RESEARCH ARTICLE

Characterizing the Syphilis-Causing Treponema pallidum ssp. pallidum Proteome Using Complementary Mass Spectrometry

Kara K. Osbak1, Simon Houston2, Karen V. Lithgow2, Conor J. Meehan3, Michal Strouhal4, David Šmajs5, Caroline E. Cameron2, Xaveer Van Ostade5, Chris R. Kenyon1,6, Geert A. Van Raemdonck1,6

1 HIV/STI Unit, Institute of Tropical Medicine, Antwerp, Belgium, 2 Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada, 3 Unit of Mycobacteriology, Institute of Tropical Medicine, Antwerp, Belgium, 4 Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic, 5 Laboratory for Protein Science, Proteomics and Epigenetic Signaling (PPES) and Centre for Proteomics (CFP), University of Antwerp, Wilrijk, Belgium, 6 Division of Infectious Diseases and HIV Medicine, University of Cape Town, Cape Town, South Africa

* kosbak@hotmail.com

Abstract

Background

The spirochete bacterium Treponema pallidum ssp. pallidum is the etiological agent of syphilis, a chronic multistage disease. Little is known about the global T. pallidum proteome, therefore mass spectrometry studies are needed to bring insights into pathogenicity and protein expression profiles during infection.

Methodology/Principal Findings

To better understand the T. pallidum proteome profile during infection, we studied T. pallidum ssp. pallidum DAL-1 strain bacteria isolated from rabbits using complementary mass spectrometry techniques, including multidimensional peptide separation and protein identification via matrix-assisted laser desorption ionization-time of flight (MALDI-TOF/TOF) and electrospray ionization (ESI-LTQ-Orbitrap) tandem mass spectrometry. A total of 6033 peptides were detected, corresponding to 557 unique T. pallidum proteins at a high level of confidence, representing 54% of the predicted proteome. A previous gel-based T. pallidum MS proteome study detected 58 of these proteins. One hundred fourteen of the detected proteins were previously annotated as hypothetical or uncharacterized proteins; this is the first account of 106 of these proteins at the protein level. Detected proteins were characterized according to their predicted biological function and localization; half were allocated into a wide range of functional categories. Proteins annotated as potential membrane proteins and proteins with unclear functional annotations were subjected to an additional bioinformatics pipeline analysis to facilitate further characterization. A total of 116 potential membrane proteins were identified, of which 16 have evidence supporting outer membrane localization. We found 8/12 proteins related to the paralogous tpr gene family: TprB,
data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

**Conclusions**

This is the most comprehensive description of the global *T. pallidum* proteome to date. These data provide valuable insights into *in vivo* *T. pallidum* protein expression, paving the way for improved understanding of the pathogenicity of this enigmatic organism.

**Author Summary**

Syphilis remains a major cause of morbidity and mortality worldwide. The bacterium causing syphilis, *Treponema pallidum* ssp. *pallidum*, has evolved into a highly distinctive organism that is only able survive (and be propagated) in mammals. In humans it can evade the immune system for decades with devastating consequences. Much remains to be learned about how it accomplishes this. Only a minority of its predicted proteins have been detected experimentally thus far. We aimed to more comprehensively characterize the proteins of this organism. Since it cannot be cultured *in vitro*, we cultured *T. pallidum* in rabbits and analyzed extracted proteins using different mass spectrometry methods, a manner of detecting proteins with high accuracy. In total, we detected more than half of the predicted number of proteins that could be expressed by this bacterium (N = 557). For approximately half of the proteins, we succeeded in characterizing their predicted cellular location using an array of bioinformatic tools and catalogued their function. This is the most comprehensive analysis of the *T. pallidum* proteome to date. This study lays the groundwork for other protein investigations of this unique organism.

**Introduction**

*Treponema pallium* ssp. *pallidum*, henceforth referred to as *T. pallidum*, is the causative agent of syphilis, a multistage chronic disease with an estimated 8 million new cases per year [1]. Recent outbreaks of syphilis infection among certain populations such as men who have sex with men (MSM) [2], together with continuing substantial perinatal morbidity and mortality attributed to congenital syphilis infections [3], highlight the need for improved diagnostics and vaccine development.

*T. pallidum* is an obligate microaerophilic bacterial pathogen [4–6] that is aptly suited to invading mammalian tissue by the use of endoflagella that produce undulations in travelling planar waves [7], thereby driving its characteristic corkscrew-like movement [8]. The membrane of *T. pallidum* lacks lipopolysaccharide (LPS) and the loosely associated fragile outer membrane contains a low amount of proteins [8–11]. Many biomedical experimental approaches such as genetic manipulation have been hampered by its lack of *in vitro* cultivability [12]. Despite these limitations, numerous studies using *T. pallidum* harvested from the experimental rabbit model have increased our basic biological understanding of this unique organism, including the description of the genome [13,14], transcriptome [15] and proteome [16,17].

The *T. pallidum* Nichols strain genome was sequenced for the first time in the late 1990’s [13], revealing only 1041 predicted open reading frames (ORFs) on a 1.14 Mb circular
chromosome, making it one of the smallest human pathogen genomes. Resequencing of the Nichols strain [14] identified 102 errors that were predicted to affect protein-coding genes and reduced the number of ORFs to 1039, 968 of which are predicted to be protein coding. Similar to other obligate pathogens such as *Mycoplasma pneumoniae* [18], *T. pallidum* is predicted to have lost many non-essential genes though genome reduction. This theory is supported by extensive genome-wide transcriptional analyses [15], which revealed the uniform expression of almost all *T. pallidum* genes during experimental rabbit infection. Consequently, *T. pallidum* has severely limited metabolic and biosynthetic capabilities, rendering it highly dependent on the host milieu and nutrients [19].

The predicted proteins found within *T. pallidum* range in size from 3,235 to 172,869 Da with an average size of 37,771 Da [13,20]. Early studies on *T. pallidum* polypeptides, including pre-MS analysis gel-based techniques and the use of recombinant DNA technology has been extensively reviewed by Norris et al. [17] and Schouls [21]. A large scale *T. pallidum* recombinant protein study included the construction of a bacterial artificial chromosome (BAC) library containing 901 of the 1039 *T. pallidum* predicted proteins for expression in *Escherichia coli* [22,23]; many of the expressed proteins were reactive with sera from syphilis-infected rabbits and/or humans [24,25] at different stages of infection as determined by serological reactivity studies. Subsequently, McGill et al. conducted a *T. pallidum* proteome investigation on *in vivo* expressed *T. pallidum* using gel-based approaches complemented with MALDI-TOF Mass Spectrometry (MS) and peptide mass fingerprinting [16]. A total of 88 polypeptides were identified and the immunoreactive potential of select proteins was characterized. Numerous bioinformatic approaches have been used to characterize *T. pallidum* proteins, including lipoprotein characterization [26], the determination of potential outer membrane proteins [27] and the re-annotation of *T. pallidum* strain SS14 hypothetical proteins [28]. However, despite rigorous analyses and major advances in genome sequencing, approximately 30% of *T. pallidum* proteins still have no known orthologues and at present cannot be assigned a biological function [13]. This ‘unknown’ category of proteins may represent an arsenal of genes encoding virulence factors specific for *T. pallidum* [20].

Progress has been made on understanding virulence and persistence strategies of this unique pathogen. Genetic sequence diversity is primarily localized in six hot spots [29] in *T. pallidum* ssp. pallidum and *T. pallidum* ssp. pertenue (the causative agents of Yaws), including regions encoding members of the paralogous *tpr* gene family consisting of 12 genes categorized into subfamilies I (*tprC, D, F and J*), II (*tprE, G and J*) and III (*tprA, B, H, K and L*). The Tpr proteins contribute to antigenic variation that aids in immune evasion [30]. Nonreciprocal gene conversion occurs between donor sites and several variable regions (V1–7) in *tprK* [31] and these variable regions in the encoded protein are targets of the host humoral response during infection [32–35]. Host immune pressure is capable of selecting against certain TprK sequence epitopes [36] and TprK sequence variability can help evade the host immune response [35] during infection. Recombinant protein studies have confirmed surface exposure, bipartite architecture and porin function related to the outer membrane proteins Tpr C/D [37] and TprI [38]. Moreover, *T. pallidum* lipoproteins, of which Tp47 is the most widely studied [39–42], play an important role in immune system activation and evasion as reviewed by Kelesidis et al. [43].

With the recent evolution of robust highly sensitive tandem MS instrumentation, the comprehensive description of bacterial proteomes, also referred to as shotgun proteomics (reviewed by Semanjski et al. [44]), is achievable. Many current state-of-the-art proteomic studies have approached 80% coverage of the predicted expressed proteome [45,46]. The study of pathogens expressed *in vivo* is of particular interest since this would be the closest approximation of human pathophysiological conditions. For example, previous studies on the *Mycoplasma tuberculosis* proteome from guinea pig infected lungs during early and chronic stages of disease...
have provided valuable insights into pathogen protein expression. However, interference of host proteins present in large excess can hinder MS detection of low abundance pathogen proteins. Thus, several strategies have been used to overcome issues of sample complexity to enrich bacteria cells, such as the use of density gradient centrifugation [16].

Using highly sensitive non-gel based complementary proteomic techniques, we sought to further elucidate the global proteome of T. pallidum in order to gain insights into the fundamental physiological state of T. pallidum during rabbit infection. Three biological replicates of in vivo cultured T. pallidum were subjected to multidimensional chromatographic separation and tandem MS/MS analysis whereby 557 T. pallidum proteins were identified at a high level of confidence, representing 54% of the predicted proteome. This is the first description of 499 T. pallidum proteins expressed in vivo, of which 106 were annotated as uncharacterized/hypothetical proteins. Detected proteins were comprehensively analysed to predict cellular localization and function. This unique ‘snapshot’ view of the T. pallidum proteome during infection extends our understanding of T. pallidum pathogenesis and forms the basis for further proteome investigations.

Methods

Rabbit inoculation and T. pallidum purification using Percoll density gradient centrifugation

Three biological samples, hereafter referred to as samples TPA-A, TPB-B and TPC-C, originated from three New Zealand White rabbits that were inoculated intra-testicullarly with T. pallidum DAL-1 strain bacteria according to established methods [48]. Inoculations originated from two different bacterial stocks of DAL-1 strain bacteria, whereby sample TPB-B and TPC-C originated from the same stock. When peak orchitis was reached, on average 11–14 days post-inoculation, rabbits were sacrificed using T61 administration according to the manufacturer’s instructions and the bacteria was extracted from the testes and purified using Percoll density gradient centrifugation as previously described [49]. Briefly, collected organisms were separated from host cellular gross debris by low-speed centrifugation at 34 800 g for 30 minutes followed by gradient separation via ultra-centrifugation at 100 000 g for 1 hour. Bacteria were quantified using darkfield microscopy and a counting chamber. For sample TPA-A, approximately 10^8–9 treponemes were re-suspended and stored in 1 mL NaCl solution and frozen at -80°C. For samples TPB-B and TPC-C, approximately 10^8–9 treponemes were re-suspended in 1 mL phosphate buffered saline (PBS) (HiMedia Laboratories, Mumbai, India) and frozen at -80°C. Two samples, TPA-A and TPB-B, were subjected to an extra thaw cycle before protein extraction due to inadvertent thawing during sample shipment. Each rabbit was tested serologically to rule out a naturally occurring infection with T. paraluiscuniculi.

Ethics statement

The treponemal DAL-1 strain was propagated in rabbits at the Veterinary Research Institute in Brno, Czech Republic. The handling of animals in the study was performed in accordance with the current Czech legislation (Animal Protection and Welfare Act No. 246/1992 Coll. of the Government of the Czech Republic). These specific experiments were approved by the Ethics Committee of the Veterinary Research Institute (Permit Number 20–2014).

