Nanoscale

PAPER

View Article Online
View Journal | View Issue

Cite this: *Nanoscale*, 2019, **11**, 587

A new DNA sensor system for specific and quantitative detection of mycobacteria†

Oskar Franch,‡[a] Xiao Han,‡[b] Lærke Bay Marcussen,‡[a] Asger Givskov,[a] Marie Bech Andersen,[a] Adwait Anand Godbole,[c] Charlotte Harmsen,[a] Niels Nørskov-Lauritsen,[d] Jonas Thomsen,[a] Finn Skou Pedersen,[a,e] Yilong Wang,[f,g] Donglu Shi,[f,g] Christian Wejse,[h] Lone Pødenphant,[d] Valakunja Nagaraja,[i] Johanna Bertl,[j] Magnus Stougaard,[k] Yi-Ping Ho,[l] Marianne Smedegaard Hede,[a] Rodrigo Labouriau[j] and Birgitta Ruth Knudsen *[a,e]

In the current study, we describe a novel DNA sensor system for specific and quantitative detection of mycobacteria, which is the causative agent of tuberculosis. Detection is achieved by using the enzymatic activity of the mycobacterial encoded enzyme topoisomerase IA (TOP1A) as a biomarker. The presented work is the first to describe how the catalytic activities of a member of the type IA family of topoisomerases can be exploited for specific detection of bacteria. The principle for detection relies on a solid support anchored DNA substrate with dual functions namely: (1) the ability to isolate mycobacterial TOP1A from crude samples and (2) the ability to be converted into a closed DNA circle upon reaction with the isolated enzyme. The DNA circle can act as a template for rolling circle amplification generating a tandem repeat product that can be visualized at the single molecule level by fluorescent labelling. This reaction scheme ensures specific, sensitive, and quantitative detection of the mycobacteria TOP1A biomarker as demonstrated by the use of purified mycobacterial TOP1A and extracts from an array of non-mycobacteria and mycobacteria species. When combined with mycobacteriophage induced lysis as a novel way of effective yet gentle extraction of the cellular content from the model *Mycobacterium smegmatis*, the DNA sensor system allowed detection of mycobacteria in small volumes of cell suspensions. Moreover, it was possible to detect *M. smegmatis* added to human saliva. Depending on the composition of the sample, we were able to detect 0.6 or 0.9 million colony forming units (CFU) per mL of mycobacteria, which is within the range of clinically relevant infection numbers. We, therefore, believe that the presented assay, which relies on techniques that can be adapted to limited resource settings, may be the first step towards the development of a new point-of-care diagnostic test for tuberculosis.

Received 27th September 2018,
Accepted 25th November 2018

DOI: 10.1039/c8nr07850e

rsc.li/nanoscale

[a]*Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark. E-mail: brk@mbg.au.dk*
[b]*School of Materials Science and Engineering, Tongji University, Shanghai, PR China*
[c]*Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore, India*
[d]*Department of Clinical Microbiology, Aarhus University Hospital, Aarhus, Denmark*
[e]*Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Aarhus, Denmark*
[f]*The Institute for Translational Nanomedicine, Shanghai East Hospital, the Institute for Biomedical Engineering & Nano Science, Tongji University School of Medicine, Shanghai, P. R. China*
[g]*Materials Science and Engineering Program, Department of Mechanical and Materials Engineering, College of Engineering and Applied Science, University of Cincinnati, Cincinnati, Ohio, USA*
[h]*Department of Infectious Diseases, Aarhus University Hospital & Department of Public Health, Aarhus University, Aarhus, Denmark*
[i]*Department of Microbiology and Cell Biology, Indian Institute of Science & Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India*
[j]*Department of Mathematics, Aarhus University, Aarhus, Denmark*
[k]*Department of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark*
[l]*Department of Biomedical Engineering, The Chinese University of Hong Kong, Hong Kong SAR*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8nr07850e

‡ These authors contributed equally to this work.

Introduction

Tuberculosis (TB), caused by infections with the *Mycobacterium tuberculosis* complex (MTBC), is among the major threats to global health. The World Health Organization reported 10.4 million new TB cases in 2016.[1] With the increased global mobility it is likely that the number of new cases may even increase in the near future. This highlights the importance of rapid diagnosis of TB in order to allow early intervention by antibiotic treatment. This will ensure an enhanced therapeutic outcome and prevent spreading of the disease.[2] Spurred on by such needs, an array of nucleic acid based molecular methods for early detection of MTBC have been presented during the past decade.[3–6] However, when considering accessibility to high incidence areas considerable challenges still remain. For instance, although the polymerase chain reaction (PCR)[5] and sequencing[7] are promising for highly sensitive identification of MTBC, both techniques are

expensive and require extensive laboratory resources, which are often not available in developing countries where TB is most prevalent.[8] In these areas, the diagnosis often relies on the easily accessible tuberculin skin test, which suffers from low sensitivity and specificity.[9,10] Even in high resource countries, complicated procedures of X-ray examination combined with culturing of the bacteria and microscopic investigations remain the gold standard for diagnosis of TB.[11] For these methods, the slow growth rate of MTBC results in a report time of up to 8 weeks, which considerably delays commencement of correct treatment and may lead to overtreatment and development of multidrug-resistant MTBC strains. Hence, the search for reliable and fast methods for MTBC detection continues. Recently, relatively simple nucleic acid based methods for MTBC detection such as Loop-Mediated Isothermal Amplification (LAMP) and Reverse transcriptase-LAMP[3,12] have been presented with some encouraging results. These methods utilize isothermal amplification techniques and may be suited for low resource settings. However, LAMP techniques are still susceptible to assay variations and require extensive sample preparation for optimal performance.[3,12–14]

Detection of pathogens based on the presence of an "inactive" biomarker (without catalytic activity) *e.g.* a nucleotide or amino acid sequence often suffers from the need for extensive signal amplification before readout. In some cases, sufficient amplification is not even possible, while in other cases extensive amplification steps may hamper the specificity of detection. To address such challenges, it is a great advantage to utilize an "active" biomarker such as an enzyme (with catalytic activity), which can add an extra step of signal amplification to the detection system without compromising specificity.[15–19] This is achieved by the inherent ability of enzymes to convert substrate molecules present in molar excess to a large number of products. Such products can then be detected either directly[16–18] or after further signal amplification.[19] We previously demonstrated how the catalytic activity of DNA modifying enzymes combined with a second signal amplification step can be exploited for the development of highly sensitive, specific and directly quantitative sensor systems for the detection of the malaria causing parasite *Plasmodium*[20–22] or the human immunodeficient virus (HIV).[23] For this purpose, we used the pathogen expressed enzymes plasmodium Topoisomerase IB (TOP1B) or HIV Integrase (IN) as biomarkers for detection. Also, we demonstrated the highly sensitive and quantitative measurement of important enzyme targets of anti-cancer therapy, namely the human TOP1B and TOP2.[24–27] In all cases, enzyme activity was detected by the generation of a specific DNA product that could be further amplified in an isothermal rolling circle amplification (RCA) reaction that allowed measurement with single-catalytic-event resolution.

The enzyme based detection systems were superior to other detection systems with regard to sensitivity and the *Plasmodium* specific assay was even demonstrated to be adaptable for low resource settings.[21] However, different enzyme activities from the same or different species are not as easily distinguishable as different DNA or RNA segments, which can be recognised based on their unique base sequence using a battery of standard methods.[28–30] The utilization of a pathogen expressed enzyme activity as a biomarker for detection depends on the identification of a unique enzymatic mechanism of action that can form the basis for specific detection. This means that the development of sensor systems for the detection of a novel enzyme biomarker is never trivial. This applies even when sensor systems exist, specific for other enzymes that at first glance appears very similar to the new biomarker enzyme.

