TESIS DE LA UNIVERSIDAD DE ZARAGOZA

2021 288

Irene Pérez Sánchez

Construction and characterization of live attenuated vaccines in modem lieages of Mycobacterium tuberculosis based on phoP an fadD26 deletions

Director/es
Gonzalo Asensio, Jesús Ángel
Martín Montañés, Carlos

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

CONSTRUCTION AND CHARACTERIZATION OF LIVE ATTENUATED VACCINES IN MODEM LIEAGES OF MYCOBACTERIUM TUBERCULOSIS BASED ON PHOP AN FADD26 DELETIONS

Autor

Irene Pérez Sánchez

Director/es

Gonzalo Asensio, Jesús Ángel
Martín Montañés, Carlos

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

2019

Universidad Zaragoza
1542

Construction and characterization of live attenuated vaccines in modern lineages of *Mycobacterium tuberculosis* based on *phoP* and *fadD26* deletions

Tesis Doctoral 2019

Irene Pérez Sánchez

FACULTAD DE MEDICINA

Departamento de Microbiología, Medicina Preventiva y Salud Pública

Construction and characterization of live attenuated vaccines in modern lineages of *Mycobacterium tuberculosis* based on *phoP* and *fadD26* deletions

Memoria para optar al grado de Doctor presentada por:

Irene Pérez Sánchez

Licenciada en Bioquímica

Directores:

Carlos Martín Montañés

Jesús Ángel Gonzalo Asensio

D. CARLOS MARTÍN MONTAÑÉS, Catedrático de Microbiología del Departamento de Microbiología, Medicina Preventiva y Salud Pública de la Universidad de Zaragoza.

D. JESÚS ÁNGEL GONZALO ASENSIO, Doctor en Bioquímica y Biología Molecular y Celular por la Universidad de Zaragoza. Profesor Contratado Doctor del Departamento de Microbiología, Medicina Preventiva y Salud Pública de la Universidad de Zaragoza.

Directores de la Tesis Doctoral presentada por **Irene Pérez Sánchez** bajo el título:

Construction and characterization of live attenuated vaccines in modern lineages of *Mycobacterium tuberculosis* based on *phoP* and *fadD26* deletions

(Construcción y caracterización de vacunas vivas atenuadas en linajes modernos de *Mycobacterium tuberculosis* basadas en deleciones de *phoP* y *fadD26*)

EXPONEN:

Que dicha Tesis Doctoral ha sido realizada bajo su dirección y reúne los requisitos necesarios para optar al grado de Doctor y a la mención de Doctorado Internacional.

Por lo anterior, emiten el presente **INFORME FAVORABLE**.

Zaragoza, 7 de junio de 2019

Fdo: Carlos Martín Montañés

Fdo: Jesús Ángel Gonzalo Asensio

Esta tesis doctoral ha sido elaborada en el Departamento de Microbiología, Medicina Preventiva y Salud Pública de la facultad de Medicina en la Universidad de Zaragoza, dentro del Programa de Doctorado en Bioquímica y Biología Molecular.

Irene Pérez es beneficiaria de una ayuda para la contratación de personal investigador predoctoral en formación concedida por el Departamento de Industria e Innovación, de la Diputación General de Aragón financiado por el Gobierno de Aragón y el Fondo Social Europeo.

El trabajo experimental ha sido realizado con la ayuda de los siguientes proyectos de investigación:

- "TBVAC2020; Advancing novel and promising TB vaccine candidates from discovery and early clinical development". Financiado por la Unión Europea (H2020-PHC-643381)
- "Innovando MTBVAC como vacuna contra la tuberculosis y nuevas aplicaciones terapéuticas contra el cáncer". Financiado por el Ministerio de Economía y Competitividad y el Fondo Europeo de Desarrollo Regional (BIO2014-52580-P).

Parte del trabajo se realizó en una estancia de tres meses en el Instituto Pasteur de París (Francia) cofinanciada por una beca para estancia de investigación de la Federación Europea de Sociedades de Microbiología (FEMS Research and Training Grant FEMS-GO-2018-118) y una beca del programa CAI-Ibercaja de Estancias de Investigación (CM 3/18).

Agradecer el uso del Servicio General de Apoyo a la Investigación-SAI de la Universidad de Zaragoza. Agradecer también a las instalaciones de Proteómica de los Servicios Científico Técnicos del CIBA (IACS-Universidad de Zaragoza) (las instalaciones de proteómica son miembro de ProteoRed, PRB2-ISCIII, apoyado por la subvención PT13/0001), al Insitute of Pharmacology, Hannover Medical School (Hannover, Alemania) y a la Fundación Medina (Granada, España) por su apoyo técnico en las determinaciones del c-di-AMP.

INDEX

CHAPTER 1

List of abbreviations

ADC	Albumin, Dextrose and Catalase
AES	Allelic exhange substrate
Ag85	Antigen 85
Amp	Ampicilin
BAC	Bacterial artificial chromosome
BCG	Bacillus Calmette-Guérin
bp	base pair
BSL	Biosafety laboratory
c-di-AMP	Cyclic di-AMP
c-DNA	Complementary deoxyribonucleic acid
CFP-10	10 KDa culture filtrate protein (*esxB/Rv3874)*
CFU	Colony forming unit
Cm	Cloramphenicol
CTAB	Cetyl Trimetyl Ammonium Bromide
dsDNA	Double strand DNA
EDTA	Ethylene Diamine Tetr-Acetic
ELISA	Enzyme-linked immunnosorbent assay
ESAT-6	6 KDa early secretory antigenic target (*esxA/Rv3875)*
fw	Forward
Gm	Gentamicin
GMP	Good Manufacturing Practices
GroEL-2	60 KDa chaperonin 2 (Rv0440)
HCl	Hydrochloric acid
HCU	Hospital Clínico Universitario Lozano Blesa
HIV	Human Immunodeficiency virus
HMS	Hospital Miguel Servet
Hyg	Hygromycin
i.n.	Intranasal
i.p.	Intraperitoneal
IFN-γ	Interferon gamma
IGRA	Interferon gamma release assay
IL-1β	Interleukin-1 beta
IS*6110*	Insertion sequence 6110
kb	kilo base
kDa	kilo dalton
Km	Kamamycin
L2-L4	Lineage 2 to 4 of *M. tuberculosis*
LB	Luria Bertani
LTBI	Latent tuberculosis infection
MDR	Multi-drug resistant
MTBC	*Mycobacterium tuberculosis* Complex
MTBVAC	*Mycobacterium tuberculosis* vaccine (lineage 4)

MTBVAC-L2	*Mycobacterium tuberculosis* vaccine lineage 2
MTBVAC-L3	*Mycobacterium tuberculosis* vaccine lineage 3
NaAc	Sodium acetate
NaCl	Sodium chloride
NaF	Sodium fluoride
NHP	Non-human primate
OADC	Oleic, albumin, dextrose and catalase
OD	Optical density
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PDIM	Phthiocerol dimycocerosates
PIM	Phosphatidylinositol mannosides
PGL	Phenol glycolipid
PKs	Polyketide synthase
RD	Region of difference
RNA	Ribonucleic acid
rv	Reverse
s.c.	Subcutaneous
SDS	Sodium Dodecyl Sulfate
Suc	Sucrose
TAT	Twin Arginine TranslocaseSecretion System
TB	Tuberculosis
TE	Tris-EDTA
WHO	World Health Organization
WT	Wild type
XDR	Extensively drug resistant

Summary

Tuberculosis (TB) is the most devastating disease caused by a single infectious agent. *Mycobacterium tuberculosis,* the main causative agent of TB in humans, is transmitted from human to human by aerosol route. The estimation is that in 2017 TB caused 1.3 million deaths in HIV-negative people aside from 300, 000 deaths in HIV positive people and 10 million of new TB cases.

Even though that there is a preventive vaccine, BCG, with a wide coverage for almost 90% and there are antibiotics available for TB treatment, the increase of drug-resistant strains and the variable protective efficacy of BCG against pulmonary TB make this disease the first cause of death due to an infectious agent today.

To overcome this problem, several vaccine candidates are being developed to obtain a more efficacious vaccine. One of them is MTBVAC, a live attenuated vaccine based on deletion of *phoP* and *fadD26* genes in a clinical isolate of lineage 4 of *M. tuberculosis*. Both genes are key virulence factors of *M. tuberculosis.*

Mycobacteria adapted to infect humans show specific geographical distribution. They can be classified in seven lineages that can be grouped in "modern" or "ancient" lineages based on *M. tuberculosis* specific deletion (TbD1). Modern lineages are lineages 2, 3 and 4 of *M. tuberculosis* whereas ancient lineages are lineages 1 and 7 of *M. tuberculosis* and lineages 5 and 6 of *Mycobacterium africanum.* Lineages 2 and 4 of *M. tuberculosis* are the most globally widespread lineages, followed by lineages 1 and 3 of *M. tuberculosis* that present an intermediate distribution and the others are restricted to specific areas.

Considering that MTBVAC was obtained in a clinical isolate of lineage 4 of *M. tuberculosis,* the same deletions as MTBVAC in *phoP* and *fadD26* genes were obtained in two clinical isolates of *M. tuberculosis* lineages 2 and 3, named MTBVAC-L2 and MTBVAC-L3, using two different genetic engineering strategies. Consequently, these three vaccine candidates based on *phoP* and *fadD26* deletions constructed in the modern lineages of *M. tuberculosis* allow the evaluation of lineage dependent protection. Western-blot analysis and qRT-PCR of the new double mutants were performed to confirm the PhoP and FadD26-deficient phenotypes previously described in MTBVAC.

Preclinical characterization of safety and protective efficacy were performed in the mouse model. In safety studies, both MTBVAC-L2 and MTBVAC-L3 were attenuated, less attenuated than MTBVAC, being only MTBVAC more attenuated than BCG Pasteur in SCID mice.

For protective efficacy evaluation, mice were vaccinated with MTBVAC, MTBVAC-L2, MTBVAC-L3, BCG Pasteur or unvaccinated as control. Bacterial burden in lungs and spleen was evaluated in protective experiments after the challenge with a virulent strain from lineage 2, 3 or 4 of *M. tuberculosis.* MTBVAC, MTBVAC-L2 and MTBVAC-L3 conferred protection against the modern strains used in the challenge without significant differences among them.

Remarkably, MTBVAC and MTBVAC-L2 conferred superior protection against lineage 2-Beijing strain in comparison with BCG Pasteur.

These results suggested lineage-independent protection in mice, and reinforce previous knowledge about MTBVAC protective efficacy in several animal models.

To further characterize MTBVAC, after genomics, transcriptomics and proteomics characterization, we focused on the study of metabolites differentially produced by MTBVAC. Production of cyclic-di-AMP (c-di-AMP) was PhoPR dependent. A *phoPR* mutant strain in H37Rv exhibited higher quantity of c-di-AMP than the wild type (WT) strain. Secretion of the metabolite was only observed in *phoPR* mutant. Increased production and secretion of c-di-AMP was also observed in MTBVAC compared to the WT strain.

c-di-AMP is a second messenger described to be involved in different processes in bacteria. In *M. tuberculosis,* bacterial c-di-AMP has been recently described to trigger type I interferon response (IFN-β) by activation of eukaryotic sensor STING. To evaluate whether c-di-AMP from MTBVAC activates this response in the host, deletions in genes that code for the c-di-AMP cyclase and the phosphodiesterase were obtained in MTBVAC. Therefore, three different MTBVAC with different levels of c-di-AMP were obtained. Despite increased bacterial c-di-AMP in MTBVAC, no IFN-β response was observed after infection of THP-1 cells. By contrast, MTBVAC and mutants in c-di-AMP cyclase and phosphodiesterase exhibited IL-1β response.

Resumen

Tuberculosis es la enfermedad más devastadora producida por un único agente infeccioso. *Mycobacterium tuberculosis,* el principal agente causal de la tuberculosis en humanos, se transmite entre humanos por el aire. Se estima que en 2017, 1.3 millones de personas (HIV-negativas) murieron a causa de la tuberculosis además de 300, 000 muertes en personas con HIV y se estima que hubo 10 millones de nuevos casos de tuberculosis.

A pesar de la existencia de una vacuna preventiva, BCG, con una amplia cobertura de casi el 90% y la disponibilidad de antibióticos para el tratamiento de la tuberculosis, el aumento de cepas resistentes a los antibióticos y la variable eficacia de protección de BCG frente a la tuberculosis pulmonar hace que esta enfermedad sea la primera causa de muerte debido a un agente infeccioso en la actualidad.

Para superar este problema, varios nuevos candidatos a vacuna están siendo desarrollados para obtener una vacuna más eficaz. Una de ellas es MTBVAC, una vacuna viva atenuada basada en deleciones en los genes *phoP* y *fadD26* en un aislado clínico perteneciente al linaje 4 de *M. tuberculosis*. Ambos genes son importantes factores de virulencia de *M. tuberculosis.*

Las micobacterias adaptadas a infectar a humanos presentan una distribución geográfica específica. Son clasificados en siete linajes que pueden diferenciarse entre "modernos" o "ancestrales" basándonos en la deleción TbD1. Los linajes 2, 3 y 4 de *M. tuberculosis* son los linajes modernos mientras que los linajes 1 y 7 de *M. tuberculosis* y los linajes 5 y 6 de *Mycobacterium africanum* son los linajes denominados ancestrales. Los linajes más ampliamente distribuidos son los linajes 2 y 4 de *M. tuberculosis* seguido de los linajes 1 y 3 de *M. tuberculosis* con una distribución intermedia y los otros linajes están presentes en determinadas zonas.

Teniendo en cuenta que MTBVAC se obtuvo en un aislado clínico del linaje 4 de *M. tuberculosis,* las mismas deleciones que en MTBVAC en los genes *phoP* y *fadD26* fueron obtenidas en dos aislados clínicos de *M. tuberculosis* de los linajes 2 y 3, denominadas MTBVAC-L2 y MTBVAC-L3. Para obtener las deleciones se utilizaron dos estrategias diferentes de ingeniería genética. Por lo tanto, estas tres vacunas candidatas basadas en las deleciones de *phoP* y *fadD26* obtenidas en los linajes modernos de *M. tuberculosis,* permiten la evaluación de la protección dependiente de linaje. Se realizaron análisis de Western-blot y qRT-PCR de los nuevos mutantes dobles para confirmar los fenotipos dependientes de PhoP y FadD26 que han sido descritos previamente en MTBVAC.

Se realizó la caracterización preclínica de experimentos de seguridad y eficacia de protección en el modelo de ratón. En experimentos de seguridad, tanto MTBVAC-L2 como MTBVAC-L3 mostraron atenuación, menor que MTBVAC, siendo solo MTBVAC más atenuada que BCG Pasteur en ratones SCID.

Para la evaluación de la eficacia en la protección, grupos de ratones fueron vacunados con MTBVAC, MTBVAC-L2, MTBVAC-L3, BCG Pasteur o no vacunados como control. Se evaluó la carga bacteriana en pulmones y bazo en estudios de protección frente a cepas virulentas de los

linajes 2, 3 o 4 de *M. tuberculosis.* MTBVAC, MTBVAC-L2 and MTBVAC-L3 confirieron protección frente a las cepas modernas utilizadas para el desafío sin diferencias significativas entre ellas. Cabe destacar que las cepas MTBVAC y MTBVAC-L2 confirieron una mayor protegieron frente a la cepa del linaje 2-Beijijng en comparación con BCG Pasteur.

Estos resultados sugirieron protección independiente de linaje en ratones y refuerza el conocimiento previo sobre la eficacia de protección de MTBVAC en varios modelos animales.

Con objeto de profundizar con la caracterización de MTBVAC, tras estudio de genómica, transcriptómica y proteómica, nos enfocamos en el estudio de metabolitos producidos diferencialmente por MTBVAC. Se observó que la producción del metabolito di-AMP-cíclico (c-di-AMP) era dependiente de PhoPR. El mutante *phoPR* en H37Rv producía mayor cantidad de c-di-AMP que la cepa silvestre. La secreción del metabolito se observó solo en el mutante *phoPR*. También se observó una mayor producción y secreción de c-di-AMP en MTBVAC en comparación con la cepa silvestre.

El c-di-AMP es un segundo mensajero que está descrito su involucración en diferentes procesos en bacterias. En *M. tuberculosis,* el c-di-AMP bacteriano es capaz de desencadenar la respuesta de interferón de tipo I (IFN-β) mediante activación del sensor eucariótico STING. Para evaluar si el c-di-AMP de MTBVAC activa esta respuesta en el hospedador, se obtuvieron deleciones en MTBVAC en los genes que codifican para la ciclasa y la fosfodiesterasa del c-di-AMP. Por lo tanto, se obtuvieron tres cepas diferentes de MTBVAC con diferentes niveles de c-di-AMP. A pesar de los aumentados niveles de c-di-AMP bacteriano en MTBVAC, no se observó respuesta de IFN-β tras la infección en células THP-1. En cambio, MTBVAC y los mutantes en la ciclasa y fosfodiesterasa del c-di-AMP mostraron respuesta IL-1β.

INTRODUCTION

Tuberculosis, a disease still present nowadays

History of tuberculosis

Tuberculosis (TB) is an ancient disease whose history has been very close to human mankind. Globally, TB is the most devastating disease from single infectious diseases (1) and is still the major cause of mortality caused by an infectious agent (2) (Figure 1).

TB is transmitted by air and caused by organisms from *Mycobacterium tuberculosis* Complex (MTBC). In humans, the disease is mainly caused by *M. tuberculosis* and *Mycobacterium africanum* although it has also been reported some human TB cases caused by *Mycobacterium bovis.* The majority of the cases are pulmonary TB, although the bacteria might also cause disease in other parts of the body, named extrapulmonary disease (3-6).

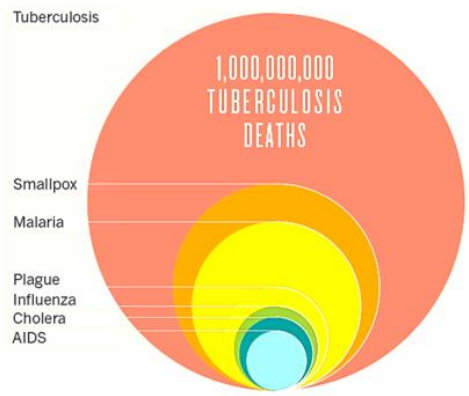


Figure 1. Estimation of deaths caused by TB and other diseases in the last 200 years. TB disease has caused more deaths than any other infectious disease. Adapted from (1).

It is considered that an early progenitor was present in East Africa 3 million years ago and coevolved with early hominids at that time (7). Presence of TB was reported in Egypt, India and China 5,000, 3,300 and 2,300 years ago respectively. In Egyptian and Peruvian mummies typical skeletal abnormalities of bony TB, including Pott´s deformities, have been found and *M. tuberculosis* DNA have been recovered from mummified tissues (8). In 1000-year old Peruvian skeletons were found mycobacterial genomes most closely related to the seal pathogen *Mycobacterium pinnipedii* (9).

Around 400 BC Hippocrates used the word "phthisis", from the Greek *phthiein* and means decaying, to define the wasting away TB patient´s body because of the pulmonary disease (10, 11). In the 17th century, the disease caused an epidemic in Europe and it was known as "Great White Plague" because of the paleness as a consequence of the disease. In 1679, Sylvius de la Boë described the characteristic lesions in the lungs named "tubercles". This term comes from the Latin word *tuberculum,* meaning protuberance. In the 19th century, the disease is known as "consumption" coming from the Latin word *consumere,* meaning to eat up or devour. This came from the idea that the body was wasted away because of the passions and beautify the paleness of TB patients. In 1834 Johann Lukas Schönlein coined the term of "tuberculosis" from the tubercles present in people suffering TB (10).

In 1865 the French military surgeon Jean-Antoine Villemin stablished the nature of TB as an infectious disease. He demonstrated that the disease could be transmitted from humans or cattle to rabbits or guinea pigs and he postulated that the disease is caused by a specific microorganism (10).

On 24th March 1882, Robert Koch made the presentation of "*Die Aetiologie der Tuberculose"* in which he identified the tubercle bacillus as the causative agent of TB in humans and announced his famous postulates: identification of the pathogen in an infected tissue, isolation of the pathogen in a pure culture *in vitro* and inoculation to a healthy organism to develop the original disease and recovery the original microorganism from lesions of the diseased organism (12) (Figure 2).

From sanatorium to vaccine and antibiotics

Around mid-19th century the sanatoriums began as places to cure TB. It was described for the first time in 1854 by Hermann Brehmer in his doctoral dissertation *"Tuberculosis is a curable disease"* where explained his healing from TB after travelling to the Himalayan Mountains. He opened the first sanatorium for TB patients and proposed that continuous fresh air, a rich diet, rest and carefully supervised exercise had successful results. The subsequent sanatoriums were built in other places with the same purpose (13, 14).

One of the major events to directly combat TB disease was given by Albert Calmette and Camille Guérin, who obtained the bacilli Calmette-Guérin (BCG) after 230 passages of a strain of *M. bovis,* the causative agent of TB in cattle. Calmette and Guerin demonstrated attenuation of BCG and protection against TB in laboratory animals. BCG immunization was carried out for the first time in 1921 and is still the current vaccine against TB (Figure 2) (10, 14).

Another event that dramatically changed the TB landscape occurred in the middle of 20th century, when antibiotics with activity against *M. tuberculosis* were discovered. In 1944, streptomycin was described by S. Waksman as the first antibiotic and bactericidal agent effective against *M. tuberculosis.* During the next decades, new antituberculous drugs were discovered, *p*-aminosalicylic acid (1949), isoniazid (1952), pyrazinamide (1954), cycloserine (1955), ethambutol (1962) and rifampicin (1963) (10). Nowadays, TB treatment is based in the combination of different antibiotics during six months. The first two months (intensive phase of the treatment), the therapy consists of the combination of rifampicin, isoniazid, pyrazinamide and ethambutol and the next four months (continuation phase) treatment is based on combination of isoniazid and rifampicin (5).

The combination of these two big advances led to a notably decrease in TB incidence. Unfortunately, the appearance of drug resistant strains, the irruption of human immunodeficiency virus (HIV), variable protection of BCG against pulmonary TB (from 20 to 80%) and drop of TB programs in industrialized countries because of the thought that the disease was close to eradication have hampered to stop TB disease and allowed its re-emergence. This re-emergence of the disease led to TB to be declared as a global health emergency in 1993 by the World Health Organization (WHO) (Figure 2) (15).

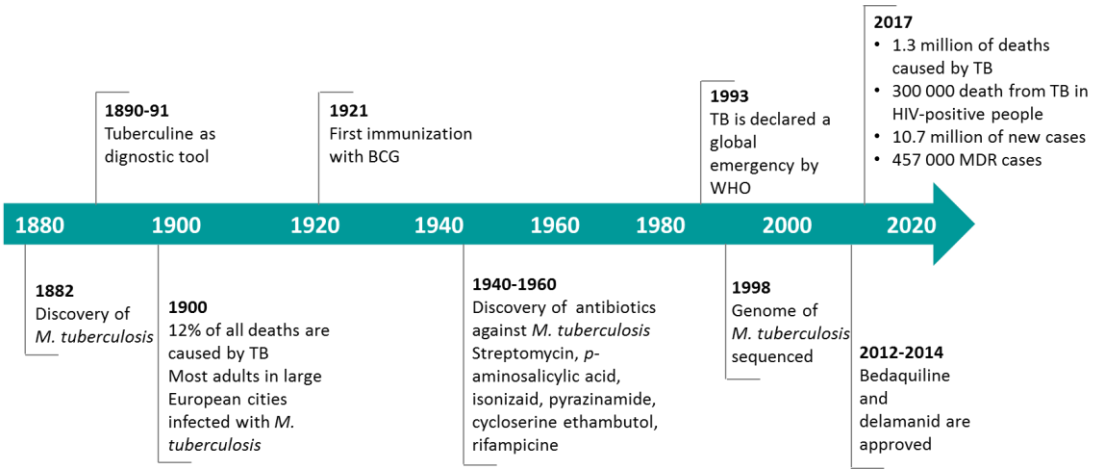


Figure 2. Timeline of *M. tuberculosis* history. Main events that have occurred in the history of TB from the discovery of the causative agent of the disease in humans in 1882 by Robert Koch until now. Adapted from (12).

Antituberculous drug resistant strains gradually appeared. RR-TB is defined for strains resistant to rifampicin. Multidrug-resistant TB (MDR-TB) is referred to *M. tuberculosis* strains resistant to isoniazid and rifampicin. It is estimated that 3.5% of new TB cases and 18% of previously treated cases are RR/MDR-TB, with wide variability depending on the area. The extensively drug-resistant strains (XDR-TB) are defined as MDR-TB and additionally resistant to one drug of the most important medicines for MDR-TB treatment: fluoroquinolone and second-line injectable agents, amikacin, capreomycin or kanamycin (16).

Evaluation of TB- and MDR-TB incidence rates in different areas exhibited that Europe is the area with almost the lowest TB-incidence rate together with America. However, incidence rate of MDR-TB in Europe is its highest because of the high incidence in Russian Federation (Figure 3A) (2).

For the last decades a big effort is focused on research to develop new vaccines and new drugs to decrease the high number of TB cases. In 2017, it is estimated that TB caused 1.3 million deaths in HIV-negative people in addition to 300,000 deaths in HIV-positive people (Figure 3B). It is estimated that during 2017, 10 million people developed TB and two thirds of the cases occurred in eight countries: India (27%), China (9%), Indonesia (8%), the Philippines (6%), Pakistan (5%), Nigeria (4%), Bangladesh (4%) and South Africa (3%). Drug resistance continues to be a public health problem, 558,000 of people who developed TB was resistant to rifampicin and the 82% of them were MDR-TB. Almost the 50% of the MDR-TB cases occurred in three countries; India (24%), China (13%) and the Russian Federation (10%) (2).

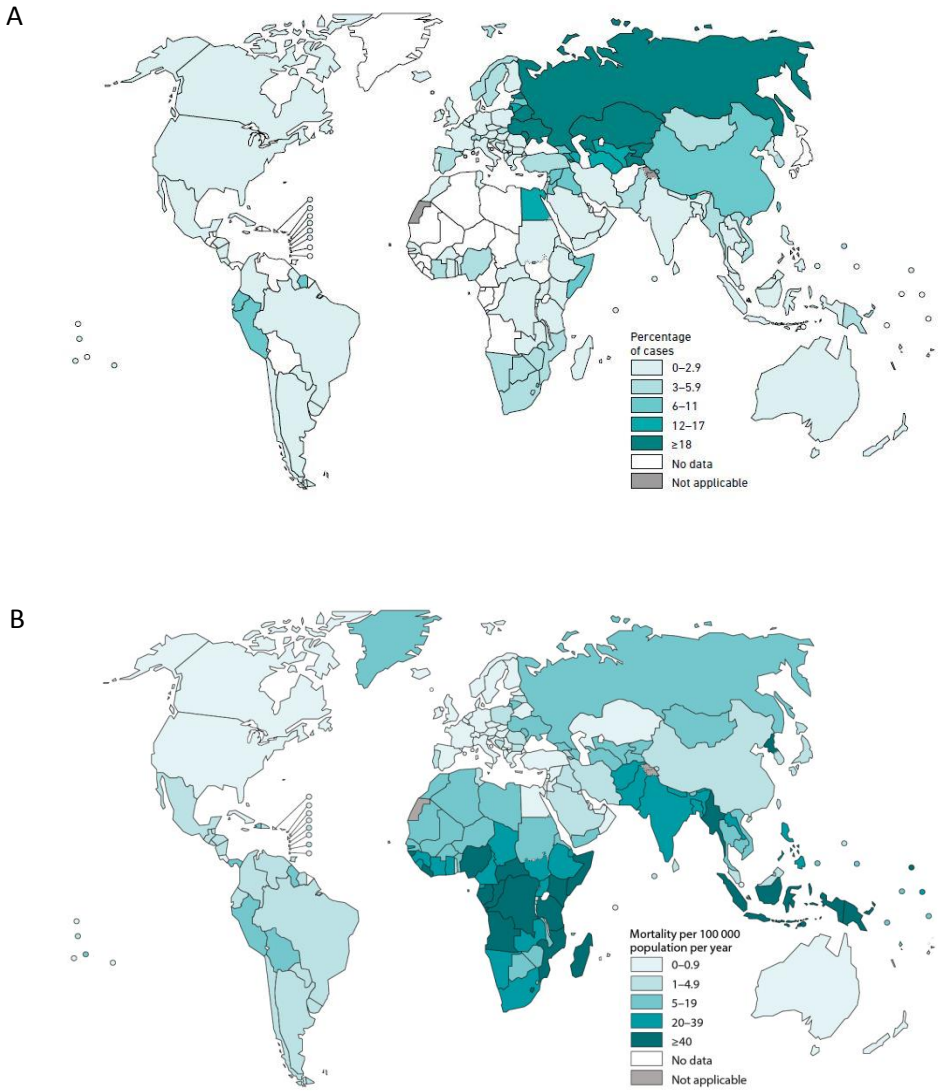


Figure 3. Estimated percentage of new TB cases with RR/MDR-TB and TB mortality rate. (A) Estimation of percentages of new RR/MDR-TB cases in 2017. (B) Estimated TB mortality rates from TB excluding TB deaths with HIV-positive people in 2017. From (2).

Because of these alarming data, the World Health Organization described a Global Plan to End TB to diminish the global TB epidemic. The main objectives are the reduction of 75% of TB deaths by 2025 and 50% of reduction of TB incidence (<55/100,000) compared with data from 2015. To achieve these goals it is required to raise the current 1.5% annual decline of global TB incidence to 10% per year (2) (17).

Therefore, new diagnostics, effective TB vaccines and drugs against *M. tuberculosis* are required. Unfortunately, there is no new vaccine licensed aside BCG to combat TB and only two new drugs have been approved since the gold era of antibiotics. These new drugs against *M. tuberculosis* are bedaquiline and delamanid approved in 2012 and 2014 respectively (18). Nevertheless, good results of clinical trials from diverse vaccine candidates and different drugs are expected to improve the global scenario of TB.

Biology of *M tuberculosis*, pathogenesis of the disease and animal model in TB research

The tubercle bacillus

The *Mycobacterium* genus belongs to the family *Mycobacteriaceae*, included into the suborder *Corynebacterineae,* order *Actinomycetales* and phylum *Actinobacteria* (19).

The *Mycobacterium* genus comprises strict and opportunistic pathogens and nonpathogenic bacteria and can be classified in fast- or slow-growth species. All strict and most opportunistic pathogens are slow growing mycobacteria, including *M. tuberculosis, Mycobacterium ulcerans* or *Mycobacterium leprae* that cause TB, Burulli ulcer and leprosy respectively (20).

M. tuberculosis is an acid-fast bacilli, aerobe, non-motile and non-spore forming bacillus of 2-5 μm length and 0.2-0.5 μm width in size and an approximate generation time of 24 h (20). The genome of *M. tuberculosis* H37Rv, the laboratory reference strain, was sequenced for the first time in 1998 and its DNA contains 4.4 megabase (Mb), which contains 4173 genes (Mycobroser database https://mycobrowser.epfl.ch/) and a high percentage of cytosine and guanine (65.6%) (21).

One of the most remarkable characteristic of mycobacteria is the cell envelope, which differs from Gram-positive and Gram-negative bacteria and mycobacteria appear as "ghosts" in Gram staining. Despite being structurally more close to Gram-positive bacteria, mycobacteria do not fit into the Gram-positive category, because the molecules attached to the cell wall are mainly lipids instead of proteins or polysaccharides. The uncommon lipids of the cell envelope of the mycobacteria lead to its distinctive Ziehl-Neelsen staining. In this stain, mycobacteria retain the dye carbol-fuchsin after a wash with an alcoholic solution in acid medium and mycobacteria is stained in red in contrast to the rest of bacteria that are decolorized (Figure 4) (22) (19).

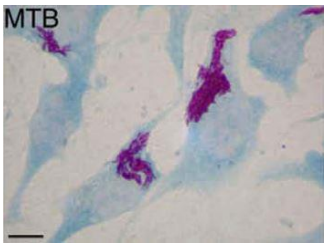


Figure 4. Ziehl-Neelsen stain. Ziehl-Neelsen staining of NRK-49F fibroblasts infected with *M.tuberculosis* (23).

The mycobacterial cell envelope

The mycobacterial envelope is unique in molecular composition and in the organization of its components. From inside to outside, the cell envelope consists on: a typical bacterial plasma membrane, periplasm space, a cell wall core that includes peptigoglycan, arabinogalactan and

mycolic acids, some specific lipids which are non-covalently associated to mycolic acids and the capsule (Figure 5) (24, 25)(26).

The detailed composition of the cell envelope components of *M. tuberculosis* are:

The innermost layer is the plasma membrane that is typical of bacterial membranes (24). The polar lipids of the mycobacterial plasma membrane are composed of hydrophilic head groups and fatty acid chains. The glycophospholipids, phosphatidylinositol mannosides (PIMs) and the related glycosylated, lipomannan (LM) and lipoarabinomannan (LAM) are non-covalently attached to the plasma membrane (27).

The cell wall of mycobacteria consists on a "cell wall core" composed of peptidoglycan, arabinogalactan and mycolic acids (26).

The peptidoglycan is composed of N-acetylglucosamine and N-acetyl/glycolylmuramic acid cross-linked by short peptides. The arabinogalactan is a heteropolysaccharide composed of arabinose and galactose residues covalently linked to the muramic acid residues of the peptidoglycan by a phosphodiester bond (28) (26). Mycolic acids are long chain (C_{60}-C_{90}) α-branched and β-hydroxylated fatty acids found esterifying the terminal pentaarabinofuranosyl units of arabinogalactan (26).

A variety of long-chain lipids and glycolipids are non-covalently associated with the mycolic acids including: trehalose monomycolates (TMM), trehalose dimycolates (TDM), phthiocerol dimycocerosate (PDIM) and closely related phenolic glycolipid (PGL) (present in some strains) (29, 30), the trehalose ester families that include sulfolipid (SL), diacyltrehalose (DAT) and polyacyltrehalose (PAT) (24). These complex lipids have been proposed to be involved in the virulence of the bacilli. This layer, together with mycolic acids, forms a highly impermeable asymmetrical bilayer, named outer membrane or "mycomembrane" and confers the natural characteristic resistance to many drugs (31) although it can also be a target for drugs (24, 32).

The outermost layer is the capsule, which is mainly composed of polysaccharides, proteins and minor amounts of lipids (24). It is thought that this layer is not present in the bacteria when bacteria are grown in presence of detergents, which is commonly used in *M. tuberculosis* laboratory growth (33).

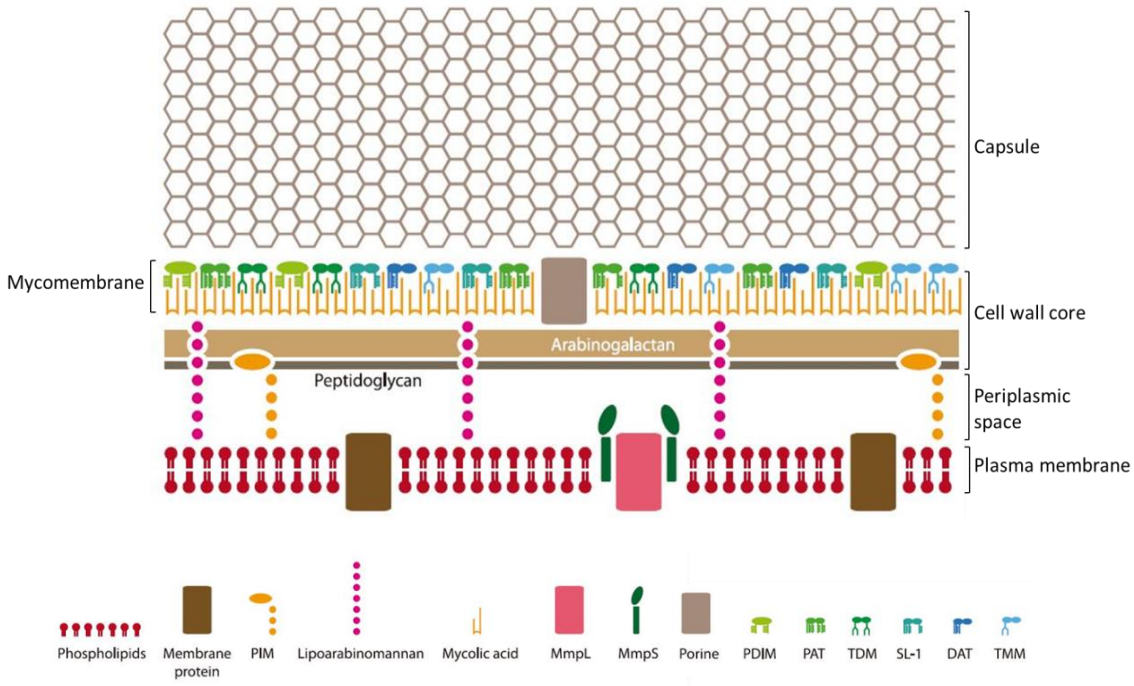


Figure 5. Schematic representation of the cell envelope of *M. tuberculosis*. From innermost to outermost layers, plama membrane, peptidoglycan, arabonigalactan, mycomembrane (composed by mycolic acids and lipids and glycolipids) and the capsule. Adapted from (34).

Pathogenesis of *M. tuberculosis*

Humans are the unique known reservoir of *M. tuberculosis*. TB is transmitted from human to human by aerosol route when small drops containing the bacilli are inhaled by an individual. In the respiratory tract, *M. tuberculosis* encounters alveolar macrophages, which are the dominant cell type that *M. tuberculosis* infects. *M. tuberculosis* is an obligate intracellular pathogen widely adapted to live inside macrophages. The majority of TB infections consist in pulmonary TB, although the pathogen can infect almost any tissue and develop extrapulmonary disease (5).