T. pallidum protein sample preparation

Cell lysis of the purified T. pallidum extract was performed by conducting three consecutive freeze-thaw cycles, followed by ultrasonication on ice with an amplitude of 50% and a pulser.
frequency of 2 seconds for 2 minutes (Sonics, Vibra cell; Newton USA). Protein concentration was determined by loading a small fraction of the lysed sample on a high performance liquid chromatographic (HPLC) reversed phase C4 system that was calibrated using a serial dilution of a protein standard mixture. Protein concentrations were determined based on the area under the curve (AUC at 214 nm). Approximately 400–500 μg of protein was extracted from each biological replicate; a large proportion of this amount was host protein in the form of albumin. Samples were acetone precipitated by adding 6 volumes of LC-MS grade acetone (Biosolve, Valkenswaard, Netherlands) and incubated overnight at -20°C. In all cases, lo-bind Eppendorf tubes (Eppendorf, Hamburg, Germany) were used to ensure high recovery rates of proteins/peptides.

**Protein enzymatic digestion**

Following protein precipitation, protein samples were re-suspended in 50 mM Tris-HCl/6 M urea/5 mM DTT/10% beta-mercaptoethanol (25 μL/100 μg protein) at pH 8.7. For the denaturation and reduction process all samples were incubated at 65°C during 1 hour. Subsequently, proteins in all fractions were diluted in 50 mM Tris-HCl/1 mM CaCl₂ (75 μL/100 μg protein) and alkylated by adding 200 mM iodoacetamide (10 μL/100 μg protein) during 1 hour at room temperature. Proteomics-grade modified trypsin (Promega, Madison, Wisconsin, United States) was added at a 30:1 protein-to-enzyme ratio. After incubation at 37°C for 18 hour the digestion was stopped by freezing the samples.

**Peptide separation by reversed phase C18 at high pH (1st dimension)**

After tryptic digestion, peptides were separated in a first dimension based on hydrophobicity at high pH by using a reversed phase C18 column (X!Select, CSH, RP-C18, 2.1 x 150 mm, 3.5 μm, Waters) connected to a Waters Alliance e2695 HPLC bio-system and a Waters 996 PDA detector (Waters Corporation, Milford, MA, USA). Solvent A contains 200 mM ammonium formate at pH 10, while solvent C contains 100% water and solvent D 100% acetonitrile (ACN) (LC-MS grade, Biosolve, Valkenswaard, Netherlands). During the chromatographic run, an ACN gradient was performed, while continuously 10% of solvent A was added to become an overall pH of 10 during the entire run. The following gradient was used at a constant flow rate of 200 μL/min: 5% to 15% D over the first 5 min, 15% to 40% D over 80 min, 40% to 90% D over 8 min, 5 min 90% D, and 90% to 5% D over 2 min. In total, 30 fractions were collected starting from 10 to 100 min with an interval of 3 min/fraction. The peptide concentration of the different fractions was determined based on the area under the curve (AUC at 214 nm). Fractions were pooled in a concatenated way (e.g. fractions 1, 11 and 21) to obtain optimal orthogonality, yielding in total 10 fractions for further analysis. Collected fractions were lyophilized and re-suspended in RP mobile phase (97% water, 3% ACN, 0.1% FA).

**Peptide separation by micro-capillary reversed phase C18 (2nd dimension)**

Peptide fractions were separated in a second dimension using an Agilent 1100 series micro-capillary HPLC system (Agilent Technologies, Waldbronn, Germany). For each fraction 15 μg of peptides was injected on a Zorbax 300SB-C18 guard column (0.3 mm x 5 mm; particle size 3.5 μm; Agilent Technologies) serially connected with a Zorbax 300SB-C18 analytical RP column (0.3 mm x 150 mm; particle size 3.5 μm; Agilent Technologies). Samples were online desalted by loading the peptides on the guard column before the ACN gradient was started. Solvent A contained 0.1% formic acid (FA) in water while solvent B contained 0.1% FA in 90% ACN /10% water. Following ACN gradient was performed using the capillary pump with a
constant flow rate at 6 μL/min: 5% to 60% B in 56.7 min, ramp to 90% B over 3.3 min persistent 90% B for 5 min, 85% B for 5 min and back to equilibrating conditions of 3% B. Starting from minute 5 until minute 51.7 of the chromatographic run, 350 spots (800 nl/spot) for each fraction were spotted on an Opti-TOF MALDI-target (28 columns x 25 rows; 8 sec interval; 700 spots; 2 runs per target) (Applied Biosystems, Inc.). Afterwards, each spot was covered with matrix (2 mg/ml α-cyano-4-hydroxycinnamic acid in 70% ACN; internal calibrant: 93 pmol/ml human [Glu1]-fibrinopeptide B) using an external syringe pump with a 4 second interval (800 nl matrix/spot) at a flow rate of 12 μL/min.

**MALDI-TOF/TOF MS/MS analysis**

Spotted fractions were offline analysed using a MALDI ABI4800 proteomics analyser (Applied Biosystems). MALDI-TOF MS-analysis (reflectron mode; laser intensity: 3400; 25 x 20 laser shots per spot; mass-range 800–3000 Da) was performed first, after which precursors were selected with a signal-to-noise (S/N) ratio above or equal to 100. [Glu1]-fibrinopeptide B (m/z 1570.667) was used as internal standard to calibrate MS-spectra. MALDI-TOF/TOF MS/MS-analysis was performed on the selected MS precursors. A maximum of 50 unique precursors per spot were selected for fragmentation, starting from the precursors with the lowest S/N-ratio. These precursors were ionized (laser intensity: 4300; 25 x 20 laser shots per spot) and fragmented in a collision cell (CID, 1 kV collision energy).

**MALDI-TOF/TOF MS/MS spectral data analysis**

Spectra from each sample were extracted by Peak Explorer and screened against a *T. pallidum* database UniProt proteomes IDs UP000014259 and UP000000811 using the MASCOT search engine (Matrix Science; version 2.1.03) based on the digestion enzyme trypsin. We chose to screen against the Nichols strain database since the DAL-1 strain proteome is not well annotated and the genetic differences between the strains are minimal and described [50]. The latter database is generally used as the treponemal reference database while the former is a more recent version. Carbamidomethylation of cysteines was listed as fixed modification, while oxidation of methionine was set as a variable modification. A maximum of two missed cleavages of trypsin was tolerated. Mass tolerance was set to 200 ppm for the precursors and 0.20 Da for the fragment ions. The MudPIT scoring algorithm of MASCOT was used. Scaffold Q+ (version Scaffold 4.0.5, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Because the *T. pallidum* proteome contains several small proteins with just one or a few detectable tryptic peptides, protein identifications based on one unique peptide were only allowed if they fulfilled certain stringent conditions; these criteria were comprised by the peptide prophet algorithm that was performed by Scaffold Q+. Protein identifications were accepted if they could be established at greater than 95.0% probability according to the protein prophet algorithm.

Protein abundances were estimated based on the spectral counts (SC) of each identified protein by calculating the normalized spectral abundance factor (NSAF) as previously described [51]. In short, this approach includes a normalisation step based on (1) the observable peptides (OP) and (2) on the total number of identified peptides. The NSAF values reflecting an average of the biological and technical runs of each detected proteins are provided in S3 Table. Pearson’s correlation test and Mann Whitney test were calculated to compare the cDNA/DNA signal data to the NSAF protein abundance data. A *P*-value of < 0.05 was considered statistically significant. All analyses were performed in Stata 12 (StataCorp LP, College Station, TX, USA).

In order to determine whether the identification methodology was stringent enough, the false discovery rate (FDR) was defined on protein level by using a concatenated database.
consisting of the target spectral database and a shuffled database. Calculation of FDR was performed as follows: 2x false positive identifications / (false positive identifications + true positive identifications) [52]. For all samples, the FDR on protein level had to be less than 5%. Spectra were also screened against the mammalian Swissprot database containing human (Homo sapiens) and rabbit (Oryctolagus cuniculus) proteomes for spectra verification to prevent assignment of peptides with a conserved amino acid sequence.

**Orbitrap Velos LTQ MS/MS analysis**

**Nano reverse phase liquid chromatography and mass spectrometry.** The peptide mixtures were separated in the second dimension by reverse phase chromatography on an Eksigent nano-UPLC system using an Acclaim C18 PepMap100 nano-Trap column (200 μm x 20 mm, 5 μm particle size) connected to an Acclaim C18 analytical column (75 μm x 150 mm, 3 μm particle size) (Thermo Scientific, San Jose, CA). Peptide fractions were dissolved in mobile phase A, containing 2% ACN and 0.1% formic acid and spiked with 20 fmol [Glu1]-fibrinopeptide B. A linear gradient of mobile phase B (0,1% FA in 98% ACN) in mobile phase A (0,1% FA in 2% ACN) from 2 to 45% B in 35 min followed by a steep increase to 95% mobile phase B in 2 min was used at a flow rate of 350 nl/min. The nano-LC was coupled online with the mass spectrometer using a PicoTip Emitter (New objective, Woburn, MA) coupled to a nanospray ion source.

The LTQ Orbitrap Velos (Thermo Scientific, San Jose, CA) was set up in a data dependent MS/MS mode where a full scan spectrum (350–5000 m/z, resolution 60000) was followed by a maximum of ten CID tandem mass spectra (100 to 2000 m/z). Peptide ions were selected as the twenty most intense peaks of the MS scan. Collision induced dissociation (CID) scans were acquired in the LTQ ion trap part of the mass spectrometer. The normalized collision energy used was 35% in CID. A dynamic exclusion list of 45 sec for data dependent acquisition was applied.

**Orbitrap Velos LTQ MS/MS spectral data analysis**

Spectra from each sample were extracted by Proteome discoverer software (Thermo Scientific, San Jose, CA) and screened against a *T. pallidum* database (UniProt ID proteomes IDs UP000014259 and UP000000811) using the MASCOT search engine (Matrix Science; version 2.1.03) based on the digestion enzyme trypsin. Carbamidomethylation of cysteines was listed as fixed modification, while methionine oxidation was set as variable modification. A maximum of two missed cleavages of trypsin was tolerated. Mass tolerance was set to 10 ppm for the precursors and 0.8 Da for the fragment ions. The MudPIT scoring algorithm of MASCOT was used. Further protein identification, quantification and validation procedures were conducted as mentioned above for the MALDI-TOF/TOF analysis. All Orbitrap LTQ mass spectrometric data are available at PeptideAtlas [53]. The identifier is PASS00903.

**Identification of known and predicted *T. pallidum* membrane proteins**

Initially, all potential membrane proteins were identified from the 557 *T. pallidum* proteins detected by mass spectrometry by: (1) analyzing annotated functions (and subcellular localizations, if available) from all published *T. pallidum* ssp. *pallidum* genome sequences (http://www.ncbi.nlm.nih.gov/genome/?term=treponema+pallidum%5Borgn%5D), (2) identification of lipoproteins based on previous bioinformatic analyses performed by Setubal et al. [26], (3) identification of rare outer membrane proteins based on previous bioinformatic analyses performed by Cox et al. [27], and (4) by additional review of experimental findings in the scientific literature. Next, all potential membrane proteins (and proteins annotated as ‘uncharacterized’,
'hypothetical' or 'conserved hypothetical') were analyzed using 5 bioinformatic prediction tools. The SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) [54] and the LipoP 1.0 server (http://www.cbs.dtu.dk/services/LipoP/) [55] were used to predict the presence and location of potential signal peptide cleavage sites and lipoprotein signal peptides, respectively. PSORTb version 3.0.2 (http://www.psort.org/psortb/) [56] was used to predict protein subcellular localization. TMHMM server version 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/) [57] and PRED-TMBB (http://bioinformatics.biol.uoa.gr/PRED-TMBB/) [58] were used for predicting the presence and location of transmembrane alpha-helices and beta strands, respectively. Proteins with unclear subcellular localization predictions using the above bioinformatic pipeline were further analyzed using the following eight prediction tools. CELLO version 2.5 (http://cello.life.nctu.edu.tw/) [59] was used to predict subcellular localization. Philius (http://www.yeastrc.org/philius/pages/philius/runPhilius.jsp) [60], Phobius (http://phobius.sbc.su.se/) [61], Octopus/Spectopus (http://octopus.cbr.su.se/index.php) [62], HMMPRED version 2.0 (http://www.enzim.hu/hmmpred/html/submit.html) [63], and TMpred (http://www.ch.embnet.org/software/TMPRED_form.html) were used for the prediction of transmembrane alpha-helices BOMP (http://services.cbu.uib.no/tools/bomp) [64] and TBETADISC-RBF (http://rbf.bioinfo.tw/~sachen/OMPpredict/TBETADISC-RBF.php) [65] were used for the prediction of beta-barrel outer membrane proteins.

