in disease pathogenesis. PLoS Negl Trop Dis. 2010;4(5):e690.

152. Gayet-Ageron A, Ninet B, Toutous-Trellu L, Lautenschlager S, Furrer H, Piguet V, et al. Assessment of a real-time PCR test to diagnose syphilis from diverse biological samples. Sex Transm Infect. 2009;85(4):264-9.

153. Centurion-Lara A, Castro C, Shaffer J, Van Voorhis W, Marra C, Lukehart S. Detection of Treponema pallidum by a sensitive reverse transcriptase PCR. J Clin Microbiol. 1997;35(6):1348-52.

154. Deutscher MP. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res. 2006;34(2):659-66.

155. Wicher K, Abbruscato F, Wicher V, Collins DN, Auger I, Horowitz HW. Identification of persistent infection in experimental syphilis by PCR. Infect Immun. 1998;66(6):2509-13.

156. Mahoney JF, Arnold RC, Harris A. Penicillin Treatment of Early Syphilis-A Preliminary Report. Am J Public Health Nations Health. 1943;33(12):1387-91.

157. Douglas JM. Penicillin treatment of syphilis: clearing away the shadow on the land. JAMA. 2009;301(7):769-71.

158. Pao D, Goh BT, Bingham JS. Management issues in syphilis. Drugs. 2002;62(10):1447-61.

159. French P, Gomberg M, Janier M, Schmidt B, van Voorst Vader P, Young H, et al. IUSTI: 2008 European Guidelines on the Management of Syphilis. Int J STD AIDS. 2009;20(5):300-9.

160. Stamm LV. Global challenge of antibiotic-resistant Treponema pallidum. Antimicrob Agents Chemother. 2010;54(2):583-9.

161. Riedner G, Rusizoka M, Todd J, Maboko L, Hoelscher M, Mmbando D, et al. Single-dose azithromycin versus penicillin G benzathine for the treatment of early syphilis. N Engl J Med. 2005;353(12):1236-44.

162. Cunningham TM, Miller JN, Lovett MA. Identification of Treponema pallidum penicillin-binding proteins. J Bacteriol. 1987;169(11):5298-300.

163. Cha JY, Ishiwata A, Mobashery S. A novel beta-lactamase activity from a penicillin-binding protein of Treponema pallidum and why syphilis is still treatable with penicillin. J Biol Chem. 2004;279(15):14917-21.

164. Katz KA, Klausner JD. Azithromycin resistance in Treponema pallidum. Curr Opin Infect Dis. 2008;21(1):83-91.

165. Fenton LJ, Light IJ. Congenital syphilis after maternal treatment with erythromycin. Obstet Gynecol. 1976;47(4):492-4.

166. Hashisaki P, Wertzberger GG, Conrad GL, Nichols CR. Erythromycin failure in the treatment of syphilis in a pregnant woman. Sex Transm Dis. 1983;10(1):36-8.

167. Stamm LV, Stapleton JT, Bassford PJ. In vitro assay to demonstrate high-level erythromycin resistance of a clinical isolate of Treponema pallidum. Antimicrob Agents Chemother. 1988;32(2):164-9.

168. Centres for Disease Control. Azithromycin treatment failures in syphilis infections--San Francisco, California, 2002-2003. Morb Mortal Wkly Rep. 2004;53(9):197-8.

169. Lukehart S, Godornes C, Molini B, Sonnett P, Hopkins S, Mulcahy F, et al. Macrolide resistance in Treponema pallidum in the United States and Ireland. N Engl J Med. 2004;351(2):154-8.

170. Matejková P, Flasarová M, Zákoucká H, Borek M, Kremenová S, Arenberger P, et al. Macrolide treatment failure in a case of secondary syphilis: a novel A2059G mutation in the

23S rRNA gene of Treponema pallidum subsp. pallidum. J Med Microbiol. 2009;58(Pt 6):832-6.

171. Mitchell SJ, Engelman J, Kent CK, Lukehart SA, Godornes C, Klausner JD. Azithromycin-resistant syphilis infection: San Francisco, California, 2000-2004. Clin Infect Dis. 2006;42(3):337-45.

172. Martin IE, Gu W, Yang Y, Tsang RS. Macrolide resistance and molecular types of Treponema pallidum causing primary syphilis in Shanghai, China. Clin Infect Dis. 2009;49(4):515-21.

173. Norgard MV, Miller JN. Plasmid DNA in Treponema pallidum (Nichols): potential for antibiotic resistance by syphilis bacteria. Science. 1981;213(4507):553-5.

174. Martin I, Gu W, Yang Y, Tsang R. Macrolide resistance and molecular types of Treponema pallidum causing primary syphilis in Shanghai, China. Clin Infect Dis. 2009;49(4):515-21.

175. Van Damme K, Behets F, Ravelomanana N, Godornes C, Khan M, Randrianasolo B, et al. Evaluation of azithromycin resistance in Treponema pallidum specimens from Madagascar. Sex Transm Dis. 2009;36(12):775-6.

176. Morshed MG, Jones HD. Treponema pallidum macrolide resistance in BC. CMAJ. 2006;174(3):349.

177. Chen XS, Yin YP, Wei WH, Wang HC, Peng RR, Zheng HP, et al. High prevalence of azithromycin resistance to Treponema pallidum in geographically different areas in China. Clin Microbiol Infect. 2012.

178. Grimes M, Sahi SK, Godornes BC, Tantalo LC, Roberts N, Bostick D, et al. Two mutations associated with macrolide resistance in Treponema pallidum: increasing prevalence and correlation with molecular strain type in Seattle, Washington. Sex Transm Dis. 2012;39(12):954-8.