In our previously described DNA sensor systems specific for two eukaryotic TOP1B enzymes (expression in human or *Plasmodium* spp.),[20,21,25,26] human TOP2[24] or HIV IN[23] we were able to identify characteristics unique to each of the target enzymes. Such characteristics were exploited for the development of a DNA substrate specific for the target enzyme even when present in crude biological samples. In the present study, we describe the development of a DNA sensor system that allows the detection of mycobacteria by using the activity of the mycobacteria expressed enzyme TOP1A as a biomarker. This enzyme belongs to the group of type IA topoisomerases. Despite what the name may imply, members of this enzyme group show no phylogenetic, structural or mechanistic similarities to any of the eukaryotic topoisomerases previously used as biomarkers for detection. Moreover, mycobacterial TOP1A does not possess any mechanistic characteristics that allow its catalytic activity to be distinguished from other type IA topoisomerases using a simple DNA substrate design. The development of a DNA biosensor system for specific detection of mycobacteria *via* the mycobacteria TOP1A catalytic activity must circumvent this problem.

Like all known type IA topoisomerases, mycobacteria TOP1A regulates DNA topology in a Mg^{2+} dependent reaction[31–34] that involves temporary nicking of the DNA and the formation of a transient covalent 5′-phosphotyrosyl cleavage complex. Restoration of intact DNA and release of the enzyme is ensured by subsequent ligation of the 3′-OH DNA end formed during cleavage. Unlike other type IA topoisomerases, however, mycobacterial TOP1A distinguishes itself from other known enzymes by cleaving single-stranded DNA in a Mg^{2+} independent and sequence specific manner at the strong topoisomerase site (STS).[35] These characteristics were utilized in a previously reported Quantum dot (QD) based DNA nanosensor[36] that measured specific cleavage at a STS. However, in this QD sensor setup it was not possible to exploit the amplification potential inherent to the mycobacterial TOP1A catalytic capability. This seriously hampered the analytical sensitivity (detection limit) of the sensor and prevented the detection of mycobacteria TOP1A in concentrations relevant for clinical testing.

In the current study, we describe a new mycobacteria specific DNA sensor system that takes advantage of the catalytic action of mycobacteria TOP1A for fast, quantitative and sensitive detection of mycobacteria in crude samples. This is achieved by the use of a solid support anchored linear DNA substrate with a STS that can support both isolation and

specific detection of mycobacteria TOP1A activity at the single catalytic event level. Isolation of mycobacteria TOP1A was accomplished by trapping the enzyme in the covalent complex (the cleavage complex) with the DNA substrate. This was followed by washing and addition of Mg^{2+} that induced the DNA ligation reaction of the enzyme. Completion of the catalytic cycle by DNA ligation released the enzyme to perform consecutive catalytic cycles, which each converted the linear DNA substrate to a circle. Each circle could subsequently be detected at the single molecule level after RCA. By allowing each enzyme to generate many DNA products that could each be detected, the analytical sensitivity of the assay was superior to that of the previously described QD-based DNA nanosensor. When combining the RCA based sensor system with efficient, yet gentle, extraction of active enzymes by mycobacteriophage induced cell lysis, we were able to detect mycobacteria spiked in small volumes of human saliva at concentrations relevant for clinical testing.

Materials and methods

Biological samples

HEK293 T cells were provided by Associate Professor Pia Møller Martensen, Department of Molecular Biology and Genetics, Aarhus University. *M. smegmatis* and all utilized mycobacteriophages were a kind gift from Professors Graham F. Hatfull and Carlos A. Guerrero Department of Biological Sciences, University of Pittsburgh, Pittsburgh, USA. The utilized *E. coli* strain, XL-Blue, was from Thermo Fisher.

Purified *M. smegmatis* TOP1A

M. smegmatis TOP1A was expressed in *E. coli* from the plasmid pPVN123 and purified essentially as previously described.[37] Briefly, sonicated cell lysates were centrifuged at 100 000*g*, subjected to PEI and ammonium sulphate precipitation (45–65% cut), and dialyzed and the enzyme was purified by HiTrap heparin Sepharose and HiTrap SP Sepharose column chromatography.

Preparation of the *E. coli* lysate

E. coli was grown in standard 2xTY media and harvested by centrifugation at 4000*g* for 10 minutes. The pellet was resuspended in 5 times volume of lysis buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM PMSF, and 1 mM DTT) resulting in a suspension containing 40×10^6 *E. coli* per μL. The bacteria suspension was incubated on ice for 1 hour and sonicated 10 times for 10 seconds. NaCl was added to a final concentration of 500 mM to the cell extract, which was then cleared by centrifugation at 10 000*g* at 4 °C for 10 minutes. 200×10^6 *E. coli* cells were used per reaction.

Preparation of the lysate using vortexing with glass beads

The mycobacterial strains (except *M. smegmatis*) and non-mycobacterial strains (except *E. coli*) were grown on agar plates with (mycobacteria) or without (non-mycobacterial bacteria) 5% blood. The cells were scraped off using a spatula to generate 30 μL of the cell pellet. The pellet was washed in PBS, and resuspended in 60 μL PBS containing 1 mM phenylmethane sulfonyl fluoride (PMSF). Lysis of the cells was performed by adding 1× vol of 212–300 μm glass beads to the resuspended cells before 8 × 1 minute vortexing followed by 1 minute's incubation on ice. Subsequently, the lysate was cleared by centrifugation at 13 000*g* for 5 minutes at 4 °C. Heat inactivation of the lysate was performed by heating to 80 °C for 5 minutes.

Mycobacterial strains were cultured, lysed, and analysed individually. Each of the non-mycobacterial strains was cultured individually and then pooled into four groups listed below. The groups were analysed in bulk using the lysis method described above.

1: *Streptococcus pneumoniae*, *Streptococcus mitis*, *Streptococcus oralis*, *Streptococcus salivarius*, *Streptococcus anginosus*, *Streptococcus pyogenes*, and *Burkholderia cepacia*.

2: *Streptococcus agalactiae*, *Streptococcus dysgalactiae*, *Enterococcus faecalis*, *Enterococcus faecium*, *Staphylococcus epidermidis*, and *Achromobacter xylosoxidans*.

3: *Staphylococcus lugdunensis*, *Nocardia cyriacigeorgica*, *Nocardia farcinica*, *Bacillus cereus*, *Moraxella catarrhalis*, and *Pseudomonas aeruginosa*.

4: *Neisseria meningitis*, *Haemophilus influenzae*, *Haemophilus parainfluenzae*, *Actinobacillus ureae*, *Escherichia coli*, and *Actinomyces naeslundii*.

Preparation of lysates using mycobacteriophages

Colony forming units (CFU) per μL determination. *M. smegmatis* mc^2155 cultures were grown in growth media (7H9 liquid medium supplemented with carbenicillin (50 mg ml^{-1}), cycloheximide (10 mg ml^{-1}), 10% ADC, 1 mM $CaCl_2$ and 0.25 vol% of TWEEN-80) to OD_{600} = 1.8. The culture was then streaked on 7H9 agar plates and a single colony was grown in 7H9 liquid medium to exponential phase (OD_{600} = 1.8), before the dilution series was prepared. 50 μL of each diluted sample was streaked out on a 7H9 agar plate and incubated until colonies were countable by the naked eye (2–3 days). Colonies were counted using the ImageJ software, to determine the number of CFU μL^{-1}.

Plague forming units (PFU) per μL determination. *M. smegmatis* was grown to reach the exponential phase (OD_{600} = 1.8) in growth media. 2 mL of the generated culture were preincubated with 500 μL of mycobacteriophage solution for 20 minutes, before 2 mL of melted 14% (w/v) agarose were added. The solution was then mixed before being poured on top of a 7H9 agar plate and the PFU was determined after incubation for 2–3 days by counting the plagues using ImageJ software.