Assignment of orthologous functional categories and cellular localization

The eggNOG version 4 database (retrieved 21/04/15) was used to assign COG and NOG categories to all genomes. First all proteins per sample were compared to the eggNOG database using USEARCH version 7.0.959 with an e-value of 1e-30 and a bit-score cut-off of 70% of the top hit to ensure only close matches were retrieved and reduce the likelihood of spurious annotations. An eggNOG membership is assigned to each protein if 70% of the UBLAST hits belong to the same eggNOG member. Distinctions are then made between proteins with no UBLAST hit to any eggNOG sequence (no hit) and over 70% of hits to a member that is not assigned an eggNOG code (none). Annotations are also clustered at the 25 higher COG functional category levels as per the eggNOG assignments. Classification of proteins according to their cellular location was achieved using data extraction from online databases (Swissprot) and the methods as outlined for the membrane localized proteins.

Results and Discussion

Mass spectrometry analysis

In short, from the three biological replicates, a total of 6033 T. pallidum peptides were detected corresponding to 557 proteins and 54% of the total predicted proteome (S1 Table). Proteins ranged in size from 6–173 kDa with a pI range of 4.15 to 12.05. Acquired spectra were screened against two Nichols strain UniProt proteomes whereby three extra proteins (TP0248, TP0651 and TP0922) were uncovered compared to when solely screened against the Nichols reference UniProt proteome (ID: UP000000811) [13]. In the resequenced proteome (ID: UP000014259) [14] three of these proteins were below the 150bp annotation limit. We found 57/102 proteins containing previously reported sequencing errors [14] compared to the original genome analysis [13], including two genes with an authentic frameshift, 14 reannotated gene fusions and 5 novel ORFs reannotated in the new proteome (S2 table).

Pertaining to the individual samples, 394/398 (TPA-A), 279/321 (TPB-B) and 217/247 (TPC-C) proteins were uniquely identified by MALDI-TOF/TOF and ESI- MS/MS analysis, respectively, of which 106 (MALDI-TOF/TOF) and 119 (ESI- MS/MS) proteins were present in all three biological samples (Fig 1A/1B). Only 31 proteins were found with less than 2
peptide identifications in one biological and one technical run (S3 Table). For the individual MS analyses (MALDI-TOF/TOF versus ESI-LTQ Orbitrap MS/MS detection), 514 proteins were detected by both methods (Fig 2C). Only one and 42 additional proteins were exclusively identified by MALDI-TOF/TOF MS/MS analysis and ESI-MS/MS analysis, respectively (Fig 2C) indicating that we are possibly approaching the upper limit of the detectable *T. pallidum* proteome and that the non-detected proteins are 1) not expressed, or 2) are expressed at a very low level. All *T. pallidum* designated spectra were rescreened against human and rabbit UniProt protein databases and no overlap was found.

A previous proteomics study of *in vivo* rabbit expressed *T. pallidum* Nichols strain bacteria [16] detected 88 proteins using MALDI-TOF MS with peptide mass fingerprinting. We detected 58 of these proteins, therefore, to date 58% of the whole *T. pallidum* predicted proteome has been detected using MS methods. We failed to detect 30 of these previously identified proteins as outlined in Table 1. The protein detection differences between the studies could be attributed to different experimental methods, for example gel-based versus liquid chromatographic separation, which may favor the detection of proteins with certain physiochemical characteristics.

---

**Fig 1.** Venn Diagrams depicting the total number of unique *T. pallidum* proteins identified per rabbit biological replicate (N = 3) analyzed by (A) Matrix-assisted laser desorption/ionization time of flight and (B) Electrospray Ionization LTQ-Orbitrap Velos MS/MS. All spectra were screened against the UniProt databases (ID: UP000000811 & UP000014259), with a peptide and protein identification confidence interval of 95%. There was considerable overlap between the complementary MS analytical methods whereby an additional 42 treponemal proteins were found in the Orbitrap analysis as depicted in diagram (C).

doi:10.1371/journal.pntd.0004988.g001

Characterizing the Syphilis-Causing *Treponema pallidum* ssp. pallidum Proteome
Although the differences on the genomic level between the two strains are minimal [50], different duplication rates or other strain characteristics could contribute to different protein expression profiles found between these studies.

Detection of possible *T. pallidum* heterogeneous sites at the protein level

All *T. pallidum* protein sequences were screened for possible heterogeneous sites by searching the spectral databases for amino acid sequences containing sites designated with ‘X’, meaning ‘undetermined amino acid site’. Heterogeneous sites were defined as differing amino acids located at the same coordinate ‘X’ in the same protein sequence. A total of 25 *T. pallidum* proteins contained sites designated as ‘X’, of which four proteins were identified with heterogeneous peptide matches at site ‘X’ (Table 2). Amino acid sequence diversity was found within one sample for three proteins, TP0082 (TPC-C), TP0248 (TPC-C) and TP0922 (TPB-B). Protein TP0692 contained two peptides with heterogeneous sites within two samples (TPA-A/TPC-C). This is the first account of sequence heterogeneity at the protein level for these particular proteins. Although the amino acid sequence designation is of high confidence (95%),
cautious interpretation of these results is warranted as de novo peptide sequencing was not utilized so these analyses could represent falsely identified sites, therefore, further research is advised.

Treponema pallidum intra-strain nucleotide sequence heterogeneity has been reported previously [14,66,67], including tprK [22,31,32,66,68,69] and heterogeneity in four DAL-1 strain genes related to chemotaxis and metabolism [66]. The functional relevance of this observed intra-strain variability in these proteins is currently unknown.

Bioinformatic characterization of detected T. pallidum proteins

Bioinformatic analyses assigned 279 detected proteins to 19 higher Clusters of Orthologous Genes (COG) functional category levels according to their eggNOG assignments. Distributional

| TP number | UniProt Accession Number | Protein Name(s) | Molecular Weight (kDa) | COG category code | cDNA/DNA signal ratio Smajs et al. [15] |
|-----------|--------------------------|-----------------|------------------------|------------------|----------------------------------------|
| TP0112    | O83150                   | Aminopeptidase C (PepC)/Bleomycin hydrolase | 51                      | none            | 1.0                                    |
| TP0139    | O83175                   | Trk family potassium (K+) transporter, NAD+ binding protein | 25                      | P               | 0.6                                    |
| TP0201    | R9UUV9                   | 50S ribosomal protein L5 | 21                      | none            | 1.8                                    |
| TP0216    | R9UVAn                   | Chaperone protein DnaK/Heat shock 70 kDa protein | 68                      | none            | 4.5                                    |
| TP0239    | R9UUu25                  | 50S ribosomal protein L10 | 20                      | none            | 1.9                                    |
| TP0259    | O83283                   | Peptidoglycan binding (LysM domain-bacterial cell wall degradation) | 29                      | S               | 5.4                                    |
| TP0298    | R9UW99                   | ABC superfamily ATP binding cassette transporter, membrane protein/Membrane lipoprotein TpN38 | 38                      | none            | 3.3                                    |
| TP0349    | R9UUA6                   | Peptidyl-prolyl cis-trans isomerase/Metallochaperone SlyD | 18                      | O               | 2.2                                    |
| TP0356    | O83375                   | Putative RNA-binding protein | 12                      | none            | 10.7                                   |
| TP0424    | R9UW5                    | Two-sector ATPase, V(1) subunit E | 19                      | none            | 4.7                                    |
| TP0435    | R9UWP1                   | Copper resistance lipoprotein NlpE/17 kDa lipoprotein | 17                      | none            | 5.1                                    |
| TP0448    | R9UUM3                   | Uracil phosphoribosyltransferase | 41                      | F               | 0.3                                    |
| TP0453    | O67998                   | Outer membrane protein/30kLp | 32                      | none            | 1.0                                    |
| TP0476    | R9UWS9                   | Acetate kinase/Acetokinase | 49                      | none            | 1.3                                    |
| TP0505    | O83518                   | Hexokinase | 48                      | none            | 1.8                                    |
| TP0537    | R9UTC2                   | Triosephosphate isomerase | 27                      | G               | 1.9                                    |
| TP0554    | O83565                   | Phosphoglycolate phosphatase | 25                      | R               | 1.1                                    |
| TP0556    | R9UWZ4                   | Aspartate—ammonia ligase/Asparagine synthetase A | 37                      | E               | 1.8                                    |
| TP0584    | R9UTG3                   | Uncharacterized protein | 54                      | none            | 0.8                                    |
| TP0611    | O83620                   | ABC superfamily ATP binding cassette transporter, ABC protein | 28                      | none            | 1.2                                    |
| TP0655    | O83661                   | Spermidine/putrescine ABC superfamily ATP binding cassette transporter, binding protein | 40                      | E               | 0.6                                    |
| TP0734    | R9UXE0                   | Purine-nucleoside phosphorylase | 25                      | F               | 1.6                                    |
| TP0769    | R9UTV1                   | Treponemal membrane protein B/Antigen TmpB | 37                      | S               | 5.2                                    |
| TP0821    | R9UV14                   | Lipoprotein/Membrane lipoprotein TpN32; 29 kDa protein | 29                      | P               | 1.5                                    |
| TP0823    | O83795                   | Desulfoferrodoxin (Rbo) | 14                      | C               | 3.2                                    |
| TP0862    | O83834                   | Peptidyl-prolyl cis-trans isomerase | 28                      | none            | 5.3                                    |
| TP0925    | R9UUA1                   | Nitrogenase (Flavodoxin) | 16                      | C               | 9.4                                    |
| TP0964    | O83930                   | ABC transporter, ATP-binding protein | 25                      | V               | 0.9                                    |
| TP0971    | R9UXB8                   | Tp34 lipoprotein/34kDa membrane antigen | 22                      | none            | 12.6                                   |
| TP1013    | R9UUH7                   | 10 kDa chaperonin/GroES protein/Protein Cpn10 | 9                       | O               | 4.6                                    |

Legend:

#: UniProt Proteome ID: UP000014259; Clusters of Orthologous Genes (COG)

doi:10.1371/journal.pntd.0004988.t001
description of these proteins and their categorical frequencies are depicted in Fig 2 and extensive
descriptions, including COG/NOG codes for all detected proteins, can be found in S2 Table. Of
the proteins that were delegated into a clear functional category, the highest representative cate-
gories were 'J' (translation, ribosomal structure and biogenesis) (17%) and 'L' (replication,
recombination and repair) (12%). High category coverage was found for the categories 'M'
(cell wall/membrane/envelope biogenesis) and 'O' (posttranslational modification, protein turnover
and chaperones) with 25/28 and 17/21 proteins found, respectively. Forty-five proteins fell
under category 'S' or 'R', indicating poor functional characterization. A total of 9 proteins had
no UBLAST hit to any eggNOG sequence (category 'no_hit'), of which 5 proteins were ribo-
somal and 4 were uncharacterized. Many proteins (N = 275) were at least 70% homologous to a
protein member not assigned an eggNOG code (category 'none') indicating that the T. pallidum
proteome is very unique compared to other organisms. Six proteins were categorized under
multiple COG categories. In almost all of the COG categories, more than half of the predicted
proteins were detected, supporting the theory that T. pallidum has shed its unnecessary genes
during its evolution [13].

The T. pallidum Nichols and SS14 strain genomes differ minimally [14], thus in the case of
genetically congruent ORFs we extrapolated recent T. pallidum strain SS14 hypothetical pro-
tein function re-annotations [28] to 22 'uncharacterized/hypothetical' proteins detected in this
analysis. In total, 114 proteins remained classified as 'uncharacterized proteins/hypothetical
proteins'. This category did not include 17 proteins with ambiguous "putative" membrane protein
descriptions. A previous study [16] detected eight of these uncharacterized proteins, mean-
ing that this is the first account of 106 'uncharacterized/hypothetical' proteins at the protein
level. This uncharacterized area of the T. pallidum proteome may contain novel proteins with
important roles in pathogenesis and even represent novel biomarker, treatment or vaccination
targets.