The bacilli enter into the macrophages by phagocytosis via specific receptors and remain in the phagosome. Activated macrophages promote the fusion of the phagosome with the lysosome forming the phagolysosome. In this compartment, the pathogen finds antimicrobial features including acidic pH, reactive oxygen and nitrogen intermediates or hydrolases. However, intracellular pathogens have developed strategies to overcome these defense mechanisms. *M. tuberculosis* arrests the maturation of the phagosome at early-stages and has the ability to reduce phagosomal acidification (35).

In addition, *M. tuberculosis* is able to escape to the cytosol in an ESX-1 (6-kDa early secretory antigenic target, ESAT-6, secretion system 1) dependent manner. ESAT-6 and CFP-10 are secreted through ESX-1 secretion system and ESAT-6 secretion is involved in the phagosome rupture. This phagosome rupture allows the bacteria to reach the cytosol. An ESX-1 functional system is required to the bacteria to disrupt the phagosome membrane and get access to the cytosol (36, 37). A recent study describes that PDIM and ESAT-6 act in concert to produce

membrane rupture of the phagosome (38). *M. tuberculosis* gaining access to the cytosol is a virulent factor of the bacilli and favors its dissemination. Deficient ESX-1 strains are unable to disrupt the phagosome membrane (39).

The progression of the infection depends on the state of the host immune system (Figure 6). In the 90% of individuals, the pathogen persists in a latent state, known as latent TB infection (LTBI), and the bacilli are contained in an immune microenvironment called granuloma. The granuloma is a hallmark of TB and is an organized structure that comprises macrophages that contains the bacilli, highly differentiated cells including multinucleated giant cells, epithelioid cells and Foamy cells, surrounded by many other cells such as neutrophils, dendritic cells, B and T cell, natural killer (NK) cells, fibroblast and cells that secrete extracellular matrix components. Most individuals can content the infection within the granuloma without symptoms (40, 41).

The containment could fail if the immune system of the host is altered, because of co-infection with HIV, malnutrition, old age or any condition that impairs the function of $CD4^+$ T cells. In these individuals, the granuloma tends to become necrotic, facilitating bacterial released into the lung (pulmonary TB) or spread to other parts of the body (42). In this situation, individuals develop the disease (active TB) and patients with pulmonary TB can transmit the disease by coughing. It is estimated that 23% of the world´s population (1.7 billion people) has latent TB infection (2) although approximately 10% of these infected individuals will develop active TB during their lifetime. Risk to develop symptomatic TB is considerably increased in HIV co-infected individuals and it has also been described to be related with age: infants and children under 5 years who are infected with *M. tuberculosis* are at relatively high risk because of their immature immune systems and usually the disease is manifested in disseminated forms of diseases (miliary TB and tuberculous meningitis). Children from 5-15 years are relatively resistant to TB and in adolescents risk of TB is increased. At this age, pulmonary TB is the most frequent form of the disease (43, 44).

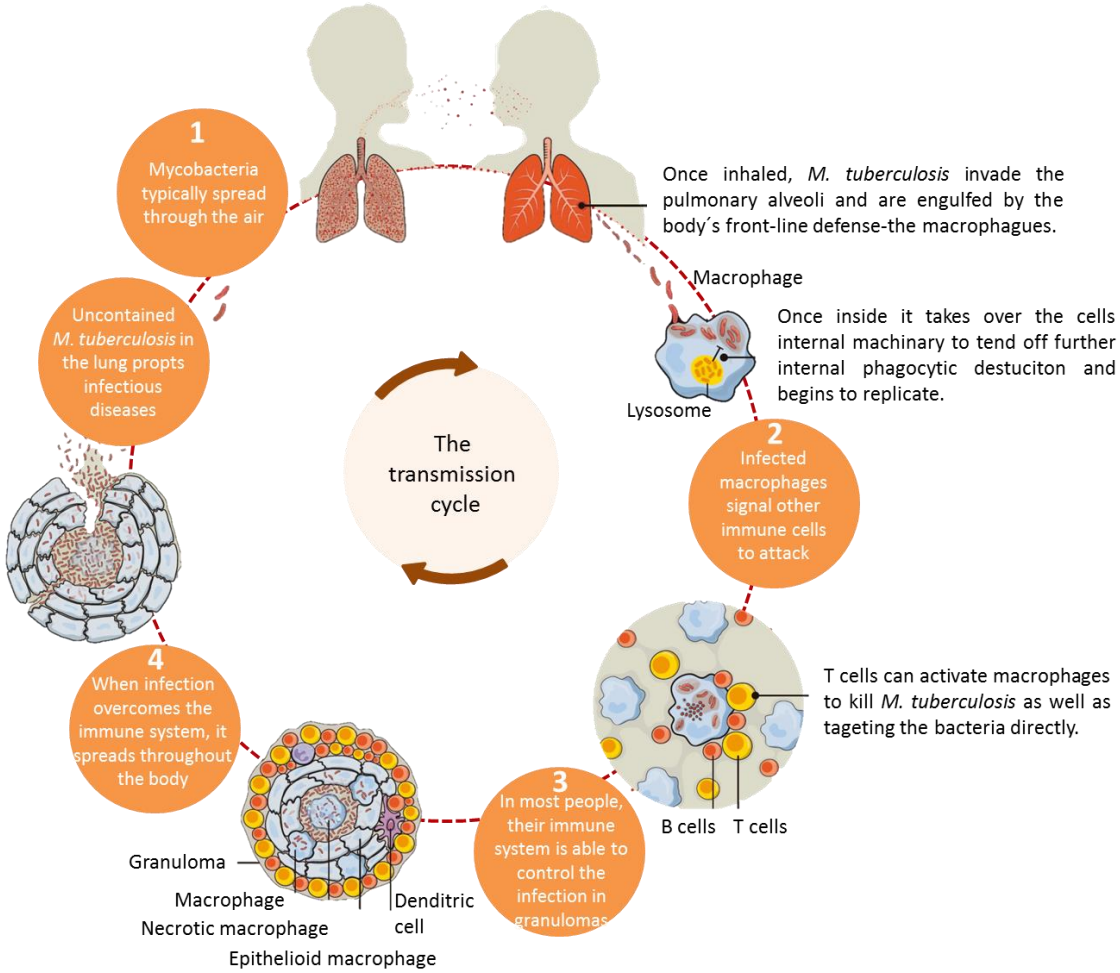


Figure 6. Cycle transmission of *M. tuberculosis*. *M. tuberculosis* is transmitted by air from people with active disease (1). In the majority of cases, the immune system can contain the bacilli in the granuloma (LTBI) (3). However, in the 10% of the cases, the granuloma breaks down and the bacilli is released to the lungs and develop the disease (4). Adapted from (45).

Animal models for TB research

Different animal models are used in TB research to understand the process of the disease process, the immune response of the host and to evaluate new vaccine candidate and new drugs. Some of the animal models used for TB research are: mice, guinea pigs, rabbits and non-human primates (NHP). Differences in susceptibility to the human pathogen and variable histopathology have been observed in the different species (Table 1).

Despite the variability of different animal models for TB, there is no a perfect one and all of them present some advantages and disadvantages, summarized in Table 1.

Table 1. Characteristics of mouse, guinea pig and non-human primate animal models in TB research. Adapted from (46-48).

Model	Lung lesion			Susceptibility to TB	Laboratory space requirements and cost	Application
	Necrosis	Caseation	Cavitation			
Mouse	No, except C3HeB/ FeJ	Usually not	No	Low	Relatively small	Evaluation of drug and vaccines.
Guinea pig	Yes	Yes	Sporadic	High	Moderate	Vaccine safety and efficacy studies. Drug evaluation.
NHP	Yes	Yes	Yes	High	Large	Drug, vaccine efficacy and treatment strategy.

Most relevant animal models for TB research are:

Mouse: is the most widely animal model used in TB research despite the low susceptibility to the human pathogen. An immune system that resembles that of humans, availability of immunology reagents, well-annotated mice genome, possibility of disruption genes in mice that can be used to decipher the role of molecules in host immune system, small size of animals, relatively low housing costs and affordable costs for biosafety level 3 (BSL3) facilities are some of the advantages of the use of this model. Different routes of TB infection (aerosol, intravenous, intranasal or intratracheal) have been successfully described in the mouse model (48-50).

Even though the immune system of mice is comparable to human immune system, lung pathology of TB differs from the histopathology observed in humans. After mice infection formation of granulomatous lesions containing macrophages, T cells and B cells are observed but the spherical shape of human granulomas is not observed. Caseous necrotic tissue is not formed because no hypoxia occurred inside the lesions and macrophages undergo apoptosis instead of necrosis (50). An exception in lung pathology was observed in C3HeB/FeJ, which was similar to humans (necrotic lesions with caseous centers) (51).

Guinea pig: from the studies of Robert Koch in which he tried to identify the causative agent of TB to the present days, guinea pig has been an important animal model for better understanding of TB. The main advantages of this model are: high susceptibility to TB in comparison with mice, similarity of human granulomas with necrosis enclosed by lymphocytes, macrophages and multinucleate giant cells and a fibrotic capsule (52) and the possibility to perform more analyses in the same animal compared to mice because of the higher size (48).

However this model is not more extensively used because of the restricted immunologic reagents available and the higher costs of this model in comparison with mice (increased space

needed for the animal and BSL3 requirements). Fortunately, the limitation of immunology reagents for this specie has been gradually changed during the last years (48, 49).

This model has been highly used to test protective efficacy of new vaccine candidates because of its high susceptibility to TB and the protection observed of BCG in this model (50). Efficacy can be evaluated by the observation of a reduction of bacterial burden in lungs and spleen from four to five weeks after challenge or by comparison of survival time over a prolonged period compared to unvaccinated animals (52).

Additionally, guinea pig model has been used to develop diagnostic skin test reagents and to novel drugs (50, 52).

NHP: are naturally susceptible to *M. tuberculosis* and develop all the states of the disease in humans, including latency and hypoxic caseous granulomas of similar structures (50). NHP are the closest to humans of all species mentioned and present high genetic similarities to humans and similar immune response (49). For all these reasons studies in NHP are usually a link between animal and humans studies although its use is limited.

The most commonly non-human primate used for TB research are rhesus macaque (*Macaca mullatta*) and cynomolgus macaques (*Macaca fascicularis*) (53), presenting considerable differenced in TB susceptibility (54), disease manifestations (55) (rhesus macaques are more susceptible to progressive TB than cynomolgus macaques) and protection of BCG (BCG protects almost completely in cynomolgus while inefficient BCG protection is observed in rhesus) (56).

Diagnostics of TB

Active TB diagnostic: four main technologies are used: imaging techniques, microscopy, culture-based techniques and molecular techniques.

Chest X-ray is a rapid technique that permits observing the lesions in the lungs. Nevertheless, the technique has poor specificity because the observation of some abnormalities consistent with pulmonary TB can be also as a consequence of other lungs pathologies. Sputum smear is one of the most widely used to detect TB disease; this technique is based on visualizing the presence of bacteria with the microscope after Ziehl-Neelsen stain or with fluorescence microscopy using the acid-fast fluorochrome dye auramine-rhodamine. Culture-based method is based on mycobacterial cultivation on liquid media, but the disadvantage is that it takes 2 or 3 weeks to obtain the result (5). Molecular techniques, such as Xpert ® MTB/RIF assay (Cepheid), are based on nucleic acid amplification and present higher sensitivity compared to conventional sputum smear microscopy (5, 57).

An Xpert MTB/RIF Ultra assayed has been developed to improve Xpert MTB/RIF, because of decreased sensitivity and some false identification of rifampicin resistance (58, 59).

Latent TB detection: tuberculin skin test (TST) and the interferon gamma release assay (IGRA) are available. IGRA is more specific than TST and can distinguish between BCG vaccinated individuals and latent TB.

TST or Mantoux test is based on an intradermal injection of tuberculin or purified protein derivative (PPD). A hypersensitivity reaction occurs after 48-72 hours in an individual who has cell-mediated immunity to these antigens. The disadvantage is that BCG vaccination or previous exposure to non-TB mycobacteria could interfere in the result.

IGRA consists of *in vitro* blood tests of cell-mediated immune response measuring T cell release of interferon gamma after stimulation with *M. tuberculosis* antigens encoded in RD1 to differentiate BCG vaccinated from *M. tuberculosis* infected individuals. Currently available IGRAs are QuantiFERON®-TB Gold In-Tube (QFT-GIT) and ELISpot (T-SPOT®.*TB*). QFT-GIT comprises ESAT-6, CFP-10 and TB7.7 antigens. Whole blood is processed and IFN-γ concentration is measured. In T-Spot, ESAT-6 and CFP-10 antigens are used, isolated PMBCs are processed and it is measured number of IFN-γ producing cells. (60)

QuantiFERON Gold Plus has been recently approved by FDA and has some differences from QFT-GIT. First, TB7.7 antigen is removed of there is a fourth tube which contains peptide configuration of ESAT-6 and CFP-10 to measure responses from $CD8^+$ cells (61). Also, a new ESAT-6 free IGRA has been developed for diagnostic for TB vaccines based on ESAT-6. The study exhibits similar levels of sensitivity and specificity compared to Quantiferon (62).

Further studies of T-SPOT.*TB* exhibited that EspC (encoded outside RD1, similar size and sequence homology than ESAT-6 and CFP-10) could be potentially used for T-cell based immunodiagnosis being at least as immunodominant as ESAT-6 and CFP-10 (63). Second generation of T-SPOT.*TB* which includes EspC, ESAT-6 and CFP-10 had significantly higher diagnostic sensitivity than T-SPOT.*TB* and QFT-GIT (64).

History of BCG, the vaccine against TB since 1921

BCG, the current vaccine against TB

BCG is the current and the only vaccine licensed to be used against TB to date. BCG was delivered orally for the first time in 1921 to a child whose mother has died a day after birth from TB and whose grandmother also was infected with *M. tuberculosis*. This first vaccination succeeded without serious side effects and without TB infection. During the next seven years more than 100,000 children had been vaccinated (10, 65, 66). The introduction of BCG vaccination notably reduced the mortality to 1.8% in vaccinated newborns in contrast to 25% in unvaccinated children (67).

BCG is still administrated, nowadays by intradermic route, with a global coverage close to 90% of newborns (2). Consequently, BCG is one of the vaccines most widely used with more than 4,000 million of doses in total and approximately 200 million of doses per year (67). The experience during this century is that BCG has a huge efficacy protection against the severe forms of the disease in children, although the protective efficacy against the pulmonary disease is adults is variable (68). Lack of central memory T cells, variability of BCG strains, different BCG preparations, loss of immunodominant antigens during the subcultivation or pre-exposure of environmental mycobacteria are some of the reasons proposed to try to explain this variability (67-69).

To overcome this variable protection of protective efficacy against pulmonary disease, new vaccine candidates are being developed based on different strategies.

BCG, a live attenuated vaccine obtained from a cattle pathogen *M. bovis*

BCG was obtained after subcultivation of *M. bovis,* the causative agent of TB in cattle, during 13 years (1908-1921) with a total of 230 passages by the physician Albert Calmette and the veterinarian Camille Guérin. The bacilli were grown in potato slides saturated in beef bile with 5% of glycerine and passages were done every three weeks. After subcultivation in the laboratory, the bacilli failed to produce TB in several animal models (65, 70).

Extensive studies of BCG exhibited that the main cause of its attenuation is the loss of the RD1 region. BCG complemented with RD1 exhibited more virulence in immunocompromised mice than BCG (71) and deletion of RD1 in *M. tuberculosis* and *M. bovis* led to an increased attenuation in immunocompromised and immunocompetent mice (72, 73). However, after complementation of BCG with RD1, virulence was not completely restored. This observation suggested that there are other mutations or insertions in the genome which may collaborate in the attenuation of BCG (71).

RD1 region encodes some components of ESX-1 protein secretion system, which is the most characterized of the five type-VII secretion systems contains in the genome of *M. tuberculosis* (Figure 7) (74, 75). This region comprises 9.5 kilobase (kb) and 9 genes are codified in this region, including the T-cell antigenic targets ESAT-6 and CFP-10. BCG complemented with the

complete RD1 recovers the secretion of ESAT-6. Secretion of ESAT-6 is related with an increased virulence in SCID mice (76) and the complementation with RD1 in BCG confers better protection in mice and guinea pigs compared to BCG after challenge with *M. tuberculosis* H37Rv (77).

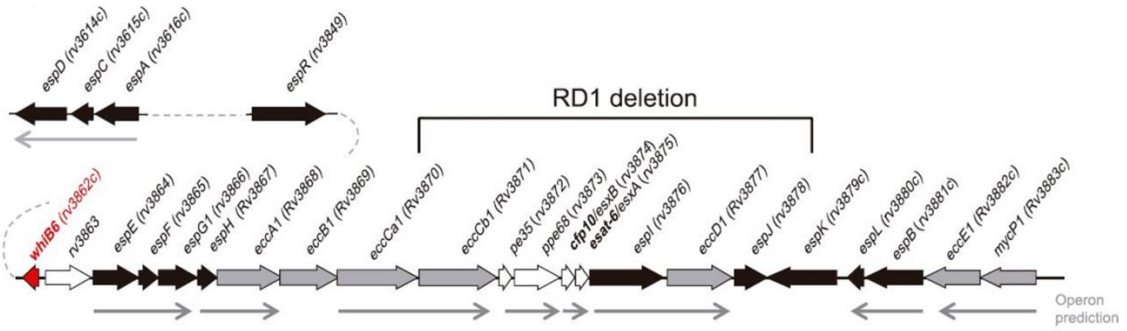


Figure 7. Genetic organization of ESX-1 and extended ESX-1 genes from *M. tuberculosis*. Scheme of ESX-1, *espACD* operon and *espR* in *M. tuberculosis*. ESAT-6 and CFP-10, also known as EsxA and EsxB proteins, are secreted through ESX-1 system. *ecc* genes are considered to encode ESX-conserved components and *esp* genes encode ESX secretion-associated proteins. RD1 is deleted in all BCG strains and considered the main cause of attenuation of BCG. From (78).

Evolution of BCG substrains as a consequence of laboratory subcultivation

After the succeeded vaccinations with BCG during the 1920 decade, BCG was distributed worldwide. Until 1960, when the lyophilized methods to conserve bacteria appeared, BCG was maintained in the different laboratories by passaging the bacilli. As a consequence of parallel BCG subcultivation in several laboratories, different substrains of BCG appeared and were named with the place of subcultivation (Figure 8) (79, 80).

Genotypic diversity is observed in BCG substrains, including region of difference (RD), specific nucleotide polymorphisms (SNPs), insertion sequences (IS*6110*), deletions and tandem duplications (65) (Figure 9). The most notable genomic modifications are the tandem duplications, DU1 and DU2. DU1, only present in BCG Pasteur, consists of 29,667 bp and spans the chromosomal origin of replication. DU2 is present in all BCG strains although this duplication is observed in four different forms, named DU2-I-IV. DU2-I is present in early BCGs strains (BCG Moreau, Russia and Japan) and consists on a duplication of ~ 20 kb. A duplication of ~140 kb and later Δint is the precursor of groups II, II and IV. DU2-II group comprised by BCG Birkhaug and Sweden have and additional deletion, ΔII, and the duplicated segment with Δint had switched position relative to groups III and IV. The duplicated fragment in DU2-III consists of 78.5 kb and is present in BCG Prague, Glaxo, Mérieux and Danish. BCG Tice, Frappier, Connaught, Phipps and Pasteur constitute the last group, containing DU2-IV which is obtained from DU2-III as a precursor and the deletion named ΔIV (81).

Other genetic modifications differs the BCG substrains, different RDs have been described in the genealogy of BCG strains, presence of IS*6110* in early BCGs (BCG Russia, Moreau and

Japan), missense mutation in BCG strains except in the early BCG strains in the start codon of *sigK* gene that lead to lose of expression of major antigens MPB70 and MPB83 (81, 82).

A remarkable variable phenotype is the presence or absence of PDIM in BCG substrains. BCG Pasteur, Denmark, Russia, Sweden, Prague, Frappier, Phipps, Tice and Birkhaug produce PDIM in contrast to BCG Japan, Moreau and Glaxo (83). This phenotype does not correlate with BCG phylogeny, suggesting these mutations were acquired in these strains independently.

Polymorphisms in the PhoPR virulence system have been also identified in BCG substrains in addition to the mutation Gly71Ile in PhoR that abolishes synthesis of SL, DAT and PAT (84). Group I of BCG has an IS*6110* insertion upstream of *phoP*, absent in the other BCG groups. Group II of BCG contains a deletion in *phoR* that truncates the C terminus. In BCG group III, BCG Prague is a natural *phoP* mutant because of a frameshift in *phoP* which eliminates C-terminal DNA binding domain. In the other three BCG from group III (BCG Glaxo, BCG Danish and BCG Mérieux) a 10 bp deletion in codon 91 in *phoR* is observed. In group IV, a frameshift in *phoR* abolishes its expression in BCG Frapier (85).

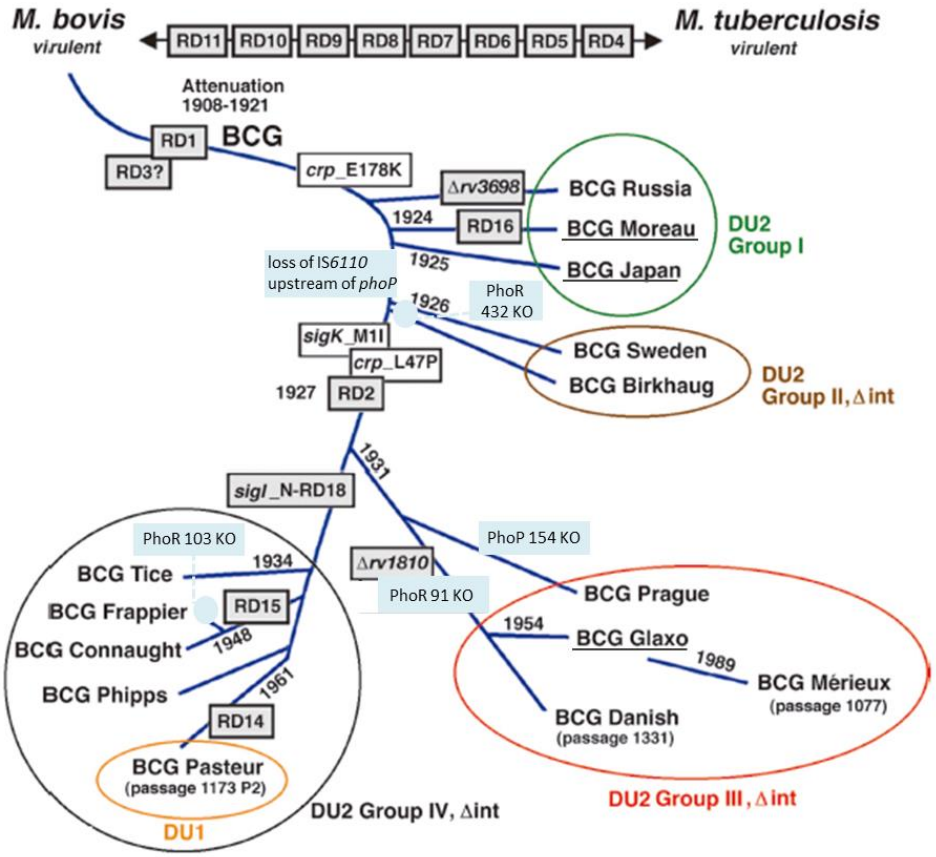


Figure 8. Genealogy of BCG substrains. The scheme shows the distribution of the BCGs in the four different groups based on the DU2 groups, different deletion or SNPs and several polymorphisms in the two components system PhoPR. Underlined BCG means absence of PDIM. Adapted from (81, 85).

Current TB vaccine pipeline

Phases of clinical trials for TB vaccine development

The existence of a TB vaccine, BCG, implies that new TB vaccines require to be as safe as BCG and to confer higher protection compared to BCG (86, 87).

After successful rigorous preclinical development, vaccine candidates could enter in clinical evaluation that consists mainly in three phases. Phase I aims to validate safety of new vaccine candidates in healthy adults unexposed to *M. tuberculosis* and later in healthy exposed adults. Phase II continues the safety control, give information about immunogenicity and selection of the dose on a target population. Phase III clinical trial is to study protective efficacy with a natural exposure of *M. tuberculosis.* (88).

TB vaccine candidates in clinical trials

New TB vaccines can be administered in different stages of the infection or the disease. Pre-exposure vaccines are administered before infection with *M. tuberculosis* and the population targeted is infants, while post-exposure vaccines are thought to target vaccinated adolescents and adults with LTBI. Therapeutic vaccines target population with active TB and the goal is to be administered in combination with chemotherapy to shorten the treatment and patients with XDR-TB (89). It is suggested that the most effective pathway to control TB disease is the combination of efficient pre- and post-exposure vaccines (90).

TB vaccines in the current pipeline have been obtained using different strategies: subunit vaccines (adjuvanted recombinant proteins or viral vectors expressing antigens), live attenuated vaccines, whole cell inactivated or fragmented mycobacteria (Figure 9) (67, 89, 91, 92).

- **Subunit vaccines**

Subunit vaccines consist of the presence of one or a few immunodominant antigens of *M. tuberculosis* and they are being developed to boost BCG, pre or post-exposure. There are two different types of subunit vaccines: recombinant proteins used with an adjuvant and viral vector systems (89).

Adjuvanted protein subunit vaccines consist of one or more antigens formulized with an adjuvant. M72/ASOE$_1$ harbors the antigens Rv1196 (Mtb39A) and Rv0125 (Mtb32A) with ASO1 adjuvant (93). Phase IIb of efficacy against progression of disease has recently concluded. HIV-negative adults from Kenya, South Africa and Zambia with LTBI and most of them BCG vaccinated were randomized to be vaccinated with M72/ASOE$_1$ or placebo. Results concluded increased efficacy against active pulmonary disease in M72/ASOE$_1$ group compared to placebo. M72/ASO1$_E$ provided 54% protection in adults infected with *M. tuberculosis* against active pulmonary TB disease (94).

H56 contains Ag85B, ESAT-6 and Rv2660c antigen and H4 harbors Ag85B with antigen TB10.4. These three fusion proteins are combined with IC31© adjuvant. Results in clinical trial of

H56:IC31 in BCG vaccinated and HIV negative adults with or without LTBI and without history of TB disease showed that the vaccine was safe and immunogenic in all doses tested (95). A trial to evaluate safety and immunogenicity of H56:IC31 and BCG revaccination was performed in HIV negative and QFT negative adolescents BCG vaccinated at birth. Neither H4:IC31 nor BCG prevented from initial QFT conversion. However, revaccination with BCG reduced the rate of sustained conversion (96). Currently a phase IIb of H56:IC31 is ongoing to evaluate safety and efficacy in reducing rate of TB disease recurrence in HIV negative adults successfully treated for drug-susceptible pulmonary TB.

ID93/GLA-SE contains Rv2608, Rv3619, Rv3620 and Rv1813 antigens formulated with GLA-SE adjuvant (97). A phase IIa trial for therapeutic use in HIV-negative TB patients has been completed and a phase IIa in BCG-vaccinated for prevention of infection is ongoing.

GamTBvac consist on Ag85A, EST-6 and CFP-10 with dextran-binding domain immobilized on dextran and with and adjuvant consisting of DEAE-dextran core (98). Phase I trial has been completed in healthy BCG vaccinated adults.

Viral vectored vaccines consist of replicon-deficient variants of virus. These subunit vaccine candidates in clinical trials, named MVA85A, Ad5Ag85A, ChadOx185A and TB/FLU-04L are viral vectors that all contains Ag85A from *M. tuberculosis*. TB/FLU-04L also contains ESAT-6. They are based on the modified vaccinia virus Ankara, adenovirus of serotype 5 vector, simian adenovirus and influenza vector respectively (67, 91). MVA85A was the first efficacy trial of a novel vaccine against TB since BCG. Unfortunately, no evidence of prevention of *M. tuberculosis* infection or TB disease was observed in the Phase IIb clinical trial in South Africa (99). Phase Ia of Ad5Ag85A is ongoing, phase I trial of ChAdOx185A alone or as part of a prime-boost strategy with MVA85A in BCG-vaccinated adults has been completed. TB/FLU-04L is currently in phase IIa and is administered by mucosal route (2) (89).

- **Live attenuated vaccines**

Two live attenuated vaccine candidates are in clinical trials. The former is a recombinant BCG (rBCGΔUreC::hly), also named VPM1002, based on deletion of *ureC* and expression of listeriolysin O (LLO) encoded by *hly* gene from *Listeria monocytogenes.* Hly perturbates the phagosome membrane in acidic pH conditions and these acid conditions are reached because of *ureC* deletion. Consequently, this recombinant BCG can reach the cytosol and its immunogenicity is improved by antigenic presentation (100). Two phases I were performed containing Hyg marker, one in healthy adults in Germany (BCG vaccinated or unvaccinated) and the other in the endemic country South Africa in healthy BCG-vaccinated adults. Results showed that the vaccine was safe in both trials (101, 102). Phases IIa were performed in HIV-unexposed newborns in South Africa and results exhibited that VPM1002 is safe and immunogenic in newborns (102). The next step, a phase IIb in HIV-unexposed and HIV-exposed uninfected, BCG-naive newborns in South Africa were performed with VPM1002 with and without Hyg marker. HIV-exposed uninfected newborns are included in the trial because it represents 30% of the newborns that will receive BCG vaccine in South Africa and the may be a high risk group of *M. tuberculosis* infection. In addition of prime vaccine at birth, VPM1002 is being developed as a post-exposure vaccine. A phase II/III trial is ongoing in India in adults that

have been successfully treated of TB. The aim is to defined if VPM1002 can prevent TB recurrence, because this population has high risk of recurrence (103).

The latter is a recombinant *M. tuberculosis* strain, named MTBVAC, based on two independent non-reverting deletions in *phoP* and *fadD26* genes without antibiotic markers, both key virulence factors of the bacilli (104). MTBVAC is being developed at the University of Zaragoza in collaboration with the industrial partner Biofabri, responsible of industrial and clinical development. Currently two phases IIa are ongoing in newborns and adults in South Africa. Molecular details, preclinical and clinical trials of MTBVAC are detailed in the next section.

- **Whole cell inactivated or fragmented mycobacteria**

Mycobacterium vaccae (Vaccae™) is a whole cell, heat killed *M.vaccae*, MIP is a whole cell heat-inactivated *mycobacterium indicus pranii* (92).

RUTI® consists on cell wall fragments from *M. tuberculosis.* Both are planned to be used as therapeutic vaccines (89, 105).

DAR-901 is a non-TB mycobacteria inactivated, *Mycobacterium obuense* to use as booster vaccine against TB. DAR-901 exhibited safety profile in healthy adults in phase Ia trial (106), and Phase II is ongoing to study the prevention of infection in adolescents and adults.

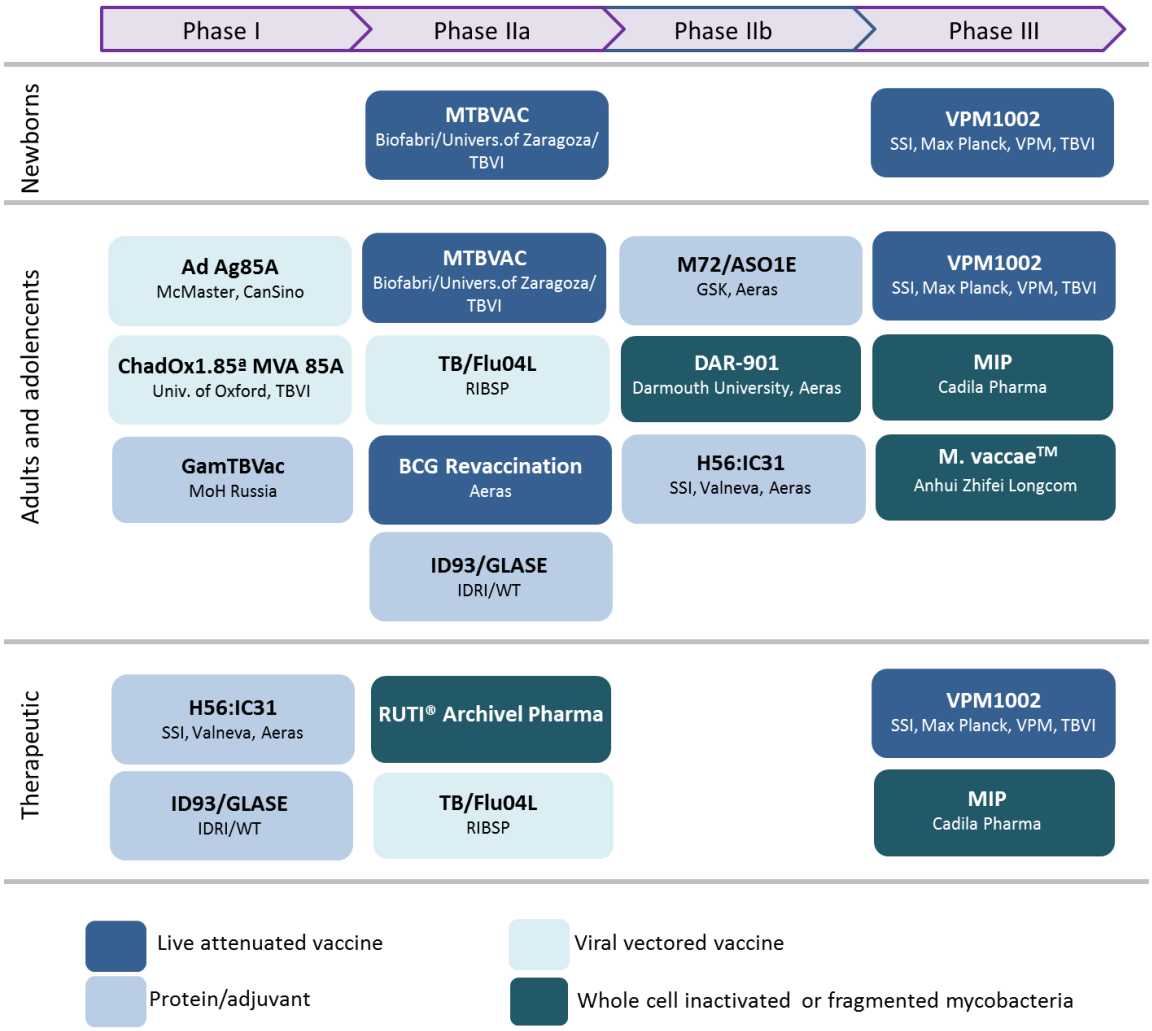


Figure 9. TB vaccine pipeline. Different box colours represent the type of vaccine: live attenuated vaccine, proteins with an adjuvant, viral vectored vaccines or whole cell inactivated or fragmented mycobacteria. Adapted from (107)

In addition to the phase IIb clinical trial of MVA85A (99), results of two clinical trials IIb have been recently published. H4:IC31 or BCG revaccination in adolescents BCG vaccinated QFT negative; none of them prevented initial QFT conversion although BCG revaccination reduced the rate of sustained conversion, opening the possibility to continue the study of revaccination with BCG (96). In a prevention of progression of disease, M72/AS01E exhibited superior efficacy against TB disease than placebo group in adults with LTBI (94).

MTBVAC, a live attenuated vaccine candidate based on *M. tuberculosis*

The rational construction of MTBVAC

New live attenuated vaccine candidates against TB should fulfill the Geneva consensus to enter into clinical trials. Two independent non-reverting mutations without antibiotic markers are recommended for live attenuated candidates (108).

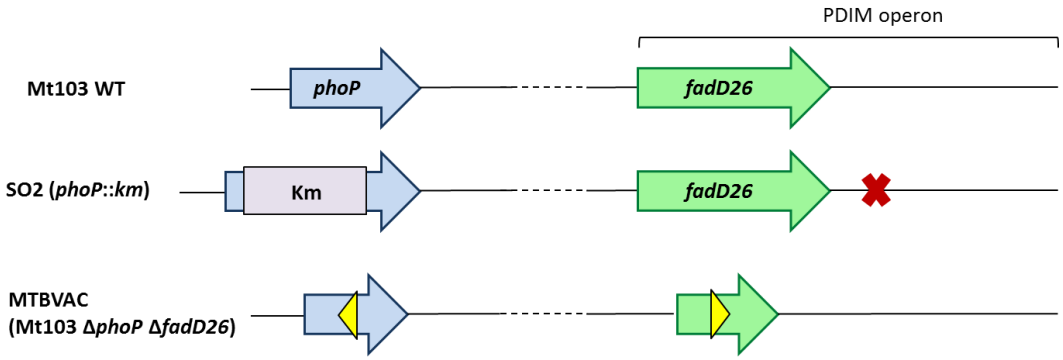
The rational construction of MTBVAC started in 1991, when an outbreak was reported in Spain caused by a *M. bovis* strain, which exceptionally was transmitted between humans (109). Genetic studies revealed the presence of two IS*6110* insertions, a rare event since *M. bovis* strains usually contains only one IS*6110*. The second copy of IS*6110* in the human outbreak strain, named *M. bovis* B, was upstream *phoP* gene. This insertion was responsible of the increased expression of this gene and this event suggested the involvement of *phoP* in virulence (110).