179. Müller EE, Paz-Bailey G, Lewis DA. Macrolide resistance testing and molecular subtyping of Treponema pallidum strains from southern Africa. Sex Transm Infect. 2012;88(6):470-4.

180. Tipple C, McClure MO, Taylor GP. High prevalence of macrolide resistant Treponema pallidum strains in a London centre. Sex Transm Infect. 2011;87(6):486-8.

181. Salazar JC, Hazlett KR, Radolf JD. The immune response to infection with Treponema pallidum, the stealth pathogen. Microbes Infect. 2002;4(11):1133-40.

182. Sellati TJ, Waldrop SL, Salazar JC, Bergstresser PR, Picker LJ, Radolf JD. The cutaneous response in humans to Treponema pallidum lipoprotein analogues involves cellular elements of both innate and adaptive immunity. J Immunol. 2001;166(6):4131-40.

183. Brightbill HD, Libraty DH, Krutzik SR, Yang RB, Belisle JT, Bleharski JR, et al. Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. Science. 1999;285(5428):732-6.

184. Wooten RM, Morrison TB, Weis JH, Wright SD, Thieringer R, Weis JJ. The role of CD14 in signaling mediated by outer membrane lipoproteins of Borrelia burgdorferi. J Immunol. 1998;160(11):5485-92.

185. Lukehart S, Baker-Zander S, Sell S. Characterization of lymphocyte responsiveness in early experimental syphilis. I. In vitro response to mitogens and Treponema pallidum antigens. J Immunol. 1980;124(1):454-60.

186. Lukehart S, Baker-Zander S, Lloyd R, Sell S. Characterization of lymphocyte responsiveness in early experimental syphilis. II. Nature of cellular infiltration and Treponema pallidum distribution in testicular lesions. J Immunol. 1980;124(1):461-7.

187. Leader BT, Godornes C, VanVoorhis WC, Lukehart SA. CD4+ lymphocytes and gamma interferon predominate in local immune responses in early experimental syphilis. Infect Immun. 2007;75(6):3021-6.

188. Salazar JC, Cruz AR, Pope CD, Valderrama L, Trujillo R, Saravia NG, et al. Treponema pallidum elicits innate and adaptive cellular immune responses in skin and blood during secondary syphilis: a flow-cytometric analysis. J Infect Dis. 2007;195(6):879-87.

189. Bryceson AD. Clinical pathology of the Jarisch-Herxheimer reaction. J Infect Dis. 1976;133(6):696-704.

190. Heyman A, Sheldon WH, Evans LD. Pathogenesis of the Jarisch-Herxheimer reaction; a review of clinical and experimental observations. Br J Vener Dis. 1952;28(2):50-60.

191. Griffin GE. Cytokines involved in human septic shock--the model of the Jarisch-Herxheimer reaction. J Antimicrob Chemother. 1998;41 Suppl A:25-9.

192. Kindmark CO, Möller H, Persson K. C-reactive protein, C3, C4 and properdin during the Jarisch-Herxheimer reaction in early syphilis. Acta Med Scand. 1978;204(4):287-90.

193. Farmer TW. Jarisch-Herxheimer reaction in early syphilis treated with crystalline penicillin G. J Am Med Assoc. 1948;138(7):480-5.

194. Ansell DA, Hu TC, Straus M, Cohen M, Sherer R. HIV and syphilis seroprevalence among clients with sexually transmitted diseases attending a walk-in clinic at Cook County Hospital. Sex Transm Dis. 1994;21(2):93-6.

195. Dickerson MC, Johnston J, Delea TE, White A, Andrews E. The causal role for genital ulcer disease as a risk factor for transmission of human immunodeficiency virus. An application of the Bradford Hill criteria. Sex Transm Dis. 1996;23(5):429-40.

196. Prevention CfDCa. Primary and Secondary Syphilis - United States. MMWR Mord Mortal Wkly Rep. 2006;55:269-73.

197. Zetola NM, Klausner JD. Syphilis and HIV infection: an update. Clin Infect Dis. 2007;44(9):1222-8.

198. Wong W, Chaw JK, Kent CK, Klausner JD. Risk factors for early syphilis among gay and bisexual men seen in an STD clinic: San Francisco, 2002-2003. Sex Transm Dis. 2005;32(7):458-63.

199. Zetola N, Klausner J. Syphilis and HIV infection: an update. Clin Infect Dis. 2007;44(9):1222-8.

200. Vallabhaneni S, Li X, Vittinghoff E, Donnell D, Pilcher CD, Buchbinder SP. Seroadaptive practices: association with HIV acquisition among HIV-negative men who have sex with men. PLoS One. 2012;7(10):e45718.

201. Cohen MS, McCauley M, Gamble TR. HIV treatment as prevention and HPTN 052. Curr Opin HIV AIDS. 2012;7(2):99-105.

202. Chesson HW, Pinkerton SD, Irwin KL, Rein D, Kassler WJ. New HIV cases attributable to syphilis in the USA: estimates from a simplified transmission model. AIDS. 1999;13(11):1387-96.

203. Theus SA, Harrich DA, Gaynor R, Radolf JD, Norgard MV. Treponema pallidum, lipoproteins, and synthetic lipoprotein analogues induce human immunodeficiency virus type 1 gene expression in monocytes via NF-kappaB activation. J Infect Dis. 1998;177(4):941-50.

204. Schacker T, Ryncarz AJ, Goddard J, Diem K, Shaughnessy M, Corey L. Frequent recovery of HIV-1 from genital herpes simplex virus lesions in HIV-1-infected men. JAMA. 1998;280(1):61-6.

205. Ghys PD, Fransen K, Diallo MO, Ettiègne-Traoré V, Coulibaly IM, Yeboué KM, et al. The associations between cervicovaginal HIV shedding, sexually transmitted diseases and immunosuppression in female sex workers in Abidjan, Côte d'Ivoire. AIDS. 1997;11(12):F85-93.