Purification of mycobacteriophages using ultracentrifugation. Mycobacteriophages (AdephagiaΔ41,Δ43; BPsΔ33; Bxx2; D29; L5; and TM4,[38–40] which were a kind gift from C.A. Guerrero and G. F. Hatfull at the University of Pittsburgh, USA) were grown on plates with *M. smegmatis* (see PhagesDB[41]). The plates were covered with 10 mL of phage buffer (10 mM Tris-

HCl (pH 7.5), 10 mM $MgSO_4$, 68 mM NaCl) and incubated at 4 °C, overnight. Next day, the buffer (now containing mycobacteriophages) was retrieved. 1 ml of this mycobacteriophage suspension was layered on top of a 12.5% (w/v) to 52.5% (w/v) sucrose gradient made on a Foxy Jr. gradient mixer (Teledyne ISCO). The sucrose was dissolved in phage buffer. Samples were centrifuged at 100 000*g* for 40 minutes in a SW41 swing bucket rotor (Beckman Coulter, Optima L-80 XP Ultracentrifuge) and stopped without applying the brakes. The gradient was divided into 12 fractions of 1 ml (Foxy Jr., Teledyne ISCO).

Preparation of lysates from *M. smegmatis*. *M. smegmatis* was grown to an exponential phase (OD_{600} = 1.8) in the growth medium and harvested by centrifugation at 6000*g* for 15 minutes. The pellet was then washed in the growth medium without Tween®80. The cultivated *M. smegmatis* was dissolved in the medium and further diluted before added to either phage buffer or human saliva as stated in the text and lysed by adding mycobacteriophages (D29 or AdephagiaΔ41,Δ43) to a final concentration of 17 000 PFU μL^{-1} with a supplement of $CaCl_2$ (2 mM final). Mycobacteriophage D29 was used for the quantitative detection of *M. smegmatis* diluted in phage buffer, whereas mycobacteriophage AdephagiaΔ41,Δ43 was used for lysis of *M. smegmatis* diluted in saliva. Samples were incubated at 37 °C for 2 hours under constant mixing. Subsequently, NaCl was added to a final concentration of 560 mM, and cell debris was precipitated by centrifugation with 20 800*g* for 5 minutes at 4 °C.

All protocols for the growth of *M. smegmatis* and Mycobacteriophages are available from PhagesDB.[41]

REEAD assay

Immobilization of the mycobacterial TOP1A DNA substrate molecule. An amine coupled DNA primer (5′-Am-CCAACCAACCAACCAAATAAGCGATCTTCACAGT-3′) was immobilized on NHS-coated microscopy slides (CodeLink® Activated Slides) as described by the supplier. The mycobacterial TOP1A DNA substrate molecule (5′-CAGTGAGCG AGCTTCCGCTTGACATCCCAATATCTCTACTGTGAAGATCGCTT-ATTCTCTCCTCAATGCACATGTTTGGCTCCTCTCTGAGCTTCCG-CT-3′) was hybridized to the immobilized primer using a buffer containing 0.1 μM of the mycobacterial TOP1A DNA substrate molecule, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 100 mM NaCl. After incubation at 37 °C for 1 hour in a humidity chamber, the slides were washed for 1 minute in wash buffer 1 (0.1 M Tris-HCl (pH 7.5), 150 mM NaCl, 0.3% SDS), 1 minute in wash buffer 2 (0.1 M Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween 20), and 1 minute in 96% ethanol. Finally, the slides were equilibrated for 10 minutes in an equilibration buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 200 mM NaCl.

Mycobacterial TOP1A reaction. After removal of the equilibration buffer, mycobacterial TOP1A mediated cleavage of the DNA substrate molecule was initiated by addition of a buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 100 mM NaCl with or without purified mycobacterial TOP1A as well as with or without 20 vol% of the cell lysate from different species as stated in the text. Slides were incubated for 60 minutes at 37 °C in a humidity chamber. After incubation, the slides were washed twice for 10 minutes in a buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 500 mM NaCl. Mycobacterial TOP1A mediated ligation was initiated by addition of a buffer containing 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 100 mM NaCl. After incubation at 37 °C for 1 hour in a humidity chamber, the reaction was terminated by washing the slides for 1 minute in wash buffer 1, 1 minute in wash buffer 2, and 1 minute in 96% ethanol.

RCA and visualization of signals. RCA was performed for 60 minutes at 37 °C in 1× Phi29 buffer supplemented with 0.2 μg μL^{-1} BSA, 250 μM dNTP, and 6 units Phi29 DNA polymerase. The reactions were stopped by addition of wash buffers 1 and 2. The RCA products were detected by hybridization of 0.2 μM of a fluorescently labeled detection probe (FAM-CCTCAATGCACATGTTTGGCTCC) in a buffer containing 20% formamide, 2 × SSC, and 5% glycerol for 30 minutes at 37 °C. The slides were subsequently washed for 15 minutes in each of the wash buffers 1 and 2, dehydrated in 96% ethanol, and mounted with Vectashield (Vector Laboratories). Fluorescently labelled RCA products were visualized using an Olympus IX73 fluorescence microscope. Finally, the results were analysed using the Image J software.

Statistical analyses

The number of signals per frame was modelled using a common statistical framework in the several independent statistical analyses performed (Fig. 2–4, Tables 1, 2, and ESI Tables S1–S3†), namely a Poisson mixed model (see ref. 42, 43 and the Supplemental statistics section). The models contained a fixed effect representing the treatments and two independent Gaussian random components: one representing the over-dispersion present in the data, and a second taking the same value for the observations arising from the same experiment or the same slide for accounting for the dependence of those observations. The models related to Tables 1 and 2 contained only the first random component representing the over-dispersion since only one experiment was involved. The models were adjusted using the software R version 3.2.5,[44] specifically using the R-package "lme4"[45] for generalized linear mixed models using the Poisson distribution and the identity link function.

The *post hoc* analyses were performed using the R-package "pairwiseComparisons" available at (http://home.math.au.dk/astatlab/software/pairwisecomparisons).

The *p*-values for *post hoc* analyses were adjusted for multiple comparisons by the method for controlling the FDR (False Discovery Rate) defined in ref. 46.

Results and discussion

Assay design

The catalytic mechanism of mycobacterial TOP1A resembles that of other type IA topoisomerases making it difficult to

design a DNA substrate that can be converted to a detectable product exclusively by mycobacterial TOP1A catalytic activity. This complicates the design of a RCA based DNA sensor system specific for mycobacterial TOP1A and several attempts to achieve this failed (see ESI Fig. 1†). However, as mentioned in the introduction the DNA cleavage step of mycobacterial TOP1A catalysis is unique and the enzyme distinguishes itself from other known enzymes by cleaving the STS sequence in single stranded DNA in a Mg^{2+} independent reaction.[34,47] This leads to the generation of a covalent 5′-phosphotyrosyl cleavage complex, which allows the enzyme subsequently to ligate a 3′-OH DNA end upon addition of Mg^{2+}.[32,34,35] By taking advantage of the unique cleavage step of mycobacterial TOP1A catalysis we designed a DNA substrate with the dual functionality of first isolating mycobacterial TOP1A from the crude sample, and thereafter allowing detection of the TOP1A activity at the single catalytic event level (Fig. 1). This DNA substrate is composed of a solid support anchored DNA oligonucleotide with a free 3′-OH end (termed primer) annealed to the primer annealing sequence of a scissile DNA oligonucleotide with a STS and a specific ID sequence (Fig. 1A). In a Mg^{2+} depleted buffer, the only enzyme known to be able to cleave and become covalently attached to the surface anchored DNA substrate is mycobacterial TOP1A. Hence, potential contaminants in a sample that may compromise the specificity of the assay such as other DNA cleaving or ligating enzymes can be removed by consecutive wash steps before addition of Mg^{2+}, which induces the DNA ligation activity of mycobacterial TOP1A. Note that Mg^{2+} is a prerequisite for the activity of most DNA modifying enzymes. Avoiding this cofactor during the initial isolation step of the assay setup prevents nonspecific enzymes to interfere with the assay. After addition of Mg^{2+}, mycobacterial TOP1A can convert the cleaved substrate DNA molecule to a closed circle by ligation of the 3′-OH end. Subsequently the enzyme is released and free to perform consecutive rounds of cleavage-ligation on other substrate DNA molecules. This step constitutes the first amplification step inherent to the developed DNA sensor system. In the second amplification step, the generated DNA circle is subjected to isothermal RCA leading to the generation of a 10^3 tandem repeat product (RCP). The RCP is detected upon hybridization of fluorescent probes with a sequence identical to the ID sequence of the scissile DNA. Thereby the RCPs can be visualized at the single molecule level (see the example in Fig. 1B) using a fluorescence micro-