Deletion of *phoP* in a clinical isolate of lineage 4 of *M. tuberculosis*, named Mt103, led to an attenuated strain named SO2. A *M. tuberculosis* clinical isolate was chosen because it is the main causative agent of TB in humans and also in order to avoid side effects of laboratory subcultivation. SO2 was constructed by insertion of kanamycin (Km) resistant cassette in *phoP* (111). Subsequent experiments with SO2, revealed a spontaneous loss of PDIM. (112). Consequently SO2 has the genetic profile of *phoP* mutant and PDIM deficient phenotype (113) (Figure 10).

This *phoP* mutant/PDIM deficient strain showed reduction of multiplication in mouse bone marrow-derived macrophages from BALB/c mice and highly attenuation in the mouse model in comparison with the parental strain and even increased attenuation compared to BCG Pasteur (111, 114). Also, SO2 conferred protection in mice, guinea pigs and NHP (114, 115).

These promising results suggested the feasibility of a new vaccine candidate based on SO2 prototype. Taking into account the recommendations of new live attenuated vaccines for TB, the Geneva consensus stablished that two independent non-reverting mutations without antibiotic markers are recommended. To fulfill the Geneva consensus, the Km resistant cassette was replaced for a resolvable hygromycin (Hyg) cassette that was subsequently removed and the second deletion was obtained in *fadD26* gene, which is required for PDIM synthesis (104).

As a result, MTBVAC is a double unmarked mutant based on deletion of *phoP* and *fadD26* genes, both involved in virulence, in the clinical isolate of *M. tuberculosis* Mt103 (Figure 10) (104). MTBVAC harbors the epitope repertoire of *M. tuberculosis* because it was obtained from an isolate of the human pathogen *M. tuberculosis.* Comparison of BCG and MTBVAC in the number of T cell epitopes showed that 433 of the 1603 experimentally demonstrated T cell epitopes (27%) are located in RD regions absent in BCG Pasteur and almost the 20% of them were located in RD1 region. Therefore MTBVAC harbors 1603 epitopes and BCG 1170 epitopes (116).



Strain	*phoP* deficiency	PDIM deficiency
SO2	Insertion of Km within *phoP* gene	Spontaneous mutation in PDIM operon
MTBVAC	Internal deletion in *phoP* gene	Internal deletion in fadD26 gene

Figure 10. Genotype of Mt103, SO2 and MTBVAC strains. Mt103 is the parental strain with *phoP* and *fadD26* genes WT. In SO2, *phoP* mutant was obtained by insertion of Km cassette resistance and is PDIM deficient (spontaneously). MTBVAC is a double mutant unmarked in *phoP* and *fadD26* genes. Adapted from (113).

Implication of *phoP* in the virulence of *M. tuberculosis*

phoP gene (Rv0757) codes for the transcription factor PhoP, a constituent of the two-component system (TCS) PhoP/PhoR from *M. tuberculosis.* TCS are bacterial sensors that respond extracellular stimuli. TCS are composed of membrane-associated sensor (histidine kinase) which detect the environmental stimuli and have the ability to phosphorylate itself and after transfer the phosphate to the response regulator which acts as effector. The response regulator alters its conformation after being phosphorylated and control the expression of particular genes (117). PhoP is the response regulator of the TCS PhoP/PhoR and controls approximately 2% of the coding capacity of the *M. tuberculosis* genome (118).

Genes positively regulated by PhoP include: genes required for hypoxia adaptation, genes involved in aerobic/anaerobic respiration, genes within the RD1, genes encoding stress proteins and genes involved in lipid metabolism (118-120) (Figure 11).

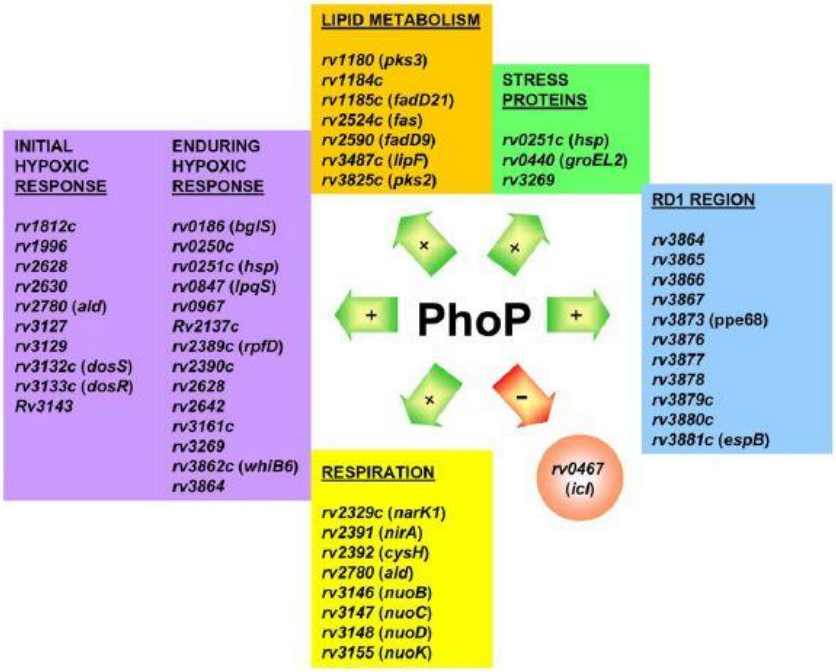


Figure 11. The PhoP regulon of *M. tuberculosis*. Positive PhoP regulated genes (with green arrows) and negative PhoP regulated genes (with red arrow), from (118).

Phenotypic molecular implications of PhoP are:

- Transcriptomics reveals that PhoP regulates *pks2/3* genes which are involved in the synthesis DAT, PAT and SL (121). Consequently, a *phoP* mutant is devoid of DAT, PAT and SL in the mycomembrane.
- PhoP controls the secretion of ESAT-6, one of the major antigens secreted by *M. tuberculosis* (122). PhoP is the key of a complex pathway that regulates ESAT-6 secretion. PhoP can directly activate expression of *espACD* operon, which controls the ESAT-6 secretion, or indirectly by activating the expression of *espR* or *whiB6*, which can also activate the expression of *espACD* operon (85). Therefore, a *phoP* mutant produce ESAT-6 but its secretion is impaired.
  A single nucleotide insertion in the promoter region of *whiB6* in H37Rv and H37Ra leads to downregulation of *whiB6* by PhoP in contrast to clinical strains in which PhoP upregulates *whiB6* expression (78).
- ChIP-seq analysis reveals that the most prominent PhoP-regulated region is the non-coding RNA (ncRNA) *mcr7*. PhoP upregulates *mcr7* expression and this ncRNA inhibits the translation of *tatC* mRNA. *tatC* gene encodes a transmembrane protein which is a component of the Twin Arginine Translocation (TatABC) apparatus which is involved in the secretion of proteins that contains a twin arginine motif (RR) in their signal peptide. Immunodominant Ag85 complex is secreted through TAT system (119). Consequently, expression of *mcr7* is notably downregulated in a *phoP* mutant and there is no inhibition of translation of *tatC* mRNA, and increased amounts of TAT substrates are secreted in *phoP* mutants.
- Metabolomics studies showed that synthesis of PIM derivatives is a novel-PhoP regulated phenotype. Production of PIMs was higher in H37Rv Δ*phoP*, GC1237 Δ*phoP*

and MTBVAC than in the parental strains. In addition 34 metabolites were differentially observed in MTBVAC compared to Mt103 (123).

Implication of *fadD26* in the virulence of *M. tuberculosis*

fadD26 gene (Rv2930) belongs to a family of 36 *fadD* genes. The gene is located in *M. tuberculosis* genome adjacent to the polyketide synthases (PKSs) genes. *fadD26* gene shows homology to acyl-CoA synthetases. FadD26 constitute one of the long-chain fatty acyl-AMP ligases (FAALs), involved in the activation of long-chain fatty acid as acyl-adenylates, subsequently transferred to multifunctional PKSs for further chain extension, for phthiocerol synthesis, indispensable for PDIM synthesis (124).

Early studies suggest the relation of PDIM in pathogenicity of the bacteria. PDIM synthesis and transport is required for growth in lungs and required for *M. tuberculosis* during acute phase infection (125, 126). Attenuation experiments demonstrated that BALB/c mice inoculated intravenously with Mt103 Δ*fadD26* showed increased survival, lower granuloma size and lower bacterial loads than WT strain (127).

A recent work has described a new function of PDIM, which seems to act in concert with ESX-1 to destabilize membranes of phagosomes containing *M. tuberculosis* and subsequent rupture of infected macrophages. Presence of PDIM was essential for phagosome disruption, which suggests that PDIM is needed for optimal ESX-1 activity (38).

- **Biosynthesis of PDIM**

The majority of genes involved in PDIM synthesis are located in a cluster of ~50 kb of the genome, with 15 open reading frames (ORF) with 35 catalytic steps.

In the first step, FadD26 activates long-chain fatty acids and transfers them to the first PKSs, PpsA. Phthiocerol is obtained after subsequent steps in the *ppsA-E* cluster. In a parallel pathway FadD28 activate fatty acids that are transferred to the PKs protein Mas to obtain mycocerosic acid. Then, PaPA5 mediates the direct transfer of the mycocerosic acids on to the phthiocerol to obtain the final product, PDIM (Figure 12). The six multifunctional PKSs proteins contain β-ketoacyl ACP synthase (KS), acyl transferase (AT), ketoreductase (KR) and ACP domains are involved in the loading, decarboxylate condensation, and ketoreduction of the acyl intermediates. Dehydratase (DH) and enoylreductase (ER) domain converts these β-hydroxy acylthioester to the reduced products (128).

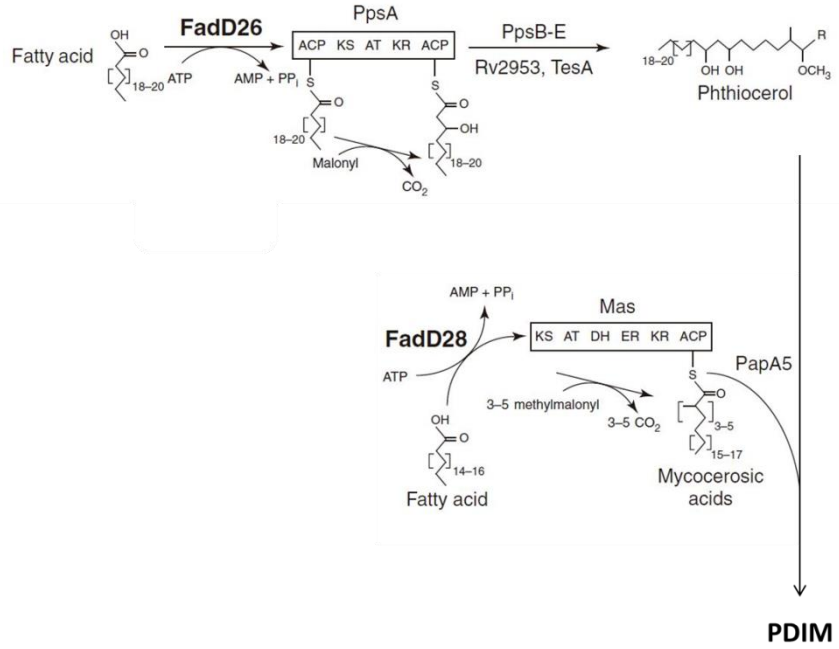


Figure 12. Byosinthesis of PDIM. Sequential steps for PDIM synthesis. Adapted from (124).

Accordingly, the phenotype of MTBVAC is the combination of the phenotypes described because of *phoP* and *fadD26* deletions (Figure 13).

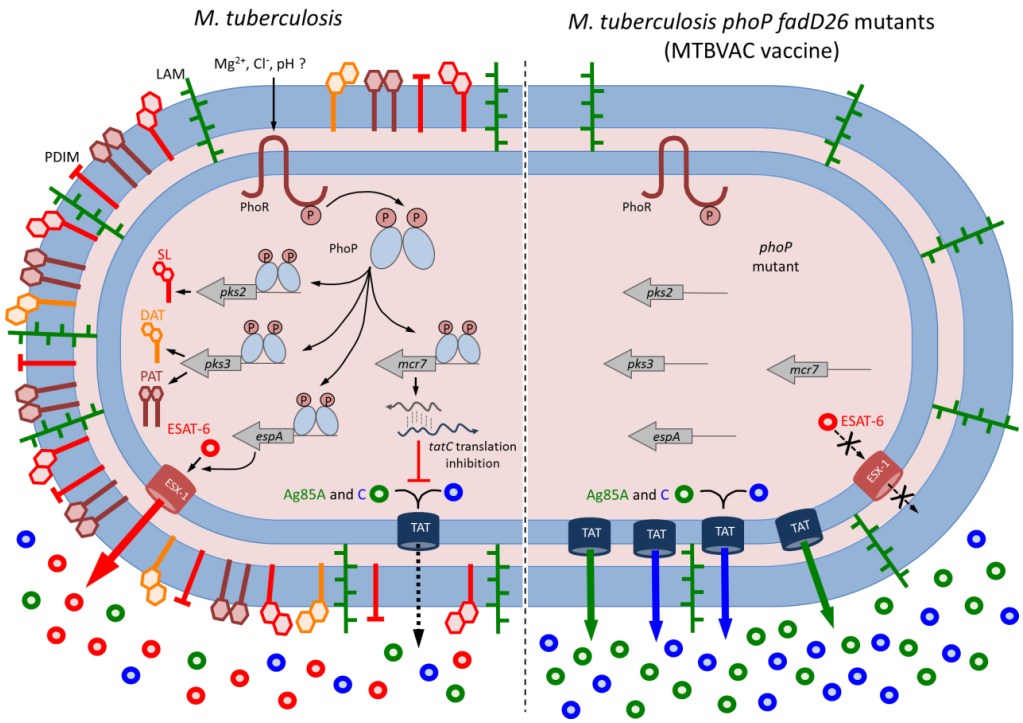


Figure 13. Comparative scheme of *M. tuberculosis* and MTBVAC. In *M. tuberculosis* (left panel) PhoR phosphorylate PhoP which controls the expression of a large set of genes. Some phenotypes have been already described to be PhoP-dependent: synthesis of SL, DAT/PAT, secretion of ESAT-6 and post-transcriptional regulation of *tatC*, mediated by *mcr7*. MTBVAC (right panel) is devoid of SL, DAT and PAT; produce but no secrete ESAT-6 and secretion of TAT-secreted proteins are increased because of *phoP* deletion. Deletion of *fadD26* leads to devoid of PDIM. Adapted from (85).

Preclinical studies of SO2 and MTBVAC

Before enter into clinical trials, new live vaccine candidates against TB require different preclinical experiments in different animal models to demonstrate protection efficacy, safety and stability. New vaccine candidates also need to be manufactured in Good Manufacturing Practice (cGMP), and preclinical and nonclinical testing needs to be performed in the GMP product (87, 108).

In general, observation of efficacy protection is recommended in more than one animal model (usually in mice, guinea-pigs, which is highly relevant model for protection efficacy studies or NHP model). Safety studies are usually performed in immunocompromised animals (mice) and stability of new vaccines is related with the maintenance of the genetic modification and non-reversion to virulent phenotype (87).

In SO2, the prototype of MTBVAC, safety studies in BALB/c mice exhibited significant increased survival of SO2 compared to Mt103 (129); in SCID mice, high attenuation of SO2 was observed compared to Mt103 and also SO2 was more attenuated than BCG Pasteur (114). Experiments in guinea pig also confirmed the safety profile of SO2, the strain was inoculated in guinea pigs and no weight differences were observed in SO2 inoculated mice compared to control and no histological lesions were observed which confirmed no toxicity effects. No reversion of *phoP* mutation was observed after 6 months of laboratory culture and after 3 months in SCID mice (130).

Comparable levels of SO2 protection to BCG-Phipps were observed with subcutaneous vaccination in BALB/c mice and intravenous or intratracheal challenge with H37Rv (114, 129). Similar protection efficacy was observed in guinea pig model, with subcutaneous vaccination of SO2 or BCG Pasteur followed by a low dose (10-50 CFU) H37Rv aerosol challenge. In high-dose challenge (500 CFU) in guinea pig, SO2 exhibited increased protection compared to BCG Pasteur or Danish vaccination followed by higher guinea pig survival, reduction of lung pathology and decreased in lung and spleen bacterial burden (114). In NHP (*rhesus macaques*) increased protective efficacy was observed in comparison with BCG (115) (Table 2).

In MTBVAC preclinical experiments, similar safety profile to BCG Pasteur and Danish was observed in SCID mice, with significant reduction in bacterial burden in lungs and spleen after 13 weeks of post-subcutaneous inoculation compared to Mt103 (104). MTBVAC also exhibited increased safety, immunogenicity and protective efficacy than BCG Danish in C57BL/6 newborn mice (131). In guinea pig model, MTBVAC exhibited protection against H37Rv aerosol challenge (104) (Table 2).

Table 2. Preclincial studies of SO2 and MTBVAC.

Safety studies	Protective efficacy
Mice (SCID and BALB/c) (104, 114, 129)	Mice (C57BL/6, BALB/c) (104, 114, 129, 131)
Guinea pig (114)	Guinea pig (104, 114)
	Non-human primates (*rhesus macaques*) (115)

Clinical trials of MTBVAC

GMP development of freeze-dried MTBVAC is the result of collaborative partnership of University of Zaragoza, Biofabri and the European non-profit Tuberculosis Vaccine Initiative (TBVI). MTBVAC entered in clinical trials in 2012, two phases I have been completed and two phases IIa are ongoing (Figure 14).

Phase Ia

Phase Ia (NCT02013245) of MTBVAC was approved in 2012 by Swissmedic to evaluate safety, local tolerance and immunogenicity testing three escalating doses. Phase Ia was performed at Lausanne University Hospital (CHUV) in Switzerland. Healthy adults without history of BCG vaccination or exposure to TB (PPD negative) and HIV negative were vaccinated with escalating doses of MTBVAC ($5x10^3$, $5x10^4$ and $5x10^5$ colony forming units, CFU) or with $5x10^5$ CFU of BCG SSI. BCG SSI is a commercial formulation of BCG of the Danish strain from Statens Serum Insitute (Copenhagen, Denmark). Successful results showed that MTBVAC is at least as safe as BCG. Promising immunogenic results were obtained: at the same dose ($5x10^5$ CFU) there was greater frequency of polyfunctional $CD4^+$ central memory T cells in MTBVAC group that in BCG group although there were no significant differences. However, until now it is unknown the immune correlation of protection against TB and consequently efficacy trials in high burden are required to validate the efficacy of new vaccine candidates (132).

Phase Ib

Data from phase Ia supported the advance to phase Ib clinical trial in an endemic country. Phase Ib (NCT02729571) was performed in South African Tuberculosis Vaccine Initiative (SATVI) in Cape Town, South Africa. This clinical trial comprised two-stage randomized and double blind clinical trial. First, 18 healthy adults BCG-vaccinated, QFT-negative and HIV-uninfected were randomized 1:1 to be vaccinated with MTBVAC or BCG SSI ($5x10^5$ CFU). A follow up during 6 months show a safety profile of MTBVAC allow the beginning of the second stage of the phase Ib in infants. 36 infants BCG naïve, HIV-unexposed were randomized 1:3 to receive BCG SSI ($2.5x10^5$ CFU) or escalating doses of MTBVAC ($2.5x10^3$, $2.5x10^4$, $2.5x10^5$ CFU).

Phase Ib finished in September 2016 and results were obtained in the first quarter of 2018 and are expected to be published during this year. Safe results allowed the progress to phase IIa.

Phase IIa

Two phases IIa have started during the first quarter of 2019 to develop MTBVAC with different population target. Both phases IIa are being performed in SATVI. The hypothesis is that the most successful strategy to achieve WHO End TB objectives is the combination of vaccination in adults and neonates.

The first strategy targets neonates and the strategy is as prime vaccine at birth. Phase IIa (NCT03536117) dose-defining trial of MTBVAC will evaluate safety, reactogenicity, immunogenicity and potential for IGRA conversion and reversion of MTBVAC in newborns in South Africa. Neonates BCG naïve, HIV-unexposed will be randomized to be vaccinated with BCG SSI ($2.5x10^5$ CFU) or escalating doses of MTBVAC ($2.5x10^4$, $2.5x10^5$ and $2.5x10^6$ CFU).

The second strategy is the use of MTBVAC as booster in BCG vaccinated adolescents and young adults. This phase IIa (NCT02933281) dose-defining enrolls healthy adults, BCG vaccinated with or without LTBI. Four different doses will be evaluated ($5x10^3$, $5x10^4$, $5x10^5$ and $5x10^6$ CFU) or BCG revaccination ($5x10^5$ CFU). Adults will receive MTBVAC vaccination or BCG Japan revaccination in eight cohorts (four cohorts of QFT-negative and four cohorts of QFT-positive).

MTBVAC vaccine is based on the attenuation of *M. tuberculosis* and harbors the whole antigen repertoire of this pathogen. Therefore, MTBVAC contains ESAT-6 and CFP-10, which are the antigens used for the diagnostic with Quantiferon and EliSpot TB. Consequently, a new diagnostic test may be required to differ between MTBVAC vaccination and *M. tuberculosis* infection (113).

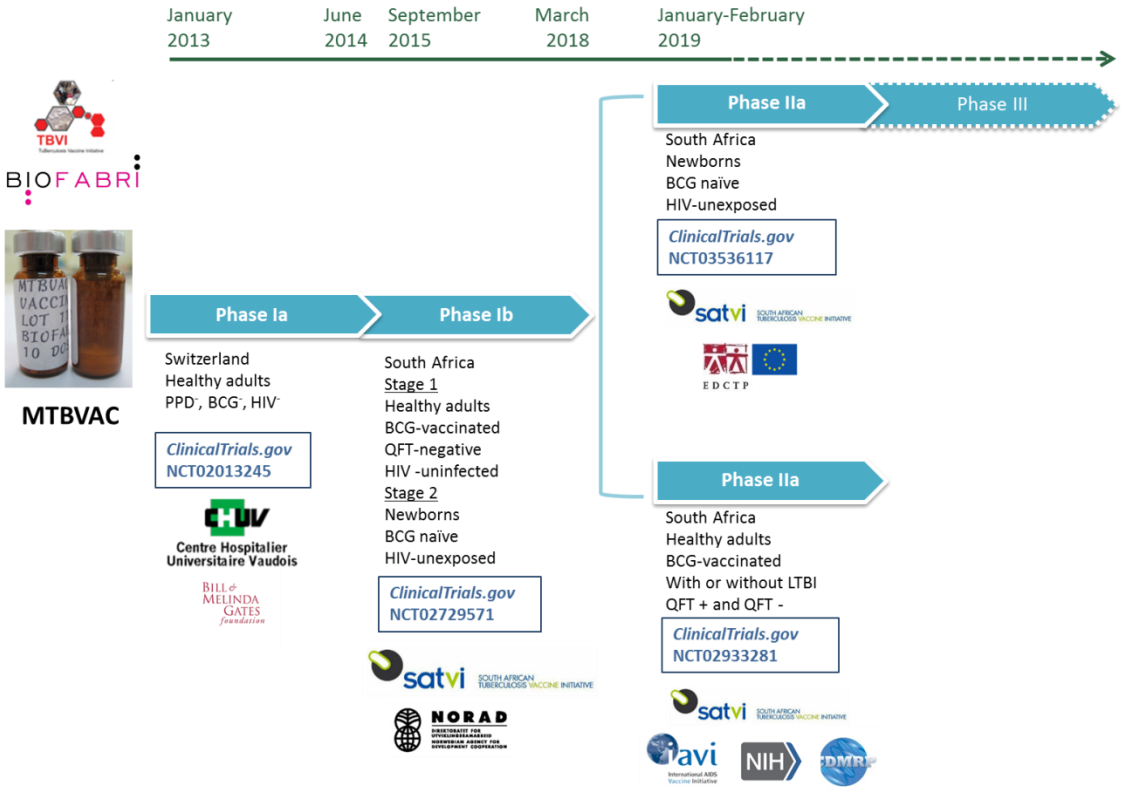


Figure 14. Clinical trials of MTBVAC. Two phases IIa are ongoing, one in newborns and the other in adults in South Africa after successful phase Ia in healthy adults in Switzerland and a phase Ib in an endemic country in healthy adults and neonates.

References

1. **Paulson T.** 2013. Epidemiology: A mortal foe. *Nature* **502:**S2-3.
2. WHO. 2018. Global Tuberculosis Report 2018.
3. **Brites D, Gagneux S.** 2017. The Nature and Evolution of Genomic Diversity in the Mycobacterium tuberculosis Complex. *Adv Exp Med Biol* **1019:**1-26.
4. **Coscolla M.** 2017. Biological and Epidemiological Consequences of MTBC Diversity. *Adv Exp Med Biol* **1019:**95-116.
5. **Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A, Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M.** 2016. Tuberculosis. *Nat Rev Dis Primers* **2:**16076.
6. **Boritsch EC, Brosch R.** 2016. Evolution of Mycobacterium tuberculosis: New Insights into Pathogenicity and Drug Resistance. *Microbiol Spectr* **4**.
7. **Gutierrez MC, Brisse S, Brosch R, Fabre M, Omais B, Marmiesse M, Supply P, Vincent V.** 2005. Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *PLoS Pathog* **1:**e5.
8. **Daniel TM.** 2006. The history of tuberculosis. *Respir Med* **100:**1862-1870.
9. **Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, Forrest SA, Bryant JM, Harris SR, Schuenemann VJ, Campbell TJ, Majander K, Wilbur AK, Guichon RA, Wolfe Steadman DL, Cook DC, Niemann S, Behr MA, Zumarraga M, Bastida R, Huson D, Nieselt K, Young D, Parkhill J, Buikstra JE, Gagneux S, Stone AC, Krause J.** 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* **514:**494-497.
10. **Herzog H.** 1998. History of tuberculosis. *Respiration* **65:**5-15.
11. **Pezzella AT.** 2019. History of Pulmonary Tuberculosis. *Thorac Surg Clin* **29:**1-17.
12. **Kaufmann SH.** 2005. Robert Koch, the Nobel Prize, and the ongoing threat of tuberculosis. *N Engl J Med* **353:**2423-2426.
13. **Barberis I, Bragazzi NL, Galluzzo L, Martini M.** 2017. The history of tuberculosis: from the first historical records to the isolation of Koch's bacillus. *J Prev Med Hyg* **58:**E9-E12.
14. **Daniel TM.** 2011. Hermann Brehmer and the origins of tuberculosis sanatoria. *Int J Tuberc Lung Dis* **15:**161-162, i.
15. WHO. 1991. TB: a global emergency. World Health Organization
16. WHO. 2018. Global Tuberculosis Report 2018
17. WHO. 2015. Global Plan to End TB, 2015
18. **Esposito S, Bianchini S, Blasi F.** 2015. Bedaquiline and delamanid in tuberculosis. *Expert Opin Pharmacother* **16:**2319-2330.
19. **Barrera L.** 2007. The Basics of Clinical Bacteriology. In Palomino, J. C., Cardoso-Leao S and Ritacco V. (ed.) Tuberculosis, From basic science to patient care
20. **Forbes BA, Hall GS, Miller MB, Novak SM, Rowlinson MC, Salfinger M, Somoskovi A, Warshauer DM, Wilson ML.** 2018. Practice Guidelines for Clinical Microbiology Laboratories: Mycobacteria. *Clin Microbiol Rev* **31**.
21. **Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG.** 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393:**537-544.

22. **Trifiro S, Bourgault AM, Lebel F, Rene P.** 1990. Ghost mycobacteria on Gram stain. *J Clin Microbiol* **28:**146-147.
23. **Ferrer NL, Gomez AB, Soto CY, Neyrolles O, Gicquel B, Garcia-Del Portillo F, Martin C.** 2009. Intracellular replication of attenuated Mycobacterium tuberculosis phoP mutant in the absence of host cell cytotoxicity. Microbes Infect **11**:115-122.
24. **Jackson M.** 2014. The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med* **4**.
25. **Chiaradia L, Lefebvre C, Parra J, Marcoux J, Burlet-Schiltz O, Etienne G, Tropis M, Daffe M.** 2017. Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane. *Sci Rep* **7:**12807.
26. **Daffé M. and Reyrat J. M.,** The Mycobacterial Cell Envelop, 2008, ASM Press, American Society for Microbiology
27. **Jankute M, Cox JA, Harrison J, Besra GS.** 2015. Assembly of the Mycobacterial Cell Wall. *Annu Rev Microbiol* **69:**405-423.
28. **Alderwick LJ, Harrison J, Lloyd GS, Birch HL.** 2015. The Mycobacterial Cell Wall--Peptidoglycan and Arabinogalactan. *Cold Spring Harb Perspect Med* **5:**a021113.
29. **Constant P, Perez E, Malaga W, Laneelle MA, Saurel O, Daffe M, Guilhot C.** 2002. Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the Mycobacterium tuberculosis complex. Evidence that all strains synthesize glycosylated p-hydroxybenzoic methyl esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. *J Biol Chem* **277:**38148-38158.
30. **Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, Kaplan G, Barry CE, 3rd.** 2004. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431:**84-87.
31. **Jarlier V, Nikaido H.** 1994. Mycobacterial cell wall: structure and role in natural resistance to antibiotics. *FEMS Microbiol Lett* **123:**11-18.
32. **Chen H, Nyantakyi SA, Li M, Gopal P, Aziz DB, Yang T, Moreira W, Gengenbacher M, Dick T, Go ML.** 2018. The Mycobacterial Membrane: A Novel Target Space for Anti-tubercular Drugs. *Front Microbiol* **9:**1627.
33. **Sani M, Houben EN, Geurtsen J, Pierson J, de Punder K, van Zon M, Wever B, Piersma SR, Jimenez CR, Daffe M, Appelmelk BJ, Bitter W, van der Wel N, Peters PJ.** 2010. Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog* **6:**e1000794.
34. **Bailo R, Bhatt A, Ainsa JA.** 2015. Lipid transport in Mycobacterium tuberculosis and its implications in virulence and drug development. *Biochem Pharmacol* **96:**159-167.
35. **Queval CJ, Brosch R, Simeone R.** 2017. The Macrophage: A Disputed Fortress in the Battle against Mycobacterium tuberculosis. *Front Microbiol* **8:**2284.
36. **Houben D, Demangel C, van Ingen J, Perez J, Baldeon L, Abdallah AM, Caleechurn L, Bottai D, van Zon M, de Punder K, van der Laan T, Kant A, Bossers-de Vries R, Willemsen P, Bitter W, van Soolingen D, Brosch R, van der Wel N, Peters PJ.** 2012. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol* **14:**1287-1298.
37. **van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, Brenner M, Peters PJ.** 2007. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129:**1287-1298.
38. **Augenstreich J, Arbues A, Simeone R, Haanappel E, Wegener A, Sayes F, Le Chevalier F, Chalut C, Malaga W, Guilhot C, Brosch R, Astarie-Dequeker C.** 2017. ESX-1 and phthiocerol dimycocerosates of Mycobacterium tuberculosis act in concert to cause phagosomal rupture and host cell apoptosis. *Cell Microbiol* **19**.
39. **Zondervan NA, van Dam JCJ, Schaap PJ, Martins Dos Santos VAP, Suarez-Diez M.** 2018. Regulation of Three Virulence Strategies of Mycobacterium tuberculosis: A Success Story. *Int J Mol Sci* **19**.

40. **Ramakrishnan L.** 2012. Revisiting the role of the granuloma in tuberculosis. *Nat Rev Immunol* **12:**352-366.
41. **Silva Miranda M, Breiman A, Allain S, Deknuydt F, Altare F.** 2012. The tuberculous granuloma: an unsuccessful host defence mechanism providing a safety shelter for the bacteria? *Clin Dev Immunol* **2012:**139127.
42. **Russell DG.** 2007. Who puts the tubercle in tuberculosis? *Nat Rev Microbiol* **5:**39-47.
43. **Harries AD, Dye C.** 2006. Tuberculosis. *Ann Trop Med Parasitol* **100:**415-431.
44. **Donald PR, Marais BJ, Barry CE, 3rd.** 2010. Age and the epidemiology and pathogenesis of tuberculosis. *Lancet* **375:**1852-1854.
45. **Scully T.** 2013. Tuberculosis. *Nature* **502:**S1.
46. **Fonseca KL, Rodrigues PNS, Olsson IAS, Saraiva M.** 2017. Experimental study of tuberculosis: From animal models to complex cell systems and organoids. *PLoS Pathog* **13:**e1006421.
47. **Gupta UD, Katoch VM.** 2005. Animal models of tuberculosis. *Tuberculosis (Edinb)* **85:**277-293.
48. **Singh AK, Gupta UD.** 2018. Animal models of tuberculosis: Lesson learnt. *Indian J Med Res* **147:**456-463.
49. **Williams A, Orme IM.** 2016. Animal Models of Tuberculosis: An Overview. *Microbiol Spectr* **4**.
50. **Myllymaki H, Niskanen M, Oksanen KE, Ramet M.** 2015. Animal models in tuberculosis research - where is the beef? *Expert Opin Drug Discov* **10:**871-883.
51. **Driver ER, Ryan GJ, Hoff DR, Irwin SM, Basaraba RJ, Kramnik I, Lenaerts AJ.** 2012. Evaluation of a mouse model of necrotic granuloma formation using C3HeB/FeJ mice for testing of drugs against Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **56:**3181-3195.
52. **Clark S, Hall Y, Williams A.** 2014. Animal models of tuberculosis: Guinea pigs. *Cold Spring Harb Perspect Med* **5:**a018572.
53. **Laddy DJ, Bonavia A, Hanekom WA, Kaushal D, Williams A, Roederer M, Seder RA, Sharpe S, Verreck FAW, Darrah PA.** 2018. Toward Tuberculosis Vaccine Development: Recommendations for Nonhuman Primate Study Design. *Infect Immun* **86**.
54. **Lin PL, Rodgers M, Smith L, Bigbee M, Myers A, Bigbee C, Chiosea I, Capuano SV, Fuhrman C, Klein E, Flynn JL.** 2009. Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. *Infect Immun* **77:**4631-4642.
55. **Maiello P, DiFazio RM, Cadena AM, Rodgers MA, Lin PL, Scanga CA, Flynn JL.** 2018. Rhesus Macaques Are More Susceptible to Progressive Tuberculosis than Cynomolgus Macaques: a Quantitative Comparison. *Infect Immun* **86**.
56. **Langermans JA, Andersen P, van Soolingen D, Vervenne RA, Frost PA, van der Laan T, van Pinxteren LA, van den Hombergh J, Kroon S, Peekel I, Florquin S, Thomas AW.** 2001. Divergent effect of bacillus Calmette-Guerin (BCG) vaccination on Mycobacterium tuberculosis infection in highly related macaque species: implications for primate models in tuberculosis vaccine research. *Proc Natl Acad Sci U S A* **98:**11497-11502.
57. **Stevens WS, Scott L, Noble L, Gous N, Dheda K.** 2017. Impact of the GeneXpert MTB/RIF Technology on Tuberculosis Control. *Microbiol Spectr* **5**.
58. **Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, Banada PP, Deshpande S, Shenai S, Gall A, Glass J, Krieswirth B, Schumacher SG, Nabeta P, Tukvadze N, Rodrigues C, Skrahina A, Tagliani E, Cirillo DM, Davidow A, Denkinger CM, Persing D, Kwiatkowski R, Jones M, Alland D.** 2017. The New Xpert MTB/RIF Ultra: Improving Detection of Mycobacterium tuberculosis and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing. *MBio* **8**.
59. **Dorman SE, Schumacher SG, Alland D, Nabeta P, Armstrong DT, King B, Hall SL, Chakravorty S, Cirillo DM, Tukvadze N, Bablishvili N, Stevens W, Scott L, Rodrigues C,**

Kazi MI, Joloba M, Nakiyingi L, Nicol MP, Ghebrekristos Y, Anyango I, Murithi W, Dietze R, Lyrio Peres R, Skrahina A, Auchynka V, Chopra KK, Hanif M, Liu X, Yuan X, Boehme CC, Ellner JJ, Denkinger CM, study t. 2018. Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study. *Lancet Infect Dis* **18:**76-84.

60. **Lalvani A, Pareek M.** 2010. Interferon gamma release assays: principles and practice. *Enferm Infecc Microbiol Clin* **28:**245-252.
61. **Bastian I, Coulter C, National Tuberculosis Advisory C.** 2017. Position statement on interferon-gamma release assays for the detection of latent tuberculosis infection. *Commun Dis Intell Q Rep* **41:**E322-E336.
62. **Ruhwald M, de Thurah L, Kuchaka D, Zaher MR, Salman AM, Abdel-Ghaffar AR, Shoukry FA, Michelsen SW, Soborg B, Blauenfeldt T, Mpagama S, Hoff ST, Agger EM, Rosenkrands I, Aagard C, Kibiki G, El-Sheikh N, Andersen P.** 2017. Introducing the ESAT-6 free IGRA, a companion diagnostic for TB vaccines based on ESAT-6. *Sci Rep* **7:**45969.
63. **Millington KA, Fortune SM, Low J, Garces A, Hingley-Wilson SM, Wickremasinghe M, Kon OM, Lalvani A.** 2011. Rv3615c is a highly immunodominant RD1 (Region of Difference 1)-dependent secreted antigen specific for Mycobacterium tuberculosis infection. *Proc Natl Acad Sci U S A* **108:**5730-5735.
64. **Whitworth HS, Badhan A, Boakye AA, Takwoingi Y, Rees-Roberts M, Partlett C, Lambie H, Innes J, Cooke G, Lipman M, Conlon C, Macallan D, Chua F, Post FA, Wiselka M, Woltmann G, Deeks JJ, Kon OM, Lalvani A, Interferon-gamma Release Assays for Diagnostic Evaluation of Active Tuberculosis study g.** 2019. Clinical utility of existing and second-generation interferon-gamma release assays for diagnostic evaluation of tuberculosis: an observational cohort study. *Lancet Infect Dis* **19:**193-202.
65. **Abdallah AM, Behr MA.** 2017. Evolution and Strain Variation in BCG. *Adv Exp Med Biol* **1019:**155-169.
66. **Calmette A.** 1931. Preventive Vaccination Against Tuberculosis with BCG. *Proc R Soc Med* **24:**1481-1490.
67. **Martin C, Aguilo N, Gonzalo-Asensio J.** 2018. Vaccination against tuberculosis. *Enferm Infecc Microbiol Clin* **36:**648-656.
68. **Fine PE.** 1995. Variation in protection by BCG: implications of and for heterologous immunity. *Lancet* **346:**1339-1345.
69. **Orme IM.** 2010. The Achilles heel of BCG. *Tuberculosis (Edinb)* **90:**329-332.
70. **Hawgood BJ.** 2007. Albert Calmette (1863-1933) and Camille Guerin (1872-1961): the C and G of BCG vaccine. *J Med Biogr* **15:**139-146.
71. **Pym AS, Brodin P, Brosch R, Huerre M, Cole ST.** 2002. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol* **46:**709-717.
72. **Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, Eisenberg D, Russell RG, Derrick SC, Collins FM, Morris SL, King CH, Jacobs WR, Jr.** 2003. The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A* **100:**12420-12425.
73. **Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, Sherman DR.** 2003. Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation. *J Infect Dis* **187:**117-123.
74. **Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, Vandenbroucke-Grauls CM, Appelmelk BJ, Bitter W.** 2007. Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5:**883-891.