**Predicted cellular localization of detected T. pallidum proteins**

The global classification of detected proteins according to their cellular localization was
achieved by screening online databases such as UniProt and by reviewing relevant literature.

| UniProt Accession Number | TP number | Protein Name | Molecular Weight (kDa) | Identified Peptide Sequence | MALDI TOF/TOF | ESI-LTQ-Orbitrap |
|--------------------------|-----------|--------------|------------------------|-----------------------------|--------------|-----------------|
| O88098                   | TP0082    | Formate hydrogenlyase transcriptional activator (FhlA) | 66                      | (R)LYPIXNAR(K)              | T            | D               |
| O83276                   | TP0248    | Uncharacterized lipoprotein | 15 | (R)AYELXER(S) | T | H |
| O83690                   | TP0692    | Protein RecA   | 44 | (K)TLRRXASR(G) | Y | H | T | P |
| O83892                   | TP0922    | Uncharacterized protein | 33 | (R)ALXGNDPSAAVR(V) | V | A |

**Legend:**

#: UniProt Proteome ID: UP000014259; Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF/TOF); electrospray ionization LTQ-Orbitrap tandem mass spectrometry (ESI-LTQ-Orbitrap), X = undetermined amino acid site

doi:10.1371/journal.pntd.0004988.t002
The cellular localization of the proteins was predicted for 292/557 proteins; these were largely localized in the cytoplasm (N = 97, 17%), membrane (N = 116, 21%), ribosome (N = 33, 6%) and flagella (N = 19, 3%) were the most represented cellular localizations.

A schematic breakdown of the predicted cellular localizations for all detected proteins can be found in Fig 3, with comprehensive information for each protein provided in S2 Table.

Detected proteins were subjected to additional bioinformatic pipeline analyses in order to identify potential membrane proteins as detailed in the methods section. In short, all potential membrane proteins (N = 131; including proteins annotated as ‘hypothetical’) were analyzed using five bioinformatic prediction tools: SignalP 4.1 [54], LipoP [55], PSORTb [56], TMHMM [57] and PRED-TMBB [58]. Proteins with unclear subcellular localization predictions using the above bioinformatics pipeline (N = 25) were further analyzed using an additional eight prediction tools including CELLO [59], Philius [60], Phobius [61], Octopus/Spoctopus [62], HMMTOP [63], Tmpred, BOMP [64] and TMBETADISC-RBF [65]. A results summary of this analysis can be found in Fig 4 and S4 Table.

Fig 3. Pie diagram depicting the subcellular localization of *T. pallidum* detected by MS analysis (N = 557). Almost half of the detected proteins (N = 265; 48%) did not have an annotated cellular location. Of the known locations, membrane (N = 116; 21%), cytoplasm (n = 99; 18%), ribosomal (N = 33; 6%) and flagella (N = 19; 3%) were the most represented cellular localizations.
In total, 116 proteins were designated as ‘membrane’ localized, with a majority (64%; N = 74) located within the inner membrane. Sixteen proteins (14%) were predicted to be located in the outer membrane (OM) (Table 3). The OM localization of five detected proteins has been experimentally investigated: (TP0117 (TprC)/TP0131 (TprD) [27,37,70], TprI (TP0620) [38], TP0126, an OmpW homologue [71] and TP0326, a BamA homologue [27,72–74].

The other 11 predicted OM proteins in this analysis were: TprB (TP0011), M23B subfamily peptidase (TP0155), TprE (TP0313), TprJ (TP0621), TP0421, TP0858, TP0324, TP0855, TP0865, TP0923 and TP0969.

A previous *in silico* prediction analysis of the *T. pallidum* genome revealed 46 predicted lipoproteins [26]. Our analysis also detected 25 lipoproteins, 23 with unknown membrane locations and two located within the periplasm (TP0796; TP0171), including the 15 kDa (Tpp15) lipoprotein (TP0171) and 47 kDa membrane antigen (TP0574). In spirochetes, lipoproteins are highly expressed molecules primarily localized in the periplasm anchored to the outer leaflet of the cytoplasmic membrane [9] where they are thought to modulate immune responses from both innate and adaptive immunity [43,77].

There were ambiguities regarding the subcellular localization of 13 proteins after analysis with the additional prediction tools (S4 Table), including TprG (TP0317), TprH (TP0610), ABC superfamily ATP binding cassette transporter (TP0786), flagellar hook length control protein Flk (TP0729) and two TolC-like proteins (TP0967 and TP0968). Of the 49/116 reported membrane proteins that could be designated to a COG category, two categories were most represented: ‘P’ (inorganic ion transport and metabolism) (N = 9) and ‘M’ (cell wall/membrane/envelope biogenesis) (N = 6). This agrees with the predicted biological functional location.

Important to note is the fact that most of the protein localization data is based on computational predictions. These types of predictions have an inherent risk of including false positives.
and also omitting real OM proteins. Further laboratory work is needed to experimentally confirm the cellular locations of these proteins.

### Relative *T. pallidum* protein abundance as determined by spectral counting

We examined the relative abundance of the proteins detected by calculating the NSAF values [78] for the proteins detected in the biological and technical runs; all values are listed in S3 Table and the log distribution of all detected proteins can be found in Fig 5A. This approach is based on the number of observable peptides and normalizes technical variability between samples [78]. A value of ‘1’ represents the mean protein level for all detected proteins. Proteins with an average NSAF value greater than 5.0 were regarded as ‘highly abundant’. A summary of the top 50 highest abundant proteins according to the spectral counting averages is provided in Table 4. High abundant proteins (N = 103) included two proteins related to redox balance, 22 proteins related to translation, two proteins related to chemotaxis and three ABC family transport proteins. Proteins related to motility were found to be high abundant, including flagellar filament proteins (TP0663; TP0792; TP0868; TP0870) and 3 proteins related to flagellar biosynthesis (TP0403; TP0658; TP0718). The fact that proteins related to motility, transport and chemotaxis are highly expressed can be indicative that these processes are essential and highly utilized for cell survival.

In terms of the cellular localizations of high abundant proteins, 18 were membrane localized. Four of these proteins were predicted lipoproteins (TP0248, TP0768, TP0895, TP0789) and two were predicted OM uncharacterized proteins (TP0858, TP0126). Surprisingly,
approximately a third of the high-abundant proteins (N = 37) were classified as uncharacterized/hypothetical and seven proteins did not have any significant (70%) match with any other EggNOG sequences indicating these are highly specializedTreponema pallidum proteins that may play an important role in unique survival and virulence tactics. The most highly represented COG category of the highly expressed proteins was category ‘J’ (translation, ribosomal structure & biogenesis). A low correlation was found between previous gel-based studies [16,17] that determined protein abundance based on silver staining and protein abundance as determined in this study. For example, some highly abundant gel-detected proteins were not detected in our analysis, such as the uncharacterized protein TP0259 and the Tp34 lipoprotein (TP0971) [16,17]. We found a low correlation between the average transcriptional rate (cDNA/DNA signal) from a previous transcriptome study [15] and the average NSAF value for each detected protein found in this study (Pearson’s correlation coefficient, r = 0.26; P = 0.000). The distribution of these data is depicted in Fig 5B. In general, flagellar proteins and proteins related to flagellar biosynthesis such as flagellar filament outer layer protein (TP0249), putative flagellar filament outer layer protein FlaA (TP0663), and flagellar biosynthetic protein Flp (TP0718) were highly expressed in both studies. There were some notable discordances between the data, such as the high gene expression level measured for lipoprotein antigen Tp47 (TP0574), galactose ABC superfamily ATP binding cassette transporter Tpp38 (TP0684) and the 60kDa chaperonin (TP0030), all of which were found in low abundance at the protein level in this study. Moreover, 27 proteins with high gene expression (cDNA/DNA signal ratios greater than 4.0) were not found in this analysis (Table 1). We theorized that the proteins we failed to detect in our analysis would have a lower mean transcription rate. There was however, no significant cDNA/DNA signal data difference between the detected and undetected proteins as determined by a Two-sample Wilcoxon rank-sum (Mann-Whitney) test (P = 0.5). Other studies have demonstrated low correlations between transcriptome and protein abundance data, as reviewed by Maier et al. [79]. Intermediary factors such as translation efficiency and protein half-life play a prominent role in accentuating the lack of a linear association between gene expression and protein abundance.
Table 4. Top 50 most abundant T. pallidum proteins detected by MS analysis.

| TP Number | UniProt Accession Number | Protein Name | Molecular Weight (kDa) | Average NSAF Value | cDNA/DNA Signal Ratio Smajs et al. [15] | GO Biological Process | Cellular Location | COG Category |
|-----------|--------------------------|--------------|------------------------|--------------------|-----------------------------------------|----------------------|------------------|--------------|
| TP0214    | R9UVZ6                   | Uncharacterized protein | 7                      | 19.6               | 1.6                                     | Unknown             | no_hit           |              |
| TP0708    | R9UXB8                   | Uncharacterized protein | 18                     | 15.7               | 1.5                                     | Inner membrane      | none             |              |
| TP0385    | O83400                   | Uncharacterized protein | 14                     | 13.6               | 1.2                                     | Unknown             | none             |              |
| TP0951    | R9UVX4                   | 50S ribosomal protein L34 | 6                      | 13.1               | 1.8                                     | Translation         | Ribosome         | no_hit       |
| TP0126a*  | R9UTR8                   | Uncharacterized protein | 23                     | 13.1               | 1.1                                     | Unknown             | none             |              |
| TP0474    | R9UUP5                   | Probable transcriptional regulatory protein | 26                     | 13.1               | 3.3                                     | Regulation of transcription, DNA-templated | Cytoplasm         | K             |
| TP0061    | R9UTJ4                   | 30S ribosomal protein S18 | 12                     | 12.5               | 1.2                                     | Translation         | Ribosome         | none         |
| TP0048    | O83087                   | Cell shape determination protein CcmA | 15                     | 12.1               | 1.1                                     | Unknown             | none             |              |
| TP0657    | R9UWD4                   | Carbon storage regulator homolog | 8                      | 12.0               | 0.9                                     | mRNA catabolic process; regulation of carbohydrate metabolic process | Unknown             | none         |
| TP0409a   | R9UT20                   | Preprotein translocase subunit YajC | 14                     | 11.4               | 1.6                                     | Inner membrane      | none             |              |
| TP0718    | R9UXC7                   | IIIISP family Type III (Virulence-related) secretory pathway protein/Flagellar biosynthesis protein FliP | 31                     | 11.2               | 0.8                                     | Bacterial-type flagellum organization; protein secretion | Flagellum/Inner membrane | N             |
| TP0808    | R9UWS1                   | Acyl carrier protein | 9                      | 10.1               | 2.8                                     | Fatty acid biosynthesis; Fatty acid metabolism | Cytoplasm           | none         |
| TP0558    | O83569                   | Nickel/cobalt efflux system | 32                     | 9.8                | 0.9                                     | Inner membrane      | none             |              |
| TP0198    | R9USD7                   | 30S ribosomal protein S17 | 10                     | 9.8                | 1.5                                     | Translation         | Ribosome         | none         |
| TP0849    | R9UJ63                   | 50S ribosomal protein L35 | 8                      | 9.7                | 7.1                                     | Translation         | Ribosome         | no_hit       |
| TP0490    | O83503                   | Uncharacterized protein | 7                      | 9.6                | 0.7                                     | Unknown             | no_hit           |              |
| TP0649    | O83655                   | Hemolysin (TlyC) | 30                     | 9.5                | 0.6                                     | Inner membrane      | P               |
| TP0213    | R9USF3                   | 50S ribosomal protein L17 | 19                     | 9.4                | 2.2                                     | Translation         | Ribosome         | J             |
| TP0072    | O83111                   | Glutaredoxin-related protein | 10                     | 9.4                | 7.6                                     | Cell redox homeostasis | Unknown           | none         |
| TP0509    | O83522                   | Alkyl hydroperoxide reductase (AhpC)/Peroxiredoxin | 21                     | 9.3                | 15.5                                    | Unknown             | O               |
| TP0919    | R9UVS2                   | Thioredoxin | 11                     | 9.2                | 8.1                                     | Cell redox homeostasis; glycerol ether metabolic process | Unknown             | none         |
| TP0362    | R9USW9                   | 50S ribosomal protein L28 | 8                      | 8.9                | 5.4                                     | Translation         | Ribosome         | none         |
| TP0368    | R9UVP4                   | Uncharacterized protein | 12                     | 8.7                | 0.8                                     | Inner membrane      | none             |              |
| TP0119    | O83156                   | Amino acid ABC transporter, permease protein (YaeE) | 23                     | 8.7                | 0.7                                     | Transport           | Inner membrane   | none         |
| TP0195    | R9USH3                   | 30S ribosomal protein S3 | 28                     | 8.5                | 1.9                                     | Translation         | Ribosome         | J             |
| TP0604    | R9UJ26                   | Ribosome-recycling factor | 21                     | 8.4                | 1.6                                     | Protein biosynthesis; translation termination | Cytoplasm           | J             |
| TP0045    | R9UUU2                   | Adenosine deaminase | 35                     | 8.3                | 0.4                                     | Hydrolase activity; metal ion binding; deaminase activity; nucleotide metabolic process | Unknown             | F             |