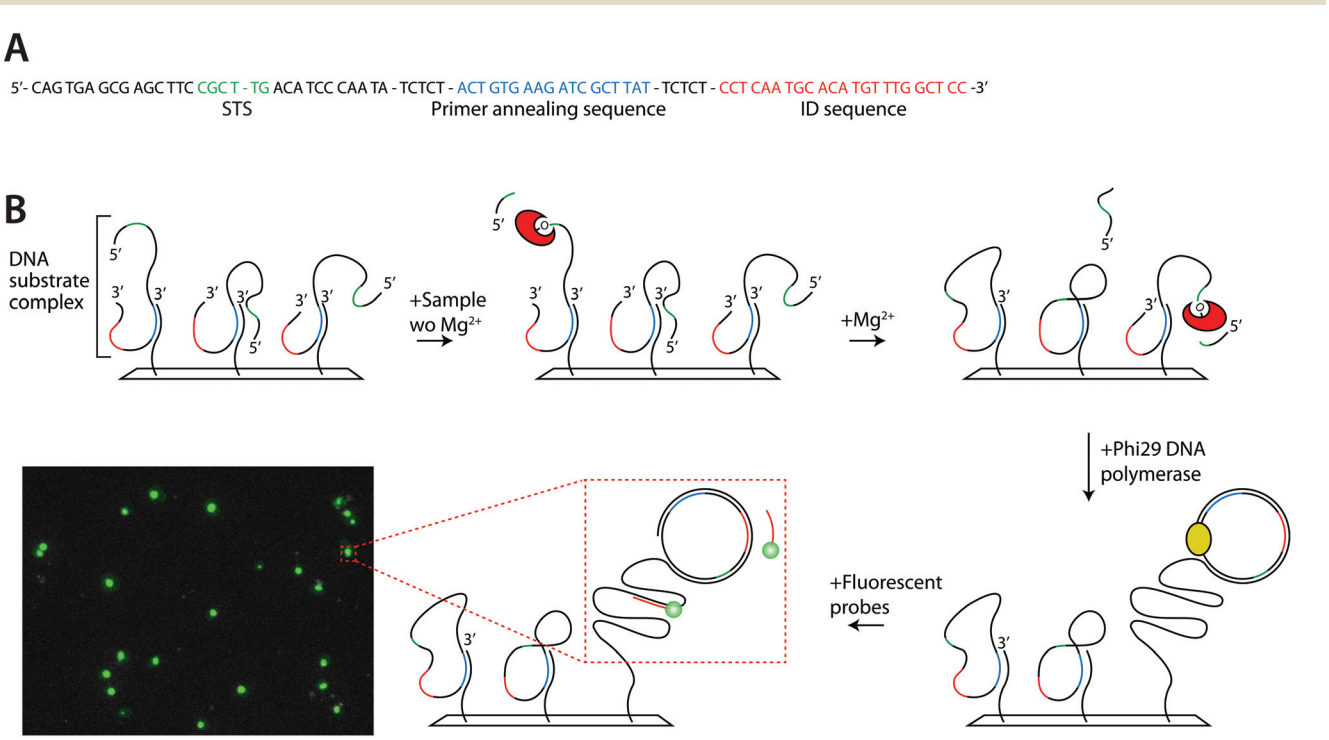


Fig. 1 (A) Sequence of the scissile DNA oligonucleotide of the mycobacterial TOP1A substrate. The recognition site for mycobacterial TOP1A, the strong topoisomerase site (STS), is marked with green text, the sequence complementary to the surface attached primer oligonucleotide is marked with blue text, and the ID sequence is marked with red text. (B) Schematic illustration of the mycobacteria sensor system. The scissile DNA oligonucleotide hybridizes to the immobilized DNA primer to generate the DNA substrate complex. Addition of a sample including mycobacterial TOP1A (red circle) in a buffer without Mg^{2+} leads to cleavage and covalent binding of the mycobacterial TOP1A enzyme to the scissile DNA molecule (to form the cleavage complex). Subsequent wash results in isolation of the covalently bound enzyme. After wash addition of Mg^{2+} induces the ligation step of the mycobacterial TOP1A catalytic cycle. Ligation results in circularization of the scissile DNA and enzyme turnover, leaving the TOP1A free to perform another catalytic cleavage-ligation cycle. The generated circle is amplified in a rolling circle amplification (RCA) reaction catalysed by Phi29 polymerase (yellow oval). The resulting rolling circle amplification product (RCP) is visualized at the single molecule level using fluorescent probes (red line with a green circle). An example of the generated RCP signals is shown in the lower panel to the left.

scope and signals can be counted as previously described.[25] Since the system relies only on isothermal amplification, each visible signal will correspond to a single circle, which will be generated by a single TOP1A cleavage-ligation reaction. Hence the assay is expected to be directly quantitative and highly sensitive.

The mycobacteria sensor system is quantitative and specific for mycobacterial TOP1A

The functionality of the mycobacteria RCA based sensor system was investigated by measuring the number of RCP signals generated by purified *M. smegmatis* TOP1A preparations with concentrations ranging from 7 to 227 nM. As a negative control 227 nM of the cleavage incompetent *M. smegmatis* TOP1A(Y339F) mutant was added instead of the wildtype enzyme (see ESI Fig. 2† for activity measurements of the utilized enzyme fractions). The number of RCP signals was counted and the results are depicted in Fig. 2 (see also ESI Table S1†) as the estimated expected number of signals per image frame adjusted for differences between experiments inferred using the Poisson mixed model with over-dispersion and a Gaussian random component representing the experiment as described in the Materials and methods section (see also the Supplementary statistics section). Note that the normalization to a standardized batch of control circles as it was reported previously[26] was complicated by the fact that the experiments were performed over an extended period of time. The control circle stock was influenced considerably by storage over time making this type of normalization unsuited for the current study.

As shown in Fig. 2 and ESI Table S1,† the expected number of signals per image frame was statistically significantly different from the negative control (*M. smegmatis* TOP1A (Y339F)) for the concentrations 114 nM and higher (227 nM), given the detection power of the present experimental design. This corresponded to a detection limit in the fmol range of purified enzymes as a total volume of 1 μL of the enzyme preparations was used for detection. Note, however, that the activity of the purified enzyme depends on many factors such as purity and storage conditions. For instance, the enzyme used in these experiments was stored as lyophilized powder, which may affect the activity. Hence, the detection limit observed in this experiment is unlikely to reflect the detection limit that can be obtained with enzymes freshly extracted from mycobacteria.