75. **Behr MA, Sherman DR.** 2007. Mycobacterial virulence and specialized secretion: same story, different ending. *Nat Med* **13:**286-287.
76. **Brodin P, Majlessi L, Marsollier L, de Jonge MI, Bottai D, Demangel C, Hinds J, Neyrolles O, Butcher PD, Leclerc C, Cole ST, Brosch R.** 2006. Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on immunogenicity and virulence. *Infect Immun* **74:**88-98.
77. **Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C, Cole ST.** 2003. Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat Med* **9:**533-539.
78. **Solans L, Aguilo N, Samper S, Pawlik A, Frigui W, Martin C, Brosch R, Gonzalo-Asensio J.** 2014. A specific polymorphism in Mycobacterium tuberculosis H37Rv causes differential ESAT-6 expression and identifies WhiB6 as a novel ESX-1 component. *Infect Immun* **82:**3446-3456.
79. **Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM.** 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284:**1520-1523.
80. **Behr MA, Small PM.** 1999. A historical and molecular phylogeny of BCG strains. *Vaccine* **17:**915-922.
81. **Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, Inwald JK, Golby P, Garcia JN, Hewinson RG, Behr MA, Quail MA, Churcher C, Barrell BG, Parkhill J, Cole ST.** 2007. Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* **104:**5596-5601.
82. **Charlet D, Mostowy S, Alexander D, Sit L, Wiker HG, Behr MA.** 2005. Reduced expression of antigenic proteins MPB70 and MPB83 in Mycobacterium bovis BCG strains due to a start codon mutation in sigK. *Mol Microbiol* **56:**1302-1313.
83. **Chen JM, Islam ST, Ren H, Liu J.** 2007. Differential productions of lipid virulence factors among BCG vaccine strains and implications on BCG safety. *Vaccine* **25:**8114-8122.
84. **Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, Laval F, Daffe M, Martin C, Brosch R, Guilhot C.** 2014. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. *Proc Natl Acad Sci U S A* **111:**11491-11496.
85. **Broset E, Martin C, Gonzalo-Asensio J.** 2015. Evolutionary landscape of the Mycobacterium tuberculosis complex from the viewpoint of PhoPR: implications for virulence regulation and application to vaccine development. *MBio* **6:**e01289-01215.
86. **Hoft DF.** 2008. Tuberculosis vaccine development: goals, immunological design, and evaluation. *Lancet* **372:**164-175.
87. **Walker KB, Brennan MJ, Ho MM, Eskola J, Thiry G, Sadoff J, Dobbelaer R, Grode L, Liu MA, Fruth U, Lambert PH.** 2010. The second Geneva Consensus: Recommendations for novel live TB vaccines. *Vaccine* **28:**2259-2270.
88. **Fruth U, Young D.** 2004. Prospects for new TB vaccines: Stop TB Working Group on TB Vaccine Development. *Int J Tuberc Lung Dis* **8:**151-155.
89. **Andersen P, Kaufmann SH.** 2014. Novel vaccination strategies against tuberculosis. *Cold Spring Harb Perspect Med* **4**.
90. **Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM, Jr., Dye C, Halloran ME.** 2009. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proc Natl Acad Sci U S A* **106:**13980-13985.
91. **Fletcher HA, Schrager L.** 2016. TB vaccine development and the End TB Strategy: importance and current status. *Trans R Soc Trop Med Hyg* **110:**212-218.
92. **Andersen P, Scriba TJ.** 2019. Moving tuberculosis vaccines from theory to practice. *Nat Rev Immunol* doi:10.1038/s41577-019-0174-z.

93. **Day CL, Tameris M, Mansoor N, van Rooyen M, de Kock M, Geldenhuys H, Erasmus M, Makhethe L, Hughes EJ, Gelderbloem S, Bollaerts A, Bourguignon P, Cohen J, Demoitie MA, Mettens P, Moris P, Sadoff JC, Hawkridge A, Hussey GD, Mahomed H, Ofori-Anyinam O, Hanekom WA.** 2013. Induction and regulation of T-cell immunity by the novel tuberculosis vaccine M72/AS01 in South African adults. *Am J Respir Crit Care Med* **188:**492-502.

94. **Van Der Meeren O, Hatherill M, Nduba V, Wilkinson RJ, Muyoyeta M, Van Brakel E, Ayles HM, Henostroza G, Thienemann F, Scriba TJ, Diacon A, Blatner GL, Demoitie MA, Tameris M, Malahleha M, Innes JC, Hellstrom E, Martinson N, Singh T, Akite EJ, Khatoon Azam A, Bollaerts A, Ginsberg AM, Evans TG, Gillard P, Tait DR.** 2018. Phase 2b Controlled Trial of M72/AS01E Vaccine to Prevent Tuberculosis. *N Engl J Med* **379:**1621-1634.

95. **Suliman S, Luabeya AKK, Geldenhuys H, Tameris M, Hoff ST, Shi Z, Tait D, Kromann I, Ruhwald M, Rutkowski KT, Shepherd B, Hokey D, Ginsberg AM, Hanekom WA, Andersen P, Scriba TJ, Hatherill M, Group HT.** 2019. Dose Optimization of H56:IC31 Vaccine for Tuberculosis-Endemic Populations. A Double-Blind, Placebo-controlled, Dose-Selection Trial. *Am J Respir Crit Care Med* **199:**220-231.

96. **Nemes E, Geldenhuys H, Rozot V, Rutkowski KT, Ratangee F, Bilek N, Mabwe S, Makhethe L, Erasmus M, Toefy A, Mulenga H, Hanekom WA, Self SG, Bekker LG, Ryall R, Gurunathan S, DiazGranados CA, Andersen P, Kromann I, Evans T, Ellis RD, Landry B, Hokey DA, Hopkins R, Ginsberg AM, Scriba TJ, Hatherill M, Team CS.** 2018. Prevention of M. tuberculosis Infection with H4:IC31 Vaccine or BCG Revaccination. *N Engl J Med* **379:**138-149.

97. **Bertholet S, Ireton GC, Ordway DJ, Windish HP, Pine SO, Kahn M, Phan T, Orme IM, Vedvick TS, Baldwin SL, Coler RN, Reed SG.** 2010. A defined tuberculosis vaccine candidate boosts BCG and protects against multidrug-resistant Mycobacterium tuberculosis. *Sci Transl Med* **2:**53ra74.

98. **Tkachuk AP, Gushchin VA, Potapov VD, Demidenko AV, Lunin VG, Gintsburg AL.** 2017. Multi-subunit BCG booster vaccine GamTBvac: Assessment of immunogenicity and protective efficacy in murine and guinea pig TB models. *PLoS One* **12:**e0176784.

99. **Tameris MD, Hatherill M, Landry BS, Scriba TJ, Snowden MA, Lockhart S, Shea JE, McClain JB, Hussey GD, Hanekom WA, Mahomed H, McShane H, Team MATS.** 2013. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *Lancet* **381:**1021-1028.

100. **Grode L, Seiler P, Baumann S, Hess J, Brinkmann V, Nasser Eddine A, Mann P, Goosmann C, Bandermann S, Smith D, Bancroft GJ, Reyrat JM, van Soolingen D, Raupach B, Kaufmann SH.** 2005. Increased vaccine efficacy against tuberculosis of recombinant Mycobacterium bovis bacille Calmette-Guerin mutants that secrete listeriolysin. *J Clin Invest* **115:**2472-2479.

101. **Grode L, Ganoza CA, Brohm C, Weiner J, 3rd, Eisele B, Kaufmann SH.** 2013. Safety and immunogenicity of the recombinant BCG vaccine VPM1002 in a phase 1 open-label randomized clinical trial. *Vaccine* **31:**1340-1348.

102. **Loxton AG, Knaul JK, Grode L, Gutschmidt A, Meller C, Eisele B, Johnstone H, van der Spuy G, Maertzdorf J, Kaufmann SHE, Hesseling AC, Walzl G, Cotton MF.** 2017. Safety and Immunogenicity of the Recombinant Mycobacterium bovis BCG Vaccine VPM1002 in HIV-Unexposed Newborn Infants in South Africa. *Clin Vaccine Immunol* **24**.

103. **Nieuwenhuizen NE, Kulkarni PS, Shaligram U, Cotton MF, Rentsch CA, Eisele B, Grode L, Kaufmann SHE.** 2017. The Recombinant Bacille Calmette-Guerin Vaccine VPM1002: Ready for Clinical Efficacy Testing. *Front Immunol* **8:**1147.

104. **Arbues A, Aguilo JI, Gonzalo-Asensio J, Marinova D, Uranga S, Puentes E, Fernandez C, Parra A, Cardona PJ, Vilaplana C, Ausina V, Williams A, Clark S, Malaga W, Guilhot**

C, Gicquel B, Martin C. 2013. Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated M. tuberculosis-based vaccine to enter clinical trials. *Vaccine* **31:**4867-4873.

105. **Cardona PJ.** 2016. The Progress of Therapeutic Vaccination with Regard to Tuberculosis. *Front Microbiol* **7:**1536.
106. **von Reyn CF, Lahey T, Arbeit RD, Landry B, Kailani L, Adams LV, Haynes BC, Mackenzie T, Wieland-Alter W, Connor RI, Tvaroha S, Hokey DA, Ginsberg AM, Waddell R.** 2017. Safety and immunogenicity of an inactivated whole cell tuberculosis vaccine booster in adults primed with BCG: A randomized, controlled trial of DAR-901. *PLoS One* **12:**e0175215.
107. TBVI, https://www.tbvi.eu
108. **Kamath AT, Fruth U, Brennan MJ, Dobbelaer R, Hubrechts P, Ho MM, Mayner RE, Thole J, Walker KB, Liu M, Lambert PH, Foundation AGTV, World Health O.** 2005. New live mycobacterial vaccines: the Geneva consensus on essential steps towards clinical development. *Vaccine* **23:**3753-3761.
109. **Samper S, Martin C, Pinedo A, Rivero A, Blazquez J, Baquero F, van Soolingen D, van Embden J.** 1997. Transmission between HIV-infected patients of multidrug-resistant tuberculosis caused by Mycobacterium bovis. *AIDS* **11:**1237-1242.
110. **Soto CY, Menendez MC, Perez E, Samper S, Gomez AB, Garcia MJ, Martin C.** 2004. IS6110 mediates increased transcription of the phoP virulence gene in a multidrug-resistant clinical isolate responsible for tuberculosis outbreaks. *J Clin Microbiol* **42:**212-219.
111. **Perez E, Samper S, Bordas Y, Guilhot C, Gicquel B, Martin C.** 2001. An essential role for phoP in Mycobacterium tuberculosis virulence. *Mol Microbiol* **41:**179-187.
112. Ainhoa Arbués, PhD Thesis, University of Zaragoza.
113. **Marinova D, Gonzalo-Asensio J, Aguilo N, Martin C.** 2017. MTBVAC from discovery to clinical trials in tuberculosis-endemic countries. *Expert Rev Vaccines* **16:**565-576.
114. **Martin C, Williams A, Hernandez-Pando R, Cardona PJ, Gormley E, Bordat Y, Soto CY, Clark SO, Hatch GJ, Aguilar D, Ausina V, Gicquel B.** 2006. The live Mycobacterium tuberculosis phoP mutant strain is more attenuated than BCG and confers protective immunity against tuberculosis in mice and guinea pigs. *Vaccine* **24:**3408-3419.
115. **Verreck FA, Vervenne RA, Kondova I, van Kralingen KW, Remarque EJ, Braskamp G, van der Werff NM, Kersbergen A, Ottenhoff TH, Heidt PJ, Gilbert SC, Gicquel B, Hill AV, Martin C, McShane H, Thomas AW.** 2009. MVA.85A boosting of BCG and an attenuated, phoP deficient M. tuberculosis vaccine both show protective efficacy against tuberculosis in rhesus macaques. *PLoS One* **4:**e5264.
116. **Gonzalo-Asensio J, Marinova D, Martin C, Aguilo N.** 2017. MTBVAC: Attenuating the Human Pathogen of Tuberculosis (TB) Toward a Promising Vaccine against the TB Epidemic. *Front Immunol* **8:**1803.
117. **Parish T.** 2014. Two-Component Regulatory Systems of Mycobacteria. *Microbiol Spectr* **2:**MGM2-0010-2013.
118. **Gonzalo-Asensio J, Mostowy S, Harders-Westerveen J, Huygen K, Hernandez-Pando R, Thole J, Behr M, Gicquel B, Martin C.** 2008. PhoP: a missing piece in the intricate puzzle of Mycobacterium tuberculosis virulence. *PLoS One* **3:**e3496.
119. **Solans L, Gonzalo-Asensio J, Sala C, Benjak A, Uplekar S, Rougemont J, Guilhot C, Malaga W, Martin C, Cole ST.** 2014. The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in Mycobacterium tuberculosis. *PLoS Pathog* **10:**e1004183.
120. **Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I.** 2006. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol* **60:**312-330.
121. **Gonzalo Asensio J, Maia C, Ferrer NL, Barilone N, Laval F, Soto CY, Winter N, Daffe M, Gicquel B, Martin C, Jackson M.** 2006. The virulence-associated two-component PhoP-

PhoR system controls the biosynthesis of polyketide-derived lipids in Mycobacterium tuberculosis. *J Biol Chem* **281:**1313-1316.

122. **Frigui W, Bottai D, Majlessi L, Monot M, Josselin E, Brodin P, Garnier T, Gicquel B, Martin C, Leclerc C, Cole ST, Brosch R.** 2008. Control of M. tuberculosis ESAT-6 secretion and specific T cell recognition by PhoP. *PLoS Pathog* **4:**e33.
123. **Diaz C, Perez Del Palacio J, Valero-Guillen PL, Mena Garcia P, Perez I, Vicente F, Martin C, Genilloud O, Sanchez Pozo A, Gonzalo-Asensio J.** 2019. Comparative Metabolomics between Mycobacterium tuberculosis and the MTBVAC Vaccine Candidate. *ACS Infect Dis* doi:10.1021/acsinfecdis.9b00008.
124. **Simeone R, Leger M, Constant P, Malaga W, Marrakchi H, Daffe M, Guilhot C, Chalut C.** 2010. Delineation of the roles of FadD22, FadD26 and FadD29 in the biosynthesis of phthiocerol dimycocerosates and related compounds in Mycobacterium tuberculosis. *FEBS J* **277:**2715-2725.
125. **Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C.** 1999. Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. *Mol Microbiol* **34:**257-267.
126. **Cox JS, Chen B, McNeil M, Jacobs WR, Jr.** 1999. Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. *Nature* **402:**79-83.
127. **Infante E, Aguilar LD, Gicquel B, Pando RH.** 2005. Immunogenicity and protective efficacy of the Mycobacterium tuberculosis fadD26 mutant. *Clin Exp Immunol* **141:**21-28.
128. **Trivedi OA, Arora P, Vats A, Ansari MZ, Tickoo R, Sridharan V, Mohanty D, Gokhale RS.** 2005. Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. *Mol Cell* **17:**631-643.
129. **Aguilar D, Infante E, Martin C, Gormley E, Gicquel B, Hernandez Pando R.** 2007. Immunological responses and protective immunity against tuberculosis conferred by vaccination of Balb/C mice with the attenuated Mycobacterium tuberculosis (phoP) SO2 strain. *Clin Exp Immunol* **147:**330-338.
130. **Cardona PJ, Asensio JG, Arbues A, Otal I, Lafoz C, Gil O, Caceres N, Ausina V, Gicquel B, Martin C.** 2009. Extended safety studies of the attenuated live tuberculosis vaccine SO2 based on phoP mutant. *Vaccine* **27:**2499-2505.
131. **Aguilo N, Uranga S, Marinova D, Monzon M, Badiola J, Martin C.** 2016. MTBVAC vaccine is safe, immunogenic and confers protective efficacy against Mycobacterium tuberculosis in newborn mice. *Tuberculosis (Edinb)* **96:**71-74.
132. **Spertini F, Audran R, Chakour R, Karoui O, Steiner-Monard V, Thierry AC, Mayor CE, Rettby N, Jaton K, Vallotton L, Lazor-Blanchet C, Doce J, Puentes E, Marinova D, Aguilo N, Martin C.** 2015. Safety of human immunisation with a live-attenuated Mycobacterium tuberculosis vaccine: a randomised, double-blind, controlled phase I trial. *Lancet Respir Med* **3:**953-962.

OBJECTIVES

The general objectives of the Thesis are:

New TB vaccines should confer protection against all lineages causing TB in humans. Three of *M. tuberculosis* lineages, 2, 3 and 4, named modern lineages, are worldwide spread. We construct two vaccine candidates based on the same deletions than MTBVAC and study lineage-dependent protection. Results obtained with these vaccines candidates could anticipate their use as either a trivalent vaccine or in the context of specific geographic distribution of the different lineages.

Additionally, the recent discovery that the secondary messenger c-di-AMP impacts on immune signaling in *M. tuberculosis*, led us to study the involvement of this molecule in the immune responses triggered by MTBVAC.

The specific objectives are:

- Construction of *phoP* and *fadD26* deletions in two clinical isolates of lineages 2 and 3 of *M. tuberculosis*, named MTBVAC-L2 and MTBVAC-L3, fulfilling the Geneva consensus.

- Molecular characterization of the new MTBVAC-L2 and MTBVAC-L3 vaccine candidates to confirm previously described PhoP- and FadD26-dependent phenotypes.

- Preclinical studies in immunocompromised mice to study the attenuation of MTBVAC, MTBVAC-L2 and MTBVAC-L3.

- Preclinical studies to evaluate the protective efficacy of MTBVAC, MTBVAC-L2 and MTBVAC-L3 after the challenge with virulent strains belonging to the three modern lineages of *M. tuberculosis*.

- Deletion of genes in MTBVAC that codes the cyclase and phosphodiesterase involved in c-di-AMP synthesis and degradation respectively. Evaluation of IFN-β and IL-1β response in infected macrophages.

Chapter 1

Construction and characterization of new double *phoP* and *fadD26* mutants in two clinical isolates of modern lineages 2 and 3 of *M. tuberculosis*

Abstract

MTBVAC is a live attenuated vaccine based on deletions in *phoP* and *fadD26* genes in the clinical isolate Mt103, which belongs to lineage 4 of *M. tuberculosis* (Euro-American lineage). Both genes are involved in virulence of the bacilli; *phoP* codes for the transcription factor PhoP which controls approximately the expression of the 2% of the genes of *M. tuberculosis*, and *fadD26* gene is required for PDIM synthesis. MTBVAC is one of the most promising TB vaccine candidates and the only live attenuated vaccine based on the human pathogen *M. tuberculosis*.

TB in humans is caused by *M. tuberculosis sensu stricto* (comprised by lineages 1-4 and 7) and *M. africanum* (lineages 5 and 6). Lineages 2, 3 and 4 (L2, L3 and L4) of *M. tuberculosis* are named "modern" lineages because of the deletion of TbD1 region while the others are named "ancient" lineages. Modern linages, in particular lineage 2 and lineage 4 are the most globally widespread according to their geographic distribution while lineage 3 presents an intermediate distribution, being present mostly in East Africa, Central- and South-Asia.

Considering the encouraging results of MTBVAC as vaccine candidate, *fadD26* and *phoP* deletions were obtained in two clinical isolates of lineages 2 (East-Asian lineage) and 3 (East-African-Indian lineage) of *M. tuberculosis*, named MTBVAC-L2 and MTBVAC-L3. Therefore, three MTBVAC vaccine candidates were obtained in clinical isolates of the modern lineages of *M. tuberculosis.* Previous phenotypes described in MTBVAC were confirmed in the new double mutants MTBVAC-L2 and MTBVAC-L3*:* inability to fix neutral red, downregulation of PhoP regulon, secretion of CFP-10 in contrast to the inability to secrete ESAT-6.

Introduction

Evolution of the *M. tuberculosis* Complex

M. tuberculosis Complex (MTBC) consists of a clonal group of mycobacterial lineages, with 99.9% of nucleotide identity and pathogenic to a wide range of mammalian hosts. MTBC comprises lineages adapted to humans for which are the unique known reservoir, called "human-adapted" mycobacteria and "animal-adapted" mycobacteria that affect wild and domestic mammalian hosts (1-3) (Figure 15A).

Mycobacteria adapted to infect humans can be classified in seven lineages and comprise lineages 1-4 and lineage 7 of *M. tuberculosis sensu strictu* and lineages 5 and 6 of *M. africanum*. The geographical distribution of these lineages differs considerably (Figure 15B); some of them are globally widespread whereas others are geographically restricted (4). Lineage 4 (Euro-American lineage) and lineage 2 (East-Asian lineage) of *M. tuberculosis* are the most globally widespread and considered the most transmissible lineages. Lineage 1 (Indo-Oceanic) and lineage 3 (East-African-Indian) of *M. tuberculosis* show intermediate distributions. These lineages are geographically associated all around the Indian Ocean and East Africa, Central- and South-Asia respectively. The remaining lineages, lineage 7 of *M. tuberculosis* and lineages 5 and 6 of *M. africanum*, called *M. africanum* West Africa 1 and West Africa 2 respectively are geographically restricted (4, 5). Lineage 7 occurs particularly in Ethiopia and lineages 5 and 6 occurs specially in West-Africa (6).

Additionally, human-adapted mycobacteria can be divided in "modern" or "ancient" lineages based on the deletion in modern lineages of the region of the genome called, *M. tuberculosis* specific deletion 1 (TbD1), which comprises *mmpS6* and *mmpL6* genes. Hence, lineages 2, 3 and 4 of *M. tuberculosis* are named modern lineages whereas lineages 5 and 6 of *M. africanum* and lineages 1 and 7 of *M. tuberculosis* are known as ancient lineages (7).

Differences in immune response and disease progression are found in modern and ancient lineages. Modern lineages produce lower proinflammatory response compared to ancient lineages in human monocyte-derived macrophages (Figure 16A) (8). Relation between low inflammatory innate immune response and increased virulence in animal models has been previously described (9, 10). The hypothesis is that the reduction of immune response observed in modern lineages could contribute to a fast progression to active disease and increased transmission. The notion is that modern lineages have evolved in the context of high population densities and these lineages have adapted to large host populations and became more virulent. By contrast, ancient lineages have evolved in low population densities until very recently and these lineages remain less virulent because of the limited host availability (Figure 16B) (8).

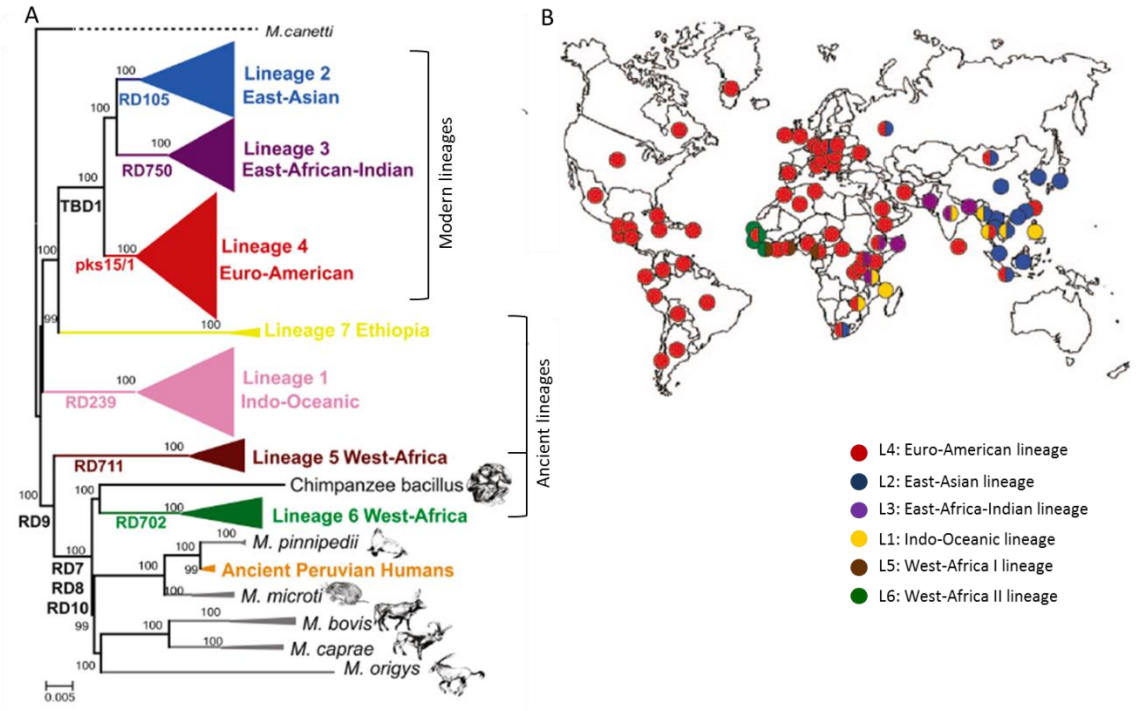


Figure 15. MTBC and geographical distribution of human adapted mycobacteria. (A) Maximum likelihood phylogeny, adapted from (11). Large Sequence Polymorphisms (LPs) indicated along branches are described in (7) (B) Geographical distribution of human adapted mycobacteria, adapted from (4).

However, differences and controversial results have been described in the immune response produced by strains belonging to lineages 2, 3 and 4. Some studies observed lower levels of proinflammatory response in lineage 2 strains compared to lineage 4 strains in monocyte-derived macrophages and peripheral blood mononuclear cells (12, 13). The opposite was observed in bone-marrow-derived murine macrophages (14). These controversial results could be explained by the different cell lines used in the experiments, and the variability of strains in the lineages.

A recent study of lineage 4 confirms the intra-lineage heterogeneity of lineage 4, which is constituted by 10 sublineages that present high differences in geographical distribution. Three sublineages (L4.1.2/Haarlem, L4.3./LAM and L4.10/PGG3) are widely widespread and named "generalists" and can cause disease in many populations. By contrast, L4.6.1/Uganda and L4.6.2/Cameroon sublineages are geographically restricted to certain countries and seem to be specialized to infect specific populations, and are named "specialists". The five remaining sublineages present an intermediate geographical distribution. MTBVAC was obtained from the clinical isolate Mt103 which belongs to sublineage L4.10/PGG3, considered a generalist sublineage and the third most frequent of the ten sublineages of *M. tuberculosis* lineage 4 (11.9%) according to Stucki and colleagues (15).

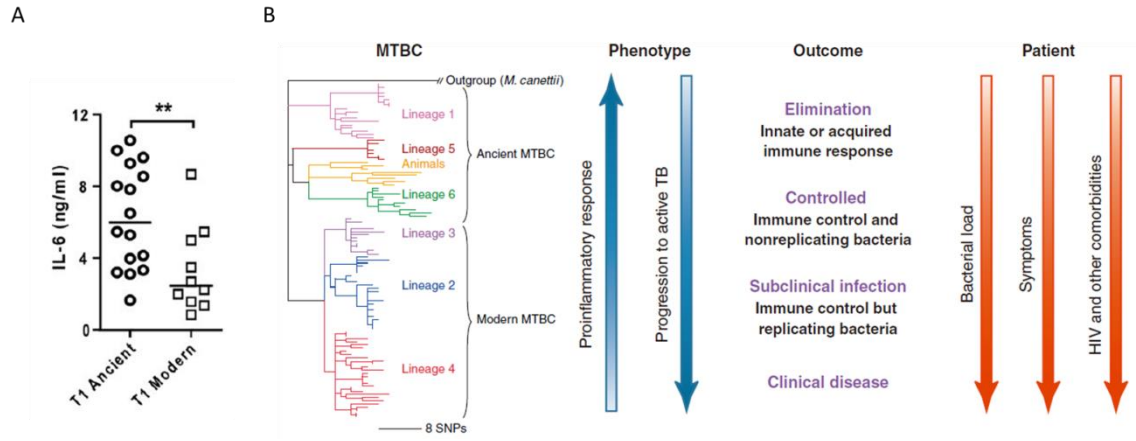


Figure 16. Spectrum of responses of human adapted mycobacteria. (A) IL-6 detection in supernatants of macrophages 24 h post-infection from 3 independent donors. Same trend was observed other cytokines including, IL-12p40/70 and TNF-α. (B) It is suggested that modern lineages elicit a delayed innate immune response and progression to an active disease is faster than in ancient lineages. From (16)

Besides human adapted mycobacteria, MTBC comprises animal-adapted mycobacteria. Animal-adapted mycobacteria can be considered as "ecotypes" and are constituted of different mycobacteria adapted to infect a specific mammal. Differences in genomics have been observed among them, with different RDs and SNPs. Mycobacteria ecotypes are: *Mycobacterium microti*, causative agent of TB in rodents, *Mycobacterium mungi*, adapted to infect mongoose, *Mycobacterium suricattae*, and cause TB in meercats, *Mycobacterium orygis, Mycobacterium caprae, M. pinnipeddi and M. bovis,* which are mainly adapted to infect oryxes, caprine, seals and cattle respectively (3).

Lineage 2 *of M. tuberculosis* – East Asian

Lineage 2 of *M. tuberculosis* (East-Asian) is one of the most widespread lineages and is geographically associated with East-Asia, Russia and South Africa. Lineage 2 has been extensively classified by eight different studies, based on insertion of IS*6110,* RDs, SNPs database, VNTR or WGS, reviewed by Shitikov and colleagues (17).

All lineage 2 strains are defined by RD105 or a mutation at position (pos.) 4974491(G>A), and the lineage can be classified in two major phylogenetic clades. The first clade is named lineage 2.1, proto-Beijing or group 1 by Luo, Shitikov or Tsolaki and Gagneux and colleagues respectively (4, 17-19). The second major clade, named lineage 2.2 by Coll, Luo and colleagues (18, 20), is composed of all Beijing strains/Beijing family, which are the majority of lineage 2 strains. Beijing strains are defined by the RD207 deletion, mutation at pos. 2505085 (G>A) (20) and pos. 4280708 (G>A) (21) and can be divided in several sublineages. The next division point is the presence or deletion of RD181 or three different SNPs described in different studies. Presence of this region defines Asian Ancestral 1 by Merker and colleagues (22) or lineage 2.2.1 by Coll and colleagues (20). In the next level, mutation at pos. 886115 (T>C) defines Asian Ancestral 2 lineage, while Asian Ancestral 3 is defined by mutation at pos. 1477522 (C>A) by

Merker and colleagues (22). All these Beijing sublineages described until now are referred as ancestral Beijing strains.

The next phylogenetic point is the insertion of IS*6110* in the noise transfer region (NTF), which defines modern Beijing clade (23) and comprise seven groups taking into account all the studies, including Asian African 1, Europe/Russia BO/W148 outbreak, Central Asia, Asian African and Pacific RD150 (22) (Figure 17).

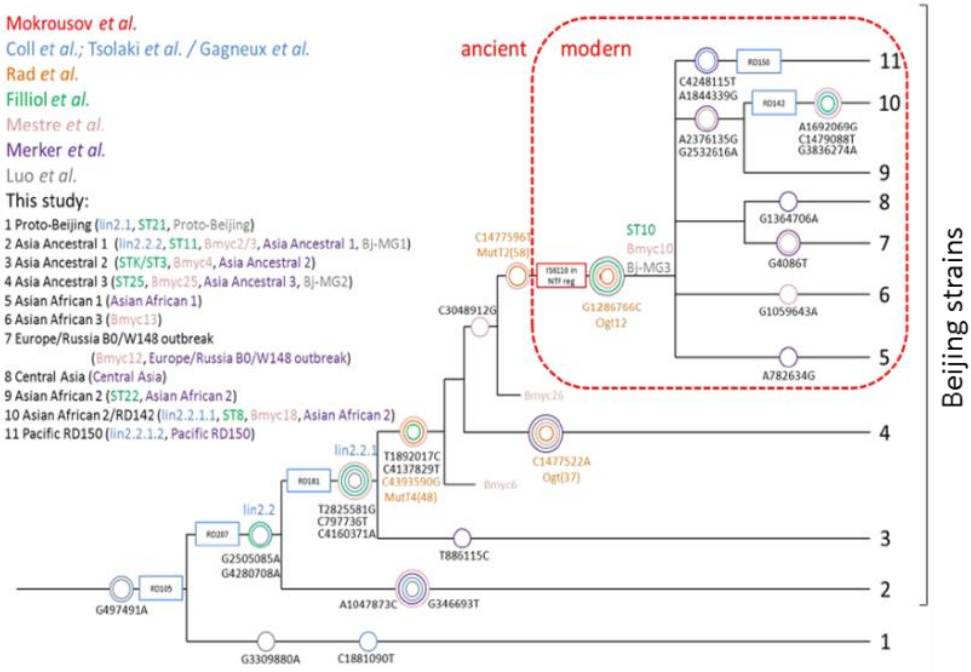


Figure 17. Scheme of sublineages of lineage 2 of *M. tuberculosis*. Representation of the major phylogenetic groups and genetic markers of lineage 2 of *M. tuberculosis* based on different studies. Adapted from (17).

Beijing genotype family represents 50% of strains in East Asia and at least 13% of strains worldwide (24). Beijing strains are usually associated with hyper-virulent phenotypes, increased pathogenicity, rapid progression to active disease, drug resistance and recent major TB outbreaks (25-28). Highest content of IS*6110* were observed in Beijing strains. Additionally, lineage-specific content of IS*6110* was observed being higher in modern lineages (29).

Some molecular characterizations have been associated with the increased virulence observed. Presence of PGL in a subset of Beijing strains, although its presence is variable in Beijing strains, is associated with increased virulence (10). Intact *pks1-15* cluster is essential for PGL biosynthesis. A 7-base-pair deletion in H37Rv and CDC5115, both from lineage 4 of *M. tuberculosis,* lead to the absence of PGL in these strains (30).

Other phenotype described is that some Beijing strains show higher levels of expression of DosR, which is a regulon of 48 co-regulated genes controlled via DosR-DosS/DosT two-component system and it is required for MTBC persistence (31, 32). Duplication gene and SNP immediately upstream of DosR are suggested to be associated with its augmented expression (33, 34).

In the mouse model, several studies have observed an increased virulence of Beijing strains. Beijing strains showed higher mortality and increased bacterial burden than H37Rv after intratracheal injection in BALB/c mice (35, 36) and lower survival was also observed in Beijing strain compared to CDC1551 in B6D2/F1 mice challenged by aerosol route (9).

One of the hypothesis of the emergence of Beijing strains is the possibility of less protection of BCG against these Beijing strains than against others or that BCG vaccination may have a selective factor that favored a genotypic and phenotypic characteristic acquired by Beijing family (37, 38). Controversial results have been obtained in the mouse model and epidemiologic studies. In mice, reduced protection of BCG was observed against Beijing strains compared to challenge with other *M. tuberculosis* strains from other lineages in BALB/c mice (36) whereas this phenotype was no observed in C57BL/6 mice (39, 40). In epidemiologic studies, disagreement conclusions were obtained. In some studies, BCG vaccinated patients were more frequently infected with Beijing strains than non-Beijing strains in Vietnam, Hong Kong and the Netherlands (41, 42). However, no association was found in other studies between BCG vaccination and the spread of Beijing strains in Vietnam and Indonesia (43, 44). Therefore, the hypothesis of BCG vaccination and emergence of Beijing strains remains unresolved (45).

- **Differences between modern and ancient Beijing**

Ancestral or modern Beijing differs in the presence or absence of IS*6110* in NTF region, (46), mentioned before in the classification of lineage 2

Differential molecular characterization, specifically organization of *ppe38* locus, has been recently described between modern and ancient Beijing strains. Modern Beijing strains contains an IS*6110* that truncates *ppe38* (mt2419) and *ppe71* (mt2422) genes and deletes *esxXY* genes (which are localized between *ppe38* and *ppe71*). By contrast, ancient Beijing strains used in the study contains an intact copy of *ppe71* and *esxXY* genes and an insertion of IS*6110* in *ppe38*. Modern Beijing strains have lost the ability to secrete PE_PGRS proteins while ancient Beijing strain can secrete them because of the differences in *ppe38* locus. Lack of secretion of these proteins is linked to increased virulence in mice (47).