206. Cameron DW, Simonsen JN, D'Costa LJ, Ronald AR, Maitha GM, Gakinya MN, et al. Female to male transmission of human immunodeficiency virus type 1: risk factors for seroconversion in men. Lancet. 1989;2(8660):403-7.

207. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex Transm Infect. 1999;75(1):3-17.

208. Otten MW, Zaidi AA, Peterman TA, Rolfs RT, Witte JJ. High rate of HIV seroconversion among patients attending urban sexually transmitted disease clinics. AIDS. 1994;8(4):549-53.

209. Rompalo A, Lawlor J, Seaman P, Quinn T, Zenilman J, Hook Er. Modification of syphilitic genital ulcer manifestations by coexistent HIV infection. Sex Transm Dis. 2001;28(8):448-54.

210. Rolfs RT, Joesoef MR, Hendershot EF, Rompalo AM, Augenbraun MH, Chiu M, et al. A randomized trial of enhanced therapy for early syphilis in patients with and without human immunodeficiency virus infection. The Syphilis and HIV Study Group. N Engl J Med. 1997;337(5):307-14.

211. Flood JM, Weinstock HS, Guroy ME, Bayne L, Simon RP, Bolan G. Neurosyphilis during the AIDS epidemic, San Francisco, 1985-1992. J Infect Dis. 1998;177(4):931-40.

212. Marra CM, Maxwell CL, Tantalo L, Eaton M, Rompalo AM, Raines C, et al. Normalization of cerebrospinal fluid abnormalities after neurosyphilis therapy: does HIV status matter? Clin Infect Dis. 2004;38(7):1001-6.

213. Rompalo AM, Cannon RO, Quinn TC, Hook EW. Association of biologic false-positive reactions for syphilis with human immunodeficiency virus infection. J Infect Dis. 1992;165(6):1124-6.

214. Kofoed K, Gerstoft J, Mathiesen LR, Benfield T. Syphilis and human immunodeficiency virus (HIV)-1 coinfection: influence on CD4 T-cell count, HIV-1 viral load, and treatment response. Sex Transm Dis. 2006;33(3):143-8.

215. Jarzebowski W, Caumes E, Dupin N, Farhi D, Lascaux AS, Piketty C, et al. Effect of early syphilis infection on plasma viral load and CD4 cell count in human immunodeficiency virus-infected men: results from the FHDH-ANRS CO4 cohort. Arch Intern Med. 2012;172(16):1237-43.

216. Sadiq ST, McSorley J, Copas AJ, Bennett J, Edwards SJ, Kaye S, et al. The effects of early syphilis on CD4 counts and HIV-1 RNA viral loads in blood and semen. Sex Transm Infect. 2005;81(5):380-5.

217. Ghanem KG, Moore RD, Rompalo AM, Erbelding EJ, Zenilman JM, Gebo KA. Antiretroviral therapy is associated with reduced serologic failure rates for syphilis among HIV-infected patients. Clin Infect Dis. 2008;47(2):258-65.

218. Seña AC, Wolff M, Martin DH, Behets F, Van Damme K, Leone P, et al. Predictors of serological cure and Serofast State after treatment in HIV-negative persons with early syphilis. Clin Infect Dis. 2011;53(11):1092-9.

219. Hook Er, Martin D, Stephens J, Smith B, Smith K. A randomized, comparative pilot study of azithromycin versus benzathine penicillin G for treatment of early syphilis. Sex Transm Dis. 2002;29(8):486-90.

220. Kiddugavu MG, Kiwanuka N, Wawer MJ, Serwadda D, Sewankambo NK, Wabwire-Mangen F, et al. Effectiveness of syphilis treatment using azithromycin and/or benzathine penicillin in Rakai, Uganda. Sex Transm Dis. 2005;32(1):1-6.

221. Yobs AR, Rockwell DH, Clark JW. Treponemal survival in humans after penicillin therapy: A preliminary report. Br J Vener Dis. 1964;40:248-53.

222. Yobs AR, Clark JW, Mothershed SE, Bullard JC, Artley CW. Further observations on the persistence of Treponema pallidum after treatment in rabbits and humans. Br J Vener Dis. 1968;44(2):116-30.

223. Yogeswari L, Chacko CW. Persistence of T. pallidum and its significance in penicillin-treated seropositive late syphilis. Br J Vener Dis. 1971;47(5):339-47.

224. Collart P, Borel L, Durel P. Significance of spiral organisms found, after treatment, in late human and experimental syphilis. Br J Vener Dis. 1964;40:81-9.

225. Lukehart SA, Marra CM. Isolation and laboratory maintenance of Treponema pallidum. Curr Protoc Microbiol. 2007;Chapter 12:Unit 12A.1.

226. Noris E, Accotto GP, Luisoni E. Advances in diagnosing tomato yellow leaf curl geminivirus infection. Mol Biotechnol. 1994;2(3):219-26.

227. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 1975;94(3):441-8.

228. Kaye S. Simple PCR-based method for synthesis of molecular calibrators and controls. Clin Chem. 2003;49(2):325-7.

229. Simmonds P, Balfe P, Peutherer JF, Ludlam CA, Bishop JO, Brown AJ. Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. J Virol. 1990;64(2):864-72.

230. Miller JN. Value and limitations of nontreponemal and treponemal tests in the laboratory diagnosis of syphilis. Clin Obstet Gynecol. 1975;18(1):191-203.

231. Personal communitation with Hanna M. 2010.

232. Risbud A, Chan-Tack K, Gadkari D, Gangakhedkar R, Shepherd M, Bollinger R, et al. The etiology of genital ulcer disease by multiplex polymerase chain reaction and relationship to HIV infection among patients attending sexually transmitted disease clinics in Pune, India. Sex Transm Dis. 1999;26(1):55-62.