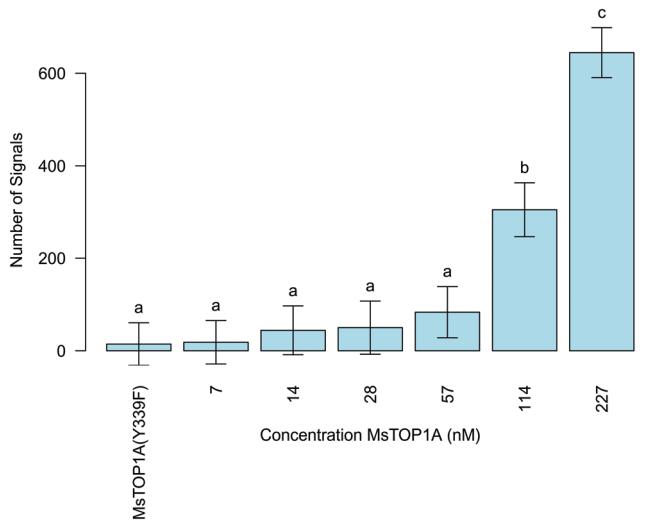


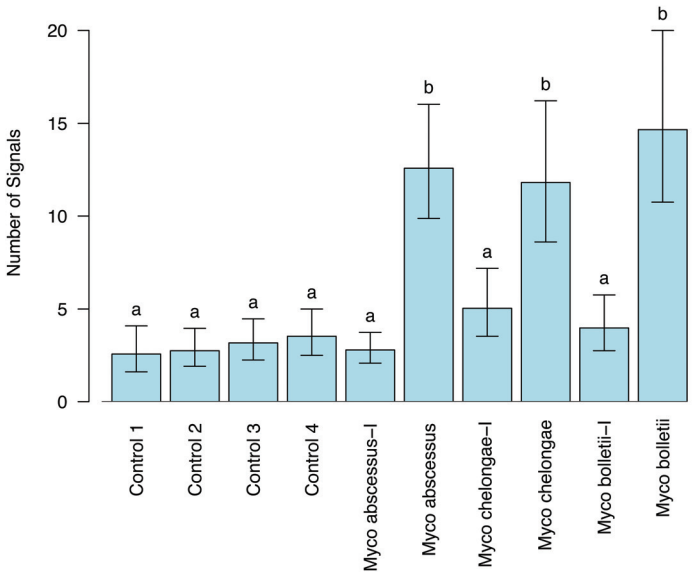
Fig. 2 Adjusted number of signals obtained when analysing increasing concentrations of purified *M. smegmatis* TOP1A and the catalytic inactive mutant *M. smegmatis* TOP1A(Y339F). Confidence intervals with 95% coverage depicted as interval bars; significant groups at 5% confidence level labelled "a", "b" and "c" (based on *p*-values adjusted for multiple testing by the FDR method, see ESI Table S1†).

As mentioned in the introduction, mycobacterial TOP1A belongs to the family of type IA topoisomerases and shares important mechanistic characteristics with members of this family although it also differs from these.[32,47,48] The prototype member of the type IA topoisomerases is *E. coli* TOP1A.[47,48] To address if *E. coli* TOP1A or other *E. coli* expressed enzymes can react with the mycobacteria RCA based sensor system the number of RCP signals generated by an *E. coli* cell extract with or without added spike-in *M. smegmatis* TOP1A was analysed. The extract from 10^{12} *E. coli* cells (500 μL cell pellet) was generated by sonication and an aliquot of the extract corresponding to 200 million cells with or without added purified *M. smegmatis* TOP1A was incubated with the sensor system. As evident from Table 1, the extracts with added spike-in *M. smegmatis* TOP1A gave rise to a number of signals significantly (*p*-value = 0.000003) above the number of background signals observed after incubation with the *E. coli* extract alone, which in turn did not differ significantly from the number of signals observed in control samples incubated with storage buffer in place of the sample (data not shown). This result suggests that the mycobacteria sensor system is specific for mycobacterial TOP1A at least in a background of the *E. coli* cellular content. Similar results were obtained when testing the specificity in a background of extracts from human HEK293 T cells (ESI Fig. 3†) and, consistently, purified human TOP1B did not produce any RCP signals when added to the mycobacteria sensor system (data not shown).

The results described above hold promise for the potential use of the mycobacteria sensor system for the specific detection of mycobacteria in human samples. However, humans may be infected by a number of bacteria other than mycobac-

Table 1 Estimated number of signals per image frame obtained from the analysis of extracts from pure *E. coli* (10 observations) and from *E. coli* containing 448 nM of spike-in purified *M. smegmatis* TOP1A using the mycobacteria sensor system (10 observations). 95% confidence interval in parenthesis. The *P*-value for equality of the mean numbers of signals per image frame is equal to 0.000003

Sample	Est. number of signals
E. coli	1.6 (0.7–3.4)
E. coli + *M. smegmatis* TOP1A	23.6 (13.2–40.1)



Bacteria species tested	
Control 1	Control 3
Streptococcus pneumoniae	*Staphylocuccus lugdunensis*
Streptococcus mitis	*Nocardia cyriacigeorgica*
Streptococcus oralis	*Nocardia farcinica*
Streptococcus salivarius	*Bacillus cereus*
Streptococcus anginosus	*Moraxella catarrhalis*
Streptococcus pyogenes	*Pseudomonas aeruginosa*
Burkholderia cepacia	
Control 2	Control 4
Streptococcus agalactiae	*Neisseria meningitis*
Streptococcus dysgalactiae	*Haemophilus influenzae*
Enterococcus faecalis	*Haemophilus parainfluenzae*
Enterococcus faecium	*Actinobacillus ureae*
Staphylocuccus epidermidis	*Escherichia coli*
Achromobacter xylosoxidans	*Actinomyces naeslundii*
Mycobacteria species	
Mycobacterium abscessus	
Mycobacterium chelongae	
Mycobacterium bolletii	

Fig. 3 Left panel: Adjusted number of signals obtained when analysing active or heat inactivated extracts (marked "-I") from three different mycobacteria species and four pools of extracts of non-mycobacteria species (marked control 1–4). Confidence intervals with 95% coverage depicted as interval bars; significant groups at 5% confidence level labelled "a" and "b" (based on *p*-values adjusted for multiple testing by the FDR method, see also ESI Table S2†). Right panel: List of non-mycobacterial and mycobacterial strains tested.

teria and some of these are likely to be found in mycobacteria relevant specimens.[49–51] Therefore, for a potential diagnostic use of the mycobacteria sensor system it is a prerequisite that such bacteria will not generate false positive signals. The specificity of the system was tested using extracts generated by vortexing 30 μL of cell pellets with glass beads. For these analyses, we used cell pellets from an array of non-mycobacteria species (mixed in four pools) and three different non-pathogenic mycobacteria species, which were used as a model for MTBC (listed in Fig. 3, right panel). As negative controls, we used extracts from the three selected mycobacteria species inactivated by heating to 80 °C for 5 minutes, which denatured DNA modifying enzymes present in the extracts and thereby prevented the catalytic activity. The results are depicted in Fig. 3, left panel and ESI Table S2.† The expected number of signals per image frame from each of the active mycobacteria extracts differed significantly from the respective heat inactivated mycobacteria extracts and the extracts of non-mycobacteria species (see *p*-values reported in ESI Table S2† and the statistical significant groups depicted in Fig. 3). Taken together the results described above demonstrate that the mycobacteria RCA based sensor system is specific for mycobacteria species when tested against other bacteria as well as human cells.