Differential induction of proinflammatory cytokines was observed between modern and ancient Beijing strains. Induction of IL-1β, IFN-γ and IL-22 was significantly lower in peripheral blood mononuclear cells (PBMCs) by modern Beijing strains compared to ancient Beijing isolates (13). Same phenotype was observed in human-monocytes derived macrophages (48). In C57BL/6 mice, three different patterns were observed when modern and ancient Beijing strains were studied. Some modern Beijing exhibited higher virulence than ancient Beijing strain. However, intermediate phenotype was observed in one modern and one ancient Beijing strain (49).

In general, modern Beijing group is more likely to exhibited increased virulence in comparison with ancient Beijing strains, although both groups can exhibit high virulent phenotypes. For example, similar or higher levels of virulence were observed in ancient Beijing strains compared to modern Beijing strains in guinea pigs (50).

Lineage 3 of *M. tuberculosis* - East-African-Indian

Lineage 3 of *M. tuberculosis* (East-African-Indian) presents and intermediate geographical distribution and is less widespread than lineages 2 and 4 of *M. tuberculosis.* This lineage is mostly geographically associated with East Africa, Central and South-Asia.

M. tuberculosis lineage 3 strains are characterized by the genomic deletion RD750. Lineage 3 includes sublineages for CAS1-Delhi, CAS1-Kili, CAS and CAS2 spoligotypes. All non-CAS1-Delhi samples are grouped in the same sublineage 3.1. CAS1-Delhi and CAS1-Kili were the most frequent sublineages of the studied strains (20).

In vitro cultures showed that lineage 3 *M. tuberculosis* strains reach an early plateau compared to strains from lineages 2 and 4 *M. tuberculosis* and the growth was lower in lineage 3 strains. Lower growth in human monocyte derived macrophages was also observed in lineage 3 strains compared to lineages 2 and 4 isolates although no differences in bacteria uptake were observed among all the modern strains (12).

Intermediate phenotype of IL-6 production was observed in human monocyte-derived macrophages in response to lineage 3. Lowest levels of IL-6 were produced by lineages 2 and 4 strains, followed by lineage 3 and strains, and the highest levels of this interleukin were observed in the ancient lineage 6 (8).

In this chapter, we describe two new *phoP* and *fadD26* deleted strains in two clinical isolates of lineages 2 and 3 of *M. tuberculosis*. In agreement with the Geneva consensus, antibiotic markers were eliminated. Finally, we characterize previously described PhoP and FadD26 phenotypes.

Material and methods

Bacterial strains, media and growth conditions

Escherichia coli MC1061, DH5α, DH10B, XL1-BLUE and BW25141 strains were grown at 37°C in Luria-Bertani (LB) broth. When required, media were supplemented with ampicillin (Amp, 100 μg/ml), kanamycin (Km, 20 μg/ml), hygromycin (Hyg, 50μg/ml) or chloramphenicol (Cm, 12.5 μg/ml); 0.15% arabinose was used for induction. For cultivation in solid media, LB agar was used.

E.coli DH10B containing a BAC library of H37Rv chromosome cloned in pBeloBAC were grown at 30°C when the thermosensitive pKD46 plasmid was contained.

Mycobacterial strains were cultured at 37 °C in Middlebrook 7H9 (Difco™) liquid medium supplemented with 10% (vol/vol) ADC (0.2% dextrose, 0.5% bovine serum albumin, 0.085% NaCl and 0.0003% beef catalase) and 0.05% of Tween 80. For cultures in solid media, Middlebrook 7H10 agar (Difco™) were used supplemented with 10% (vol/vol) of ADC and 0.05% Tween 80 with the exception of HMS13037 lineage 3 strain in which was supplemented with OADC (ADC with 0.005% oleic acid). When necessary Hyg 20 μg/ml, Km 20 μg/ml or Gm 5 μg/ml were added to media. Catechol solution was obtained to test XylE activity: 0.55 g of catechol were dissolved in 47.5 ml of water and 2.5 ml of PBS. 2% (w/v) sucrose (Suc) was added to 7H10 supplemented with ADC or OADC for the negative selection in mutant constructions. 0.2% acetamide was used for induce the expression of *gp60/61* contained in the pJV53 plasmid.

Mycobacterial strains used and mutants obtained are detailed in Table 3.

Manipulation of *E. coli* and mycobacterial strains was performed in a biosafety level 1 laboratory (BSL1), in a biosafety level 2 laboratory (BSL2) or biosafety level 3 laboratory (BSL3), with facilities notification A/ES/10/I-05, A/ES/06/I-02 and A/ES/04/I-05 respectively.

Mycobacterial DNA extraction

Mycobacteria were resuspended in 400 μl of TE (100 mM Tris/HCl, 10 mM ethylenediaminetetraacetic acid, EDTA, pH=8) and heated at 85 °C during 20 min. Samples were incubated at 37 °C overnight after addition of 0.5 mg of lysozyme. Subsequently, 0.05 mg of proteinase K and 72.5 μl of 10% SDS were added and incubated at 65 °C for 10 min. 100 μl of CTAB (Cetyltrimetylammonium bromide)/sodium chloride, NaCl. 10% CTAB in 0.7 M of NaCl, pre-warmed at 65 °C was added to the samples and incubated at 65 °C for 10 min. The aqueous phase were collected after adding 750 μl of chloroform:isoamyl alcohol 24:1 (v/v), vortex for 10 s and centrifuged for 5 min at 12000 rpm. Genomic DNA was precipitated by adding 0.6 volumes of isopropanol and an incubated of 30 min at -20 °C. Precipitated nucleic acids were collected by centrifugation (5 min at 12000 rpm) and washed with 70% ethanol by tube inversion. Nucleic acids were collected again by centrifugation (12000 rpm for 5 min). Pellets were dissolved in 50

µl of double-distilled water after the complete evaporation of the ethanol. DNA was quantified by Abs_{260}/Abs_{280} readings using a ND-1000 spectrophotometer (NanoDrop)

Plasmid extraction (mini-prep)

Bacteria were resuspended in 500 µl of TE. Samples were incubated at room temperature after addition of 100 µl of Solution I (50 mM Glucose, 10 mM EDTA, 25 mM Tris/HCl pH=8 and 4 mg/ml of lysozyme). Subsequently, 200 µl of fresh Solution II (0.2 M sodium hydroxide, NaOH and 1% of sodium dodecyl sulfate, SDS) were added and tubes were inverted several times to favor the lysis of the bacteria. After 5 min of incubation on ice, 150 µl of cooled Solution III (3 M potassium acetate, KAc, 11.5% v/v acetic acid glacial) were added and incubated on ice during 5 min after vortex the mixture.

The aqueous phase (upper) was mixed with 500 µl of chloroform:isoamylic alcohol 24:1 (v/v) by inversion. After centrifugation (12000 rpm for 5 min), the upper phase was added to a fresh tube containing 0.1 volumes of 3 M KAc and 2 volumes of absolute ethanol and incubated at -20 °C for 30 min. Plasmid DNA was pelleted by centrifugation (12000 rpm for 30 min at 4 °C) and washed with 70% ethanol. Hereafter, pellet was dried in a vacuum drier and dissolved in 50 µl of double-distilled water. To remove co-purified RNA, samples were incubated at 37 °C for 15 min with 1 µl of RNase A (1 mg/ml).

Plasmid extraction was also performed using High Pure Plasmid Purification Kit (Roche).

Polymerase chain reaction (PCR)

PCR amplifications for constructions and genomic analyses were performed in a final volume of 10µl or 50µl with 0.38 and 1.87 U of Mytaq DNA polymerase respectively (Bioline), 25 µM of each primer (Table 4) and MyTaq Reaction Buffer 5x. PCR program consisted on an initial step of 1 min at 95 °C, followed by 30 cycles (95 °C 15 s, 58 °C 15 s and 72 °C 1 min/kb) and a final step of 7 min at 72 °C. For colony PCR, initial step was 95 °C for 10 min. PCR amplification products were analyzed in 0.8-2% agarose gel electrophoresis. For molecular weight comparison, DNA Molecular Weight 100 bp or λ *PstI* were used. PCR products were purified by High Pure PCR product purification Kit (Roche).

Table 3. Mycobacterial strains used classified by lineages.

L	Strains	Description	Reference
Lineage 4	Mt103	Clinical isolate of *M. tuberculosis* lineage 4	(51)
	MTBVAC	Double unmarked mutant of *phoP* and *fadD26* in Mt103 strain	(52)
Lineage 2	GC1237	Clinical isolate of *M. tuberculosis* lineage 2	(53)
	GC1237 Δ*fadD26::hyg*	Single mutant marked of *fadD26* in the clinical isolate GC1237	This study
	GC1237 Δ*fadD26*	Single mutant unmarked of *fadD26* in the clinical isolate GC1237	This study
	GC1237 Δ*fadD26* Δ*phoP::hyg* (MTBVAC-L2::*hyg*)	Double mutant, *phoP::hyg* and unmarked *fadD26* in the clinical isolate GC1237	This study
	GC1237 Δ*fadD26* Δ*phoP* (MTBVAC-L2)	Double mutant unmarked of *phoP* and *fadD26* in the clinical isolate GC1237	This study
Lineage 3	HMS13037	Clinical isolate of *M. tuberculosis* lineage 3, from Hospital Universitario Miguel Servet,	This study
	HMS13037 Δ*fadD26::km*	Single mutant marked of *fadD26* in the clinical isolate HMS13037	This study
	HMS13037 Δ*fadD26*	Single mutant unmarked of *fadD26* in the clinical isolate HMS13037	This study
	HMS13037Δ*fadD26* Δ*phoP::hyg* (MTBVAC-L3::*hyg*)	Double mutant, *phoP::hyg* and unmarked *fadD26* in the clinical isolate HMS13037	This study
	HMS13037 Δ*fadD26* Δ*phoP* (MTBVAC-L3)	Double mutant unmarked of *phoP* and *fadD26* in the clinical isolate HMS13037	This study
	BCG Pasteur	Vaccine against TB, attenuated strain of *M. bovis*	(54)

Protein Extraction

Mycobacteria were grown in Middlebrook 7H9 (Difco™) supplemented with Tween 80 and ADC or OADC. Bacteria were pelleted and washed twice with PBS to remove albumin and resuspended in PBS. Bacteria were transferred to Middlebrook 7H9 (Difco™) supplemented with 0.05% Tween 80 Dextrose/NaCl with catalase (dextrose, NaCl and catalase in the absence of albumin) to avoid interference of albumin in the secreted fraction. 50 ml of bacterial cultures were grown until late log-phase (OD_{600nm} ~0.8). Bacteria were pelleted by centrifugation.

The supernatant fractions were filtered through a 0.22 μm-pore-size low binding protein and 10% (v/v) of trichloroacetic acid was added and incubated on ice for 2 h to favor protein precipitation. Hereafter, samples were centrifuged (1 h at 4000 rpm at 4°C) and pellets were washed with acetone. Supernatants were discarded after centrifugation (4000 rpm for 10 min) and pellets were air dried and dissolved in 250 μl of Tris 150 mM pH=8.8.

For extraction of cellular proteins, bacteria were resuspended in 1 ml of PBS containing 1% Triton X-100 and transferred to tubes containing glass beads (MP Biomedicals). Mycobacterial suspensions were disrupted by Fast-Prep (6.5 m/s, 45 s) twice and samples were incubated on ice for 5 min between the cycles. Supernatants containing the whole-cell lysate were collected after centrifugation (12000 rpm for 5 min at 4 °C).

Proteins samples were quantified using the QuantiPro BCA assay (Sigma Aldrich) and stored at -20 °C.

Gel electrophoresis of proteins

Samples were heated at 100 °C for 10 min in presence of laemmli sample buffer (150 mM Tris/HCl pH=7.4, 3% SDS, 0.3 mM sodium molybdate, 30 mM sodium pyrophosphate, 30 mM sodium floride, NaF, 30% glycerol, 30% mercaptoethanol and 0.06% bromophenol blue). Equivalent quantities of proteins were separated in 12-17% polyacrylamide gels containing 0.1% SDS in running buffer (25 mM Tris, 192 mM glycine, 3.4 mM SDS) at constant amperage of 30 mA per gel.

PageRuler Plus Prestained Protein Ladder (Thermo Scientific) was used as molecular weight marker (kDa).

- **Western blot analysis**

Proteins were transferred to a PDFV membrane (preactivated with methanol) in transfer buffer (48 mM Tris/HCl pH=8.3, 39 mM glycine, 0.037% SDS, 20% methanol) for 1 h at 20 V with a semi-dry electrophoretic transfer cell (Trans-Blot® Semi-Dry Transfer cell, Bio-Rad).

Proteins were blocked with 5% (w/v) of skimmed milk in TBS-T buffer (25 mM Tris pH=7.5, 150 mM NaCl, 0.05% Tween 20) for 30 min and incubated overnight with primary antibodies. Subsequently membranes were washed with TBS-T buffer three times before incubation with human serum adsorbed secondary antibodies conjugated with peroxidase for 1 h. Signals were detected using chemiluminescent substrate (Western Bright™ Quantum, Advansta).

Immunodetection was carried out using PhoP-antiserum (1:5000), antibodies anti-sigA (1:5000) and anti-CFP-10 (Thermo Scientific) (1:5000) with the incubation with secondary antibody anti-rabbit IgG human serum adsorbed conjugate (1:5000) (KPL).

Incubation with monoclonal antibody anti-ESAT-6 (1:2500) (abcam), antibody anti-GroEL2 (Hsp65) (1:2500) (Invitrogen) or anti-PE_PGRS (1:2000) followed by incubation with anti-mouse IgG Human Serum Adsorbed conjugate (1:5000) (KPL) were used to detect ESAT-6, GroEL2 and PE_PGRS proteins respectively.

To reprobe blots, membranes were incubated at room temperature for 15 min with ReBlot Plus Strong Antibody Stripping Solution (Millipore).

- **Coomassie staining**

Gels were stained with Colloidal Blue Staining kit NOVEX®, Invitrogen.

Preparation of electrocompetent cells of *E.coli*

50 ml of bacterial culture were grown until $OD_{600\ nm}$ 0.4-0.6 at 37 °C (in case of *E.coli* containing pKD46, bacteria were grown at 30 °C and induced with 0.15% arabinose w/v at $OD_{600\ nm}$ 0.1). When the optimal OD was reached, bacteria cultures were incubated on ice for 30 min. After two washes with chilled water and one with chilled 10% glycerol, bacteria were resuspended in 500 µl of 10% glycerol. Electroporation of 500 ng of double strand DNA or 250 ng of plasmid DNA were performed in 50 µl of aliquoted competent cells in a 0.2 cm cuvette (Bio-Rad) with a single pulse in a GenePulser Xcell™ (Bio-Rad)(2500 V, 25 µF, 200 Ω). Cells were recovered in 1 ml of LB and incubated at 37 °C for 1 h to allow the expression of the resistance to the antibiotic and then serial dilutions were plated onto LB containing the required antibiotic.

Preparation of electrocompetent cells in *Mycobacterium tuberculosis*

200 ml of bacterial cultures were grown until OD_{600nm} 0.6-0.8 and 20 ml of 2 M glycine were added and incubated overnight at 37 °C. After two washes with water-0.05% Tween-80 and one with 10% glycerol-0.05% Tween-80, bacteria were resuspended in 2 ml of 10% glycerol-0.05% Tween-80. Aliquots of 400 µl were transformed with plasmid DNA in 0.2 cm cuvettes (Bio-Rad) with a single pulse in a GenePulser Xcell™ (Bio-Rad) (2500 V, 25 µF, 1000 Ω). Bacteria were recovered in 1 ml of 7H9-ADC-0.05% Tween 80 and incubated at 37 °C overnight to allow the expression antibiotic resistance. Serial dilutions were plated onto 7H10-ADC containing the required antibiotic for selection.

Preparation of electrocompetent cells in *Mycobacterium tuberculosis* for recombineering

200 ml of *M. tuberculosis* strains containing pJV53H plasmid were grown until OD_{600nm} 0.6-0.8. Acetamide were added to a final concentration of 0.2% (w/v) and incubated at 37 °C overnight. Bacteria were resuspended in 2 ml of 10% glycerol-0.05% Tween 80 after four washes with 10% glycerol-0.05%. Aliquots of 400 µl were transformed with 500 ng of purified PCR amplification product in 0.2 cm cuvettes (Bio-Rad) with a single pulse in a GenePulser Xcell™ (Bio-Rad) (2500 V, 25 µF, 1000 Ω). Bacteria were recovered in 1 ml of 7H9-ADC-0.05% Tween 80, incubated overnight at 37 °C and plated onto 7H10-ADC with the required antibiotic.

Neutral red staining

Bacteria were grown in 7H10-ADC and transferred to a tube containing 4 ml of 50% methanol in water and incubated for 1 h at 37°C. Bacteria were centrifuged (4000 rpm for 5 min) and resuspeded in 750 µl of barbital buffer (1% sodium barbital in 5% NaCl, pH=9.8). Hereafter, bacteria were resuspended in 4 ml of barbital buffer after centrifugation (4000 rpm for 5 min). 150 µl of a solution of 0.05% neutral red in barbital buffer were added and incubated at 37 °C for 1 h. Fixation of neutral red was evaluated after the supernatants were discarded.

RNA isolation from Mycobacteria

Mycobacteria cultures were grown in 7H9-ADC-0.05% Tween-80 at 37 °C until log-phase (OD_{600nm} 0.6-0.7). Bacteria were centrifuged (3000 rpm for 10 min) and pellets were resuspended in 250 µl in wash buffer (aqueous solution containing 0.137 M NaCl and 0.5% Tween 80) and 500 µl of RNA protect reagent (Quiagen) to avoid RNA degradation. Suspensions were centrifuged (14000 rpm for 5 min) after incubation at room temperature for 5 min. Pellets were dissolved in lysis buffer (20 mM NaAc, 0.5% SDS, 0.1 mM EDTA) and 1 ml of phenol:chloroform (5:1), pH=4.5. Bacterial suspensions were placed in tubes containing glass beads (MP Biomedicals) and cells were lysed by mechanical fraction (Fast-Prep) (2 cycles, 45 s at speed 6.5 m/s each cycle and samples were cooled between cycles by incubation on ice). Tubes were centrifuged and aqueous phases (upper) were transferred to a tube containing 900 µl of chilled chloroform:isoamylic alcohol (24:1). After centrifugation (14000 rpm for 5 min at 4 °C) upper phases were transferred to a fresh tube containing 900 µl of isopropanol and 90 µl of 3 M sodium acetate, NaAc pH=5.5 and samples were incubated over night at -20 °C. Precipitated nucleic acids were collected by centrifugation (14000 rpm for 1 h at 4 °C). Supernatants were discarded and pellets were washed with 500 µl of 70% ethanol. Samples were again centrifuged and pellets were dissolved in diethylpyrocarbonate-treated water (RNase free). DNA was removed by two consecutive incubations with Turbo DNA-free (Ambion), by addition of 1 µl of DNase and incubating at 37 °C for 1 h. RNA was purified by adding phenol:chloroform (5:1), pH=4.5, and the previous steps were repeated to precipitate, collect, dry and dissolve the RNA in diethylpyrocarbonate-treated water (RNase free). Concentration of RNA was estimated by reading of Abs_{260}/Abs_{280} with the ND-1000 spectrophotometer (NanoDrop). RNA integrity was confirmed by agarose gel electrophoresis and RNA extractions were stored at -80 °C. Absence of DNA was confirmed by lack of amplification product by PCR using BCG2A and BCG2B primers (Table 4).

Quantitative Real time PCR (qRT-PCR)

Reverse transcription was performed using 500 ng of RNA, 2 µl of PrimeScript™ RT Master Mix (Takara) and RNase Free dH2O (Takara) until a volume of 10 µl per reaction. cDNA was obtained after 15 min at 37 °C and 15 s at 85 °C for enzyme inactivation.

For qRT-PCR reaction, TB Green Premix Ex Taq™ (Tli Rnase H Plus) (Takara) was used. Gene-specific primers at final concentration of 0.25 µM and cDNA obtained as described above diluted one tenth were used. The reaction was performed in the StepOne Plus Real Time PCR System (Applied Biosystems). Meting curves were performed for each pair of primers (Table 5) to verify that they produce a single specific product. PCR amplification program used was 95 °C for 10 min; 40 cycles of 95 °C for 15 s, 60 °C for 1 min).

Table 4. Primers used in the chapter for PCR

Primers	Sequence 5´→3´
fadD26up	CAACGCAAGACGACATGG
fadD26down	GCACCGTCTTGATGAAGC
res1 (CG)	CTAGAGCAACCGTCCGAAATATTATAA
res2 (CG)	GATCTCATAAAAATGTATCCTAAATCAAATATC
fadD26I1	CACGAATGTCATTGCCAATG
fadD26I2	GCTTGAGCATGACCTCTTCG
fadD26-P1-Fw	GCCTGACAGCACTGCATATACGTACATCGACTACGGATCCGTGTAGGCTGGAGC TGCTTC
fadD26-P2-rv	AACGAGATCGGCCACCCGCAGGCTGTGTGACTTCGATATCCATATGAATATCCTC CTTAGT
Confirm-fadD26-Fw	GGCTCGGGGGTCAACACCTG
Confirm-fadD26-Rv	GCTTCCCAGGCCACTTCCAGC
fadD26F	CTCGAGTTCTCTATCCGTGTATTC
fadD26R	CTCGAGGTTGGTCTTGACAG
KO-fadD26-Fw	GACACACACGGGGGCGGTG
KO-fadD26-Rv	CCGCGTCGAAGGCGTCGATG
P1-inv	GAAGCAGCTCCAGCCTACAC
P2-inv-long	CTTCGGAATAGGAACTAAGGAGGATATTCATATG
phoPup	ACATCGTCAGGCATCACC
phoPR	GATCACCCGAACGTAGAACC
KO-phoP-rv	GGTGGTCAAGTAGCCATCCGAC
RT-phoP-Fw	GCCTCAAGTTCCAGGGCTTT
phoP-I1	GCTCGACGAGGAGACCCAC
phoP-I2	GGTGAATTCGGTGGGCGACAG
phoRF	GGGCAAGGGCAACAAGGAAC
PPE38F	TTTTCGGTGTGGATTGTCT
PPE38R	GCCAGGGATTTCCAACGAC
PPE38IntF	ATGTCGGCGGAGTTGGGTAAG
PPE38IntR	TAGCCTGACCAGCCGACAACT
pKD46-gam-fw	CCTGTTTTCCTAATCAGCCCGGC
pKD46-bet-rv	AAATGCCGTCTGGCGAAGAGTG
Km-pKD4-out-1 (pKD13-2R)	CCACGATAGCCGCGCTGCCTCG
Km-pKD4-out2 (pKD13-1F)	GGGCTGACCGCTTCCTCGTGCT
gp60fw	ATCCGGCTCTACGCCGAC
gp61rv	CGGCAAATGACTCTTGCGT
BCG-2A	GCCGTCCATCCCGGGCATC
BCG-2B	CCATGTTCAAACCGGTGTC

Table 5. Primers used for RT-PCR

Primers	Sequence 5´→3´
RT-sigA-Fw	CCGATGACGACGAGGAGATC
RT-sigA-Rv	CGGAGGCCTTGTCCTTTTC
RT-mcr7-Fw	ACGCCGCGAGGACATG
RT-mcr7-Rv	AGGGAGCTGCTTGGACAGAA
RT-pks2-Fw	GCATCGGTGAAGACCAACTTC
RT-pks2-Rv	GATTACGTGGAACCACACCATGT
RT-pks3-Fw	GACGCTCGCTGAATCACAAA
RT-pks3-Rv	TCGCCGTGTGTCAGTCCTAC
ppsA1	GCCGAGAGGTCGGTGTGATGC
ppsB1	AGCATTCGCTGCTGCGGGTCC
ppsB-se1	TGTGGATGGACCAGAGAG

Plasmids for unmarked deletion in *phoP* and *fadD26* genes by suicide plasmid technique

Suicide plasmids pAZ5 and pAZ18 contains a copy of *fadD26* and *phoP* respectively disrupted by *res-Ωhyg-res* cassette from pWM27 (55). The counter-selectable marker *sacB* and the reporter gene *xylE* are also contained in both plasmids (52) (Figure 18).

To obtain unmarked mutants, the replicative plasmid pAZ20 containing γδ-resolvase was used. pAZ20 also contains Km and Gm resistance cassettes, the counter-selectable *sacB*, and *E.coli* and mycobacteria origins of replication (52) (Figure 19).

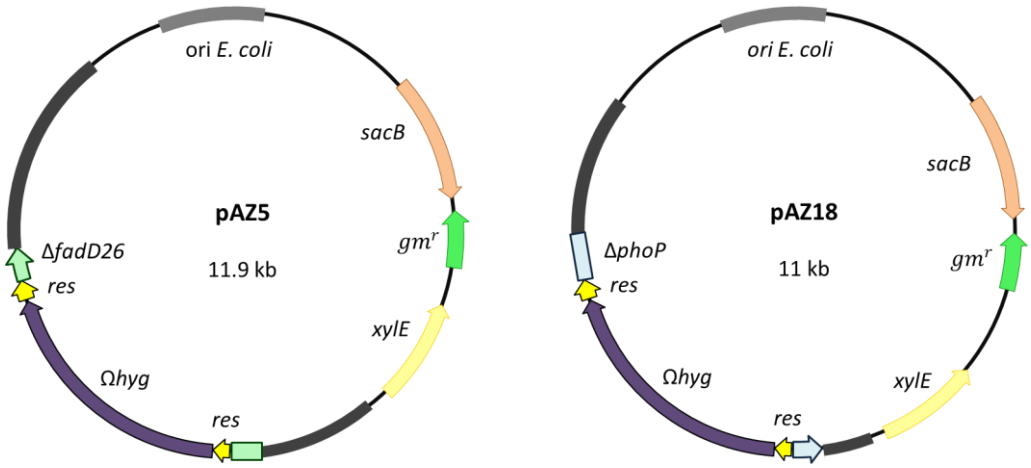


Figure 18. Suicide plasmids pAZ5 and pAZ18 for *fadD26* and *phoP* deletion. pAZ5 contains *fadD26* gene disrupted by Hyg gene resistance flanked by res sites while in pAZ18 the disrupted gene is *phoP*. Both plasmids contain the reporter gene *xylE* and the counter selectable marker. *sacB*. Adapted from (56).

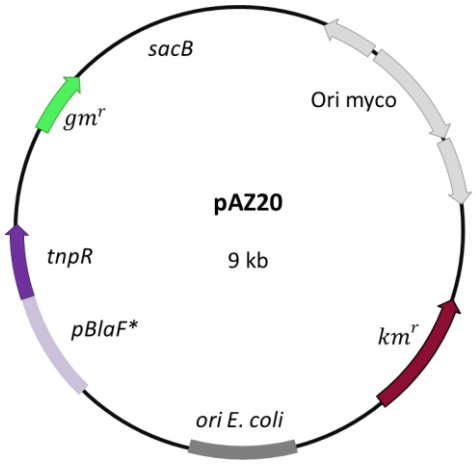


Figure 19. Components of pAZ20 plasmid. pAZ20 replicative plasmid contains *tnpR* gene, which codes for γδ-resolvase, the counter-selectable marker *sacB* and the resistance markers of Gm and Km. Adapted from (56).

Plasmids for unmarked gene deletion by BAC-recombineering

Thermosensitive plasmid pKD46 contains the Red system which contains three genes: γ, β and *exo*, whose products are Gam, Bet and Exo. Bet and Exo gain access to DNA ends to promote recombination and Gam inhibits host exonucleases. The expression of these genes is controlled by the presence of arabinose. pKD46 also contains Amp resistance cassette, RepA101 temperature sensitive origin in *E. coli,* and origin of replication for *E. coli* (57) (Figure 20).

pKD4 plasmid contains replicative origin for *E. coli* and Km cassette flanked with FLP recognition target (FRT) sequences. This plasmid is used as template for PCR amplification of FRT-*km*-FRT flanked with 40 bp homology arms to the target gene present in the BAC (57).

pJV53H, is a replicative plasmid that contains the mycobacteriophage Che9c-encoded proteins gp60 and gp61. Expression of gene product 60 (gp60) and gp61, phage exonuclease and recombinase, increase the recombination frequencies. The expression of these genes is induced by acetamide. Resistant cassette to Hyg and replicative origin of *E. coli* and mycobacteria are also contained in pJV53H (58, 59) (Figure 20).

pRES-FLP-Mtb is a replicative plasmid used to remove the Km cassette. This plasmid contains the *flp* recombinase from *Saccharomyces cerevisiae* with adapted codon usage of *M. tuberculosis* (60). The FLP recombinase acts on the directly repeated FRT that flanked the antibiotic cassette. Additionally Hyg resistance cassette and origins for *E. coli* and mycobacteria are present in the plasmid (Figure 21).

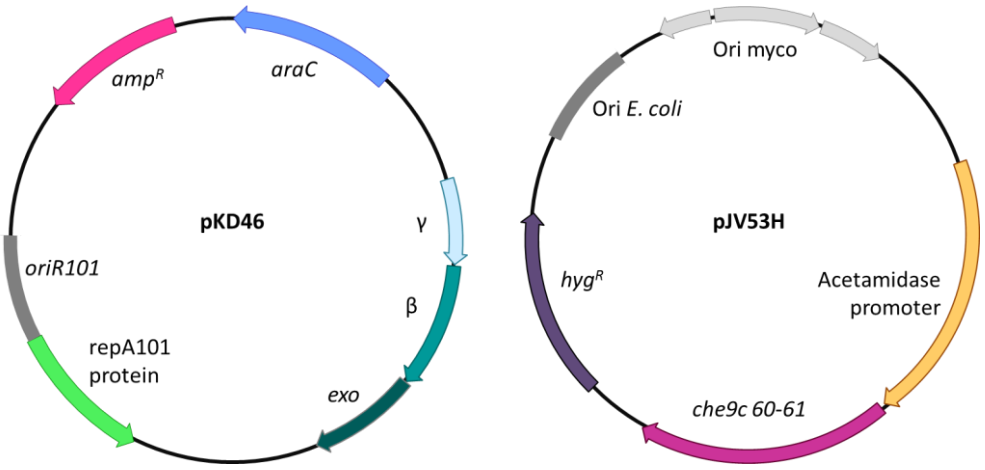


Figure 20. Components of pKD46 and pJV53H plasmids. (A) The replicative plasmid pKD46 contains the Red system that includes γ, β and *exo* preceded by arabinose-inducible promoter, termosensitive origin of *E.coli* and cassette of Amp resistance. (B) pJV53H, is a replicative plasmid that contains origins of replication of *E. coli* and mycobacteria, Hyg resistant cassette and mycobacteriophage Che9c gene products gp60 and gp61 whose expression is controlled by the presence of acetamide.

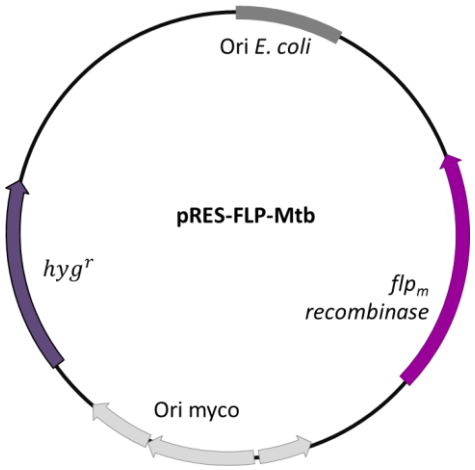


Figure 21. Components of pRES-FLP-Mtb plasmid. This replicative plasmid for *E.coli* and mycobacteria consist of the gene *flp* that codifies for the FLP recombinase, origins for *E.coli* and mycobacteria, and Hyg antibiotic resistance marker.

Results

Genotyping of *M. tuberculosis* strains belonging to lineages 2 and 3 for construction of double mutants

The construction of the double mutant in a lineage 2 strain was performed in the clinical isolate GC1237. GC1237 has been extensively described by the group and used in the laboratory. This strain was responsible of an outbreak in Canary Island in the 1990 decade. RD105, RD207 and RD181 are deleted in GC1237. Any copy of IS*6110* is present in NTF region and the study of *pks1/15* region showed intact gene in this strain suggesting that GC1237 could retain the ability to synthesize functional PGL (53). Consequently this clinical isolate is an ancestral Beijing strain that belongs concretely to Asian Ancestral 3 sublineage according to Merker and colleagues (22) or 2.2.1 according to Coll and colleagues (20).

For double mutant construction in a lineage 3 *M. tuberculosis* strain, the clinical isolate HMS13037 was used. HMS13037 was isolated in Aragón in 2013 from a Pakistani patient. This strain belongs to the sublineage CAS1-Delhi and this is the first time that this strain is used for laboratory manipulation.

For the classification of these strains three genotyping methods were performed in collaboration with Dr. Sofía Samper: IS*6110* RFLP revealed the presence of 19 copies of IS*6110* in GC1237 while HMS13037 contained 13 copies. Spoligotyping was performed for both strains; GC1237 showed deletion of spacers 1-34 in the spoligoprofile which is typical of Beijing genotype, and HMS13037 showed spoligo international type SIT26; 15 or 24 MIRU-VNTR (variable number of DNA tandem repeats) was also performed for GC1237 and HMS13037 respectively. Genotyping methods and previous studies confirmed that GC1237 strain belongs to lineage 2 of *M. tuberculosis* (East-Asian) and specifically to an ancestral Beijing sublineage. HMS13037 belongs to lineage 3 of *M. tuberculosis* belonging to sublineage CAS1-Delhi.

Construction of MTBVAC-L2, a double unmarked mutant Δ*fadD26* Δ*phoP* in GC1237

Suicide plasmids and BAC-recombineering strategies were used in parallel to obtain *phoP* and *fadD26* deletions in GC1237. First deletion was obtained in both genes *phoP* and *fadD26* by suicide plasmid strategy whereas no recombinants were obtained using BAC-recombineering strategy. We continued with single mutants GC1237 Δ*fadD26* and GC1237 Δ*phoP* to try to obtain the second deletion in *phoP* and *fadD26* respectively with the suicide plasmids pAZ18 and pAZ5 (Figure 22).

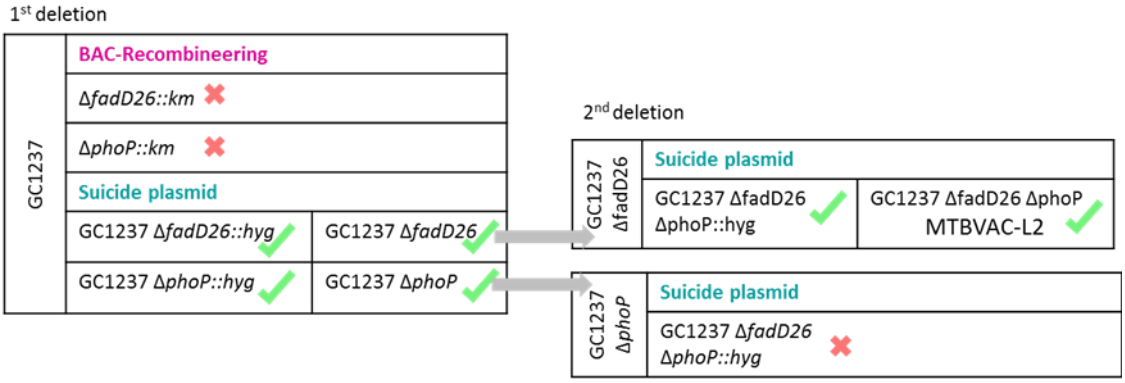


Figure 22. Results of BAC-recombineering and suicide plasmids strategies to obtained MTBVAC-L2. Single mutants of *fadD26* and *phoP* were obtained using suicide plasmids.MTBVAC-L2 was obtained using the suicide plasmid pAZ18 in GC1237 Δ*fadD26*.

Consecutive steps that led to MTBVAC-L2 construction (*fadD26* and *phoP* deletions obtained using pAZ5 and pAZ18 suicide plasmids respectively) are described in the next sections.

- **Construction of the single mutant marked strain, GC1237 Δ*fadD26::hyg***

GC1237 clinical isolate was transformed with 1 µg pAZ5 and plated onto 7H10-ADC containing Hyg. Allelic exchange was selected in two steps: a positive selection followed by a negative selection (52) (Figue 23A).

In the first step, colonies resulting from the pAZ5 transformation were tested for XylE activity. Catechol solution was added to the colonies and the ones that stained yellow (positive XylE activity) confirmed single recombination event and the integration of the plasmid into the genome of the mycobacteria (Figure 23B).

In the second step, single recombinants were grown in 7H9-ADC-0.05% Tween 80 containing Hyg to allow the second recombination event to occur. Serial dilutions were plated onto 7H10-ADC containing 2% Suc and Hyg to select the colonies in which the second recombination had occurred. Consequently, *sacB* was lost and double cross over recombinants were able to grow in presence of Suc. Considerable decrease of colony number was observed in the plates containing Suc compared to plates without Suc, because of the *sacB* selection (Figure 23C). Replica plating in 7H10-ADC-Hyg of colonies grown in 2% of Suc were performed to confirm the second recombination by evaluation of XylE activity. Unstained colonies after catechol addition (XylE negative activity) confirmed that the second recombination had occurred (Figure 23D).