(Continued)
### Table 4. (Continued)

| TP number | UniProt Accession Number | Protein Name | Molecular Weight (kDa) | Average NSAF value | cDNA/DNA signal ratio Smajs et al. [15] | GO Biological Process | Cellular location | COG category |
|-----------|--------------------------|--------------|------------------------|--------------------|------------------------------------------|-----------------------|------------------|--------------|
| TP0700    | R9UV74                   | Uncharacterized protein | 14                     | 8.0                | 3.7                                      | Unknown              | none            | none         |
| TP0315    | R9USP4                   | Uncharacterized protein | 23                     | 8.0                | 0.9                                      | Unknown              | none            | none         |
| TP0922    | O83892*                  | Uncharacterized protein | 33                     | 7.8                | 1.6                                      | Unknown              | none            | none         |
| TP0887    | R9UV5P                   | 30S ribosomal protein S15 | 10                     | 7.8                | 4.3                                      | Translation          | Ribosome        | none         |
| TP0707    | R9UTN7                   | Putative membrane protein | 17                     | 7.6                | 0.6                                      | Inner membrane       | none            | none         |
| TP0202    | R9UTZ5                   | 30S ribosomal protein S14 type Z | 7                    | 7.6                | 1.8                                      | Translation          | Ribosome        | no_hit       |
| TP0433    | R9UT05                   | Acidic repeat protein | 66                     | 7.6                | 1.4                                      | Unknown              | none            | none         |
| TP0663    | R9UV45                   | Putative flagellar filament outer layer protein FlaA; Tromp-2 | 27                    | 7.5                | 13.9                                     | Bacterial-type flagellum-dependent cell motility | Flagellum        | none         |
| TP0076    | R9UXU9                   | Sugar ABC superfamily ATP binding cassette transporter, membrane protein | 30                    | 7.4                | 0.4                                      | Transport            | Inner membrane  | G            |
| TP1029    | R9UV97                   | Uncharacterized protein | 25                     | 7.4                | 0.9                                      | Unknown              | none            | none         |
| TP0249    | R9UVD9                   | Flagellar filament outer layer protein (FlaA1) | 39                    | 7.4                | 16.1                                     | Flagellum            | N               | none         |
| TP1012    | O83976                   | RNA polymerase sigma factor | 35                     | 7.4                | 0.5                                      | DNA-templated transcription; initiation | Polymerase        | K            |
| TP0206a   | R9UU00                   | 50S ribosomal protein L30 | 7                     | 7.4                | 2.0                                      | Translation          | Ribosome        | no_hit       |
| TP0194    | R9UVW4                   | 50S ribosomal protein L22 | 14                    | 7.3                | 1.8                                      | Translation          | Ribosome        | J            |
| TP0234    | R9UJU20                  | 50S ribosomal protein L33 | 7                     | 7.2                | 4.2                                      | Translation          | Ribosome        | no_hit       |
| TP0007    | R9UZ22                   | Uncharacterized protein | 35                     | 7.1                | 1.8                                      | Unknown              | none            | none         |
| TP0412    | O83427                   | Uncharacterized protein | 12                     | 7.1                | 1.2                                      | Unknown              | S               | none         |
| TP0792    | R9UU06                   | Flagellin/Flagellar filament 33 kDa core protein; Class B | 31                    | 7.0                | 13.8                                     | Bacterial-type flagellum-dependent cell motility | Flagellum        | N            |
| TP0654    | O83660                   | Spermidine/putrescine ABC superfamily ATP binding cassette transporter, membrane protein | 30                    | 6.9                | 1.0                                      | Transport            | Inner membrane  | none         |
| TP0563    | O83574                   | Uncharacterized protein | 16                     | 6.9                | 0.8                                      | Unknown              | O               | none         |
| TP0248    | O83276*                  | Uncharacterized lipoprotein TP_0248 | 15                    | 6.8                | 1.5                                      | Lipoprotein-unknown membrane localization | S               | none         |
| TP0403    | O83418                   | Flagellar biosynthesis protein FlJ | 18                    | 6.8                | 0.8                                      | Bacterial-type flagellum-dependent cell motility; chemotaxis | Flagellum        | none         |
| TP0842    | R9UU24                   | Methionine aminopeptidase | 27                    | 6.8                | 0.8                                      | Protein initiator methionine removal | Unknown         | J            |

Legend:

* UniProt Proteome ID: UP000014259; NA: not available; NSAF normalized spectral abundance factor; no_hit: proteins with no UBLAST hit to any eggNOG sequence; none: over 70% of hits to a member that is not assigned an eggNOG code.

*: the C-terminal 49 amino acids of TP0126a correspond to the N-terminal 49 amino acids of TP0126 due to a genome insertion [29].

doi:10.1371/journal.pntd.0004988.0004
**T. pallidum** proteins confirmed or predicted to be related to virulence

Thirty-nine proteins implicated in **T. pallidum** virulence [20] were detected, including eight members of the *tpr* gene family and a protein related to a beta-barrel assembly machinery (BAM) complex. Brief descriptions of these proteins are detailed below.

**Tpr proteins**

Regarding the *tpr* gene family implicated in host-immune evasion [30], 8/12 of these proteins were detected in this analysis, including proteins TprB (TP0011), TprC/D (TP0117/TP0131), TprE (TP0313), TprG (TP0317), TprH (TP0610), TprI (TP0620) and TprJ (TP0621). Proteins TprA (TP0009), TprK (TP0897) and TprL (TP1031) were not detected. There was no unique TprF peptide sequence found in this analysis, although three peptides were uncovered that are homologous for TprC/D, F and I (Table 5). The ORF origin of these peptides cannot be determined. The *tprC* and *tprD* loci contain two identical coding sequences in the reference Nichols and DAL-1 strain genome [13,70], therefore we included the detection of both TprC and TprD since no distinction could be made between the coding ORF origin of these proteins. Even though *tprK* was previously shown to exhibit the highest level of transcription among *tpr* family genes [80], the fact that *tprK* displays high sequence variability [36] makes the likelihood of detecting this protein minimal due to rigid MS analytical criteria.

| Protein | Peptide Sequence Identified |
|---------|----------------------------|
| Tpr protein B | NELAAQMR |
| Tpr protein B | VKGKGTSNR |
| Tpr protein C | APMNALIDALLRMQWK |
| Tpr protein C/F/I | GEARSGVWAQLQLK |
| Tpr protein C/F/I | SGVWAQLQLK |
| Tpr protein C/F/I | TALLWGVGGR |
| Tpr protein D | TNGTQVNVIDTVPVNRQSPVR |
| Tpr protein E | VEQAVQENIR |
| Tpr protein E/G/J | AGISASLIEK |
| Tpr protein E/G/J | DKLLWNVGGR |
| Tpr protein E/G/J | IPVQDYGWKSVTWASTNRAHNAPAGGAVGATLYTK |
| Tpr protein E/G/J | TTNTYGVSPFLVMR |
| Tpr protein G | KKDIALDAGQQR |
| Tpr protein G | KTDIALDAGQQR |
| Tpr protein H | TDIALDAGQQR |
| Tpr protein H | AGDAYTHLIDGLEAGMDVR |
| Tpr protein H | RVRSGTWALLFMSSAAGLCAETR |
| Tpr protein H | LHTLASTPR |
| Tpr protein H | RTLLLPSAAVR |
| Tpr protein H | TKVTPGPGVAYAIAQR |
| Tpr protein I | FIQMALVK |
| Tpr protein I | VATDSEGDR |
| Tpr protein J | MVGEALIKQQLSR |
| Tpr protein J | NNANMQAVGGSLGDTARMVGEALIK |
| Tpr protein J | NNNGNPLPSGGSGGHIGLPVVGK |
| Tpr protein J | QDLADLVPMMR |

Table 5. Peptide Sequences of Proteins TprB, C/F/I, E/G/J and H, Detected by MS analysis.
BAM-complex

Outer membrane beta-barrel proteins (OMPs) are commonly involved in cellular process such as small molecule efflux (such as antibiotics) and nutrient acquisition [81,82] in bacteria. The beta-barrel assembly machinery (BAM) complex [83] is thought to facilitate OMP assembly, insertion and folding and in Gram-negative bacteria this complex is typically composed of five proteins: BamA, which is an integral membrane protein and four accessory lipoproteins, BamB-BamE [84]. The insertion and assembly of proteins into the outer membrane is controlled through interactions with periplasmic chaperones (SurA, Skp, and DegP) [85]. Studies [72,86] have demonstrated the presence of a BAM complex in *T. pallidum* which is similar to that of *Escherichia coli* [72]. We detected the BamA orthologue (TP0326) [72,87,86,74]. Peptides identified corresponded to the POTRA 2 & 3 domains and a transmembrane domain/extra-cellular Loop L3 [72,86] (Table 6).

Other detected proteins implicated in *T. pallidum* virulence

In our analyses we detected a selection of additional proteins that have been previously implicated in *T. pallidum* virulence and pathogenesis, as reported in Table 7.

Exploring the undetected *T. pallidum* proteins

Of the predicted protein coding ORFs, 482/968 proteins were not detected in this study. Most of the undetected proteins are classified as ‘uncharacterized proteins/hypothetical proteins’ (N = 197), ‘conserved hypothetical integral membrane proteins (N = 10), or ‘conserved hypothetical protein’ (N = 1). The most plausible explanations for not detecting half of the proteome are i) very low protein abundance could evade MS detection, ii) lack of protein expression during *in vivo* expression during some or all stages of infection, iii) small proteins are less viable to detection since they contain fewer peptides and/or these protein sequences lack arginine or lysine tryptic digestion sites or iv) the presence of (partial) sequence heterogeneity that would thwart peptide/database matching. Certain caveats of MS analyses will always preclude the detection of the whole proteome of organisms. A non-exhaustive list of other technical

| Identified peptide sequence | Sequence location | (Topological) Domain Location* |
|-----------------------------|------------------|-------------------------------|
| MKVDQESLR                  | 141–149          | POTRA 2                       |
| VDQESLR                    | 143–150          | POTRA 2                       |
| AFTESEVLK                  | 197–203          | POTRA 3                       |
| KVLSTQEAR                  | 205–213          | POTRA 3                       |
| VEGVAKTVDK                 | 246–255          | POTRA 3                       |
| AGSYGNGLPHPYTSR            | 514–529          | Trans-membrane beta-strand 5/ Extra-cellular Loop L3 |

*based on experimental predictions [72,86]

doi:10.1371/journal.pntd.0004988.t006

| ORF Number(s) | Reported/predicted functional role | Supporting reference |
|---------------|-----------------------------------|----------------------|
| TP0967, TP0968, TP0969 | ToIC-like proteins | [88,89] |
| TP0155        | Fibronectin binding protein       | [75]                |
| TP1038        | Bacterioferrin/ TpF1              | [90]                |
| TP0027, TP0649 | Putative hemolysins               | [91]                |

doi:10.1371/journal.pntd.0004988.t007
limitations include: i) hydrophobic peptides do not elute from LC columns during the applied gradient, ii) spectral masking of low abundant proteins by the presence of high abundant protein spectra, iii) co-elution and ion suppression that may prevent the ionization or detectability of some peptides by MS and iv) some peptides are unable to ionize sufficiently on the MS platform.