Quantitative detection of mycobacteria in small samples

The presented sensor system detects the activity of an intracellular mycobacterial enzyme, TOP1A and, hence, lysis of the cells is a prerequisite for detection. In the experiments described above extracts were prepared from 30–500 μL pellets of bacteria that were lysed by sonication or vortexing with glass beads. For diagnosis of TB the clinical samples are likely to be too small, diluted or in other ways unsuited for ultrasound or mechanical lysis.[52] Moreover, lysis by such techniques requires special equipment that may be unavailable in low resource settings where TB is highly prevalent. Standard methods for the extraction of small samples for diagnosis rely on treatment with guanidine hydrochloride or other harsh chemicals.[53,54] Such methods are acceptable for the preparation of DNA for *e.g.* PCR analyses but will destroy or seriously hamper enzyme activity, which is a prerequisite for the here presented RCA based sensor system. We therefore investigated the possibility of using mycobacteriophage induced lysis as a novel approach to prepare mycobacteria extracts suitable for the measurement of active enzymes in small samples. As a model mycobacteria, we used the non-pathogenic *M. smegmatis* for these studies. We tested the applicability of various strains of mycobacteriophages including two purely lytic mutants (AdephagiaΔ41,Δ43 and BPsΔ33).[38] As evident from ESI Fig. 4,† all tested bacteriophages were able to lyse *M. smegmatis* and generate extracts that allowed detection of mycobacteria TOP1A using the RCA based sensor system. Of the mycobacteriophages available, we chose to use D29 and the lytic mutant AdephagiaΔ41,Δ43 that both resulted in the intermediate number of RCP signals in the initial test for further analysis. When tested in the RCA based sensor system, extracts prepared by mycobacteriophage infection allowed quantitative detection of *M. smegmatis* in small samples (25 μL cell suspension). A graphical depiction of the results obtained from analyses of samples containing *M. smegmatis* ranging from 450 to 14 400 colony forming units (CFU) per μL is shown in Fig. 4. As evident from the figure and ESI Table S3†, the expected number of signals per image frame was statistically significant different (p value <0.019) from the negative control for the *M. smegmatis* concentrations of 900 CFU μL^{-1} and higher (3600–14 400 CFU μL^{-1}). This

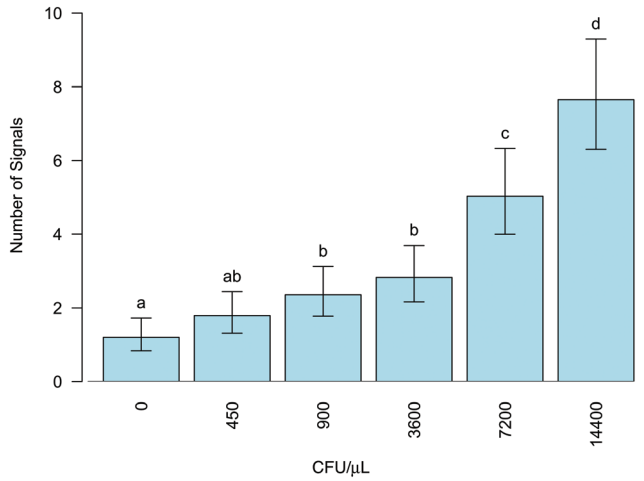


Fig. 4 Adjusted number of signals obtained when analysing increasing concentrations of *M. smegmatis* (CFU μL^{-1}). Confidence intervals with 95% coverage depicted as interval bars; significant groups at 5% confidence level labelled "a", "b", "c" and "d" (based on *p*-values adjusted for multiple testing by the FDR method). See also ESI Table S3.†

suggests that the detection limit under the present setup is between 450 and 900 CFU μL^{-1}.

Note that the negative control contained the mycobacteriophage batch that was used for lysis without added *M. smegmatis*. The mycobacteriophages were produced *via* infection of *M. smegmatis* and consequently it was difficult to completely avoid traces of mycobacterial TOP1A in the preparations, although the mycobacteriophages were purified prior to use. To avoid false positive results, the utilized mycobacteriophage preparation was therefore used as a negative reference sample in all experiments that relied on lysis by mycobacteriophage infection.

The clinical specimen for diagnosis of TB is sputum. However, some patients including HIV positive individuals and children often have trouble producing sufficient volumes of sputum for diagnosis.[55,56] Moreover, sputum can be a troublesome specimen to work with and may not be suitable for point-of-care testing.[57] Accordingly, there is a need for diagnosing TB using non-sputum samples. Since TB spread *via* droplets of saliva[58,59] the use of such specimens seems as an obvious choice. The aim with the presented RCA based sensor system, which is based on a new principle for detection, is to take the first steps in the direction of the development of a TB test suitable for quantitative measurement of TB in saliva from infected patients at low resource settings. For safety reasons, it was not possible to test the method using clinical samples in the current study. As a model sample, we therefore used saliva from uninfected individuals with or without spike-in *M. smegmatis*. The samples were treated with purified mycobacteriophages before the lysate was analysed using the sensor system. As evident from Table 2, the estimated number of signals per image frame obtained when analysing saliva samples containing 650 CFU μL^{-1} spike-in *M. smegmatis* was significantly (one tailed *p*-value <0.000001) higher than the

Table 2 Estimated number of signals per image frame obtained from the analysis of saliva and saliva with 650 CFU μL^{-1} of *M. smegmatis* (95% confidence interval in parenthesis; *p*-value for equality of the mean numbers of signals per image frame <0.000001)

Sample	Est. number of signals
Saliva	6.6 (4.9–8.8)
Saliva + *M. smegmatis*	82.5 (66.3–102.1)

signals observed for the saliva sample without added *M. smegmatis*. This result supported the feasibility of using the RCA based sensor system for the detection of mycobacteria in human saliva.

The detected concentration of 650 CFU *M. smegmatis* per μL saliva was consistent with the detection limit expected from the results shown in Fig. 4. Sputum from TB patients typically contains 0.5–1 million CFU mL^{-1}.[60–64] The observed detection of approximately 0.6 million CFU mL^{-1} under the utilized assay conditions therefore holds promise for the future development of a novel TB diagnostic test based on the RCA based sensor system specific for mycobacteria TOP1A activity.

Conclusion

TB is among the major threats to global health with one third of the world population being infected and more than 10 million new cases being reported each year.[1] Reliable diagnosis remains a problem especially for low resource countries where TB is most prevalent. With the aim of taking the first steps in the direction of the development of a new diagnostic test for TB we here present a novel type of sensor system specific for mycobacteria in crude samples. This is achieved by measuring the activity of the mycobacterial expressed enzyme TOP1A at the single catalytic event level in an assay setup that relies only on isothermal amplification steps. The current signal readout was performed using a fluorescence microscope to detect RCPs hybridized with fluorescent labelled probes. However, we and others previously demonstrated detection of RCPs using a simple colorimetric readout procedure that could be performed even at limited resource settings.[21,65] The exploitation of such a readout procedure may enable the future development of a point-of-care TB test based on the here presented sensor system. The choice of a mycobacterially expressed biomarker predicts that the RCA based sensor system will be insensitive towards previous infections and/or co-infections with immunosuppressing microbes such as HIV. This is in contrast to current point-of-care tests such as *e.g.* the tubulin skin test, which measures the human immune response to TB infections.[9,10]

We demonstrated that the mycobacteria RCA based sensor system was specific towards mycobacterial TOP1A in a background of *E. coli* or human cell extracts and that the setup detected mycobacteria exclusively when tested against an array of non-mycobacteria often found in human specimen.

Moreover, we were able to quantify and detect the model mycobacteria *M. smegmatis* in small samples (25 μL cell suspension) of *e.g.* human saliva when combining the sensor system with lysis caused by mycobacteriophage infection. The lowest concentration of mycobacteria detected by the RCA based system was approximately 0.6 million CFU mL^{-1}, which is within the range of concentrations relevant for clinical samples.[60,64] Hence, we believe that the presented DNA sensor system holds promise for the development of a potential new diagnostic test for TB that may even be suitable for low resource settings. Note that although the presented assay setup, that includes multiple steps, at first glance may appear complicated it requires limited sample preparation, can be performed under isothermal conditions and is adaptable to a simple colorimetric readout. A future TB test based on this technology may be envisioned as a semi or fully automated microfluidics setup that does not rely on electricity supply.