233. O'Farrell N, Hoosen A, Coetzee K, van den Ende J. Genital ulcer disease in men in Durban, South Africa. Genitourin Med. 1991;67(4):327-30.

234. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987;162(1):156-9.

235. Esser KH, Marx WH, Lisowsky T. Nucleic acid-free Matrix: Regeneration of DNA binding columns. Biotechniques. 2005;39(2):270-71.

236. Applied Biosystems. Tempus Blood RNA Tube and Tempus Spin RNA Isolation Kit Protocol. Applied Biosystems; 2008.

237. Duale N, Brunborg G, Rønningen KS, Briese T, Aarem J, Aas KK, et al. Human blood RNA stabilization in samples collected and transported for a large biobank. BMC Res Notes. 2012;5:510.

238. Matheson LA, Duong TT, Rosenberg AM, Yeung RS. Assessment of sample collection and storage methods for multicenter immunologic research in children. J Immunol Methods. 2008;339(1):82-9.

239. Wilson IG. Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol. 1997;63(10):3741-51.

240. Palmer H, Higgins S, Herring A, Kingston M. Use of PCR in the diagnosis of early syphilis in the United Kingdom. Sex Transm Infect. 2003;79(6):479-83.

241. Wittwer C, Kusakawa N, Persing D, Tenover F, Versalovic J, Tang J, et al. Molecular Microbiology: Diagnostic Principles and Practice: Washington ASM Press; 2004.

242. Wong ML, Medrano JF. Real-time PCR for mRNA quantitation. Biotechniques. 2005;39(1):75-85.

243. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnol Lett. 2004;26(6):509-15.

244. Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, Mathieu C. An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. Methods. 2001;25(4):386-401.

245. Mannhalter C, Koizar D, Mitterbauer G. Evaluation of RNA isolation methods and reference genes for RT-PCR analyses of rare target RNA. Clin Chem Lab Med. 2000;38(2):171-7.

246. Wicher K, Wicher V, Nakeeb SM, Dubiski S. Studies of rabbit testes infected with Treponema pallidum. I Immunopathology. Br J Vener Dis. 1983;59(6):349-58.

247. Wheeler HL, Agarwal S, Goh BT. Dark ground microscopy and treponemal serological tests in the diagnosis of early syphilis. Sex Transm Infect. 2004;80(5):411-4.

248. Burton MJ, Holland MJ, Jeffries D, Mabey DC, Bailey RL. Conjunctival chlamydial 16S ribosomal RNA expression in trachoma: is chlamydial metabolic activity required for disease to develop? Clin Infect Dis. 2006;42(4):463-70.

249. Schöfer H, Imhof M, Thoma-Greber E, Brockmeyer N, Hartmann M, Gerken G, et al. Active syphilis in HIV infection: a multicentre retrospective survey. The German AIDS Study Group (GASG). Genitourin Med. 1996;72(3):176-81.

250. Lynn W, Lightman S. Syphilis and HIV: a dangerous combination. Lancet Infect Dis. 2004;4(7):456-66.

251. Tang YW, Procop GW, Persing DH. Molecular diagnostics of infectious diseases. Clin Chem. 1997;43(11):2021-38.

252. Williams I, Churchill D, Anderson J, Boffito M, Bower M, Cairns G, et al. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012. HIV Med. 2012;13 Suppl 2:1-85.

253. Rockstroh JK, Bhagani S, Benhamou Y, Bruno R, Mauss S, Peters L, et al. European AIDS Clinical Society (EACS) guidelines for the clinical management and treatment of chronic hepatitis B and C coinfection in HIV-infected adults. HIV Med. 2008;9(2):82-8.

254. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 2009;55(4):611-22.

255. Dixon S, Horti J, Guo Y, Reed E, Figg W. Methods for extracting and amplifying genomic DNA isolated from frozen serum. Nat Biotechnol. 1998;16:91-4.

256. de Kok JB, Hendriks JC, van Solinge WW, Willems HL, Mensink EJ, Swinkels DW. Use of real-time quantitative PCR to compare DNA isolation methods. Clin Chem. 1998;44(10):2201-4.

257. Hamilton MS, Otto M, Nickell A, Abel D, Ballam Y, Schremmer R. High frequency of competitive inhibition in the Roche Cobas AMPLICOR multiplex PCR for Chlamydia trachomatis and Neisseria gonorrhoeae. J Clin Microbiol. 2002;40(11):4393.

258. LifeTechnologies. Real-time PCR handbook, second edition. 2012.

259. Kramvis A, Bukofzer S, Kew MC. Comparison of hepatitis B virus DNA extractions from serum by the QIAamp blood kit, GeneReleaser, and the phenol-chloroform method. J Clin Microbiol. 1996;34(11):2731-3.

260. Pound MW, May DB. Proposed mechanisms and preventative options of Jarisch-Herxheimer reactions. J Clin Pharm Ther. 2005;30(3):291-5.

261. Yang CJ, Lee NY, Lin YH, Lee HC, Ko WC, Liao CH, et al. Jarisch-Herxheimer reaction after penicillin therapy among patients with syphilis in the era of the hiv infection epidemic: incidence and risk factors. Clin Infect Dis. 2010;51(8):976-9.

262. Rolfs R, Joesoef M, Hendershot E, Rompalo A, Augenbraun M, Chiu M, et al. A randomized trial of enhanced therapy for early syphilis in patients with and without human immunodeficiency virus infection. The Syphilis and HIV Study Group. N Engl J Med. 1997;337(5):307-14.