Variable *T. pallidum* genomic sequences as modulators of protein expression

To address the possibility that the presence of variable sequences may have affected proteome coverage, either by altered gene expression or by precluding MS detection, we searched for known and predicted heterogeneous sequences in the *T. pallidum* genome. Within this analysis we looked for sequences containing elements indicative of phase variation (homopolymeric tracts) or antigenic variation through gene conversion (tandem repeats, *tprK* donor sites and quadruplex forming G-rich sequences (G4FS)). Previous investigations have identified and characterized 19 genes with variable sequence elements, of which 9 proteins were detected in this analysis. Aside from the 12 aforementioned Tpr family proteins there are seven additional genes shown to contain variable sequence elements including: *tprK* donor sequences to promote gene conversion (*tp0130; tp0129; tp0128*), homopolymeric G-tracts (poly-G tracts) in promoter regions to alter transcription (*TP0126*), poly-G tracts in the ORF to induce phase variation (*TP0127*), or G4FS cis-acting DNA elements that form guanine quadruplexes to induce recombination and gene conversion (*TP0104; TP0136*) [32,36,67,70,92]. Notably, *TP0136*, a fibronectin binding protein implicated in *T. pallidum* virulence not detected in this analysis, harbors two G4FS sequences localized within tandem repeats in the ORF [93]. Surprisingly, the paralogues of *TP0136*: *TP0133, TP0134, TP0462 and TP0463* were also not detected. Among these seven additional variable sequences only an OmpW homologue (*TP0126*) was detected in our analyses.

We also searched for predicted variable sequences in the *T. pallidum* genome. A previous *T. pallidum* genomic study predicted the presence of G4FS which may be involved in generation of *tprK* variants in pathogenic treponemes [32]. Similar G4 DNA structures have been implicated in the host immune evasion tactics of *Neisseria gonorrhoeae* where they function as recombination activation elements to regulate gene conversion and the expression of cell surface pilin proteins (PilE) [94]. Giacani *et al.* (2012) identified 46 putative G4FS sequences located in 33 different genes and eight unique intergenic regions (IGRs) of *T. pallidum*; 21 of the 33 predicted G4FS-containing ORFs were detected in this analysis. Among the eight putative G4FS residing within unique IGRs, only two of the downstream proteins were detected in this study (*TP0104; TP0549*). Additionally, we searched for the presence of tandem repeats [95,96] in ORFs and IGRs of genes for which peptides had previously been detected using MS [16] or exhibited high transcript abundance [15]. The presence of highly mutable sites, or contingency loci, such as tandem repeats have been suggested to represent a mechanism for rapid environmental adaptation and virulence within a host [97]. This has been explored in a recent study involving serial *in vivo* passage of *Campylobacter jejuni* in mice that resulted in increased phases in the contingency loci and virulence [98]. This analysis identified three additional genes harboring tandem repeats (*tp0470; tp0424; tp0769*), providing a possible rationale for why these proteins remained undetected in this study. Our study detected 30 proteins out of a total of 60 proteins with known and predicted variable sequences. Remarkably, four proteins discovered in this analysis were only annotated in the original *T. pallidum* genome [13], mostly due to the fact that sequences below 150 base pairs were not annotated as protein coding in the resequenced genome [14]. Perhaps there is a need for deeper mining of the *T. pallidum* genome.
and re-evaluation of the definition of protein coding sequences, especially in light of the recent attention brought to classes of endogenous polypeptides called ‘SEPs’ (sORF-encoded polypeptides). These polypeptides are encoded by short open reading frames (small ORFs or smORFs) (generally <150 amino acids in length) in bacteria and eukaryotic organisms and are thought to play an important function in biological functions [99] such as cell survival under conditions of glucose toxicity as studied in \textit{E. coli} [100]. Interestingly, in \textit{M. pneumonia}, 53% of all smORFs are deemed essential to cell survival whilst another 11% affect the fitness of the organism [101], indicating that these may also play a large (unknown) role in \textit{T. pallidum} biological function.

In general, proteins in small genomes are more likely to function as proficient “multitaskers” and have been shown to interact with other proteins from a wider range of functions in comparison to their orthologues in larger genomes [102]. It is possible that many \textit{T. pallidum} proteins perform multiple biological functions, especially under different environmental conditions. A growing area in proteomics is the concept of ‘protein moonlighting’, defined as a single protein that displays multiple functions that are not related to multiple RNA splice variants, multiple proteolytic fragments or gene fusions [103]. Many bacterial species employ protein moonlighting and the role of this phenomenon in bacteria virulence has been excellently reviewed by Henderson \textit{et al.} [104,105]. Some bioinformatic approaches have been suggested to approach genome wide annotation of potential moonlighting proteins [106,107]; these may be useful for future \textit{T. pallidum} proteome studies.

One of the many intriguing aspects of \textit{T. pallidum} is the fact so many proteins lack homology with proteins from other bacteria. This is exemplified by the fact that only 59% (581/968) of the \textit{T. pallidum} protein coding genes were allocated into COG or NOG categories. With the demonstrated expression of 114 uncharacterized/ hypothetical \textit{T. pallidum} proteins in this study, some even at high abundance, further experimental analysis is needed to elucidate the functions of these proteins such as looking at protein binding partners. Periodic re-evaluations of ‘uncharacterized’ \textit{T. pallidum} proteins are warranted, especially with the rapid sophistication of bioinformatics tools and the growing repertoire of proteins with known predicted functions from other organisms.

We are confident in the quality and extent of the protein coverage of this analysis. For example, we performed analysis on three biological replicates, employed multidimensional peptide separation techniques together with complementary MS analyses in order to improve the dynamic range and coverage of the analyses. Nevertheless, there are a number of limitations related to this study.

\section*{Limitations}

Even though our experimental approaches aimed to closely mimic the physiological conditions of human infection, a distinct advantage over the artificial conditions of \textit{in vitro} studies, we cannot exclude the effects of inter-rabbit variability. Different rabbits may exert unique immune pressures, which in turn may influence gene expression. The fact that infected rabbits typically do not transition into the secondary stage of syphilis [108] and there is no tertiary stage in rabbits [109] may suggest that the infectious dynamics of rabbit syphilis may differ from that of humans. Moreover, there may be differential gene expression depending upon the tissue environment [15], therefore the analysis of intradermal rather than intratesticular infections of rabbits, or sampling of human syphilitic lesions (pending ethical consent) could provide interesting insights into differing protein expression profiles. Lastly, technical handling after testicular extraction and treponemal purification may ‘stress’ the bacteria into a non-characteristic infectious expression state and some proteins may degrade quickly after extraction.
since individual protein half-life ranges can vary from several seconds to tens of hours [110]. Gentle and prompt sampling and handling of treponemal extract samples may help to alleviate these potential interferences.

Despite purification efforts through Percoll density gradient centrifugation, the high abundant rabbit albumin may have masked the spectra of some lower abundant \textit{T. pallidum} proteins. Additional purification or pre-fractionalization steps could be added to facilitate the detection of low abundant proteins, however, there is a risk of inadvertently depleting treponemal proteins through methods such as albumin depletion. Possible experimental method improvements include altering the LC-MS/MS settings to be focused on either small or large proteins and/or using alternative protease and/or multi protease protein digestion [111]. Regarding the use of spectral counting, this method remains a semi-quantitative estimation of protein abundance since proteins are not measured compared to a reference. More absolute and precise protein quantification methods could be used in the future such as isobaric tags for relative and absolute quantification (iTRAQ) or selected reaction monitoring (SRM) as reviewed by Maaß and Becher [112].

**Conclusions**

This study makes a number of contributions to the characterization of the \textit{T. pallidum} proteome: i) we detected 557 \textit{T. pallidum} proteins expressed during \textit{in vivo} experimental rabbit infection using complementary mass spectrometry detection techniques; this is the first account of 499 proteins at the protein level using these methods, ii) protein abundance semi-quantified by spectral counting showed a low correlation with previous gene expression transcriptome data, iii) 116 predicted membrane localized proteins were detected, of which 16 have evidence supporting outer membrane localization and iv) a number of virulence factors were detected, including 8/12 Tpr proteins.

**Supporting Information**

S1 Table. Mass spectrometry data reports for biological (N = 3) and technical runs (N = 2). (XLSX)

S2 Table. Extensive descriptions of all unique \textit{T. pallidum} proteins identified in this study by mass spectrometry analysis. (XLSX)

S3 Table. Peptide identifications per biological (N = 3) and MS run (N = 2) and corresponding calculated spectral counting NSAF values. (XLSX)

S4 Table. \textit{T. pallidum} membrane protein bioinformatic pipeline prediction analyses. (XLSX)

**Acknowledgments**

We are grateful to David L. Cox (Centers for Disease Control and Prevention, Atlanta, GA, USA) and Dr. Zákoucká (National Institute of Public Health/National Reference Laboratory for the Diagnostics of Syphilis, Prague, Czech Republic) for the DAL-1 strain bacteria samples.

**Author Contributions**

Conceptualization: CRK XVO KKO GAVR SH KVL CEC.
Data curation: GAVR.
Formal analysis: KKO SH KVL CJM GAVR.
Funding acquisition: DS CEC XVO CRK.
Investigation: KKO MS GAVR.
Methodology: GAVR CRK KKO CJM SH XVO.
Project administration: DS CEC XVO CRK.
Resources: MS DS CJM SH XVO.
Software: CJM.
Supervision: XVO CRK CEC DS.
Validation: CEC.
Visualization: KKO GAVR.
Writing – original draft: KKO GAVR.
Writing – review & editing: KKO SH KVL CJM MS DS CEC XVO CRK GAVR.

References

1. Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. PLoS One. 2015; 10: e0143304. doi:10.1371/journal.pone.0143304 PMID: 26646541

2. Read P, Fairley CK, Chow EPF. Increasing trends of syphilis among men who have sex with men in high income countries. Sex Health. CSIRO PUBLISHING; 2015; 12: 155–163. doi:10.1071/SH14153

3. Newman L, Kamb M, Hawkes S, Gomez G, Say L, Seuc A, et al. Global estimates of syphilis in pregnancy and associated adverse outcomes: analysis of multinational antenatal surveillance data. PLoS Med. Public Library of Science; 2013; 10: e1001396. doi:10.1371/journal.pmed.1001396

4. Cox DL, Riley B, Chang P, Sayahatheri S, Tassell S, Hevelone J. Effects of molecular oxygen, oxidation-reduction potential, and antioxidants upon in vitro replication of Treponema pallidum subsp. pallidum. Appl Environ Microbiol. 1990; 56: 3063–72. PMID:2285317

5. Cox CD, Barber MK. Oxygen uptake by Treponema pallidum. Infect Immun. 1974; 10: 123–7. PMID:4366918

6. Cover WH, Norris SJ, Miller JN. The microaerophilic nature of Treponema pallidum: enhanced survival and incorporation of tritiated adenine under microaerobic conditions in the presence or absence of reducing compounds. Sex Transm Dis. 9: 1–8. PMID:10328016

7. Harman M, Vig DK, Radolf JD, Wolgemuth CW. Viscous Dynamics of Lyme Disease and Syphilis Spirochetes Reveal Flagellar Torque and Drag. Biophys J. 2013; 105: 2273–2280. doi:10.1016/bj.2013.10.004 PMID:24268139

8. Izard J, Renken C, Hsieh C-E, Desrosiers DC, Dunham-Ems S, La Vake C, et al. Cryo-electron tomography elucidates the molecular architecture of Treponema pallidum, the syphilis spirochete. J Bacteriol. 2009; 191: 7566–80. doi:10.1128/JB.01031-09 PMID:19820083

9. Liu J, Howell JK, Bradley SD, Zheng Y, Zhou ZH, Norris SJ. Cellular architecture of Treponema pallidum: novel flagellum, periplasmic cone, and cell envelope as revealed by cryo electron tomography. J Mol Biol. 2010; 403: 546–61. doi:10.1016/j.jmb.2010.09.020 PMID:20850455

10. Radolf JD, Norgard MV, Schulz WW. Outer membrane ultrastructure explains the limited antigenicity of virulent Treponema pallidum. Proc Natl Acad Sci U S A. 1989; 86: 2051–5. PMID:2648388

11. 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: 5005–11. PMID:2679902

12. Norris SJ, Edmondson DG. Factors affecting the multiplication and subculture of Treponema pallidum subsp. pallidum in a tissue culture system. Infect Immun. 1986; 53: 534–539. PMID:3091504
13. 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: 375–88. PMID: 9665876

14. Pětrošová H, Pospíšilová P, Strouhal M, Čejková D, Zobaníková M, Mikalová L, et al. Resequencing of Treponema pallidum ssp. pallidum strains Nichols and SS14: correction of sequencing errors resulted in increased separation of syphilis treponeme subclusters. PLoS One. 2013; 8: e74319. doi: 10.1371/journal.pone.0074319 PMID: 24058545

15. Smajs D, McKeivitt M, Howell JK, Norris J, Cai W, Palzkill T, et al. Transcriptome of Treponema pallidum: Gene Expression Profile during Experimental Rabbit Infection Transcriptome of Treponema pallidum: Gene Expression Profile during Experimental Rabbit Infection. 1. 2005;

16. McGill MA, Edmondson DG, Carroll J a, Cook RG, Orkiszewski RS, Norris SJ. Characterization and serologic analysis of the Treponema pallidum proteome. Infect Immun. 2010; 78: 2631–43. doi: 10.1128/IAI.00173-10 PMID: 20385758

17. Norris SJ. Treponema Pallidum Polypeptide Research Group. Polypeptides of Treponema pallidum: Progress toward Understanding Their Structural, Functional, and Immunologic Roles. Microbiol Rev. 1993; 57: 750–779.