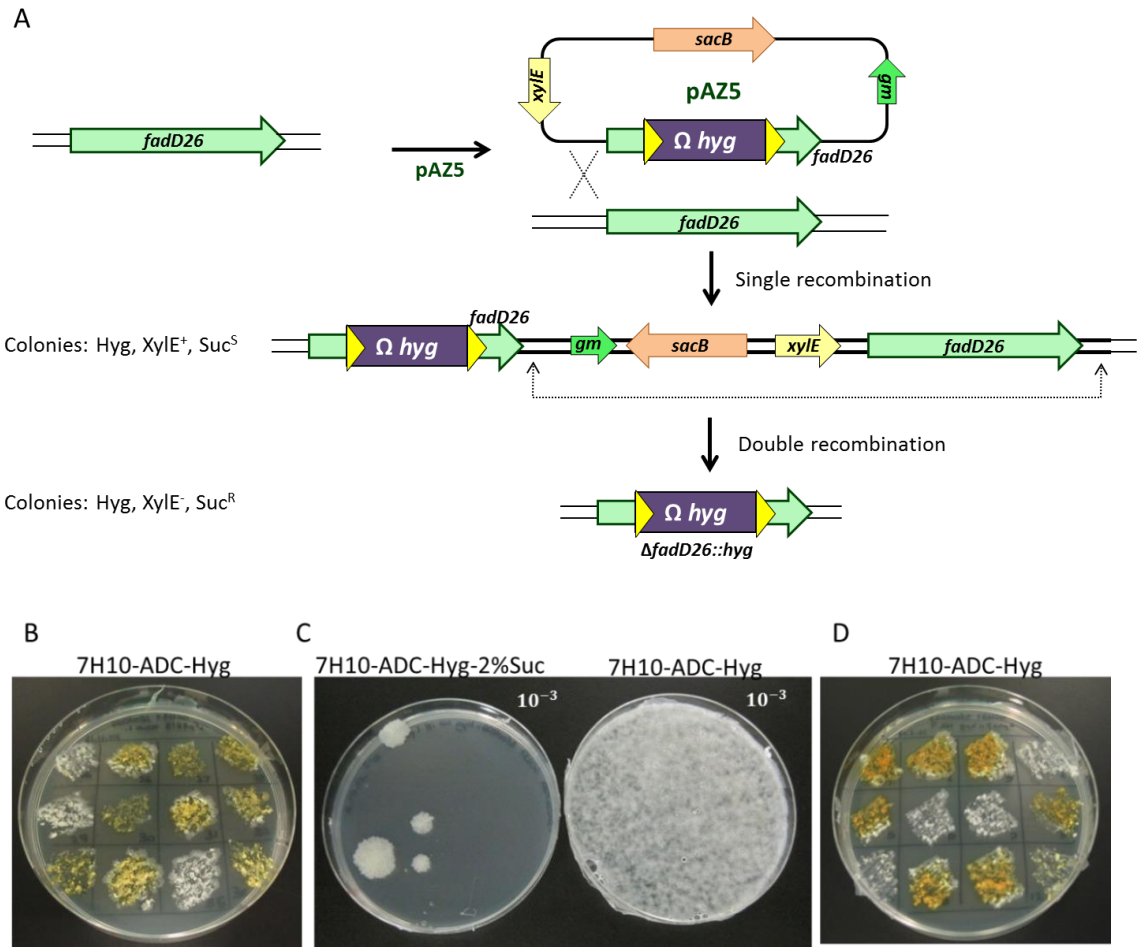


Figure 23. Schematic representation of *fadD26* deletion using pAZ5 suicide plasmid. (A) Suicide plasmid pAZ5 was transformed in WT strain. Single recombinants (phenotype Hyg, XylE$^+$, Suc^S) were selected in plates in presence of Hyg and XylE positive activity. Double cross-over recombinants were selected in plates containing Hyg, 2% Suc and negative to XylE activity. Phenotype of double cross-over colonies was Hyg, XylE$^-$, Suc^R. Adapted from (56) (B) Study of XylE activity in single cross-over recombinants. Positive colonies colored yellow after catechol addition confirmed the integration of *xylE* (and the plasmid) in the chromosome. (C) Double cross-over recombinants were selected in plates containing 2% of Suc and Hyg. Colonies that grow in presence of Suc is because of the loss of *sacB* gene and consequently loss of the plasmid backbone. (D) Colonies grown in 2% Suc plates were tested for XylE activity. Double recombinants remained unstained because of the loss of *xylE* gene in the second recombination.

A final confirmation of allelic exchange was obtained with PCR amplification using different combination of primers (Figure 24A). One of the primers is designed 5´upstream or 3´downstream of the genomic region cloned in the plasmid (fadD26up, fadD26down respectively) and the other in the *res* sites (res1, res2). The amplification will occur if the double recombination has occurred in the correct site of the mycobacteria genome. fadD26I1 and fadD26I2 primers hybridized in *fadD26* deleted region. This PCR amplified in the WT strain (*fadD26* WT) but no amplification was observed in *fadD26::hyg* strains (Figure 24B).

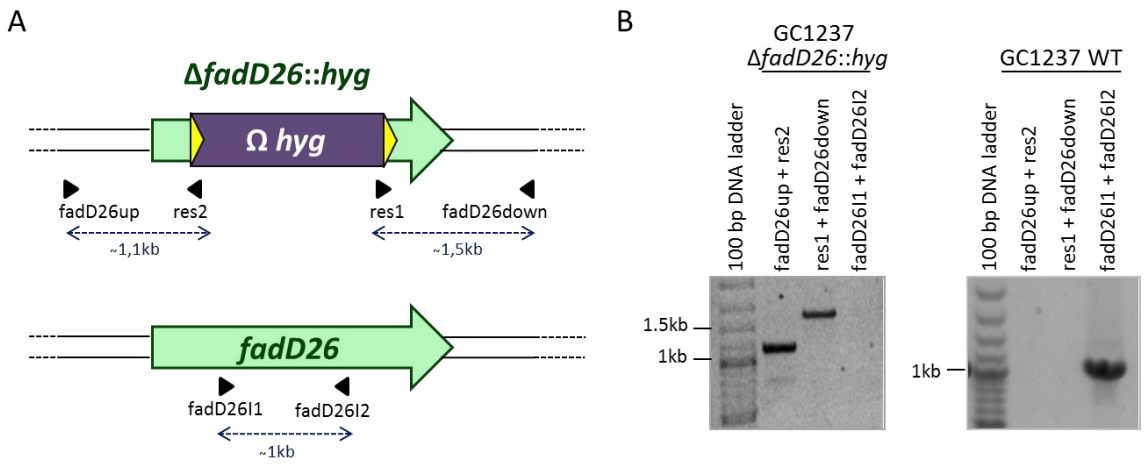


Figure 24. Analysis of GC1237 Δ*fadD26*::*hyg* by PCR. (A) Representation of genetic structures of *fadD26* mutant disrupted with *res-Ωhyg-res* cassette and WT *fadD26* gene. Primers used are indicated in the scheme. fadD26I1 and fadD26I2 primers hybridized in deleted *fadD26* region. fadD26up and fadD26down hybridized upstream and downstream of recombination sites respectively. (B) PCR analysis of the GC1237 Δ*fadD26*::*hyg* mutant and the WT strain with different combinations of the primers in panel A.

- **Construction of the single mutant unmarked, GC1237 Δ*fadD26***

GC1237 Δ*fadD26*::*hyg* was transformed with 500 ng of pAZ20, a replicative plasmid that contains γδ resolvase, TnpR, that recognize *res* sites and allows the elimination of antibiotic resistance marker from *res-Ωhyg-res* cassette (52, 55) (Figure 25A).

Transformed bacteria were recovered in 5 ml of 7H9-ADC-0.05% Tween 80 and incubated overnight at 37 °C to allow expression of the antibiotic resistance. The next day, Km was added to the liquid media to select positive colonies after the transformation and incubated for five days to allow expression of the resolvase and consequent removal of the antibiotic cassette. Then, serial dilutions were plated onto 7H10-ADC containing 2% Suc to select colonies that had lost pAZ20.

To confirm the elimination of Hyg cassette and the loss of pAZ20 plasmid (Km), colonies were replica plating without antibiotic or containing Hyg or Km. The sensibility to Hyg and Km of the 95% of tested colonies confirmed the loss of both antibiotic cassettes (Figure 25B).

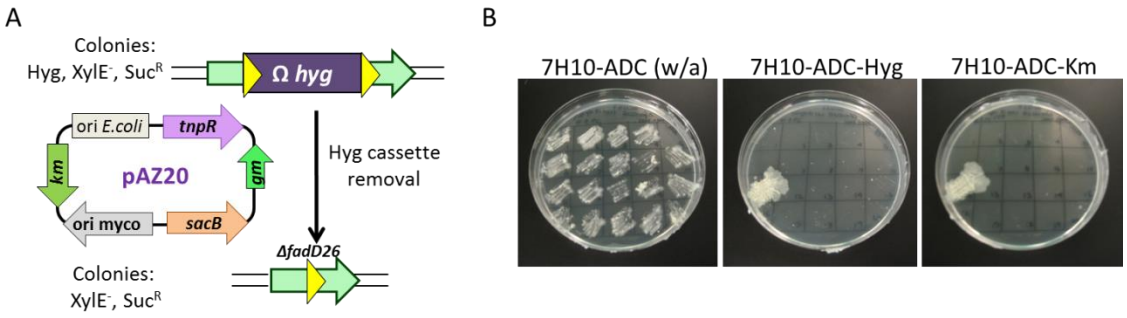


Figure 25. Removal of Hyg resistant cassette. (A) Schematic representation of removal of Hyg cassette by transformation of pAZ20 plasmid. (B) Elimination of Hyg cassette and loss of pAZ20 was confirmed by replica plating in plates without antibiotic or containing Hyg or Km.

An additional confirmation of elimination of Hyg cassette was performed by PCR amplification (Figure 26A). fadD26F and fadD26R primers were used to verify the elimination of Hyg cassette in GC127 Δ*fadD26* (2.5 kb) compared to GC1237 WT (3.9 kb). Lack of amplification in GC1237 Δ*fadD26* with fadD26I1 and fadD26I2 primers confirmed the deletion in *fadD26* gene (Figure 26B).

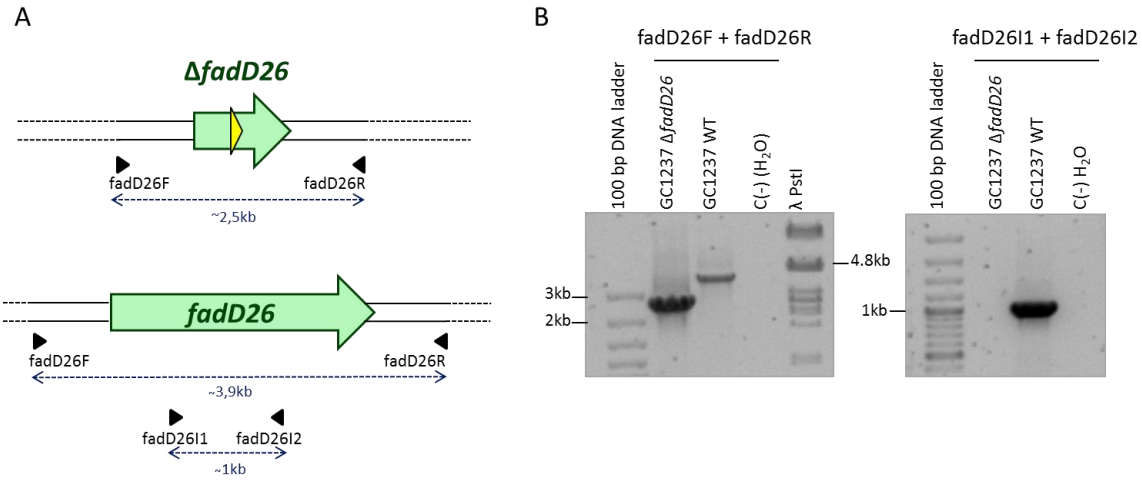


Figure 26. Confirmation of unmarked single mutant, GC1237 Δ*fadD26*. (A) Representation of genetic structures of deleted *fadD26* and *fadD26* WT. Primers used in the PCR are indicated in the scheme. (B) PCR analysis of the mutant GC1237 Δ*fadD26* and GC1237 WT with fadD26F, fadD26R and fadD26I1, fadD26I2 primers.

- **Construction of the double marked mutant, GC1237 Δ*fadD26* Δ*phoP::hyg* (MTBVAC-L2::*hyg*)**

To obtain the double mutant, GC1237 Δ*fadD26* was transformed with 1 µg of pAZ18, a suicide plasmid harboring *phoP* gene disrupted by *res-Ωhyg-res* cassette. Selection in two steps was performed as described before in construction of the single mutant marked strain, GC1237 Δ*fadD26*::*hyg* section.

After the positive and negative selection final confirmation by PCR was performed to confirm that double marked mutant GC1237 Δ*fadD26* Δ*phoP::hyg* had been obtained.

Several PCR were performed (Figure 27A). One of the primers was designed outside the genomic region cloned in the pAZ18 plasmid (phoPup and KO-phoP-rv) and the other hybridized in the *res* sites (res1 and res2). The amplification with primers confirmed the specific recombination and the insertion of *res-Ωhyg-res* cassette in *phoP*. phoPI1 and phoPI2 primers hybridize in the region deleted in *phoP*. Therefore, no amplification was obtained in phoPI1/phoPR primers and RT-phoP-fw/phoPI2 reactions in *phoP* mutants whereas both PCR amplified in *phoP* WT strain (Figure 27B).

After PCR confirmation, a Western-blot was performed to confirm the absence of PhoP in GC1237 Δ*fadD26* Δ*phoP*::*hyg* clones. Double marked mutant GC1237 Δ*fadD26* Δ*phoP*::*hyg* was named MTBVAC-L2::*hyg* (Figure 28).

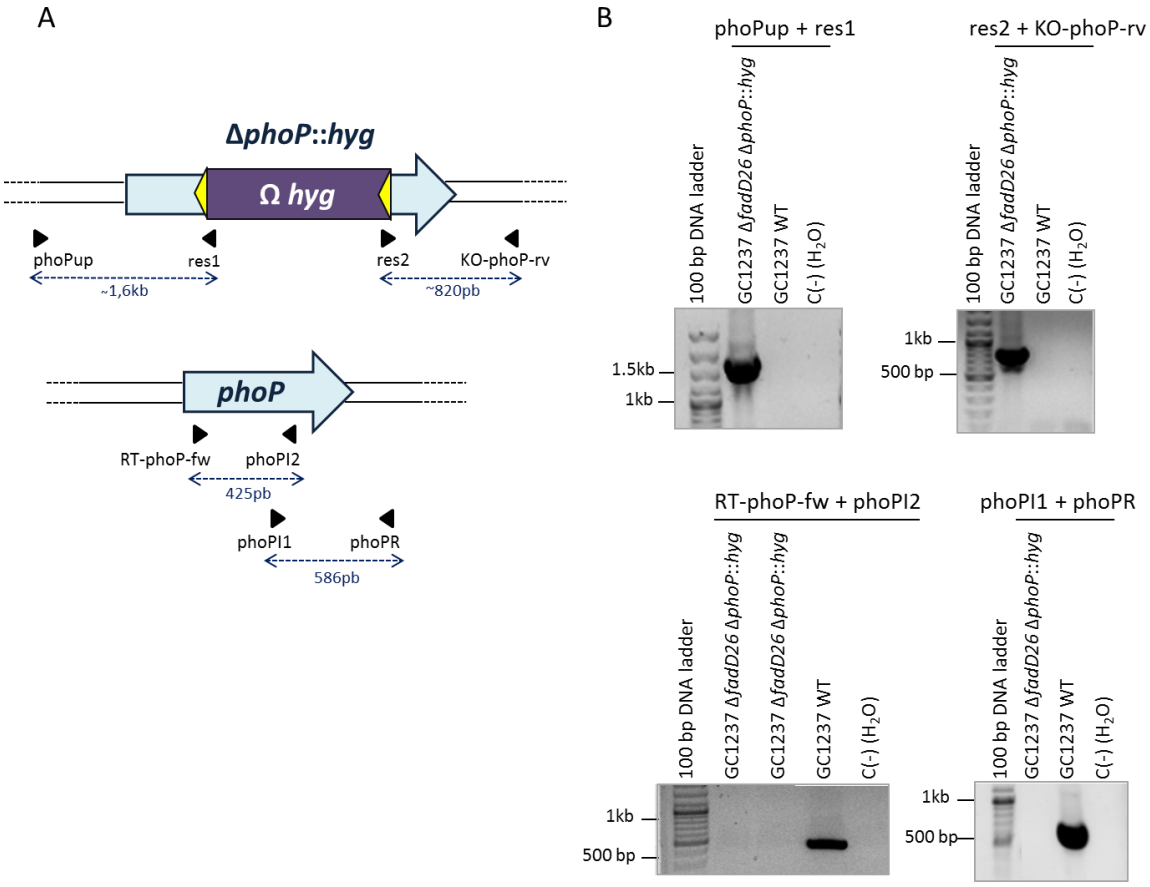


Figure 27. **Analysis of GC1237 Δ*fadD26* Δ*phoP*::*hyg* by PCR**. (A) Representation of genetic structures of *phoP* mutant disrupted with Hyg cassette and *phoP* WT. Primers used are indicated in the scheme. PhoPI1 and PhoPI2 primers hybridized in deleted *phoP* region. (B) PCR analysis of the mutant and the GC1237 WT with different combinations of the primers in panel A.

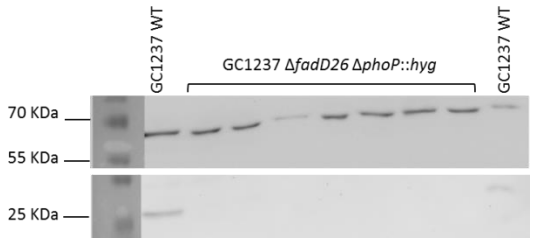


Figure 28. Confirmation of lack of PhoP by Western-blot. Detection of PhoP (~27 kDa) in GC1237 and GC1237 Δ*fadD26* Δ*phoP::hyg* (MTBVAC-L2::*hyg*) in whole cell lysate fraction. PhoP was detected in WT strain and no detection was observed in *phoP* mutants, as expected. GroEL2 (~57 kDa) was detected as load control.

- **Construction of the double unmarked mutant, GC1237 Δ*fadD26* Δ*phoP*, MTBVAC-L2**

GC1237 Δ*fadD26* Δ*phoP::hyg* was transformed with pAZ20 to remove the antibiotic resistance cassette by recognition of the flanking *res* sites by the γδ-resolvase encoded in pAZ20 plasmid.

After transformation of pAZ20 in GC1237 Δ*fadD26* Δ*phoP::hyg*, bacteria were plated onto 7H10-ADC-Km and after colonies were replica plating onto 7H10-ADC-Km and 7H10-ADC-Gm to confirm transformation of pAZ20. Colonies grown in presence of Km and Gm were grownin 7H9-ADC-0.05% Tween 80 to allow the expression of the recombinase. Cultures were plated onto 7H10-ADC containing 2% Suc to select colonies which had lost pAZ20 plasmid. To confirm the removal of Hyg cassette and the loss of pAZ20, replica plating of the colonies were performed in 7H10-ADC without antibiotic or containing Km or Hyg.

Double mutants unmarked GC1237 Δ*fadD26* Δ*phoP* grew in 7H10-ADC without antibiotic and did not grow in presence of Km or Hyg antibiotics.

Confirmation of elimination of Hyg cassette was verified by PCR amplification using phoRF and phoPR primers (Figure 29A). Amplification of in Δ*phoP* with these primers is a fragment of 705 bp compared to 630 bp in the *phoP* WT gene (Figure 29B).

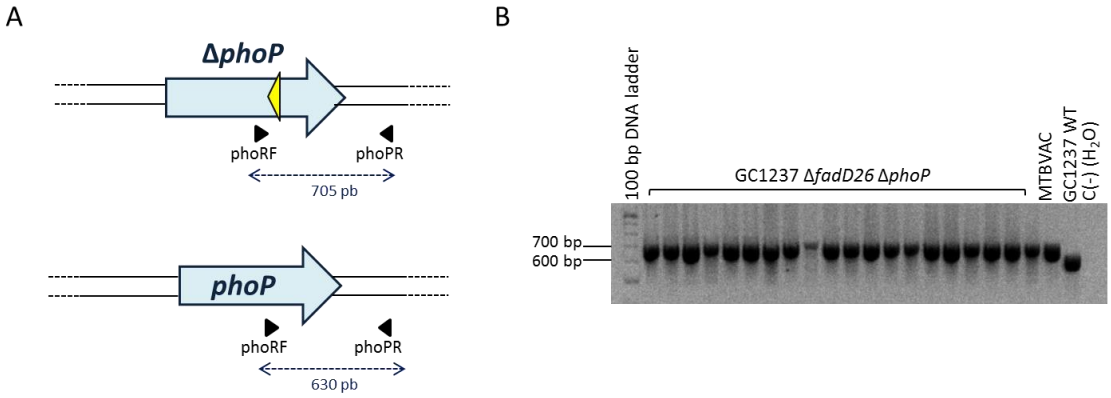


Figure 29. Confirmation of elimination of Hyg resistance cassette. Differences in the size product of the PCR using phoRF and phoPR primers confirmed the removal of antibiotic cassette by the resolvase. Elimination of Hyg cassette had ccurred in all colonies tested. MTBVAC was used as control positive (unmarked *phoP* mutant).

General scheme of the steps performed to construct MTBVAC-L2 is represented in the following figure (Figure 30).

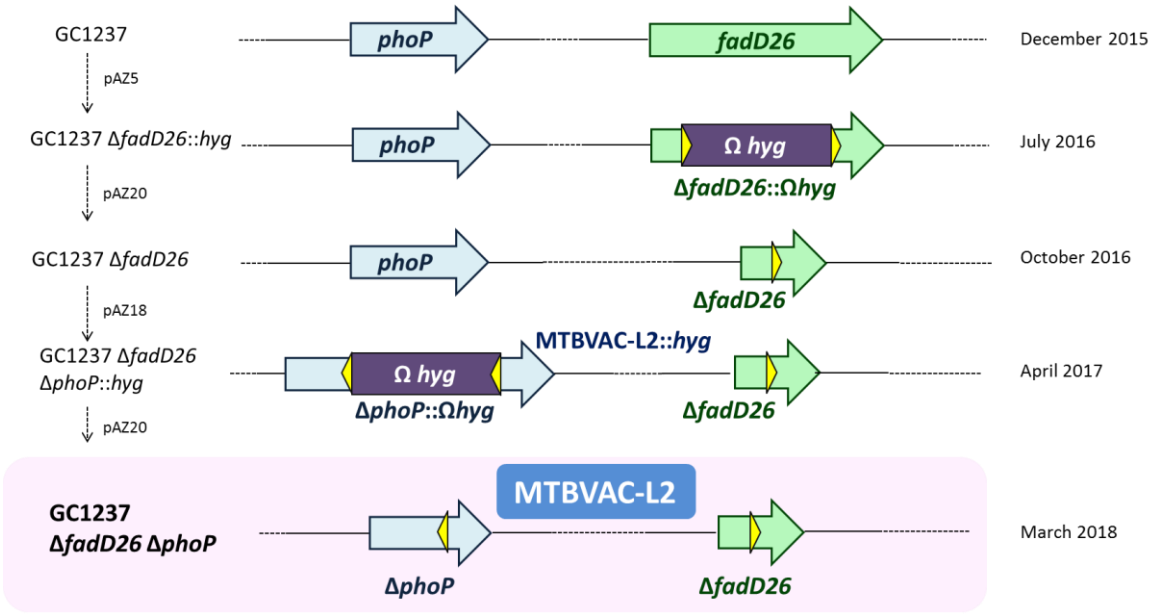


Figure 30. Schematic representation of double deletion in GC1237 strain. Sequential steps to obtain double mutant in *phoP* and *fadD26* genes using suicide plasmids pAZ5 and pAZ18.

Construction of MTBVAC-L3, a double unmarked mutant Δ*fadD26* Δ*phoP* in HMS13037

Likewise in GC12237 strain, suicide plasmid and BAC-recombineering strategies were performed in parallel to obtain deletions in *phoP* and *fadD26* genes. First deletion was obtained in *fadD26* gene by Bac-recombineering strategy and no mutants were obtained using the suicide plasmids. Second deletion in *phoP* gene was obtained using suicide plasmid pAZ18 and no recombinants were obtained with BAC-recombineering strategy (Figure 31).

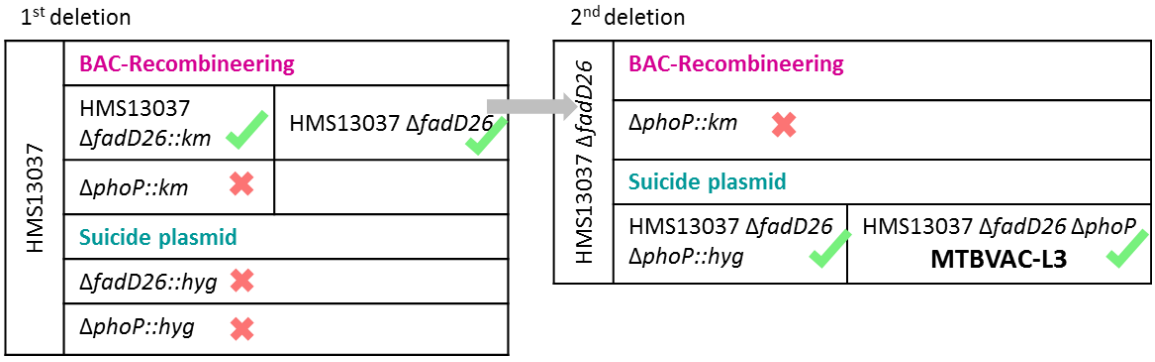


Figure 31. Results of BAC-recombineering and suicide plasmids strategies to obtain MTBVAC-L3. Single mutant of *fadD26* was obtained with Bac-recombineering strategy. Second deletion in *phoP* was obtained with pAZ18 suicide plasmid.

Sequential steps that led to MTBVAC-L3 construction (*fadD26* deletion with Bac-recombineering strategy and *phoP* deletion with pAZ18 suicide plasmid) are detailed in the following sections.

- **Growth of HMS13037 in media supplemented with ADC or OADC**

HMS13037 was grown in the laboratory for the first time and we encountered difficulties since very small size colonies were observed in 7H10-ADC plates (usual solid media used in the laboratory).

To improve the growth of the strain, HMS13037 was grown in different solid media to test differences in *in vitro* growth. 7H10 supplemented with ADC or OADC and Sauton with ADC, OADC or no supplemented were tested. Significant differences were observed. HMS13037 growth was remarkably better in media supplemented with OADC compared to the supplementation with ADC (Figure 32).

In liquid media, we also changed to 7H9 supplemented with OADC. At the beginning the bacteria grown, but at some point the growth in liquid was also impaired. Then, we observed that in liquid cultures, the bacteria grew in liquid media supplemented with ADC.

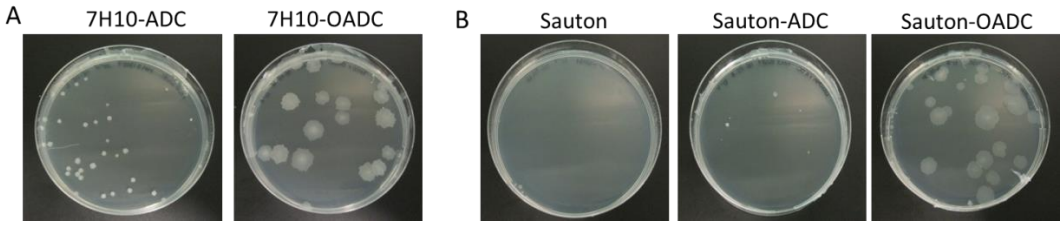


Figure 32. Growth of HMS13037 in solid media. (A) 7H10 supplemented with ADC or OADC. (B) Sauton non supplementated or supplemented with ADC or OADC.

- **Construction of the single mutant marked strain, HMS13037 Δ*fadD26::km***

BAC recombineering method consists of two steps; the first is performed in *E. coli* to obtain the allelic exchange substrate (AES) that will be used in *M. tuberculosis* to obtain gene deletion. The entire process is summarized in Figure 33. Each step of the process is detailed below.

A bacterial artificial chromosome (BAC) from H37Rv library constructed in pBeloBAC11 vector and contained into *E. coli* DH10B was used for this methodology (61).

Steps performed in *E. coli*:

pKD46 plasmid was transformed into DH10B containing BAC Rv209 which contains a fragment of the *M. tuberculosis* H37Rv genome including *fadD26* gene (3243697- 3245448). pKD46 is a thermosensitive plasmid that contains λ Red system which includes three genes that codes for Gam, Bet and Exo proteins. These proteins inhibit host exonucleases and promote recombination. Expression of the three genes is induced by arabinose (57). Transformation was plated onto LB-Amp-Cm and incubated overnight at 30 °C. Transformation of pKD46 was confirmed by PCR using pKD46-gam-fw and pKD46-bet-rv (Figure 34A).

DH10B BAC Rv209 pKD46 induced with arabinose was transformed with a PCR product containing FRT-*km*-FRT cassette flanked with 40 bp of identity arms to the site where

recombination is directed. FRT-*km*-FRT cassette was amplified from pKD4 plasmid (57) and identity arms were added in the primers (fadD26-P1-Fw and fadD26-P2-Rv) (Figure 34B). Transformation was incubated for 1 h at 37 °C to allow expression of resistance cassette and to cure pKD46 plasmid. Afterwards, serial dilutions were plated onto LB-Km at 37 °C. Recombinants were confirmed by PCR amplification using one primer outside from the 40 bp of identity (KO-fadD26-fw and KO-fadD26-rv) and the other inside the FRT-*km*-FRT cassette (km-pKD4-out-1 and Km-pKD4-out-2) (Figure 34C).

DH10B BAC Rv209 Δ*fadD26::km* were grown overnight at 37 °C in LB-Km and BAC Δ*fadD26*::*km* was purified. BAC Rv209 Δ*fadD26*::*km* was used as template to obtain AES to transform in *M. tuberculosis* (Figure 34D). AES product, consists in FRT-*km*-FRT flanked with 670 bp of identity arms for site specific recombination.

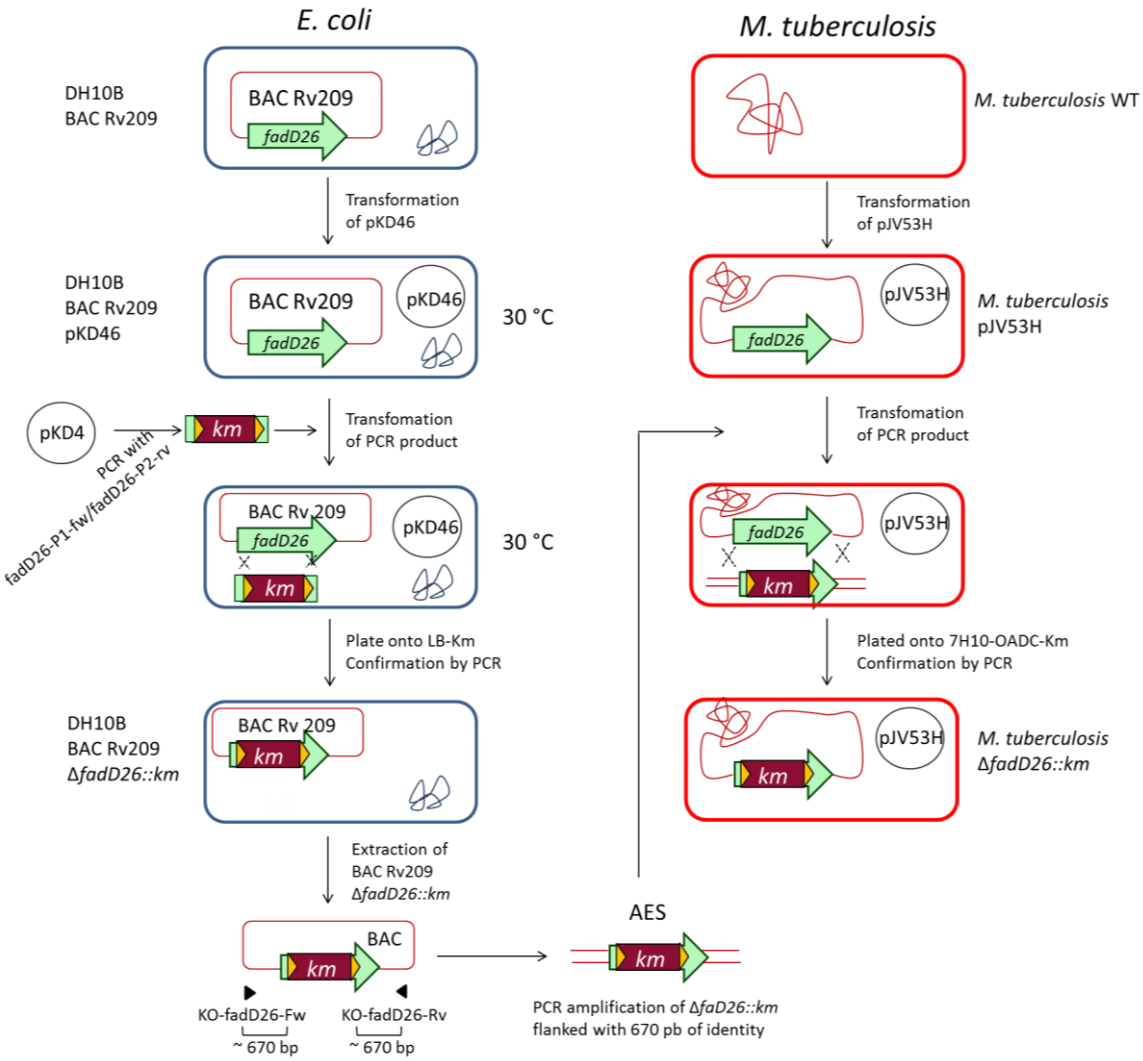


Figure 33. Deletion of *fadD26* gene by BAC-recombineering. Left column represents the steps performed in *E.coli* (blue) to obtain AES for transformation in *M. tuberculosis* (right column, bacteria represented in red).

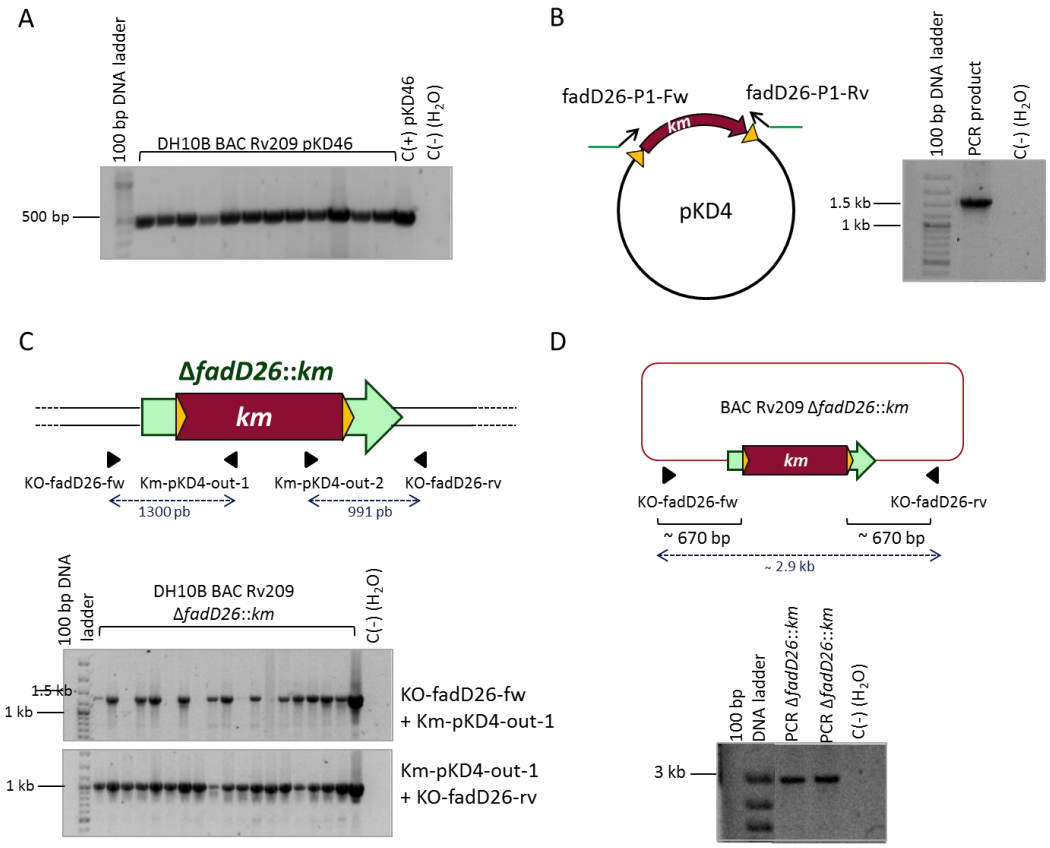


Figure 34 . Disruption of *fadD26* with FRT-*km*-FRT cassette in BAC Rv209. (A) Confirmation of pKD46 in DH10B BAC Rv209 with pKD46-gam-fw and pKD46-bet-rv primers. (B) PCR product amplification using fadD26-P1-fw and fadD26-P2-rv primers. pKD4 plasmid was used as template to obtain the PCR product (FRT-*km*-FRT flanked with 40 bp of identity arms to site specific recombination) to transform in DH10B BAC Rv209 pkD46 bacteria induced by arabinose. (C) Scheme of genomic structure of *fadD26* disrupted by FRT-*km*-FRT cassette. PCR analysis to confirm recombination in *E.coli* and disruption of *fadD26* gene. (D) AES to transform in *M. tuberculosis* were obtained using BAC Rv209 Δ*fadD26::km* as template.

Steps performed in *M. tuberculosis*:

M. tuberculosis HMS13037 was transformed with pJV53H plasmid that contains gp60 and gp61 from the mycobacteriophage Che9c. Che9c gp60 and gp61 encode exonuclease and DNA-binding activities, respectively. The expression of these proteins increases recombination facilitating allelic exchange in *M. tuberculosis* (58).