263. Song JH, Chang HH, Suh JY, Ko KS, Jung SI, Oh WS, et al. Macrolide resistance and genotypic characterization of Streptococcus pneumoniae in Asian countries: a study of the Asian Network for Surveillance of Resistant Pathogens (ANSORP). J Antimicrob Chemother. 2004;53(3):457-63.

264. Liu H, Rodes B, George R, Steiner B. Molecular characterization and analysis of a gene encoding the acidic repeat protein (Arp) of Treponema pallidum. J Med Microbiol. 2007;56(Pt 6):715-21.

265. Henke W, Herdel K, Jung K, Schnorr D, Loening SA. Betaine improves the PCR amplification of GC-rich DNA sequences. Nucleic Acids Res. 1997;25(19):3957-8.

266. Marra CM, Colina AP, Godornes C, Tantalo LC, Puray M, Centurion-Lara A, et al. Antibiotic selection may contribute to increases in macrolide-resistant Treponema pallidum. J Infect Dis. 2006;194(12):1771-3.

267. World Health Organisation Guidelines: prevention and treatment of HIV and other sexually transmitted infections among men who have sex with men and transgender people: recommendations for a public health approach. 2011.

268. Florindo C, Reigado V, Gomes JP, Azevedo J, Santo I, Borrego MJ. Molecular typing of treponema pallidum clinical strains from Lisbon, Portugal. J Clin Microbiol. 2008;46(11):3802-3.

269. Personal Communication with Greene L. 2012.

270. Yamamoto Y. PCR in diagnosis of infection: detection of bacteria in cerebrospinal fluids. Clin Diagn Lab Immunol. 2002;9(3):508-14.

271. Evans CA (2011) GeneXpert—A Game-Changer for Tuberculosis Control? PLoS Med 8(7): e1001064. doi:10.1371/journal.pmed.1001064

272. Workowski K, Berman S, Prevention CfDCa. Sexually transmitted diseases treatment guidelines, 2006. MMWR Recomm Rep. 2006;55(RR-11):1-94.

273. Greenblatt RM, Lukehart SA, Plummer FA, Quinn TC, Critchlow CW, Ashley RL, et al. Genital ulceration as a risk factor for human immunodeficiency virus infection. AIDS. 1988;2(1):47-50.

274. Arroll TW, Centurion-Lara A, Lukehart SA, Van Voorhis WC. T-Cell responses to Treponema pallidum subsp. pallidum antigens during the course of experimental syphilis infection. Infect Immun. 1999;67(9):4757-63.

275. Lalvani A, Brookes R, Hambleton S, Britton WJ, Hill AV, McMichael AJ. Rapid effector function in CD8+ memory T cells. J Exp Med. 1997;186(6):859-65.

Appendices

Appendix 1: List of Organisms used to assess species specificity of assay

Gram Characteristics	Organisms
Gram positive bacteria	*α-haemolytic streptococci, Streptococcus milleri, Streptococcus pneumoniae, Streptococcus pyogenes, Streptococcus mitis, Streptococcus agalacticae, Streptococcus faecalis, Micrococcus luteus, Diptheroids, Coagulase-negative Staphylococci, Staphylococcus aureus, Methicillin-resistant Staphylococcus aureus, lactobacilli*
Gram negative bacteria	*Neisseria gonorrhoeae, Neisseria meningitidis, Campylobacter jejuni, Coliforms, Proteus sp, Escherichia coli, Klebsiella sp, Haemophilus sp, Enterobacter, enterobacteracae, Chlamydia trachomatis, Treponema denticola, Treponema pallidum* subspecies *pertenue (Gauthier strain)*
Viruses	*Epstein-Barr virus, Cytomegalovirus, Herpes simplex virus, HIV-1, HTLV-1*
Yeast	*Candida albicans*

Appendix 2: Primers and Probes

Primer	Target gene	Sequence (5'-3')	Method	Product Size (bp)
TP1 and TP 2	*tpp047*	TP1 (sense) CGAGGAATACAAGATTACGAACG TP2 (anti-sense) ACGTGCAGAAAAACTATCCTCAG	qPCR	178
TP3 and TP4	*tpp047*	TP3 (sense) CCCTCTATACGGAGGTGTAAT TP4 (anti-sense) TTAAGACACACGGGATAGGAC	Splice-overlap PCR	1357
TP5	*tpp047*	CGCATTCTGTGACTACTCTACCCAGACAAAC TTCATACCGTGAA	Splice-overlap PCR	577
TP6	*tpp047*	CTGGGTAGAGTAGTCACAGAATGCGAAAGG GATTGCAAAGGAGA	Splice-overlap PCR	805
Tpnest4F and Tpnest4R	*tpp047*	Tpnest4F (sense) ATGGTGCGTGCCTCTGCGTC Tpnest4R (anti-sense) AGGTCGTGCGGGCTCTCCAT	pTp47 construction	326
A2058G_F and A2058G_R (169)	*23S rRNA*	A2058G_F (sense) GTACCGCAAACCGACACAG A2058G_R (anti-sense) AGTCAAACCGCCCACCTA	pTp2058 and pTp2059 construction 2nd round amplification in *23S rRNA* mutation analysis	629
TP0226_F and TP0226_R (169)	*23S rRNA (copy 1)*	TP0226_F GTACCGCAAACCGACACAG TP0226_R GCGCGAACACCTCTTTTTAC	1st round amplification of *23SrRNA* copy 1	1593
TP0267_F and TP0267_R (169)	*23S rRNA (copy 2)*	TP0267_F GTACCGCAAACCGACACAG TP0267_R GAACCGTCCCTGAAAACTCA	1st round amplification of *23SrRNA* copy 2	1593
ARP F and ARP R (82)	*arp*	ARP_F CAAGTCAGGACGGACTGTCC ARP_R GGTATCACCTGGGGATGC	*arp* amplification for strain-typing	variable
TP548 AS_3 and TP548_S (86)	*tpp0548*	TP548AS_3 CGTTTCGGTGTGTGAGTCAT TP548_S GGTCCCTATGATATCGTGTTCG	*tpp0548* amplification for strain-typing	300
TP548_AS (274)	*tpp0548*	TP548_AS GTCATGGATCTGCGAGTGG	*tpp0548* gene sequence analysis	300
EGJ_F and EGJ_R (82)	*tpr*	EGJ_F CAGGTTTTGCCGTTAAGC EGJ_R AATCAAGGGAGAATACCGTC	*tpr* amplification for strain typing	