Authors contribution

BRK conceived the project and wrote the major part of the manuscript. OF contributed to the experiments shown in Fig. 4 and ESI Fig. 3,† writing of the manuscript, data analyses and preparation of figures. HX contributed significantly to the experiments shown in Fig. 4, ESI Fig. 3† and to data analyses. LBM contributed to most of the experiments behind Fig. 2, 3 and ESI Fig. 2.† AG contributed to the experiments shown in Fig. 2. MBA contributed to Fig. 4 and ESI Fig. 3.† AAG contributed to the preparation of mycobacterial TOP1A and to ESI Fig. 1.† CH contributed to optimizations and characterization of mycobacterial TOP1A. NNL and CW contributed to the planning of the experiments shown in Fig. 3 and the preparation of the manuscript. JT and FSP contributed to purification and preparation of mycobacteriophages. YW and DS contributed to the planning of the experiments and preparation of the manuscript. LP contributed to the experiments shown in Fig. 3. VN, YPH and MS contributed to substrate design and writing of the manuscript. MSH assisted with planning of the experiments and contributed significantly to the preparation of the manuscript. RL and JB did all the statistical analyses. Moreover, RL contributed to the preparation of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful to Noriko Hansen for skillful technical assistance and to Professors Graham F. Hatfull and Carlos A. Guerrero Department of Biological Sciences, University of Pittsburgh, Pittsburgh, USA, for kindly providing the *M. smegmatis* and mycobacteriophages used in this study. We also thank Associate Professor Pia Møller Martensen, Department of Molecular Biology and Genetics, Aarhus University for kindly providing the HEK293 T cell line used in this study. BRK would like to acknowledge the support from the Novo Nordisk Foundation (grant 7653), Familien Erichsens mindefond, Aase & Ejnar Danielsen's foundation, Aage & Johanne Louis-Hansens Foundation, Marie & M.B. Richters Foundation, Minister Erna Hamiltons Legat for Videnskab og Kunst, Civilingeniør Frode V. Nyegaard og hustrus Fond, Karen Elise Jensens Foundation, Arvid Nilssons Foundation, Ludvid og Franciska Andersens Legat. YPH would like to acknowledge the support from the Start-Up fund provided by the Chinese University of Hong Kong.

References

1 G. G. T. r. World Health Organisation, WHO/HTM/TB/2017.23, http://www.who.int/tb/publications/factsheet_global.pdf?ua=1.
2 C. Wejse, *Int. J. Infect. Dis.*, 2015, **32**, 152–155.
3 C. K. Kim, E. A. Cho, D. M. Shin, S. W. Choi and S. Y. Shin, *Ann. Lab. Med.*, 2018, **38**, 119–124.
4 S. Chakravorty, A. M. Simmons, M. Rowneki, H. Parmar, Y. Cao, J. Ryan, P. P. Banada, S. Deshpande, S. Shenai, A. Gall, J. Glass, B. Krieswirth, S. G. Schumacher, P. Nabeta, N. Tukvadze, C. Rodrigues, A. Skrahina, E. Tagliani, D. M. Cirillo, A. Davidow, C. M. Denkinger, D. Persing, R. Kwiatkowski, M. Jones and D. Alland, *mBio*, 2017, **8**, pii: e00812-17.
5 K. R. Steingart, I. Schiller, D. J. Horne, M. Pai, C. C. Boehme and N. Dendukuri, *Cochrane Database Syst. Rev.*, 2014, CD009593, DOI: 10.1002/14651858.CD009593.pub3.
6 L. Yan, Q. Zhang and H. Xiao, *BMC Infect. Dis.*, 2017, **17**, 545.
7 D. M. Cirillo, A. M. Cabibbe, M. R. De Filippo, A. Trovato, T. Simonetti, G. M. Rossolini and E. Tortoli, *Int. J. Mycobacteriol.*, 2016, **5**(Suppl 1), S252–S253.
8 M. Raviglione and G. Sulis, *Infect. Dis. Rep.*, 2016, **8**, 6570.
9 S. Nayak and B. Acharjya, *Indian Dermatol. Online J.*, 2012, **3**, 2–6.
10 N. S. Ali, K. Jamal and A. K. Khuwaja, *Asia Pac. Family Med.*, 2010, **9**, 8.
11 D. M. Lewinsohn, M. K. Leonard, P. A. LoBue, D. L. Cohn, C. L. Daley, E. Desmond, J. Keane, D. A. Lewinsohn, A. M. Loeffler, G. H. Mazurek, R. J. O'Brien, M. Pai, L. Richeldi, M. Salfinger, T. M. Shinnick, T. R. Sterling, D. M. Warshauer and G. L. Woods, *Clin. Infect. Dis.*, 2017, **64**, 111–115.
12 D. Wu, J. Kang, B. Li and D. Sun, *J. Clin. Lab. Anal.*, 2017, DOI: 10.1002/jcla.22326.
13 M. F. Lee, Y. H. Chen and C. F. Peng, *J. Microbiol. Methods*, 2009, **76**, 174–180.
14 S. Mitarai, M. Okumura, E. Toyota, T. Yoshiyama, A. Aono, A. Sejimo, Y. Azuma, K. Sugahara, T. Nagasawa,

N. Nagayama, A. Yamane, R. Yano, H. Kokuto, K. Morimoto, M. Ueyama, M. Kubota, R. Yi, H. Ogata, S. Kudoh and T. Mori, *Int. J. Tuberc. Lung Dis.*, 2011, **15**, 1211–1217, i.

15 C. Y. Lee, H. Jang, K. S. Park and H. G. Park, *Nanoscale*, 2017, **9**, 16149–16153.

16 H. H. Deng, X. Q. Shi, H. P. Peng, Q. Q. Zhuang, Y. Yang, A. L. Liu, X. H. Xia and W. Chen, *ACS Appl. Mater. Interfaces*, 2018, **10**, 5358–5364.

17 J. Zajda, L. Gorski and E. Malinowska, *Bioelectrochemistry*, 2016, **109**, 63–69.

18 L. Zuccaro, C. Tesauro, T. Kurkina, P. Fiorani, H. K. Yu, B. R. Knudsen, K. Kern, A. Desideri and K. Balasubramanian, *ACS Nano*, 2015, **9**, 11166–11176.

19 J. Huang, X. Y. Li, Y. C. Du, L. N. Zhang, K. K. Liu, L. N. Zhu and D. M. Kong, *Biosens. Bioelectron.*, 2017, **91**, 417–423.

20 S. Juul, C. J. Nielsen, R. Labouriau, A. Roy, C. Tesauro, P. W. Jensen, C. Harmsen, E. L. Kristoffersen, Y. L. Chiu, R. Frohlich, P. Fiorani, J. Cox-Singh, D. Tordrup, J. Koch, A. L. Bienvenu, A. Desideri, S. Picot, E. Petersen, K. W. Leong, Y. P. Ho, M. Stougaard and B. R. Knudsen, *ACS Nano*, 2012, **6**, 10676–10683.

21 M. S. Hede, S. Fjelstrup, F. Lötsch, R. M. Zoleko, A. Klicpera, M. Groger, J. Mischlinger, L. Endame, L. Veletzky, R. Neher, A. K. W. Simonsen, E. Petersen, G. Mombo-Ngoma, M. Stougaard, Y.-P. Ho, R. Labouriau, M. Ramharter and B. R. Knudsen, *Sci. Rep.*, 2018, **8**, 4122.

22 A. Givskov, E. L. Kristoffersen, K. Vandso, Y. P. Ho, M. Stougaard and B. R. Knudsen, *Sensors*, 2016, **16**, pii: E1916.