18. Catren I, Hermann R. The proteome of Mycoplasma pneumoniae, a supposedly “simple” cell. Proteomics. 2011; 11: 3614–32. doi: 10.1002/pmic.201100076 PMID: 21751371

19. Lafond RE, Lukehart SA. Biological basis for syphilis. Clin Microbiol Rev. American Society for Microbiology; 2006; 19: 29–49. doi: 10.1128/CMR.19.1.29-49.2006

20. Weinstock GM, Hardham JM, McLeod MP, Sodergren EJ, Norris SJ. The genome of Treponema pallidum: new light on the agent of syphilis. FEMS Microbiol Rev. 1998; 22: 323–32. PMID: 9862125

21. Schous L. Recombinant DNA technology in syphilis research. In: Wright DJM, Archard L, editors. Molecular biology of sexually transmitted diseases. London: Chapman & Hall; 1992.

22. Smajs D, McKeivitt M, Wang L, Howell JK, Norris SJ, Palzkill T, et al. BAC library of T. pallidum DNA in E. coli. Genome Res. 2002; 12: 515–22. doi: 10.1101/gr.207302 PMID: 11875041

23. McKeivitt M, Patel K, Smajs D, Marsh M, McLoughlin M, Norris SJ, et al. Systematic cloning of Treponema pallidum open reading frames for protein expression and antigen discovery. Genome Res. 2003; 13: 1665–74. doi: 10.1101/gr.288103 PMID: 12805273

24. McKeivitt M, Brinkman MB, McLoughlin M, Perez C, Howell JK, Weinstock GM, et al. Genome Scale Identification of Treponema pallidum Antigens. 2005; 73: 4445–4450.

25. Brinkman MB, McKeivitt M, Perez C, Howell J, George M, Norris SJ, et al. Reactivity of Antibodies from Syphilis Patients to a Protein Array Representing the Treponema pallidum Proteome Reactivity of Antibodies from Syphilis Patients to a Protein Array Representing the Treponema pallidum Proteome. 1. 2006;

26. Setubal JC, Reis M, Matsunaga J, Haake DA. Lipoprotein computational prediction in spirochaetal genomes. Microbiology. 2006; 152: 113–21. doi: 10.1099/mic.0.28317-0 PMID: 16385121

27. Cox DL, Luthra A, Dunham-Em's S, Desrosiers DC, Salazar JC, Caimano MJ, et al. Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of Treponema pallidum. Infect Immun. 2010; 78: 5178–94. doi: 10.1128/IAI.00834-10 PMID: 20876295

28. Naqvi AAT, Shahbaaz M, Ahmad F, Hassan MI. Identification of functional candidates amongst hypothetical proteins of Treponema pallidum ssp. pallidum. PLoS One. Public Library of Science; 2015; 10: e0124177. doi: 10.1371/journal.pone.0124177

29. Mikalová L, Strouhal M, Čejková D, Zobaníková M, Pospíšilová P, Norris SJ, et al. Genome analysis of Treponema pallidum subsp. pallidum and subsp. pertenue strains: most of the genetic differences are localized in six regions. PLoS One. Public Library of Science; 2010; 5: e15713. doi: 10.1371/ journal.pone.0015713

30. Palmer G, Bankhead T, Lukehart S. “Nothing is permanent but change”—antigenic variation in persistent bacterial pathogens. Cell Microbiol. 2009; 11: 1697–1705. doi: 10.1111/j.1462-5822.2009.01366.x PMID: 19709057

31. LaFond RE, Centurion-Lara A, Godomes C, Van Voorhis WC, Lukehart SA. TprK sequence diversity accumulates during infection of rabbits with Treponema pallidum subsp. pallidum Nichols strain. Infect Immun. 2006; 74: 1896–906. doi: 10.1128/IAI.74.3.1896-1906.2006 PMID: 16495565

32. Giacani L, Brandt SL, Puray-Chavez M, Reid TB, Godomes 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: 4208–25. doi: 10.1128/JB.00863-12 PMID: 22661689
33. 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: 6054–6057. doi: 10.1128/IAI.71.10.6054-6057.2003 PMID: 14500529

34. Morgan CA, Molini BJ, Lukehart SA, Van Voorhis WC. Segregation of B and T cell epitopes of Treponema pallidum repeat protein K to variable and conserved regions during experimental syphilis infection. J Immunol. 2002; 169: 952–7. doi: 10.4049/jimmunol.169.2.952 PMID: 12097401

35. Reid TB, Molini BJ, Fernandez MC, Lukehart SA. Antigenic variation of TprK facilitates development of secondary immunity. Infect Immun. 2014; 82: 4959–67. doi: 10.1128/IAI.02236-14 PMID: 25225245

36. Giacani L, Molini BJ, Kim EY, Godornes BC, Leader BT, Tantalo LC, et al. Antigenic variation in Treponema pallidum: TprK sequence diversity accumulates in response to immune pressure during experimental syphilis. J Immunol. American Association of Immunologists; 2010; 184: 3822–9. doi: 10.4049/jimmunol.0902788

37. Anand A, Luthra A, Dunham-Emms S, Caimano MJ, Karanian C, LeDoyt M, et al. TprC/D (Tp0117/131), a trimeric, pore-forming rare outer membrane protein of Treponema pallidum, has a bipartite domain structure. J Bacteriol. 2012; 194: 2321–33. doi: 10.1128/JB.00101-12 PMID: 22389487

38. Anand A, LeDoyt M, Karanian C, Luthra A, Koszelak-Rosenblum M, Małkowski MG, et al. Bipartite Topology of Treponema pallidum Repeat Proteins C/D and I: OUTER MEMBRANE INSERTION, MMERIZATION, AND PORIN FUNCTION REQUIRE A C-TERMINAL β-BARREL DOMAIN. J Biol Chem. 2015; 290: 12313–18. doi: 10.1074/jbc.M114.629188 PMID: 25805501

39. Weigel LM, Radolf JD, Norgard M V. The 47-kDa major lipoprotein immunogen of Treponema pallidum is a penicillin-binding protein with carboxypeptidase activity. Proc Natl Acad Sci U S A. 1994; 91:11611–5. PMID: 7972112

40. Cha JY, Ishiwata A, Mobashery S. A novel beta-lactamase activity from a penicillin-binding protein of Treponema pallidum. J Biol Chem. 2004; 279: 14917–21. doi: 10.1074/jbc.M400666200 PMID: 14747460

41. Deka RK, Machius M, Norgard M V, Tomchick DR. Crystal structure of the 47-kDa lipoprotein of Treponema pallidum reveals a novel penicillin-binding protein. J Biol Chem. 2002; 277: 41857–64. doi: 10.1074/jbc.M207402200 PMID: 12196548

42. Radolf JD, Arndt LL, Akins DR, Curetty LL, Levi ME, Shen Y, et al. Treponema pallidum and Borrelia burgdorferi lipoproteins and synthetic lipopeptides activate monocytes/macrophages. J Immunol. 1999; 154: 2866–77. PMID: 7876555

43. Kelesidis T. The cross-talk between spirochetal lipoproteins and immunity. Front Immunol. 2014; 5: 85–88.

44. Semanjski M, Macek B. Shotgun proteomics of bacterial pathogens: advances, challenges and clinical implications. Expert Rev Proteomics. Taylor & Francis; 2015; doi: 10.1586/14789450.2016.1132168

45. Becher D, Hempel K, Sievers S, Zühlke D, Pané-Farré J, Otto A, et al. A proteomic view of an important human pathogen—towards the quantification of the entire Staphyloccoccus aureus proteome. PLoS One. Public Library of Science; 2009; 4: e8176. doi: 10.1371/journal.pone.0008176

46. Soufi B, Krug K, Harst A, Macek B. Shotgun proteomics of bacterial pathogens: advances, challenges and clinical implications. Expert Rev Proteomics. Taylor & Francis; 2015; doi: 10.1586/14789450.2016.1038/nmeth.1019

47. Kruh NA, Troudt J, Izzo A, Prenni J, Dobos KM. Portrait of a pathogen: the Mycobacterium tuberculosis proteome in vivo. PLoS One. Public Library of Science; 2010; 5: e13938. doi: 10.1371/journal. pone.0013938

48. Lukehart SA, Marra CM. Isolation and laboratory maintenance of Treponema pallidum. Curr Protoc Microbiol. 2007; Chapter 12: Unit 12A. 1. doi: 10.1002/9780471729259.mc12a01s7

49. Hartl PA, Norris SJ, Lovett MA, Miller JN. Purification of Treponema pallidum, Nichols strain, by Percoll density gradient centrifugation. Sex Transm Dis. 11: 275–86. PMID: 6098033

50. Zobaníková M, Mikolka P, Cejková D, Pospíšilová P, Chen L, Strouhal M, et al. Complete genome sequence of Treponema pallidum strain DAL-1. Stand Genomic Sci. 2012; 7: 12–21. doi: 10.4056/ sgs.2615838 PMID: 23449808

51. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Flores L, Washburn MP. Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. J Proteome Res. American Chemical Society; 2006; 5: 2339–47. doi: 10.1021/pr060161n

52. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. Nature Publishing Group; 2007; 4: 207–14. doi: 10.1038/nmeth1019
53. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, et al. The PeptideAtlas project. Nucleic Acids Res. Oxford University Press; 2006; 34: D655–D658. doi: 10.1093/nar/gkj040

54. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. Nature Publishing Group; 2011; 8: 785–6. doi: 10.1038/nmeth.1701

55. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. Cold Spring Harbor Laboratory Press; 2003; 12: 1652–62. doi: 10.1110/ps.0303703

56. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. Oxford University Press; 2010; 26: 1608–15. doi: 10.1093/bioinformatics/btq249

57. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001; 305: 567–80. doi: 10.1006/jmbi.2000.4315 PMID: 1152613

58. Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ. PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. Nucleic Acids Res. Oxford University Press; 2004; 32: W400–4. doi: 10.1093/nar/gkh417

59. Yu C-S, Chen Y-C, Lu C-H, Hwang J-K. Prediction of protein subcellular localization. Proteins. Wiley Subscription Services, Inc., A Wiley Company; 2006; 64: 643–51. doi: 10.1002prot.21018

60. Reynolds SM, Käll L, Riffle ME, Bilmes JA, Noble WS. Transmembrane topology and signal peptide prediction using dynamic bayesian networks. PLoS Comput Biol. Public Library of Science; 2008; 4: e1000213. doi: 10.1371/journal.pcbi.1000213

61. Käll L, Krogh A, Sonnhammer ELL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. Oxford University Press; 2007; 35: W429–32. doi: 10.1093/nar/gkm256

62. Viklund H, Bernsel A, Skwark M, Elofsson A. SPOCTOPUS: a combined predictor of signal peptides and membrane protein topology. Bioinformatics. Oxford University Press; 2008; 24: 2928–9. doi: 10.1093/bioinformatics/btn550