16s_F(seattle) (153) 16s_R7	*16S rRNA*	16s_F(seattle) CTCTTTTGGACGTAGGTCTTTGAG 16s_R7 TCACCCTCTCAGGTCGGATA	*16S rRNA* amplification and quantification	131
16S_P3 Probe	*16S rRNA*	FAM -CGGCCTCGCTCAGAGATGAGC -TAMRA	*16S rRNA* qPCR	-
16srRNA_F2 16srRNA_R2	*16S rRNA*	AAGCGAATTTTCGTTTCT GTCGGATACCCATCGA	*16S rRNA qPCR*	225
16S_P2		TCTTTTGGACGTAGGTCTTTGA		
FLAB2_F FLAB2_R	*flaB2*	FLAB2_F GCGCTATGTTCTCTCAAA FLAB2_R TTCAGAAACTGCAAGACC	*flaB2* qPCR	128
FlaB2 _P Probe	*flaB2*	FAM-ACATTGAGAAGTTGTCTTCAGG-TAMRA	*flaB2* qPCR	-
TP_P Probe	*tpp047*	FAM-CCCAGTTGCGGTTCCTCATGAATTA-TAMRA	*tpp047* qPCR	
TPP-QC probe (228)	*modified tpp047*	Yakima Yellow-CTGGGTAGAGTAGTCACAGAATGCG-BHQ-1	pTpQC detection and quantification	
TPP-QCR probe	*modified tpp047*	ROX-CTGGGTAGAGTCGTCACAGAATGCG-BHQ1	pTpQC detection and quantification	
TYLCV_F3 and TYLCV_R3	*TYLCV*	TYLCV_F3 CGTCCGTCGATCTGGAAAGT TYLCV_R3 TAGCTCCCTGAATGTTCGGAT	Internal control of *T. pallidum* DNA extraction and quantification	100bp
P_TYLCV	*TYLCV*	Yakima Yellow-TTGTCGATATAGGACTTGACGTCGGAGC-BHQ1	TYLCV detection and quantification	

Primers and probes were synthesised either by Eurogentec, Seraing, Belgium; Invitrogen, Paisley, UK or Sigma, Gillingham, UK.

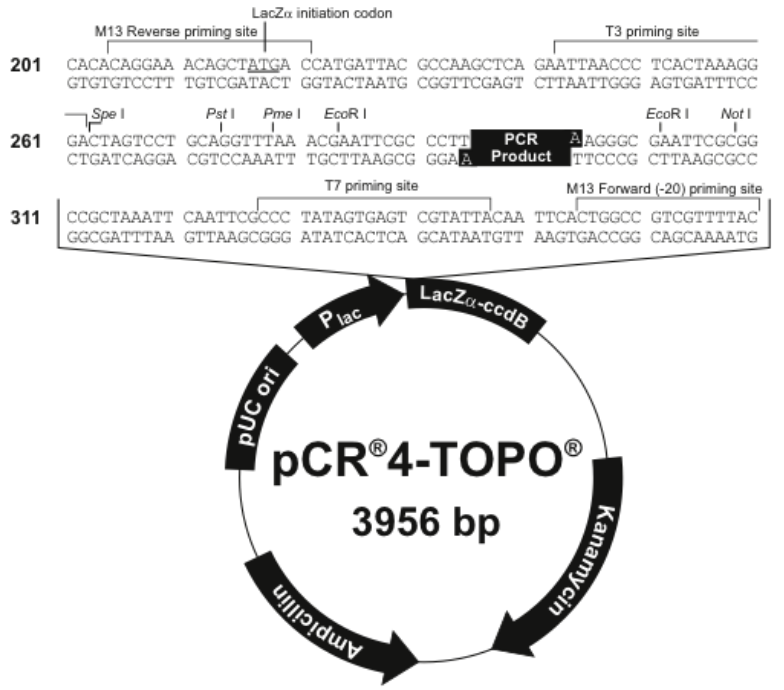
Appendix 3: Restriction endonucleases used in the project

Enzyme	Used for	Recognition Sequence	Conditions
Nco-I (New England Biolabs, Ipswich, MA)	Linearisation of plasmids pTpQC, pTp47, p16SRNA, pTYLCV	5'…C^{V}CATGG…3' 3'…GGTAC$_{\Lambda}$C…5'	1 hour at 37°C NEB Buffer 3
Eco-RI (New England Biolabs, Ipswich, MA)	Excision of *T. pallidum* sequence from pTp2058, pTp2059 and pTp47	5'…G^{V}AATTC…3' 3'…CTTAA$_{\Lambda}$G…5'	1 hour at 37°C NEB Buffer 1
Mbo-II (New England Biolabs, Ipswich, MA)	A2058G mutation analysis	5'…GAAGA(N)$_{8}$ ^{V}…3' 3'…CTTCT (N)$_{7}$ $_{\Lambda}$…5'	1 hour at 37°C NEB Buffer 4
Bsa-I (New England Biolabs, Ipswich, MA)	A2059G mutation analysis	5'…GGTCTC(N)$_{1}$ ^{V}…3' 3'…CCAGAG (N)$_{5}$ $_{\Lambda}$…5'	1 hour at 50°C NEB Buffer 4 1x BSA
Mse-I (New England Biolabs, Ipswich, MA)	Strain-typing. *tpr* sub-family 2 gene analysis.	5'…T^{V}TAA…3' 3'…AAT$_{\Lambda}$T…5'	1 hour at 37°C NEB Buffer 4 1x BSA