23 J. Wang, J. Liu, J. Thomsen, D. Selnihhin, M. S. Hede, F. C. Kirsebom, O. Franch, S. Fjelstrup, M. Stougaard, Y. P. Ho, F. S. Pedersen and B. R. Knudsen, *Nanoscale*, 2017, **9**, 440–448.

24 E. L. Kristoffersen, A. Givskov, L. A. Jorgensen, P. W. Jensen, W. B. JA, N. Osheroff, A. H. Andersen, M. Stougaard, Y. P. Ho and B. R. Knudsen, *Nucleic Acids Res.*, 2017, **45**, 7855–7869.

25 M. Stougaard, J. S. Lohmann, A. Mancino, S. Celik, F. F. Andersen, J. Koch and B. R. Knudsen, *ACS Nano*, 2009, **3**, 223–233.

26 F. F. Andersen, M. Stougaard, H. L. Jorgensen, S. Bendsen, S. Juul, K. Hald, A. H. Andersen, J. Koch and B. R. Knudsen, *ACS Nano*, 2009, **3**, 4043–4054.

27 S. Juul, J. M. Obliosca, C. Liu, Y. L. Liu, Y. A. Chen, D. M. Imphean, B. R. Knudsen, Y. P. Ho, K. W. Leong and H. C. Yeh, *Nanoscale*, 2015, **7**, 8332–8337.

28 A. C. Schurch, S. Arredondo-Alonso, R. J. L. Willems and R. V. Goering, *Clin. Microbiol. Infect.*, 2018, **24**, 350–354.

29 A. J. Sabat, E. van Zanten, V. Akkerboom, G. Wisselink, K. van Slochteren, R. F. de Boer, R. Hendrix, A. W. Friedrich, J. W. A. Rossen and A. Kooistra-Smid, *Sci. Rep.*, 2017, **7**, 3434.

30 M. Janczarek and M. Palusinska-Szysz, *J. Appl. Genet.*, 2016, **57**, 251–261.

31 A. A. Godbole, M. N. Leelaram, A. G. Bhat, P. Jain and V. Nagaraja, *Arch. Biochem. Biophys.*, 2012, **528**, 197–203.

32 T. Bhaduri, T. K. Bagui, D. Sikder and V. Nagaraja, *J. Biol. Chem.*, 1998, **273**, 13925–13932.

33 V. Nagaraja, D. Sikder and P. Jain, *Curr. Pharm. Des.*, 2002, **8**, 1995–2007.

34 A. G. Bhat, M. N. Leelaram, S. M. Hegde and V. Nagaraja, *J. Mol. Biol.*, 2009, **393**, 788–802.

35 D. Sikder and V. Nagaraja, *Nucleic Acids Res.*, 2000, **28**, 1830–1837.

36 M. L. Jepsen, C. Harmsen, A. A. Godbole, V. Nagaraja, B. R. Knudsen and Y. P. Ho, *Nanoscale*, 2016, **8**, 358–364.

37 P. Jain and V. Nagaraja, *J. Mol. Biol.*, 2006, **357**, 1409–1421.

38 Z. O. Petrova, G. W. Broussard and G. F. Hatfull, *Microbiology*, 2015, **161**, 1539–1551.

39 S. Singh, K. Rockenbach, R. M. Dedrick, A. P. VanDemark and G. F. Hatfull, *J. Mol. Biol.*, 2014, **426**, 318–331.

40 M. E. Ford, G. J. Sarkis, A. E. Belanger, R. W. Hendrix and G. F. Hatfull, *J. Mol. Biol.*, 1998, **279**, 143–164.

41 D. A. Russell and G. F. Hatfull, *Bioinformatics*, 2017, **33**, 784–786.

42 B. Jørgensen, R. Labouriau and S. Lundbye-Christensen, *J. R. Stat. Soc.*, 1996, **B58**, 573–592.

43 P. Madsen, J. Jensen, R. Labouriau, O. F. Christensen and G. Sahana, *DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics. 10th World Congress on Genetics Applied to Livestock Production (WCGALP)*, Vancouver, Canada, 2014.

44 R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2016, URL https://www.R-project.org/.

45 D. Bates, M. Maechler, B. Bolker and S. Walker, *J. Stat. Software*, 2015, **67**, 1–48, DOI: 10.18637/jss.v067.i01.

46 Y. Benjamini and D. Yekutieli, *Ann. Math. Stat.*, 2001, **29**, 1165–1188.

47 J. C. Wang, *J. Biol. Chem.*, 1991, **266**, 6659–6662.

48 J. C. Wang, *J. Mol. Biol.*, 1971, **55**, 523–533.

49 F. Baquero, M. E. Alvarez and R. Canton, *Clin. Microbiol. Infect.*, 1996, **1**(Suppl 2), S10–S15.

50 F. H. Kayser, *Postgrad. Med. J.*, 1992, **68**(Suppl 3), S17–S23.

51 T. A. Claeys and R. T. Robinson, *J. Bacteriol.*, 2018, **11**, pii: e00739-17.

52 I. Rashid, L. H. Mabuza, I. Govender and D. Pretorius, *Afr. J. Primary Health Care Family Med.*, 2011, **3**, 240.

53 R. Thakur, S. Sarma and R. Goyal, *J. Global Infect. Dis.*, 2011, **3**, 353–356.

54 L. W. Anson, K. Chau, N. Sanderson, S. Hoosdally, P. Bradley, Z. Iqbal, H. Phan, D. Foster, S. Oakley, M. Morgan, T. E. A. Peto, M. Modernizing Medical Microbiology Informatics Group, D. W. Crook and L. J. Pankhurst, *J. Med. Microbiol.*, 2018, **67**, 347–357.

55 L. R. Grant, L. L. Hammitt, D. R. Murdoch, K. L. O'Brien and J. A. Scott, *Clin. Infect. Dis.*, 2012, **54**(Suppl 2), S140–S145.

56 D. Vargas, L. Garcia, R. H. Gilman, C. Evans, E. Ticona, M. Navincopa, R. F. Luo, L. Caviedes, C. Hong, R. Escombe and D. A. Moore, *Lancet*, 2005, **365**, 150–152.
57 J. Guiot, S. Demarche, M. Henket, V. Paulus, S. Graff, F. Schleich, J. L. Corhay, R. Louis and C. Moermans, *J. Visualized Exp.*, 2017, **130**, e56612.
58 A. R. Namuganga, N. N. Chegou, P. Mubiri, G. Walzl and H. Mayanja-Kizza, *BMC Infect. Dis.*, 2017, **17**, 600.
59 Centers for Disease Control and Prevention (CDC), 2016, https://www.cdc.gov/tb/topic/basics/howtbspreads.htm.
60 D. Sloan and G. Davies, *Prog. Respir. Res.*, 2011, **40**, 227–234.
61 D. J. Sloan, E. L. Corbett, A. E. Butterworth, H. C. Mwandumba, S. H. Khoo, A. Mdolo, D. Shani, M. Kamdolozi, J. Allen, D. A. Mitchison, D. J. Coleman and G. R. Davies, *J. Clin. Microbiol.*, 2012, **50**, 2315–2320.
62 J. V. Rogers and Y. W. Choi, *J. Microb. Biochem. Technol.*, 2012, **4**, 147–151.
63 R. Brindle, J. Odhiambo and D. Mitchison, *BMC Pulm. Med.*, 2001, **1**, 2.
64 G. L. Hobby, A. P. Holman, M. D. Iseman and J. M. Jones, *Antimicrob. Agents Chemother.*, 1973, **4**, 94–104.
65 J. Yan, S. Song, B. Li, Q. Zhang, Q. Huang, H. Zhang and C. Fan, *Small*, 2010, **6**, 2520–2525.