Transformation was plated onto 7H10-OADC-Hyg and PCR amplification (gp60fw and gp61rv primers) was used to confirm the presence of the plasmid (Figure 35A). HMS13037 pJV53H was induced with 0.2% acetamide, to allow the expression of Che9c gp60 and gp61, and HMS13037 pJV53H was transformed with AES (Figure 34D) and plated onto 7H10-OADC-Km. AES product, FRT-*km*-FRT flanked with identity arms for site specific recombination.

Recombination was confirmed by PCR amplification with primers designed to hybridize outside of the genomic region amplified as AES (comfirm-fadD26-Fw and confirm-fadD26-Rv) and primers inside the FRT-*km*-FRT cassette (Figure 35B and C). HMS13037 pJV53H Δ*fadD26*::*km* was grown in 7H9-OADC-0.05% Tween 80 containing Km and in absence of Hyg.

Several passages were performed to lose pJV53H and bacteria were plated onto 7H10-OADC containing Km. Loss of pJV53H was confirmed by replica plating in plates containing or not Hyg and single mutant HMS13037 Δ*fadD26*::*km* was obtained (Figure 36).

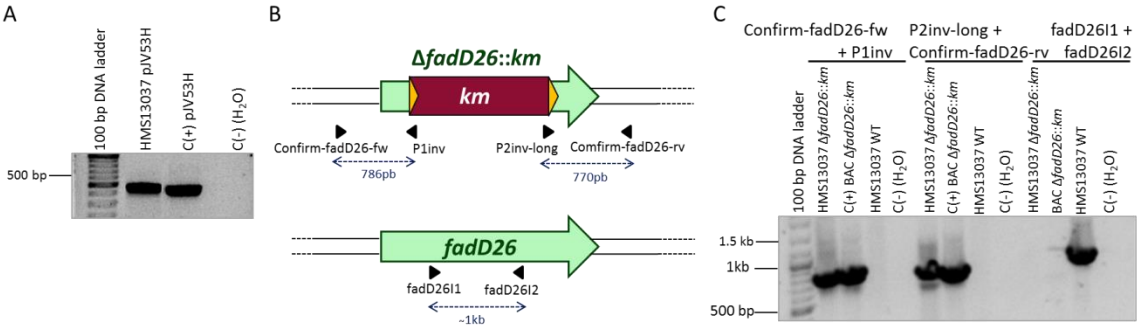


Figure 35. Confirmation of HMS13037 Δ*fadD26*::*km*. (A) Confirmation of pJV53H plasmid in HMS13037 strain using gp60fw and gp61rv primers. pJV53H was used as template in the positive control of the reaction. (B) Schematic representation of *fadD26* disrupted by FRT-*km*-FRT cassette and *fadD26* WT gene and primers used. Primers confirm-fadD26-fw and confirm-fadD26-rv hybridized upstream and downstream of the recombination site, P1inv and P2inv-long hybridized in the resistance cassette and fadD26I1/fadD26I2 hybridized in *fadD26*deleted region. (C) PCR analysis of the mutant and HMS13037 WT strain using different combinations of the primers in panel B. BAC Rv209 Δ*fadD26*::*km* was used as control.

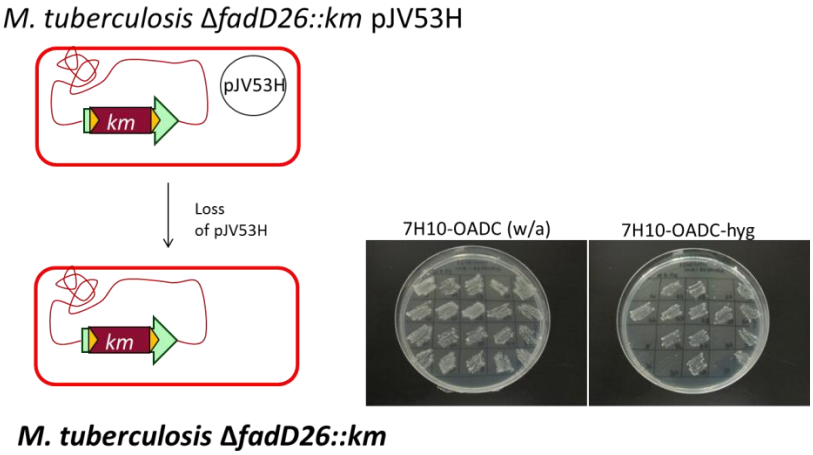


Figure 36. Loss of pJV53H. After several passages without Hyg, replica plating in 7H10-OADC and 7H10-OADC-Hyg were performed to evaluate pJV53H loss.

- **Construction of the single mutant unmarked, HMS13037 Δ*fadD26***

HMS13037 Δ*fadD26*::*km* was transformed with pRES-FLP-Mtb (Hyg), a replicative plasmid that contains *flp* recombinase gene from *S. cerevisiae* with codon usage optimized to *M. tuberculosis* that recognizes the FRT sequence. FLP recombinase acts on the directly repeated FRT sites flanking the Km resistance marker and it allows obtaining the mutant unmarked.

After pRES-FLP-Mtb transformation, bacteria were plated onto 7H10-OADC containing Hyg. Replica plating either with Km or without antibiotic were performed to confirm removal of Km resistance marker. The simple mutant HMS13037 *ΔfadD26* was obtained, although the strain still harbored pRES-FLP-Mtb plasmid (Figure 37).

Final confirmation of Km cassette removal was verified by PCR amplification (Figure 38 A and B). To cure pRES-FLP-Mtb plasmid, several passages were performed in 7H9-ADC-0.05% Tween 80 without antibiotic and plated onto 7H10-OADC without antibiotic. Loss of pRES-FLP-Mtb plasmid was confirmed by replica plating in plates without antibiotic or with Hyg (Figure 37).

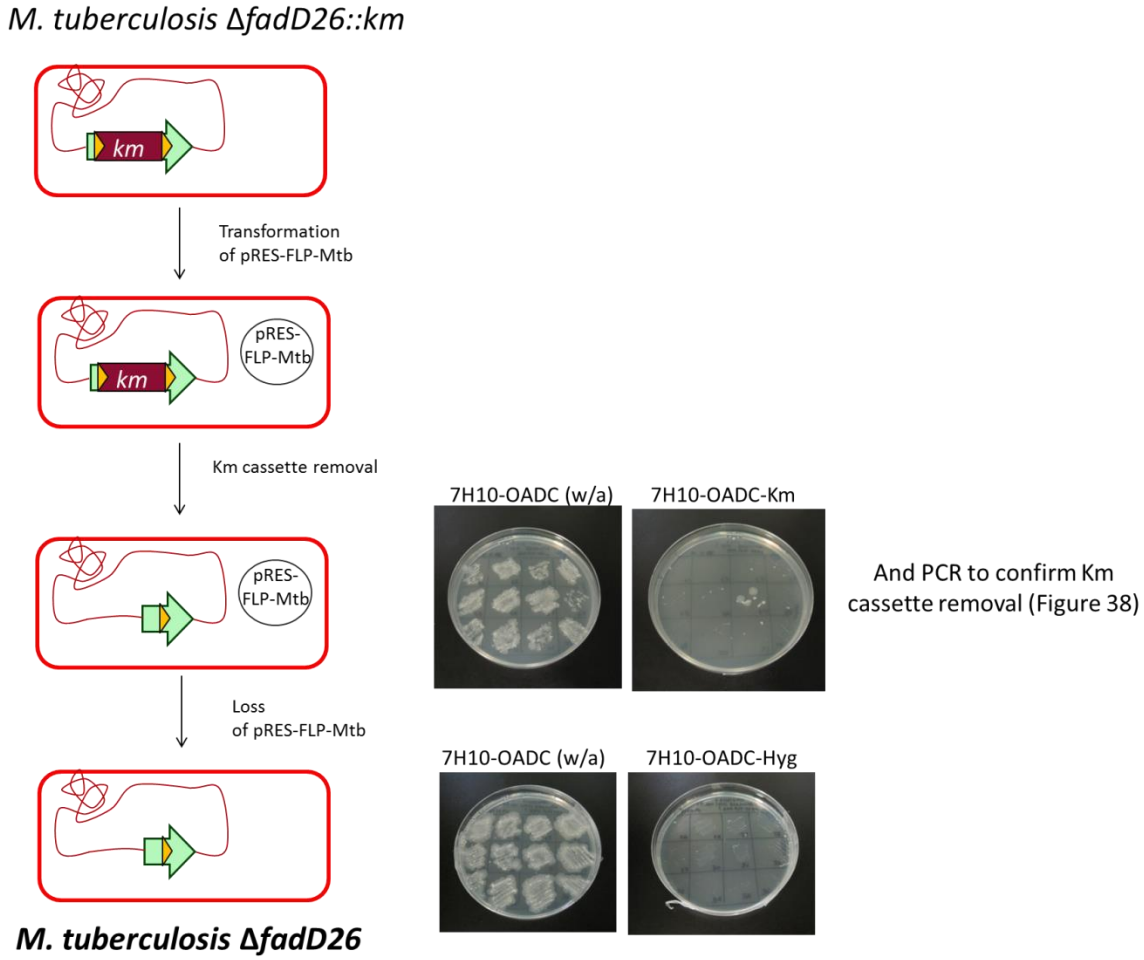


Figure 37. Sequential steps for antibiotic removal. Evaluation of Km cassette removal by replica plating either with or without Km. After several passages, loss of pRES-FLP-Mtb was confirmed by replica plating in plates without antibiotic or with Hyg.

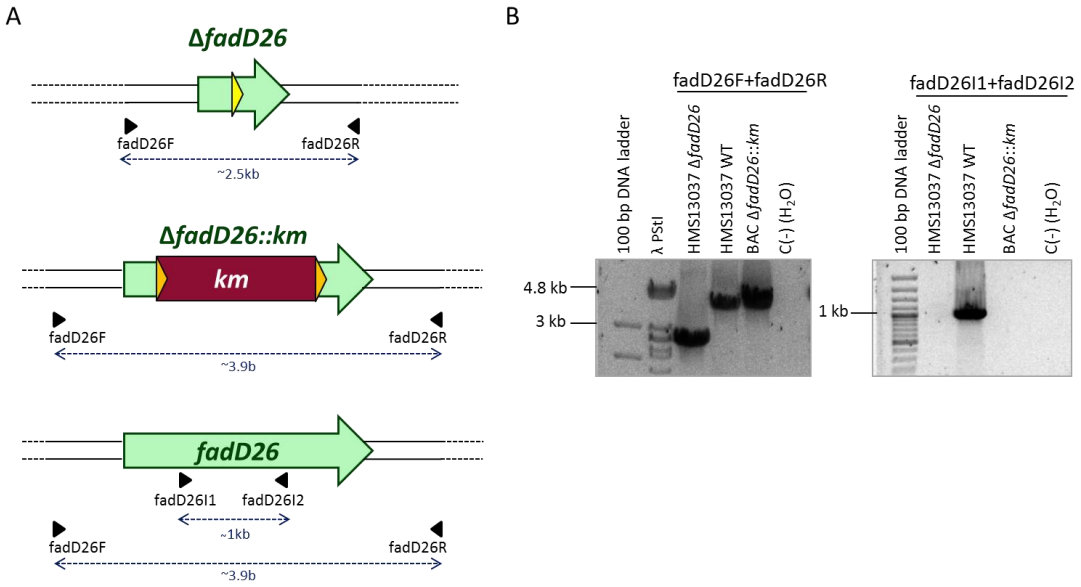


Figure 38. Removal of Km cassette. (A) Schematic representation of unmarked *fadD26*, *fadD26* disrupted with Km cassette and *fadD26* WT, and primers used. (B) PCR analysis of unmarked mutant, HMS13037 WT and BAC Δ*fadD26::km* using fadD26F/fadD26R and fadD26I1/fadD26I2 primers.

- **Construction of the double marked mutant, HMS13037 Δ*fadD26* Δ*phoP::hyg* (MTBVAC-L3::*hyg*)**

For *phoP* deletion, 1 μg of pAZ18 was transformed in HMS13037 Δ*fadD26* strain. pAZ18 contains a *phoP* deleted copy distupted by *res*-Ω*hyg*-*res* cassette. Selection in two steps, positive selection followed by a negative selection was performed as described above (see section Construction of the single mutant marked strain, GC1237 Δ*fadD26*::*hyg*). Double mutant HMS13037 Δ*fadD26* Δ*phoP*::*hyg*, named MTBVAC-L3::*hyg*, was obtained and verification of *phoP* deletion by PCR was performed as described in construction of GC1237 Δ*fadD26* Δ*phoP*::*hyg* section (Figure 39A and B). Western-blot of PhoP was also preformed to confirm *phoP* deletion (Figure 40).

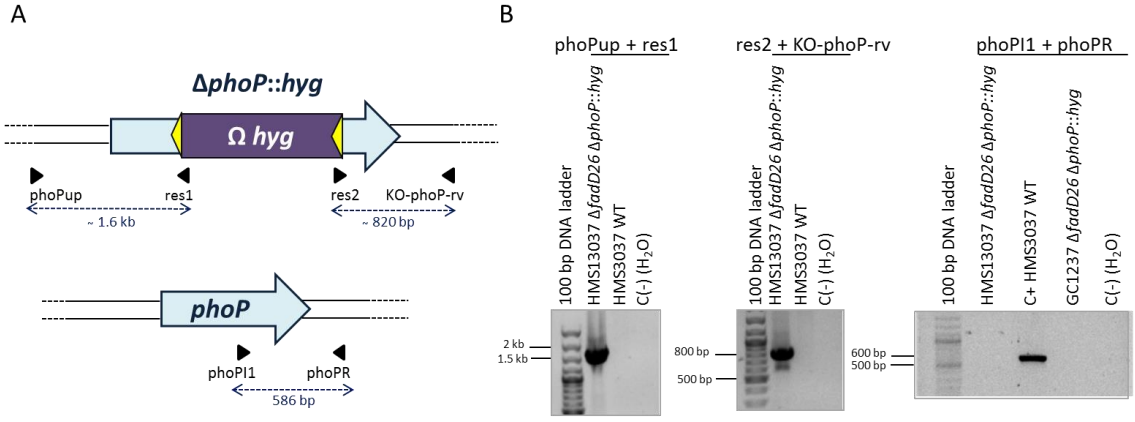


Figure 39. **PCR analysis of HMS13037 Δ*fadD26* Δ*phoP*::*hyg***. (A) Representation of genetic structures of *phoP* mutant disrupted with Hyg cassette and *phoP* WT. Primers used are indicated in the scheme. PhoPI1 primer hybridized in deleted *phoP* region. (B) PCR analysis of the mutant and the WT with different combinations of the primers in panel A.

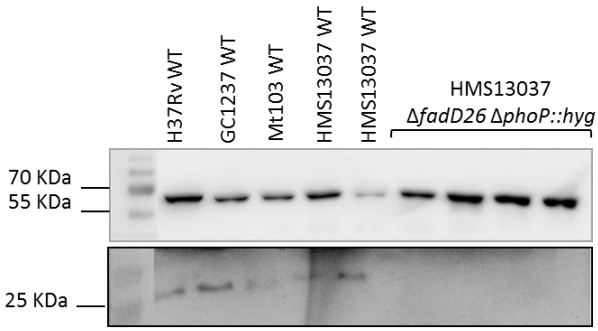


Figure 40. Confirmation of *phoP* deletion by Western-blot. Detection of PhoP (~27 KDa) in WT strains (H37Rv, GC1237, Mt103 and HMS13037) and in the double mutant HMS13037 Δ*fadD26* Δ*phoP*::*hyg*, MTBVAC-L3::*hyg* in whole cell lysate fraction. PhoP was detected in WT strain and no detection was observed in *phoP* mutants, as expected. GroEL2 (~57 KDa) was detected as load control.

- **Construction of the double unmarked mutant, HMS13037 Δ*fadD26* Δ*phoP*, MTBVAC-L3**

For Hyg cassette removal, pAZ20 (harboring the gene that codes for γδ-resolvase) was transformed in HMS13037 Δ*fadD26* Δ*phoP*::*hyg* and consequent steps were performed as described before (see section Construction of the double unmarked mutant, GC1237 Δ*fadD26* Δ*phoP*, MTBVAC-L2).

PCR amplification for confirmation of antibiotic elimination was verified by phoRF and phoPR primers. Amplification of Δ*phoP* with these primers is 705 bp compared to 630 bp in *phoP* WT gene (Figure 41).

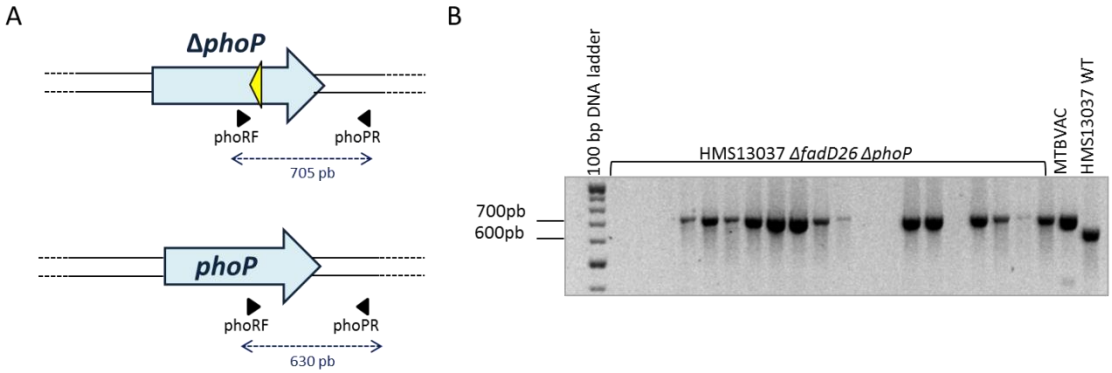


Figure 41. Confirmation of *phoP* unmarked construction. (A) Schematic representation of *phoP* deleted gene and *phoP* WT. (B) PCR analysis using phoRF/phoPR primers confirmed removal of Hyg cassette. MTBVAC was used as positive control (*phoP* deleted) and HMS13037 was used as *phoP* WT control.

Steps performed for *fadD26* and *phoP* deletions in HMS13037 lineage 3 strain are represented in the following figure (Figure 42).

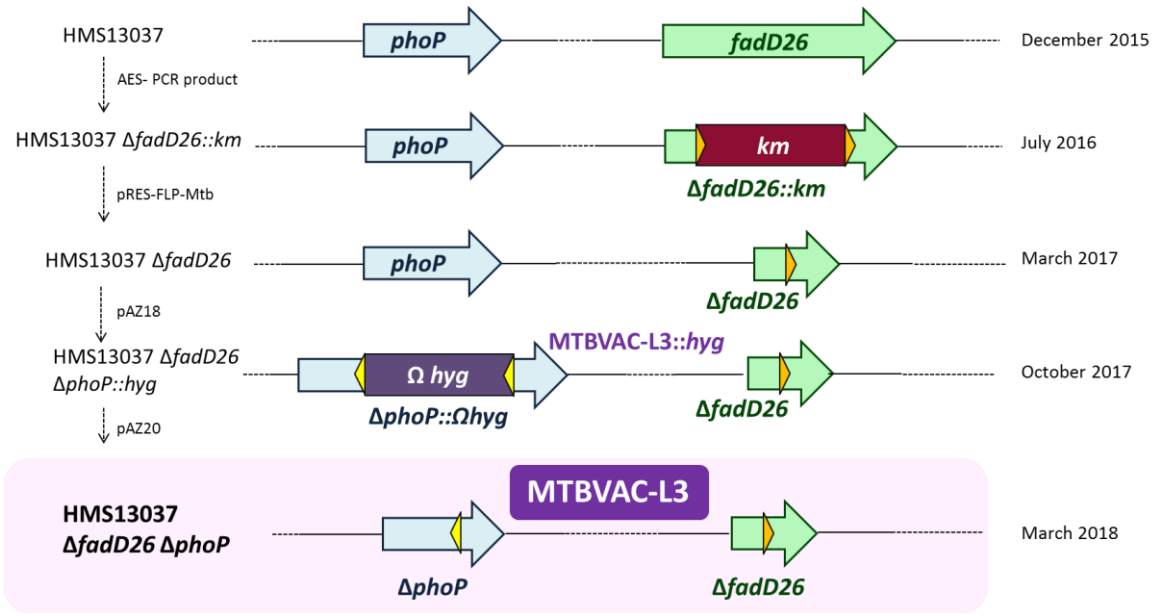


Figure 42. Schematic representation of *fadD26* and *phoP* deletions in HMS13037. Steps performed to obtain double mutant MTBVAC-L3 using Bac-recombineering and suicide plasmid strategies.

Three *phoP*, *fadD26* mutants constructed in three clinical isolates of the modern lineages of *M. tuberculosis*

In conclusion, in this study, together with MTBVAC (52), three *phoP* and *fadD26* vaccine candidates have been constructed (MTBVAC, MTBVAC-L2 and MTBVAC-L3) in the modern lineages of *M. tuberculosis*.

New double unmarked mutants obtained in this study in clinical isolates from lineages 2 and 3 of *M. tuberculosis* are named MTBVAC-L2 and MTBVAC-L3 respectively. All gene deletions were obtained by suicide plasmid strategy with the exception of deletion of *fadD26* gene in the clinical isolate from lineage 3 which was obtained by BAC-recombineering (Figure 43).

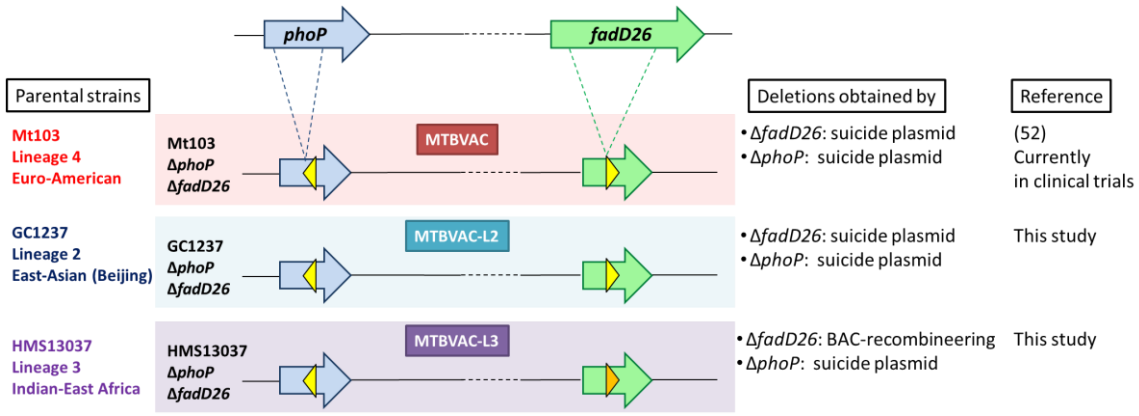


Figure 43. Schematic representation of the three *phoP*, *fadD26* strains obtained in clinical isolates of the modern lineages of *M. tuberculosis*. Deletion gene strategy and parental strains are detailed in the figure.

BAC–recombineering allows construction of marked mutants in less time compared to suicide plasmid strategy. However, the method can be improved adding counter-selectable markers to pJV53H and pRES-FLP-Mtb, to reduce the time taken to cure these plasmids. In addition to the time required to cure the strain from the plasmids, we have obtained a very low efficiency with the BAC-recombineering method. We suggest that this is because of the identity arms of the AES were approximately 700 bp and the increase of the homology of the flanked arms (up to 1 kb) could enhance the number of recombinants obtained.

After the construction of the double mutants, the following experiments consisted on the characterization of the three double mutants in the secretion of PE_PGRS proteins and to confirm the maintenance in MTBVAC-L2 and MTBVAC-L3 of the PhoP and FadD26 phenotypes previously described in MTBVAC.

MTBVAC-L3 and MTBVAC, but no MTBVAC-L2, secrete PE_PGRS proteins

PE_PGRS proteins are a subfamily of the protein family with the characteristic Pro-Glu (PE) motif and its secretion depends on *ppe38* locus organization. Disruption of *ppe38* and *ppe71* genes by an insertion of IS*6110* has been described in some Beijing strains and entailed a lack of secretion of PE_PGRS proteins, whereas intact of *ppe38, esxXY* and *ppe71* genes result in secretion of PE_PGRS proteins. Lack of secretion of PE_PGRS proteins has been linked with increased virulence of *M. tuberculosis* in mice (47). However, no involvement of these proteins in protective efficacy was observed in the mouse model (62).

Study of *ppe38* locus organization in the parental strains of the three *phoP* and *fadD26* double mutants constructions (Mt103, GC1237 and HMS13037) by PCR amplification with primers described in (63) (Figure 44A) showed that Mt103 and HMS13037 contained intact *ppe38, esxXY and ppe71* genes (Figure 44B).

No amplification PCR product was obtained in GC1237 strain using PPE38F/PPE38R and PPE38IntF/PPE38IntR primers (Figure 44B). To further study the *ppe38* locus in GC1237, we evaluated the reads from two clones with the same genotype of GC1237 isolated from the same outbreak whose genomes were sequenced by the group. There were no reads in *esxXY* genes suggesting lack of these genes in GC1237.

PE_PGRS proteins in total and secreted fraction were detected by Western-blot in the three *phoP, fadD26* mutants (MTBVAC, MTBVAC-L2 and MTBVAC-L3) and their parental strains. Results exhibited that all strains contained PE_PGRS proteins, whereas MTBVAC-L3 and MTBVAC secrete PE_PGRS proteins in contrast to MTBVAC-L2 (Figure 45).

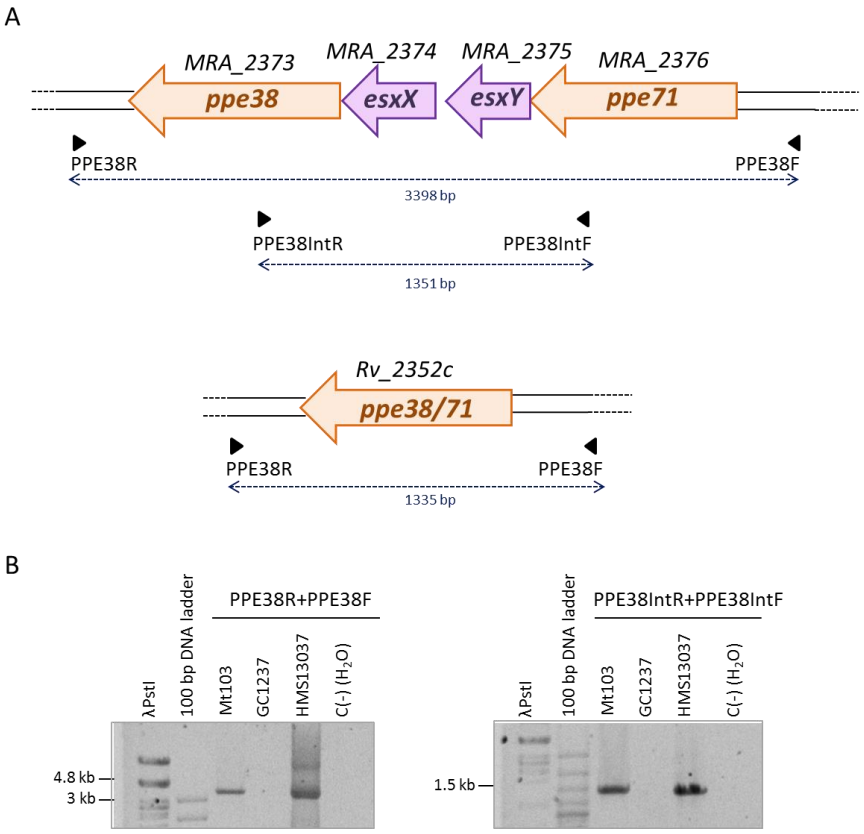


Figure 44. Genomic analysis of *ppe38* locus. (A) Schematic structure of *ppe38* locus and primers used, adapted from (63). (B) PCR amplification of Mt103, GC1237 and HMS13037 using combination of primers of panel A.

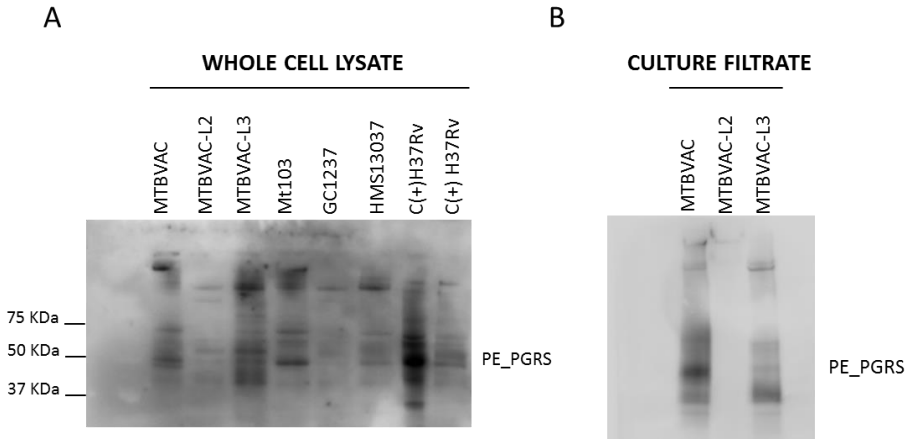


Figure 45. Detection of PE_PGRS proteins with anti-PE_PGRS antibody. (A) Detection in the whole cell lysate of MTBVAC, MTBVAC-L2, MTBVAC-L3 strains and WT strains. (B) PE_PGRS detection in supernatant fraction of MTBVAC, MTBVAC-L2, MTBVAC-L3.

Lack of secretion of PE_PGRS proteins in GC1237 and MTBVAC-L2 strains is controversial with the hypothesis of Ates and colleagues (47). They suggest that *ppe38/71* truncation may occurred in the branching point of the modern Beijing lineages, because they observed this truncation in *ppe38* locus in all modern Beijing strains studies in contrast to ancestral Beijing strains tested (from F31, previously of RD181 branching point).

GC1237 is an ancestral Beijing strain which belongs to Asia Ancestral 3 sublineage and according to Ates and colleagues hypothesis, secretion of PE_PGRS would be observed in GC1237 (47). However, secretion of these proteins was no detected. Further studies are required to decipher whether GC1237 is an exception in this phenotype of ancestral Beijing strains or other mechanism could interfere in the secretion of these proteins. Another possibility is that the insertion of IS*6110* that truncates *ppe38* and *ppe71* may have occurred in an earlier branching point from ancient/modern Beijing strains, and some ancestral Beijing strains could present lack of secretion of PE_PGRS proteins while others can secrete them.

MTBVAC-L2 and MTBVAC-L3 has lost the ability to fix the neutral red

The ability of *M. tuberculosis* to fix neutral red has been associated with virulence for more than six decades. In 1948, Dubos and Middlebrook observed that the virulent strain H37Rv fixed the dye in an alkaline solution of neutral red in contrast to the avirulent H37Ra (64).

A more recent study, negative neutral red staining was related with loss of synthesis of PDIM or other cell wall methyl-branched lipids such as SL and PAT. Strains were neutral red negative when they were deficient in more than one type of methyl-branched lipids. Consequently, neutral red staining could be a marker of virulence and shows significant changes in the outer membrane of *M. tuberculosis* (65).

phoP positively regulates the synthesis of methyl-branched fatty acid-containing acyltrehaloses (DAT, PAT and SL). *phoP* mutants exhibit neutral red negative staining in comparison to WT strains (66). Same phenotype is observed for MTBVAC compared to Mt103 (56).

Neutral red staining was performed in MTBVAC strains (in lineages 2, 3 and 4) and parental strains. MTBVAC-L2 and MTBVAC-L3 (marked and unmarked) showed inability to fix neutral red, in accordance with neutral red phenotype previously described in MTBVAC (Figure 46).

All MTBVAC strains are expected to be DAT, PAT and SL deficient because of *phoP* deletion and comparing with previous bibliography, neutral red negative.

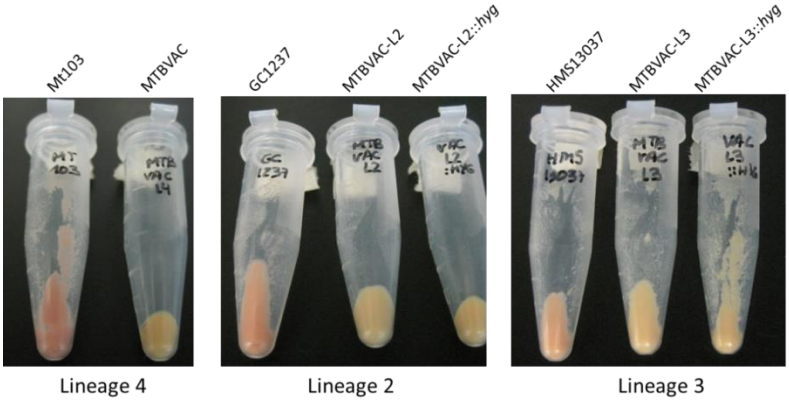


Figure 46. Neutral red staining of WT strains and mutants. Neutral red staining of double mutants (marked and unmarked) and WT strains showed the inability to fix neutral red in MTBVAC strains (*phoP, fadD26* deleted).

Secretion of ESAT-6 and CFP-10 in the three MTBVAC and WT strains

ESAT-6 secretion has been previously characterized as a PhoP dependent event (67). MTBVAC produced ESAT-6 but is unable to secrete it because of *phoP* deletion (52). Surprisingly, it was recently described the ability of MTBVAC to secrete CFP-10 (68). This result was unexpected because it has been thought that ESAT-6 and CFP-10 are co-secreted as 1:1 dimers through ESX-1 system (69). However, two studies described uncoupled ESAT-6 and CFP-10 secretion under some conditions. CFP-10 could be secreted independently of ESAT-6 in the presence of an aberrant ESX-1 system (70, 71).

ESAT-6 and CFP-10 have been described to be secreted through ESX-1 system and ESAT-6 secretion is involved in phagosome membrane disruption. This phagosome membrane disruption allows the bacilli to reach the cytosol and this process requires a functional ESX-1 (72, 73).

ESAT-6 and CFP-10 detection were performed by Western-blot in whole cell lysates and culture filtrate fractions of the three MTBVAC strains and parental strains. In whole cell lysate, ESAT-6 and CFP-10 were detected in all strains (Figure 47).

In secreted fraction, ESAT-6 secretion was detected in WT strains in contrast to MTBVAC strains. CFP-10 was observed in MTBVAC strains and WT strains. Some lysis was observed in MTBVAC and HMS13037 strains since SigA was detected in the supernatant fraction (Figure 47). Coomassie staining was performed of secreted fractions to ensure presence of protein (Figure 48).

Therefore, same results were obtained in MTBVAC-L2 and MTBVAC-L3 according to previous results described in MTBVAC: MTBVAC-L2 and MTBVAC-L3 produce ESAT-6 but ESAT-6 is not secreted in both strains (secretion of ESAT-6 is controlled by PhoP) and MTBVAC-L2 and MTBVAC-L3 produce and secrete CFP-10.

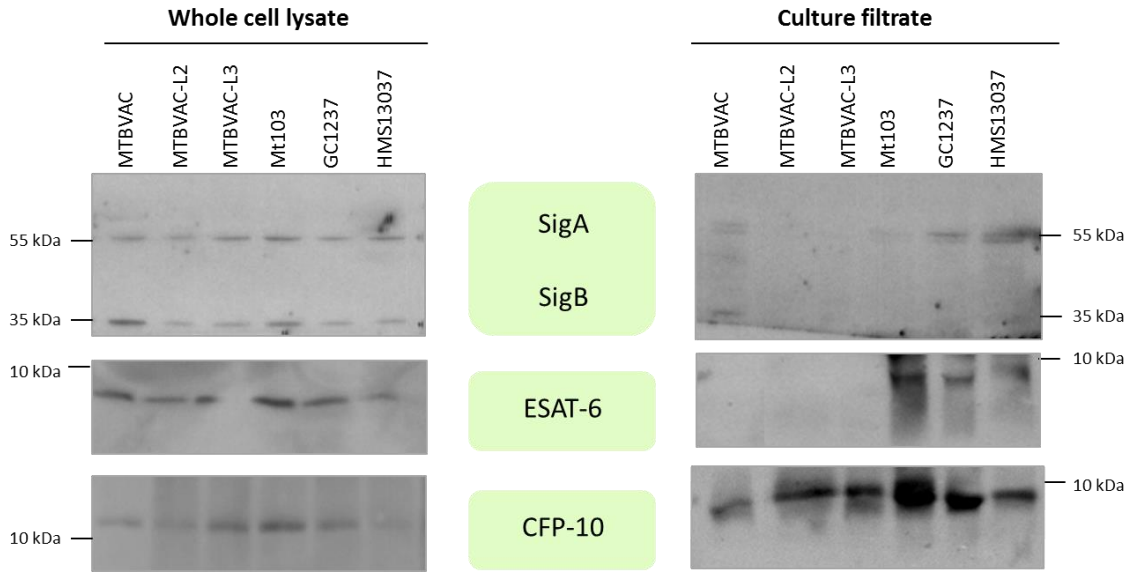


Figure 47. Western-blot analysis. ESAT-6 and CFP-10 detection in whole cell lysate and culture filtrate of MTBVAC, MTBVAC-L2, MTBVAC-L3 and WT strains. SigA was detected as load control in whole cell lysate and lysis control in the culture filtrate.

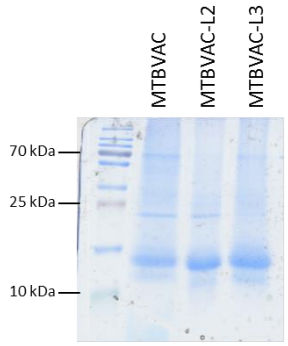


Figure 48. Coomassie staining. Coomassie staining of supernatant fractions of MTBVAC, MTBVAC-L2 and MTBVAC-L3.