Appendix 4: Map of pCR4-TOPO cloning vector

Map of pCR®4-TOPO®

pCR®4-TOPO® Map The map below shows the features of pCR®4-TOPO® and the sequence surrounding the TOPO® Cloning site. Restriction sites are labeled to indicate the actual cleavage site. **The complete sequence of pCR®4-TOPO® is available for downloading from our Web site (www.invitrogen.com) or by contacting Technical Service (page 27).**



Structure map of the 3956 bp cloning vector used for the construction of the quantification standards for *16SrRNA*; *tpp047* (wild-type and mutated) and TYLCV in addition to plasmids containing the mutated *T. pallidum 23S rRNA* sequences A2058G and A2059G. The cloning site (bases 294-295) is marked as 'PCR product' and is flanked by *EcoRI* restriction sites. The T7 priming site is at bases 328-347 and the Ampicillin (*bla*) resistance gene and promoter are located at bases 2203-3160.

Appendix 5: Case definitions for cross-sectional studies of *T. pallidum* quantification

Stage/Disease	Definition
HSV	Typical single or multiple genital ulcers from which HSV-1 or HSV-2 DNA is isolated using an in house PCR.
Primary syphilis	Diagnosis requires the presence of a chancre on examination and microbiological confirmation of *T. pallidum* with any of the following: dark ground microscopy; polymerase chain reaction provided by the sexually transmitted bacteria reference laboratory, Health Protection Agency, London or positive treponemal serology (as per national guideline) (104). Primary diagnoses to include dark ground negative ulcers and initially negative serology, which becomes positive up to 3 months later with no sign of secondary disease.
Secondary syphilis	Defined as syphilis within the first two years of infection and characterised by both clinical and microbiological findings. Clinical findings include: typical rash; condylomata lata; muco-cutaneous lesions; generalised lymphadenopathy; anterior uveitis; hepatitis and splenomegaly. Microbiological findings must include serological tests consistent with secondary disease and may, if available, include PCR of secondary lesions (104). If ano-genital examination reveals persisting chancres, the case is still counted as secondary.
Other syphilis	*Latent syphilis* Positive serological tests in the absence of any symptom or sign. Classified as early within the first 2 years of infection and late thereafter. *Symptomatic late syphilis* Positive serology including a negative or low and unchanging RPR titre together with symptoms and clinical signs of cardiovascular or

	gummatous disease. Late disease with neurological involvement is classed as neurosyphilis.
Neurosyphilis	Any stage of disease with: clinical neurological involvement[†] (including meningitis, cranial nerve palsies, parenchymatous general paresis, tabes dorsalis); serological findings consistent with the stage of disease and a relevant CSF abnormality (as per national guidelines) (104). NB. For analysis, cases with neurological involvement were counted as neurosyphilis not by stage. [†]Patients without neurological symptoms do not routinely undergo CSF examination at our centre.
Asymptomatic contact of syphilis	Patient reports sexual contact with a suspected or confirmed case of syphilis. Examination reveals no sign of disease and serological results (baseline, 6 weeks and 3 months after exposure) are not consistent with current infection.
Non-specific genital ulceration	Genital ulceration is present on examination, but no microbiological diagnosis is made. PCR for HSV is negative; DGM is negative; syphilis serology is not consistent with current infection and diagnostic *T. pallidum* PCR (if performed) is negative.

Appendix 6: Standard clinical investigation and study results for all participants