63. Tuszády GE, Simon I. The HMMTOP transmembrane topology prediction server. Bioinformatics. 2001; 17: 849–50. doi: 10.1093/bioinformatics/17.9.849 PMID: 11590105

64. Berven FS, Flikka K, Jensen HB, Eidhammer I. BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. Nucleic Acids Res. Oxford University Press; 2004; 32: W394–9. doi: 10.1093/nar/gkh351

65. Ou Y-Y, Chen S-A, Wu S-C. ETMB-RBF: discrimination of metal-binding sites in electron transporters based on RBF networks with PSSM profiles and significant amino acid pairs. PLoS One. Public Library of Science; 2013; 8: e46572. doi: 10.1371/journal.pone.0046572

66. Čejková D, Strouhal M, Norris SJ, Weinstock GM, Šmaj D. A Retrospective Study on Genetic Heterogeneity within Treponema Strains: Subpopulations Are Genetically Distinct in a Limited Number of Positions. PLoS Negl Trop Dis. Public Library of Science; 2015; 9: e0004110. doi: 10.1371/journal.pntd.0004110

67. 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: 191–202. doi: 10.1016/j.meegid.2011.12.001 PMID: 22198325

68. Stamm L V, 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: 6482–6. PMID: 11035764

69. Centurion-Lara A, Godornes C, Castro C, Van Voorhis WC, Lukehart SA. The tprK gene is heterogeneous among Treponema pallidum strains and has multiple alleles. Infect Immun. 2000; 68: 824–31. PMID: 10639452

70. Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. Fine analysis of genetic diversity of the tpr gene family among treponemal species, subspecies and strains. PLoS Negl Trop Dis. Public Library of Science; 2013; 7: e2222. doi: 10.1371/journal.pntd.0002222

71. Giacani L, Brandt SL, Ke W, Reid TB, Molini BJ, Iverson-Cabral S, et al. Transcription of TP0126, Treponema pallidum putative OmpW homolog, is regulated by the length of a homopolymeric guanosine repeat. Infect Immun. American Society for Microbiology; 2015; 83: 2275–89. doi: 10.1128/AI.00360-15

72. Desrosiers DC, Anand A, Luthra A, Dunham-Embs SM, Ledoyt M, Cummings M a D, et al. TP0326, a Treponema pallidum β-barrel assembly machinery A (BamA) orthologue and rare outer membrane protein. Mol Microbiol. 2011; 80: 1496–515. doi: 10.1111/j.1365-2958.2011.07662.x PMID: 21488980
73. Luthra A, Anand A, Radolf JD. Treponema pallidum in Gel Microdroplets: A Method for Topological Analysis of BamA (TP0326) and Localization of Rare Outer Membrane Proteins. Methods Mol Biol. 2015; 1329: 67–75. doi: 10.1007/978-1-4939-2871-2_6 PMID: 26427677

74. Cameron CE, Lukehart SA, Castro C, Molini B, Godornes C, Van Voorhis WC. Opsonic potential, protective capacity, and sequence conservation of the Treponema pallidum subspecies pallidum Tp92. J Infect Dis. 2000; 181: 1401–13. doi: 10.1086/315399 PMID: 10762571

75. Cameron CE, Brown EL, Kuroiwa JMY, Schnapp LM, Brouwer NL. Treponema pallidum fibronectin-binding proteins. J Bacteriol. 2004; 186: 7019–22. doi: 10.1128/JB.186.20.7019-7022.2004 PMID: 15466055

76. Giacani L, Sambri V, Marangoni A, Cavrini F, Storni E, Donati M, et al. Immunological evaluation and cellular location analysis of the Tprrl antigen of Treponema pallidum subsp. pallidum. Infect Immun. 2005; 73: 3817–22. doi: 10.1128/IAI.73.6.3817-3822.2005 PMID: 15908421

77. Salazar JC, Pope CD, Moore MW, Pope J, Kiely TG, Radolf JD. Lipoprotein-dependent and -independent immune responses to spirochetal infection. Clin Diagn Lab Immunol. American Society for Microbiology; 2005; 12: 949–58. doi: 10.1128/CDLI.12.8.949-958.2005

78. Zybailov BL, Florens L, Washburn MP. Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. Mol Biosyst. The Royal Society of Chemistry; 2007; 3: 354–60. doi: 10.1039/b701483j

79. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS Lett. 2009; 583: 3966–73. doi: 10.1016/j.febslet.2009.03.036 PMID: 19850042

80. Giacani L, Molini B, Godornes C, Barrett L, Van Voorhis W, Centurion-Lara A, et al. Quantitative analysis of tpr gene expression in Treponema pallidum isolates: Differences among isolates and correlation with T-cell responsiveness in experimental syphilis. Infect Immun. 2007; 75: 104–12. doi: 10.1128/IAI.01124-06 PMID: 17030565

81. Kim KH, Aulakh S, Paetzel M. The bacterial outer membrane β-barrel assembly machinery. Protein Sci. 2012; 21: 751–68. doi: 10.1002/pro.2069 PMID: 22549918

82. Pagès J-M, James CE, Winterhalter M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. Nat Rev Microbiol. Nature Publishing Group; 2008; 6: 893–903. doi: 10.1038/nrmicro1994

83. Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat Rev Microbiol. Nature Publishing Group; 2009; 7: 206–14. doi: 10.1038/nrmicro2069

84. Selkirk J, Leyton DL, Webb CT, Lithgow T. Assembly of β-barrel proteins into bacterial outer membranes. Biochim Biophys Acta. 2014; 1843: 1542–50. doi: 10.1016/j.bbamcr.2013.10.009 PMID: 24135059

85. Lyu ZX, Zhao XS. Periplasmic quality control in biogenesis of outer membrane proteins. Biochem Soc Trans. 2015; 43: 133–8. doi: 10.1042/BST20140217 PMID: 25849907

86. Luthra A, Anand A, Hawley KL, LeDoyt M, La Vake CJ, Caimano MJ, et al. A Homology Model Reveals Novel Structural Features and an Immunodominant Surface Loop/Opsonic Target in the Treponema pallidum BamA Ortholog TP0326. J Bacteriol. 2015; 197: 1906–14. doi: 10.1128/JB.02820-14 PMID: 25825429

87. Tomson FL, Conley PG, Norgard M V, Hagman KE. Assessment of cell-surface exposure and vaccinogenic potentials of Treponema pallidum candidate outer membrane proteins. Microbes Infect. 2007; 9: 1267–75. doi: 10.1016/j.micinf.2007.05.018 PMID: 17890130

88. Bunkis I, Denker K, Ostberg Y, Andersson C, Benz R, Bergström S. An RND-type efflux system in Borrelia burgdorferi is involved in virulence and resistance to antimicrobial compounds. PLoS Pathog. 2008; 4: e1000009. doi: 10.1371/journal.ppat.1000009 PMID: 18389081

89. Eshghi A, Cullen PA, Cowen L, Zuemer RL, Cameron CE. Global Proteome Analysis of Leptospira interrogans research articles. 2009; 4578.

90. Babolin C, Amedei A, Ozolinis D, Zilevica A, D’Elia MM, de Bernard M. TpF1 from Treponema pallidum activates inflamasome and promotes the development of regulatory T cells. J Immunol. 2011; 187: 1377–84. doi: 10.4049/jimmunol.1100615 PMID: 21709157

91. Weinstock GM, Smajs D, Hardham J, Norris SJ. From microbial genome sequence to applications. Res Microbiol. 2000; 151: 151–158. doi: 10.1016/S0923-2508(00)01115-7 PMID: 10865961

92. Giacani L, Lukehart S, Centurion-Lara A. Length of guanosine homopolymeric repeats modulates promoter activity of subfamily II tpr genes of Treponema pallidum ssp. pallidum. FEMS Immunol Med Microbiol. The Oxford University Press; 2007; 51: 289–301. doi: 10.1111/j.1574-695X.2007.00303.x
93. Brinkman MB, McGill MA, Pettersson J, Rogers A, Matejková P, Smajs D, et al. A novel Treponema pallidum antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect Immun. American Society for Microbiology; 2008; 76: 1848–57. doi:10.1128/IAI.01424-07

94. Cahoon LA, Seifert HS. An alternative DNA structure is necessary for pilin antigenic variation in Neisseria gonorrhoeae. Science. American Association for the Advancement of Science; 2009; 325: 764–7. doi:10.1126/science.1175653

95. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999; 27: 573–80. PMID:9862982

96. Kolpakov R, Bana G, Kucherov G. mreps: Efficient and flexible detection of tandem repeats in DNA. Nucleic Acids Res. 2003; 31: 3672–8. PMID:12824391

97. Moxon R, Bayliss C, Hood D. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. Annu Rev Genet. Annual Reviews; 2006; 40: 307–33. doi:10.1146/annurev.genet.40.110405.090442

98. Jerme JP, Bell JA, Plovanich-Jones AE, Barrick JE, Brown CT, Mansfield LS. Standing genetic variation in contingency loci drives the rapid adaptation of Campylobacter jejuni to a novel host. PLoS One. Public Library of Science; 2011; 6: e16399. doi: 10.1371/journal.pone.0016399

99. Chu Q, Ma J, Saghatelian A. Identification and characterization of sORF-encoded polyepitides. Crit Rev Biochem Mol Biol. Informa Healthcare; 2015; 50: 134–41. doi: 10.3109/10409238.2015.1016215

100. Wadler CS, Vanderpool CK. A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. Proc Natl Acad Sci. 2007; 104: 20454–20459. doi: 10.1073/pnas.0708102104 PMID: 18042713

101. Kelkar YD, Ochman H. Genome reduction promotes increase in protein functional complexity in bacteria. Genetics. Genetics; 2013; 193: 303–7. doi: 10.1534/genetics.112.145656 PMID: 23114380

102. Jeffery CJ. Moonlighting proteins: old proteins learning new tricks. Trends Genet. Elsevier; 2003; 19: 415–7. doi: 10.1016/S0168-9525(03)00167-7

103. Henderson B, Martin A. Bacterial moonlighting proteins and bacterial virulence. Curr Top Microbiol Immunol. Springer Berlin Heidelberg; 2013; 358: 155–213. doi: 10.1007/82_2011_188

104. Henderson B. An overview of protein moonlighting in bacterial infection. Biochem Soc Trans. 2014; 42: 1720–7. doi: 10.1042/BST20140236 PMID: 25399596

105. Khan I, Chen Y, Dong T, Hong X, Takeuchi R, Mori H, et al. Genome-scale identification and characterization of moonlighting proteins. Biol Direct. BioMed Central; 2014; 9: 30. doi: 10.1186/s13062-014-0030-9

106. Khan IK, Kihara D. Computational characterization of moonlighting proteins. Biochem Soc Trans. Portland Press Limited; 2014; 42: 1780–5. doi: 10.1042/BST20140214

107. Carlson JA, Dabiri G, Cribier B, Sell S. The immunopathobiology of syphilis: the manifestations and course of syphilis are determined by the level of delayed-type hypersensitivity. Am J Dermatopathol. 2011; 33: 433–60. doi: 10.1097/DAD.0b013e3181e8b587 PMID: 21694502

108. Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, Beynon RJ. Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. J Proteome Res. American Chemical Society; 2009; 8: 104–12. doi: 10.1021/pr800641v

109. Strugnell RA, Drummond L, Faine S. Secondary lesions in rabbits experimentally infected with Treponema pallidum. Genitourin Med. 1986; 62: 4–8. PMID: 3949349

110. Carlson JA, Dabiri G, Cribier B, Sell S. The immunopathobiology of syphilis: the manifestations and course of syphilis are determined by the level of delayed-type hypersensitivity. Am J Dermatopathol. 2011; 33: 433–60. doi: 10.1097/DAD.0b013e3181e8b587 PMID: 21694502

111. Tsiatsiani L, Heck AJR. Proteomics beyond trypsin. FEBS J. 2015; 282: 2612–26. doi: 10.1111/febs.13287 PMID: 25823410

112. Maaß S, Becher D. Methods and applications of absolute protein quantification in microbial systems. J Proteomics. 2016; 136: 222–33. doi: 10.1016/j.jprot.2016.01.015 PMID: 26825536