Final Diagnosis	Patient	HIV-1 Status	CD4 count	DGM result[§]	HSV PCR Result[§]	Syphilis EIA result at enrolment	RPR result at enrolment	Timing of DNA extraction	WB *tpp047* (copies/ml whole blood)	Ulcer *tpp047* Copy Number (copies/strip)
Primary Syphilis (n=14)	6	-	-	+	-	+	8	immediate	0	620
	9	+	390	+		+	16	immediate	127	2108
	11	+	480	+		+	16	immediate	0	12440
	18	+	270	-	-	+	0	immediate	0	757
	23	-	-	+		+	4	immediate	0	2742
	24	-	-	+		-	16	immediate	0	2577
	26	+	920	+		-		immediate	0	7426
	32	+	720	-		+	4	immediate	168	14244
	33	-	-	-	-	+	64	immediate	0	427
	52	+	320	+		+	2	later	0	2115
	69	-	-	+	-	+	32	later	0	251
	89	+	420	+		-		later	0	1555
	90	-	-	-		+	4	later	127	1064
	93	-	-	+	-	+	4	later	0	648
Secondary Syphilis (n=19)	3	+	660	unavailable		+	64	immediate	0	2
	16	+	240	-	-	+	128	immediate	886	630
	17	-	-	+	-	+	128	immediate	1026	422
	22	-	-	+		+	64	immediate	518	211*
	27	+	680			+	8	immediate	632	
	29	-	-			+	16	immediate	58	
	34	+	390	unavailable	-	+	128	immediate	495	251*
	45	+	350			+	64	later	6	
	46	+	330			+	16	later	0	
	47	-	-			+	256	later	0	
	49	+	630			+	128	later	10	
	53	+	400			+	64	later	28	
	61	+	470			+	512	later	91	
	63	+	790			+	32	later	0	
	74	+	410			+	128	later	0	
	75	+	470			+	256	later	0	
	77	+	370			+	32	later	0	
	81	+	430			+	128	later	65	
	83	+	420			+	128	later	0	unavailable
Other Syphilis (latent or symptomatic late) (n=8)	1	-	-			+	0	immediate	0	
	12	+	400			+	1	immediate	0	
	28	+	220			+	256	immediate	70	
	44	-	-			+	16	later	0	
	59	+	610			+	0	later	0	
	67	-	-			+	0	later	0	
	84	+	570			+	0	later	0	
	95	-	-			+	1	later	0	
Asymptomatic Contact of Syphilis (n=17)	2	-	-			+	1	immediate	0	
	4	-	-			-		immediate	0	
	8	-	-			-		immediate	0	
	15	+	730			-		immediate	0	
	25	-	-			-		immediate	0	
	30	+	320			-		immediate	0	
	38	+	490			+	0	immediate	0	
	51	-	-			-		later	0	
	57	-	-			-		later	0	
	58	+	470			+	0	later	0	
	65	+	490			-		later	0	
	73	-	-			-		later	0	
	76	+	670			+	0	later	0	
	78	+	510			-		later	0	
	88	+	320			+	0	later	0	
	94	+	640			-		later	0	
	99	-	-			-		later	0	
HSV Infection (n=19)	10	-	-	-	+	-		immediate	0	0
	19	-	-	-	+	-		immediate	0	0
	20	-	-	unavailable	+	-		immediate	0	0
	31	-	-	-	+	-		immediate	0	0
	37	-	-	-	+	-		immediate	0	0
	39	-	-	-	+	-		immediate	0	0
	40	-	-	-	+	+	0	later	0	0
	41	-	-	-	+	-		later	0	0
	42	+	780	-	-	-		later	0	0
	43	-	-	-	+	unavailable		later	0	0
	55	+	550	-	-	-		later	0	0
	70	-	-	-	+	-		later	0	0
	71	+	680	-	+	-		later	0	0
	72	-	-	-	+	-		later	0	0
	79	-	-	-	+	-		later	0	0
	86	-	-	-	+	-		later	0	0
	91	+	370	unavailable	+	-		later	0	unavailable
	92	-	-	+	+	-		later	0	0
	98	-	-	+	+	-		later	0	0
Non-specific genital ulceration (n=19)	5	-	-	-	-	-		immediate	0	0
	7	+	540	+		+	0	immediate	0	0
	13	-	-	-	-	-		immediate	0	0
	14	-	-	-	-	-		immediate	0	0
	21	-	-	-	-	-		immediate	0	0
	35	-	-	-	-	-		immediate	0	0
	48	-	-	-	-	-		later	0	0
	50	-	-	+	-	+	0	later	0	0
	54	+	210			-		later	0	
	60	+	750	-	-	-		later	0	0
	62	+	400	-	-	-		later	0	46Ψ
	64	+	130			+	0	later	0	
	66	+	590	-	-	-		later	0	0
	68	+	580			-		later	0	
	80	-	-	-	-	-		later	0	0
	82	+	340	-	-	-		later	0	0
	85	+	120	-	-	+	1	later	0	0
	87	+	610	-	-	-		later	0	0
	97	+	500			-		later	0	
Neurosyphilis (n=3)	36	+	290			+	32	immediate	0	
	56	-	-			+	64	later	0	
	96	+	660			+	8	later	0	

* mucocutaneous oral lesion

[§] Test only performed when clinically indicated

Ψ Patient with a one-week history of multiple non-tender genital ulcers thought to be atypical genital herpes and treated with acyclovir. Convalescent syphilis serology and PCR for *herpes simplex virus* types one and two were negative.

Appendix 7: Sample collection schedule for clinical study of treatment kinetics

Day	Time point		Samples taken
1 (in the GU medicine clinic)	1a	(pre-treatment)	Ulcer snostrip (if primary disease) Blood: 1x 6ml rust/red top tube for syphilis serology (EIA, RPR, TPPA), 1x 4ml gold top tube LFTs, 1x 4ml EDTA (purple top), 1x3ml Tempus™ sample (blue tube) blood sample. Pregnancy testing for female patients of child-bearing potential.
	1b		Intravenous Cannula inserted.
	2	Treatment	Patient administered antibiotic selected by treating physician by nursing team at GU Medicine clinic.
1 (inpatient)	3	2 hours post-treatment	2ml EDTA blood sample; 3ml Tempus™ sample +/- ulcer sample
	4-9	4-hourly sampling (total 6 samples) i.e 00:00, 04:00 08:00, 12:00, 16:00, 20:00	2ml EDTA blood sample; 3ml Tempus™ sample +/- ulcer sample
2 (inpatient)	10-14	4-hourly sampling (total 5 samples) i.e. 00:00, 08:00, 12:00, 16:00, 20:00	2ml EDTA blood sample; 3ml Tempus™ sample +/- ulcer sample
3 (inpatient)	15-18	6-hourly sampling (total 4 samples) i.e. 00:00, 06:00, 12:00, 18:00	2ml EDTA blood sample; 3mL Tempus™ sample +/- ulcer sample
4-7 (outpatient)	19-27	12 hourly sampling (8 samples) i.e. 08:00 and 20:00	2ml EDTA blood sample; 3mL Tempus™ sample +/- ulcer sample

*The amount to blood taken for the study will not exceed 125ml.

Appendix 8: Publications relating to the thesis

For copyright reasons, the publications relating to this thesis have been removed. They can be found at: http://sti.bmj.com/

The references are:

Tipple C, McClure MO, Taylor GP. High prevalence of macrolide resistant *T. pallidum* strains in a London centre. Sex Transm Infect 2011;87:486e488. doi:10.1136/sextrans-2011-050082

Tipple C, Hanna MOF, Hill S, et al. Getting the measure of syphilis: qPCR to better understand early infection. Sex Transm Infect (2011). doi:10.1136/sti.2011.049494