PhoP regulon is downregulated in MTBVAC strains compared to WT strains

The transcription factor PhoP controls the expression of approximately 2% of *M. tuberculosis* genes (74-76). Expression of three PhoP-regulated genes in MTBVAC strains was compared with the WT strains by qRT-PCR (*mcr7*, *pks2* and *pks3* genes) (Figure 49).

mcr7 gene, which codes for the nc-RNA *mcr7*, is the most prominent site of PhoP regulation (75), and complete downregulation was observed in the three MTBVAC strains compared to their parental strains. Similar profile was observed in *pks2* gene, while a controversial result was observed in expression of *pks3* in lineage 3 (higher expression in *phoP* mutant than in the WT strain was observed). New RNA extractions of lineage 3 strains should be obtained to evaluate whether the expression profile is maintained. Further studies, including sequence the binding region of PhoP, will be required to understand differential expression of *pks3* in HMS13037 and the possibility to be differentially regulated compared to MTBVAC-L2 and MTBVAC.

Although the controversial result of increased expression of *pks3* in MTBVAC-L3, completely downregulation of the nc-RNA *mcr7* and *pks2* gene confirmed the downregulation of PhoP-regulated genes in MTBVAC-L2 and MTBVAC-L3.

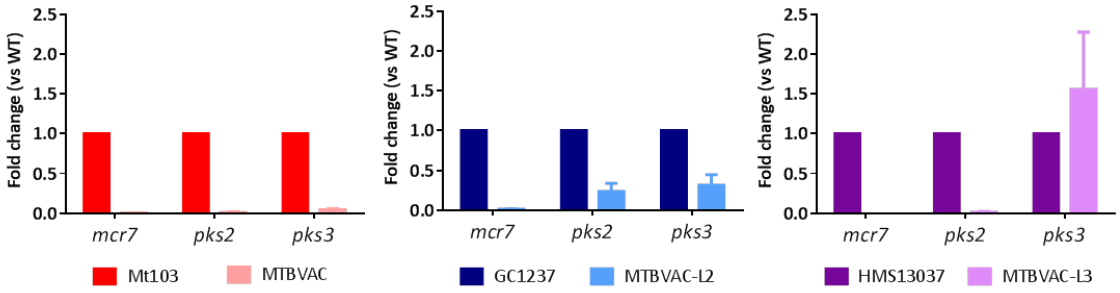


Figure 49. Expression of PhoP-regulated genes by qRT-PCR. *mcr7*, *pks2* and *pks2* were studied to compare their expression in MTBVAC, MTBVAC-L2 and MTBVAC-L3 strains compared to parental strains.

Expression of PDIM operon in the MTBVAC strains

fadD26 gene is expressed in an operon. Evaluation of expression of *ppsA* and *ppsB* were performed in the double mutants (MTBVAC, MTBVAC-L2 and MTBVAC-L3) and WT strains by qRT-PCR (Figure 50). *ppsA* and *ppsB* are 3′ downstream of *fadD26* in *M. tuberculosis* genome and code two PKSs involved in PDIM synthesis.

Results exhibited a remarkable downregulation of *ppsA* and *ppsB* in MTBVAC and MTBVAC-L2 compared to their parental strains (Mt103 and GC1237). By contrast, their expression was higher in MTBVAC-L3 than its WT strain (approximately 2 fold) (Figure 50).

Different scars are present in MTBVAC-L3 compared to MTBVAC and MTBVAC-L2 as the result of different removable cassettes used in the two methods used for *fadD26* deletion. *Res* sequence (55) is present in MTBVAC-L2 and MTBVAC while FRT sequence (57) is present in *fadD26* deleted gene in MTBVAC-L3 (60).

In MTBVAC, previous results of transcriptomics revealed the absence of expression of PDIM operon (*fadD26*, *ppsA-D*). A start transcriptional point is observed before *ppsE* gene and expression of *ppsE* and 3′downstream genes were observed in MTBVAC (Figure 51).

This previous result was confirmed by observation of complete downregulation of *ppsA* and *ppsB* genes by qRT-PCR. *Res* sequence is present in *fadD26* deleted gene after removal of antibiotic resistance gene from *res-Ωhyg-res* cassette. Same *res* scar is present after *fadD26* deletion in MTBVAC-L2 and comparable result of *ppsA* and *ppsB* downregulation was observed in this strain. Therefore, mutations obtained with the *res-Ωhyg-res* cassette disrupt the expression of the operon if the deleted gene is expressed in an operon.

In contrast to MTBVAC and MTBVAC-L2, downregulation of *ppsA* and *ppsB* was not observed in MTBVAC-L3 compared to the WT strain. This observation can be explained considering that *fadD26* deletion was obtained by BAC-recombineering strategy, after removal of antibiotic

resistance gene from FRT-*km*-FRT cassette (in pKD4 plasmid). This cassette has been described to create nonpolar gene deletions within operons (57).

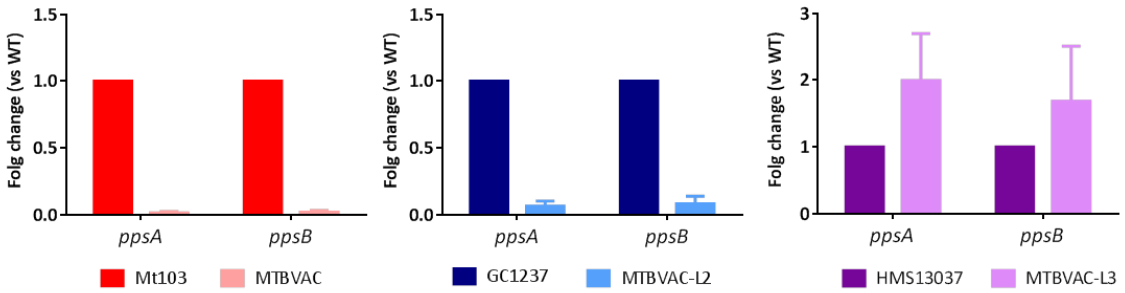


Figure 50. Expression of PDIM operon in *fadD26* mutants and WT strains. Expression of *ppsA* and *ppsB* was compared in MTBVAC, MTBVAC-L2 and MTBVAC-L3 to their respective parental strain by qRT-PCR.

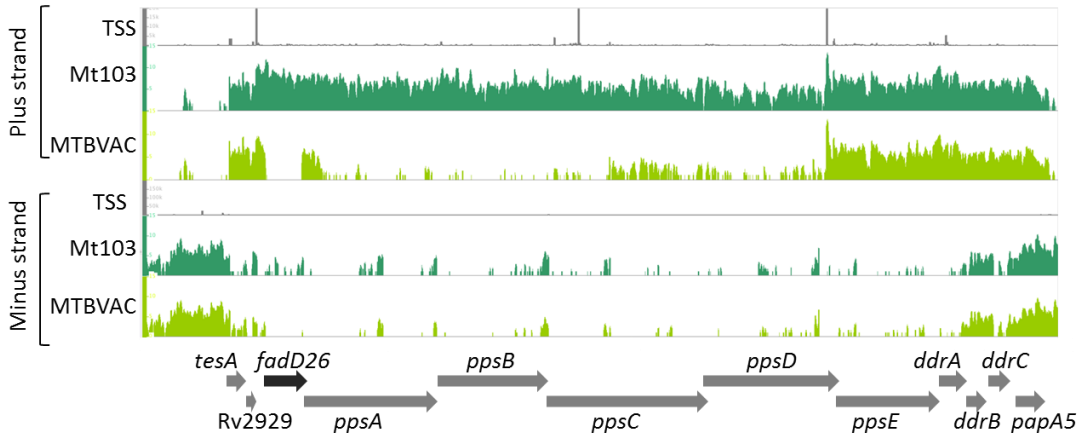


Figure 51. PDIM operon in Mt103 and MTBVAC. Start transcription sites (TSS) in plus and minus strands. RNA-seq of Mt103 and MTBVAC in plus and minus strains. Deletion of *fadD26* led to lack of transcription of *fadD26* and *ppsA-D* genes.

In vitro growth of MTBVAC-L2, MTBVAC-L3 and WT and strains

Growth curves were obtained to study possible differences in *in vitro* growth of the mutants compared to the WT strains. $OD_{590\ nm}$ was followed during three weeks in liquid medium (7H9-ADC-0.05% Tween 80).

Similar growth profile was observed in MTBVAC-L2::*hyg*, MTBVAC-L2, MTBVAC-L3::*hyg*, MTBVAC-L3 compared to WT strains GC1237, HMS13037 and Mt103. (Figure 52A). Different colony morphology in 7H10-ADC were observed in MTBVAC-L2 compared to WT strain GC1237 (Figure 52B) in agreement with small size colony observed in SO2 and MTBVAC in comparison with Mt103(52) (77). However, similar size colony was observed when HMS13037 and MTBVAC-L3 were compared (Figure 52C). Difficulties encountered to grow HMS13037 and different media culture (supplementation with OADC instead of ADC) could explain this result but further studies are required.

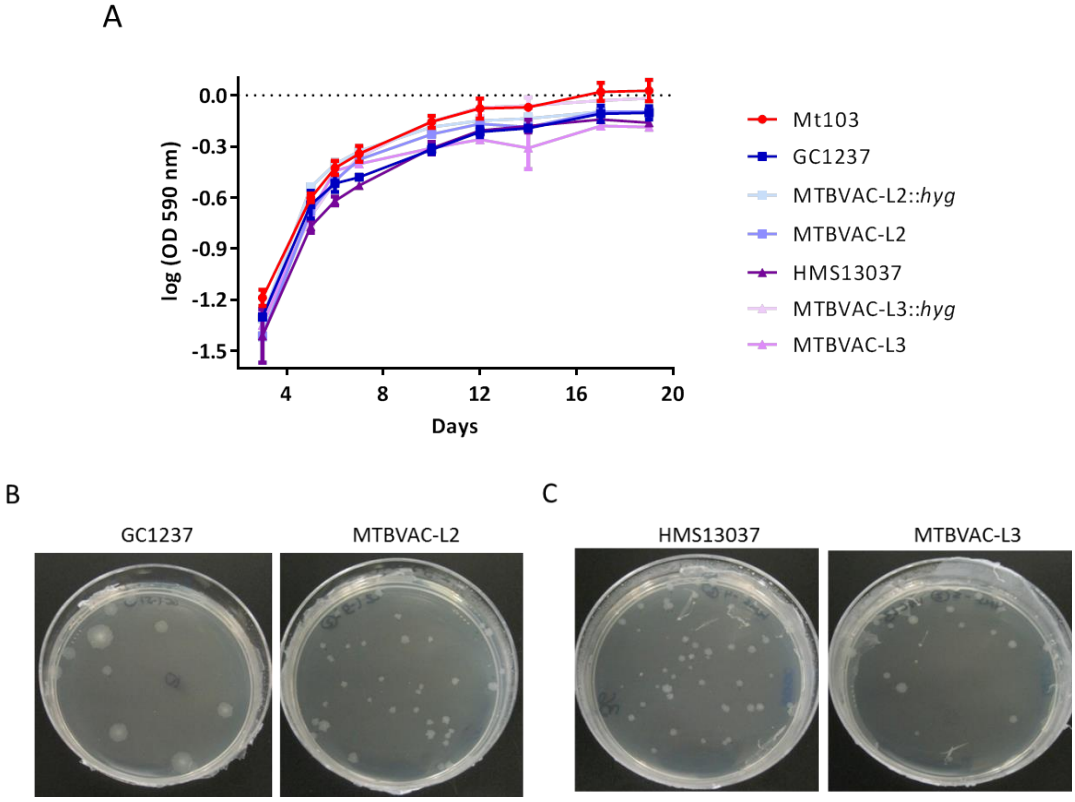


Figure 52. ***In vitro*** **growth of strains.** (A) Growth in liquid media of lineage 4 strain Mt103, lineage 2 strains, (GC1237, MTBVAC-L2::*hyg* and MTBVAC-L2) and lineage 3 strains (HMS13037, MTBVAC-L3::*hyg* and MTBVAC-L3). (B) Differential size colony of GC1237 and MTBVAC-L2 in 7H10-ADC. (C) Similar size colony of HMS13037 and MTBVAC-L3.

Conclusions

Two different techniques of genetic-engineering, suicide plasmids and BAC-recombineering, were used in parallel to obtain *phoP* and *fadD26* deletions in the clinical isolates GC1237 and HMS13037 from lineages 2 and 3 from *M. tuberculosis*. In MTBVAC-L2 both deletions were obtained with suicide plasmid strategy (same method that was used to obtain MTBVAC). In MTBVAC-L3, *fadD26* deletion was obtained by BAC-recombineering and *phoP* deletion was obtained by suicide plasmid.

Novel molecular characterization of PE_PGRS secretion was performed. MTBVAC and MTBVAC-L3 secrete PE_PGRS proteins in contrast to MTBVAC-L2. Lack of secretion of this subfamily of proteins that belongs to PE proteins has been recently related to increase *M. tuberculosis* virulence in mice although their role in protective efficacy was not observed in the mouse model.

Additional molecular characterization of MTBVAC-L2 and MTBVAC-L3 were performed to confirm maintenance of previous MTBVAC phenotypes described: MTBVAC-L2 and MTBVAC-L3 are unable to fix neutral red, both are unable to secrete ESAT-6 although CFP-10 is secreted and downregulation of PhoP regulon (*mcr7* and *pks2*). Further analyses are required to understand increased expression of *pks3* gene in MTBVAC-L3 compared to its WT strain.

Lack of expression *ppsA* and *ppsB* genes contained in PDIM operon were observed in MTBVAC and MTBVAC-L2 whereas this phenotype was not observed in MTBVAC-L3 comparing to their respective parental strains. Different methods and different cassettes used (*res-Ωhyg-res* or FRT-*km*-FRT) to obtain *fadD26* deletion could explain this result. FRT-*km*-FRT cassette can be used to create nonpolar gene deletions within operons while disruption of expression of the PDIM operon has been previously observed by RNA-seq when *res-Ωhyg-res* cassette was used and *res* scar was contained in *fadD26* deleted gene.

Thanks to the know-how in mycobacterial genetics, genetic engineering methods have been developed to reduce time to obtain knock-outs in *M. tuberculosis.*

References

1. **Boritsch EC, Brosch R.** 2016. Evolution of Mycobacterium tuberculosis: New Insights into Pathogenicity and Drug Resistance. *Microbiol Spectr* **4**.
2. **Brites D, Gagneux S.** 2017. The Nature and Evolution of Genomic Diversity in the Mycobacterium tuberculosis Complex. *Adv Exp Med Biol* **1019:**1-26.
3. **Malone KM, Gordon SV.** 2017. Mycobacterium tuberculosis Complex Members Adapted to Wild and Domestic Animals. *Adv Exp Med Biol* **1019:**135-154.
4. **Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM.** 2006. Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **103:**2869-2873.
5. **Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW, Gagneux S.** 2008. High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography. *PLoS Biol* **6:**e311.
6. **de Jong BC, Antonio M, Gagneux S.** 2010. Mycobacterium africanum--review of an important cause of human tuberculosis in West Africa. *PLoS Negl Trop Dis* **4:**e744.
7. **Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST.** 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A* **99:**3684-3689.
8. **Portevin D, Gagneux S, Comas I, Young D.** 2011. Human macrophage responses to clinical isolates from the Mycobacterium tuberculosis complex discriminate between ancient and modern lineages. *PLoS Pathog* **7:**e1001307.
9. **Manca C, Tsenova L, Freeman S, Barczak AK, Tovey M, Murray PJ, Barry C, Kaplan G.** 2005. Hypervirulent M. tuberculosis W/Beijing strains upregulate type I IFNs and increase expression of negative regulators of the Jak-Stat pathway. *J Interferon Cytokine Res* **25:**694-701.
10. **Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, Kaplan G, Barry CE, 3rd.** 2004. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431:**84-87.
11. **Coscolla M, Gagneux S.** 2014. Consequences of genomic diversity in Mycobacterium tuberculosis. *Semin Immunol* **26:**431-444.
12. **Sarkar R, Lenders L, Wilkinson KA, Wilkinson RJ, Nicol MP.** 2012. Modern lineages of Mycobacterium tuberculosis exhibit lineage-specific patterns of growth and cytokine induction in human monocyte-derived macrophages. *PLoS One* **7:**e43170.
13. **van Laarhoven A, Mandemakers JJ, Kleinnijenhuis J, Enaimi M, Lachmandas E, Joosten LA, Ottenhoff TH, Netea MG, van Soolingen D, van Crevel R.** 2013. Low induction of proinflammatory cytokines parallels evolutionary success of modern strains within the Mycobacterium tuberculosis Beijing genotype. *Infect Immun* **81:**3750-3756.
14. **Krishnan N, Malaga W, Constant P, Caws M, Tran TH, Salmons J, Nguyen TN, Nguyen DB, Daffe M, Young DB, Robertson BD, Guilhot C, Thwaites GE.** 2011. Mycobacterium tuberculosis lineage influences innate immune response and virulence and is associated with distinct cell envelope lipid profiles. *PLoS One* **6:**e23870.
15. **Stucki D, Brites D, Jeljeli L, Coscolla M, Liu Q, Trauner A, Fenner L, Rutaihwa L, Borrell S, Luo T, Gao Q, Kato-Maeda M, Ballif M, Egger M, Macedo R, Mardassi H, Moreno M, Tudo Vilanova G, Fyfe J, Globan M, Thomas J, Jamieson F, Guthrie JL, Asante-Poku A, Yeboah-Manu D, Wampande E, Ssengooba W, Joloba M, Henry Boom W, Basu I,**

Bower J, Saraiva M, Vaconcellos SEG, Suffys P, Koch A, Wilkinson R, Gail-Bekker L, Malla B, Ley SD, Beck HP, de Jong BC, Toit K, Sanchez-Padilla E, Bonnet M, Gil-Brusola A, Frank M, Penlap Beng VN, Eisenach K, Alani I, Wangui Ndung'u P, et al. 2016. Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. *Nat Genet* **48:**1535-1543.

16. **Comas I, Gagneux S.** 2011. A role for systems epidemiology in tuberculosis research. *Trends Microbiol* **19:**492-500.
17. **Shitikov E, Kolchenko S, Mokrousov I, Bespyatykh J, Ischenko D, Ilina E, Govorun V.** 2017. Evolutionary pathway analysis and unified classification of East Asian lineage of Mycobacterium tuberculosis. *Sci Rep* **7:**9227.
18. **Luo T, Comas I, Luo D, Lu B, Wu J, Wei L, Yang C, Liu Q, Gan M, Sun G, Shen X, Liu F, Gagneux S, Mei J, Lan R, Wan K, Gao Q.** 2015. Southern East Asian origin and coexpansion of Mycobacterium tuberculosis Beijing family with Han Chinese. *Proc Natl Acad Sci U S A* **112:**8136-8141.
19. **Tsolaki AG, Gagneux S, Pym AS, Goguet de la Salmoniere YO, Kreiswirth BN, Van Soolingen D, Small PM.** 2005. Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of Mycobacterium tuberculosis. *J Clin Microbiol* **43:**3185-3191.
20. **Coll F, McNerney R, Guerra-Assuncao JA, Glynn JR, Perdigao J, Viveiros M, Portugal I, Pain A, Martin N, Clark TG.** 2014. A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. *Nat Commun* **5:**4812.
21. **Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, Fyfe J, Garcia-Garcia L, Rastogi N, Sola C, Zozio T, Guerrero MI, Leon CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendon A, Sifuentes-Osornio J, Ponce de Leon A, Cave MD, Fleischmann R, Whittam TS, Alland D.** 2006. Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J Bacteriol* **188:**759-772.
22. **Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, Blum MG, Rusch-Gerdes S, Mokrousov I, Aleksic E, Allix-Beguec C, Antierens A, Augustynowicz-Kopec E, Ballif M, Barletta F, Beck HP, Barry CE, 3rd, Bonnet M, Borroni E, Campos-Herrero I, Cirillo D, Cox H, Crowe S, Crudu V, Diel R, Drobniewski F, Fauville-Dufaux M, Gagneux S, Ghebremichael S, Hanekom M, Hoffner S, Jiao WW, Kalon S, Kohl TA, Kontsevaya I, Lillebaek T, Maeda S, Nikolayevskyy V, Rasmussen M, Rastogi N, Samper S, Sanchez-Padilla E, Savic B, Shamputa IC, Shen A, Sng LH, Stakenas P, Toit K, Varaine F, Vukovic D, et al.** 2015. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nat Genet* **47:**242-249.
23. **Mokrousov I, Ly HM, Otten T, Lan NN, Vyshnevskyi B, Hoffner S, Narvskaya O.** 2005. Origin and primary dispersal of the Mycobacterium tuberculosis Beijing genotype: clues from human phylogeography. *Genome Res* **15:**1357-1364.
24. **Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, Allix C, Aristimuno L, Arora J, Baumanis V, Binder L, Cafrune P, Cataldi A, Cheong S, Diel R, Ellermeier C, Evans JT, Fauville-Dufaux M, Ferdinand S, Garcia de Viedma D, Garzelli C, Gazzola L, Gomes HM, Guttierez MC, Hawkey PM, van Helden PD, Kadival GV, Kreiswirth BN, Kremer K, Kubin M, Kulkarni SP, Liens B, Lillebaek T, Ho ML, Martin C, Martin C, Mokrousov I, Narvskaia O, Ngeow YF, Naumann L, Niemann S, Parwati I, Rahim Z, Rasolofo-Razanamparany V, Rasolonavalona T, Rossetti ML, Rusch-Gerdes S, Sajduda A, Samper S, Shemyakin IG, et al.** 2006. Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol* **6:**23.

25. **Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN.** 2002. Global dissemination of the Mycobacterium tuberculosis W-Beijing family strains. *Trends Microbiol* **10:**45-52.
26. **Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, Corander J, Bryant J, Parkhill J, Nejentsev S, Horstmann RD, Brown T, Drobniewski F.** 2014. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet* **46:**279-286.
27. **Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, Bentley SD, Parkhill J, Nejentsev S, Hoffner SE, Horstmann RD, Brown T, Drobniewski F.** 2012. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res* **22:**735-745.
28. **Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, Johnston JC, Gardy J, Lipsitch M, Fortune SM.** 2013. Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat Genet* **45:**784-790.
29. **Gonzalo-Asensio J, Perez I, Aguilo N, Uranga S, Pico A, Lampreave C, Cebollada A, Otal I, Samper S, Martin C.** 2018. New insights into the transposition mechanisms of IS6110 and its dynamic distribution between Mycobacterium tuberculosis Complex lineages. *PLoS Genet* **14:**e1007282.
30. **Constant P, Perez E, Malaga W, Laneelle MA, Saurel O, Daffe M, Guilhot C.** 2002. Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the Mycobacterium tuberculosis complex. Evidence that all strains synthesize glycosylated p-hydroxybenzoic methyl esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. *J Biol Chem* **277:**38148-38158.
31. **Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK.** 2001. Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha -crystallin. *Proc Natl Acad Sci U S A* **98:**7534-7539.
32. **Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, Schoolnik GK.** 2003. Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *J Exp Med* **198:**705-713.
33. **Domenech P, Kolly GS, Leon-Solis L, Fallow A, Reed MB.** 2010. Massive gene duplication event among clinical isolates of the Mycobacterium tuberculosis W/Beijing family. *J Bacteriol* **192:**4562-4570.
34. **Rose G, Cortes T, Comas I, Coscolla M, Gagneux S, Young DB.** 2013. Mapping of genotype-phenotype diversity among clinical isolates of mycobacterium tuberculosis by sequence-based transcriptional profiling. *Genome Biol Evol* **5:**1849-1862.
35. **Dormans J, Burger M, Aguilar D, Hernandez-Pando R, Kremer K, Roholl P, Arend SM, van Soolingen D.** 2004. Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different Mycobacterium tuberculosis genotypes in a BALB/c mouse model. *Clin Exp Immunol* **137:**460-468.
36. **Lopez B, Aguilar D, Orozco H, Burger M, Espitia C, Ritacco V, Barrera L, Kremer K, Hernandez-Pando R, Huygen K, van Soolingen D.** 2003. A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. *Clin Exp Immunol* **133:**30-37.
37. **Parwati I, van Crevel R, van Soolingen D.** 2010. Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *Lancet Infect Dis* **10:**103-111.
38. **Abebe F, Bjune G.** 2006. The emergence of Beijing family genotypes of Mycobacterium tuberculosis and low-level protection by bacille Calmette-Guerin (BCG) vaccines: is there a link? *Clin Exp Immunol* **145:**389-397.
39. **Jeon BY, Derrick SC, Lim J, Kolibab K, Dheenadhayalan V, Yang AL, Kreiswirth B, Morris SL.** 2008. Mycobacterium bovis BCG immunization induces protective immunity

against nine different Mycobacterium tuberculosis strains in mice. *Infect Immun* **76:**5173-5180.

40. **Ordway DJ, Shang S, Henao-Tamayo M, Obregon-Henao A, Nold L, Caraway M, Shanley CA, Basaraba RJ, Duncan CG, Orme IM.** 2011. Mycobacterium bovis BCG-mediated protection against W-Beijing strains of Mycobacterium tuberculosis is diminished concomitant with the emergence of regulatory T cells. *Clin Vaccine Immunol* **18:**1527-1535.
41. **Kremer K, van-der-Werf MJ, Au BK, Anh DD, Kam KM, van-Doorn HR, Borgdorff MW, van-Soolingen D.** 2009. Vaccine-induced immunity circumvented by typical Mycobacterium tuberculosis Beijing strains. *Emerg Infect Dis* **15:**335-339.
42. **Le TK, Bach KH, Ho ML, Le NV, Nguyen TN, Chevrier D, Guesdon JL.** 2000. Molecular fingerprinting of Mycobacterium tuberculosis strains isolated in Vietnam using IS6110 as probe. *Tuber Lung Dis* **80:**75-83.
43. **Anh DD, Borgdorff MW, Van LN, Lan NT, van Gorkom T, Kremer K, van Soolingen D.** 2000. Mycobacterium tuberculosis Beijing genotype emerging in Vietnam. *Emerg Infect Dis* **6:**302-305.
44. **van Crevel R, Nelwan RH, de Lenne W, Veeraragu Y, van der Zanden AG, Amin Z, van der Meer JW, van Soolingen D.** 2001. Mycobacterium tuberculosis Beijing genotype strains associated with febrile response to treatment. *Emerg Infect Dis* **7:**880-883.
45. **Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM.** 2011. Mycobacterium tuberculosis Beijing genotype: a template for success. *Tuberculosis (Edinb)* **91:**510-523.
46. **Mokrousov I.** 2017. Revisiting the Hunter Gaston discriminatory index: Note of caution and courses of change. *Tuberculosis (Edinb)* **104:**20-23.
47. **Ates LS, Dippenaar A, Ummels R, Piersma SR, van der Woude AD, van der Kuij K, Le Chevalier F, Mata-Espinosa D, Barrios-Payan J, Marquina-Castillo B, Guapillo C, Jimenez CR, Pain A, Houben ENG, Warren RM, Brosch R, Hernandez-Pando R, Bitter W.** 2018. Mutations in ppe38 block PE_PGRS secretion and increase virulence of Mycobacterium tuberculosis. *Nat Microbiol* **3:**181-188.
48. **Chen YY, Chang JR, Huang WF, Hsu SC, Kuo SC, Sun JR, Dou HY.** 2014. The pattern of cytokine production in vitro induced by ancient and modern Beijing Mycobacterium tuberculosis strains. *PLoS One* **9:**e94296.
49. **Ribeiro SC, Gomes LL, Amaral EP, Andrade MR, Almeida FM, Rezende AL, Lanes VR, Carvalho EC, Suffys PN, Mokrousov I, Lasunskaia EB.** 2014. Mycobacterium tuberculosis strains of the modern sublineage of the Beijing family are more likely to display increased virulence than strains of the ancient sublineage. *J Clin Microbiol* **52:**2615-2624.
50. **Kato-Maeda M, Shanley CA, Ackart D, Jarlsberg LG, Shang S, Obregon-Henao A, Harton M, Basaraba RJ, Henao-Tamayo M, Barrozo JC, Rose J, Kawamura LM, Coscolla M, Fofanov VY, Koshinsky H, Gagneux S, Hopewell PC, Ordway DJ, Orme IM.** 2012. Beijing sublineages of Mycobacterium tuberculosis differ in pathogenicity in the guinea pig. *Clin Vaccine Immunol* **19:**1227-1237.
51. **Jackson M, Raynaud C, Laneelle MA, Guilhot C, Laurent-Winter C, Ensergueix D, Gicquel B, Daffe M.** 1999. Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the Mycobacterium tuberculosis cell envelope. *Mol Microbiol* **31:**1573-1587.
52. **Arbues A, Aguilo JI, Gonzalo-Asensio J, Marinova D, Uranga S, Puentes E, Fernandez C, Parra A, Cardona PJ, Vilaplana C, Ausina V, Williams A, Clark S, Malaga W, Guilhot C, Gicquel B, Martin C.** 2013. Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated M. tuberculosis-based vaccine to enter clinical trials. *Vaccine* **31:**4867-4873.

53. **Alonso H, Aguilo JI, Samper S, Caminero JA, Campos-Herrero MI, Gicquel B, Brosch R, Martin C, Otal I.** 2011. Deciphering the role of IS6110 in a highly transmissible Mycobacterium tuberculosis Beijing strain, GC1237. *Tuberculosis (Edinb)* **91:**117-126.
54. **Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, Inwald JK, Golby P, Garcia JN, Hewinson RG, Behr MA, Quail MA, Churcher C, Barrell BG, Parkhill J, Cole ST.** 2007. Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* **104:**5596-5601.
55. **Malaga W, Perez E, Guilhot C.** 2003. Production of unmarked mutations in mycobacteria using site-specific recombination. *FEMS Microbiol Lett* **219:**261-268.
56. Ainhoa Arbués, PhD Thesis, Univeristy of Zaragoza
57. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97:**6640-6645.
58. **van Kessel JC, Hatfull GF.** 2007. Recombineering in Mycobacterium tuberculosis. *Nat Methods* **4:**147-152.
59. **van Kessel JC, Marinelli LJ, Hatfull GF.** 2008. Recombineering mycobacteria and their phages. *Nat Rev Microbiol* **6:**851-857.
60. **Song H, Niederweis M.** 2007. Functional expression of the Flp recombinase in Mycobacterium bovis BCG. *Gene* **399:**112-119.
61. **Brosch R, Gordon SV, Billault A, Garnier T, Eiglmeier K, Soravito C, Barrell BG, Cole ST.** 1998. Use of a Mycobacterium tuberculosis H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. *Infect Immun* **66:**2221-2229.
62. **Ates LS, Sayes F, Frigui W, Ummels R, Damen MPM, Bottai D, Behr MA, van Heijst JWJ, Bitter W, Majlessi L, Brosch R.** 2018. RD5-mediated lack of PE_PGRS and PPE-MPTR export in BCG vaccine strains results in strong reduction of antigenic repertoire but little impact on protection. *PLoS Pathog* **14:**e1007139.
63. **McEvoy CR, van Helden PD, Warren RM, Gey van Pittius NC.** 2009. Evidence for a rapid rate of molecular evolution at the hypervariable and immunogenic Mycobacterium tuberculosis PPE38 gene region. *BMC Evol Biol* **9:**237.
64. **Dubos RJ, Middlebrook G.** 1948. Cytochemical reaction of virulent tubercle bacilli. *Am Rev Tuberc* **58:**698.
65. **Cardona PJ, Soto CY, Martin C, Giquel B, Agusti G, Andreu N, Guirado E, Sirakova T, Kolattukudy P, Julian E, Luquin M.** 2006. Neutral-red reaction is related to virulence and cell wall methyl-branched lipids in Mycobacterium tuberculosis. *Microbes Infect* **8:**183-190.
66. **Gonzalo Asensio J, Maia C, Ferrer NL, Barilone N, Laval F, Soto CY, Winter N, Daffe M, Gicquel B, Martin C, Jackson M.** 2006. The virulence-associated two-component PhoP-PhoR system controls the biosynthesis of polyketide-derived lipids in Mycobacterium tuberculosis. *J Biol Chem* **281:**1313-1316.
67. **Frigui W, Bottai D, Majlessi L, Monot M, Josselin E, Brodin P, Garnier T, Gicquel B, Martin C, Leclerc C, Cole ST, Brosch R.** 2008. Control of M. tuberculosis ESAT-6 secretion and specific T cell recognition by PhoP. *PLoS Pathog* **4:**e33.
68. **Aguilo N, Gonzalo-Asensio J, Alvarez-Arguedas S, Marinova D, Gomez AB, Uranga S, Spallek R, Singh M, Audran R, Spertini F, Martin C.** 2017. Reactogenicity to major tuberculosis antigens absent in BCG is linked to improved protection against Mycobacterium tuberculosis. *Nat Commun* **8:**16085.
69. **Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, Williamson RA, Carr MD.** 2002. Conclusive evidence that the major T-cell antigens of the Mycobacterium tuberculosis complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6*CFP-10 complex. Implications for pathogenesis and virulence. *J Biol Chem* **277:**21598-21603.

70. **Gao LY, Guo S, McLaughlin B, Morisaki H, Engel JN, Brown EJ.** 2004. A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol Microbiol* **53:**1677-1693.
71. **Pang X, Samten B, Cao G, Wang X, Tvinnereim AR, Chen XL, Howard ST.** 2013. MprAB regulates the espA operon in Mycobacterium tuberculosis and modulates ESX-1 function and host cytokine response. *J Bacteriol* **195:**66-75.
72. **van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, Brenner M, Peters PJ.** 2007. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129:**1287-1298.
73. **Houben D, Demangel C, van Ingen J, Perez J, Baldeon L, Abdallah AM, Caleechurn L, Bottai D, van Zon M, de Punder K, van der Laan T, Kant A, Bossers-de Vries R, Willemsen P, Bitter W, van Soolingen D, Brosch R, van der Wel N, Peters PJ.** 2012. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol* **14:**1287-1298.
74. **Gonzalo-Asensio J, Mostowy S, Harders-Westerveen J, Huygen K, Hernandez-Pando R, Thole J, Behr M, Gicquel B, Martin C.** 2008. PhoP: a missing piece in the intricate puzzle of Mycobacterium tuberculosis virulence. *PLoS One* **3:**e3496.
75. **Solans L, Gonzalo-Asensio J, Sala C, Benjak A, Uplekar S, Rougemont J, Guilhot C, Malaga W, Martin C, Cole ST.** 2014. The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in Mycobacterium tuberculosis. *PLoS Pathog* **10:**e1004183.
76. **Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I.** 2006. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol* **60:**312-330.
77. **Perez E, Samper S, Bordas Y, Guilhot C, Gicquel B, Martin C.** 2001. An essential role for phoP in Mycobacterium tuberculosis virulence. *Mol Microbiol* **41:**179-187.

Chapter 2

Preclinical studies of new live attenuated vaccines based on *phoP, fadD26* deletions in the mouse model

Abstract

New live attenuated vaccine candidates against TB should succeed tough preclinical studies of safety, protective efficacy and stability in different animal models before entering into clinical trials.

Following the Geneva consensus we have constructed two new vaccine candidates in two clinical isolates of lineages 2 and 3 of *M. tuberculosis*, based on the same deletions of MTBVAC (deletion of *phoP* and *fadD26* genes). We performed initial preclinical characterization of MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* of protective efficacy and safety in the mouse model.

The main objective is to evaluate the protective efficacy of MTBVAC, MTBVAC-L2 and MTBVAC-L3 vaccines against representative virulent isolates of the three modern lineages of *M. tuberculosis*. In case that lineage dependent protection is observed, it would be possible to combine the three vaccine candidates (MTBVAC, MTBVAC-L2 and MTBVAC-L3) as a trivalent vaccine, and evaluate the protective efficacy of the three strains in combination against modern lineages of *M. tuberculosis.* Since they contain identical deletions, recombination events could not revert the phenotype.

In protective efficacy experiments to evaluate lineage dependent protection, we observed that all double mutants (MTBVAC, MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg*) protected against the strains tested belonging to the modern lineages. However, no lineage dependent protection was observed.

Safety experiment in SCID mice of MTBVAC, MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* showed that they were attenuated but with significant differences. MTBVAC was the most attenuated while MTBVAC-L2::*hyg* was the least attenuated of the strains tested. BCG Pasteur and MTBVAC-L3::*hyg* exhibited an intermediate attenuated profile between MTBVAC and MTBVAC-L2 .

Introduction

Mouse model in TB research

Mice are the most widely model used in TB research. Small size, wide range of immunology reagents, possibility of knock-out animals and an immune system that resembles that of humans are the main advantages of this animal model (1-3).

Available knock-out mice helped in the study of TB disease, confirming the important role of IFN-γ, TNF and IL-12 in the control of *M. tuberculosis* infection (4-6). The mouse model have been also used to evaluate new antimicrobial drugs to study new therapies against TB (3).

To study safety and protective efficacy of new vaccine candidates, the mouse model is usually the first choice. A decrease of 0.7 $\log_{10}$ in CFU should be observed in new vaccine candidates, as in BCG-vaccinated mice in comparison with unvaccinated mice (7).

Despite the general notion that mice are resistant to TB, it has been described diversity of resistance to TB in different mice strains. Survival curves of different mice strains after an intravenously inoculation of 10^5 CFU have been described. Results exhibited two differentiated phenotypes, classified as susceptible or resistant mice strains (Figure 53). Major histocompatibility complex (MHC) haplotype was shown that can influence in susceptibility although other genes are more important (8). This susceptibility profiles in the different mouse strains were maintained after vaccination with BCG (9).

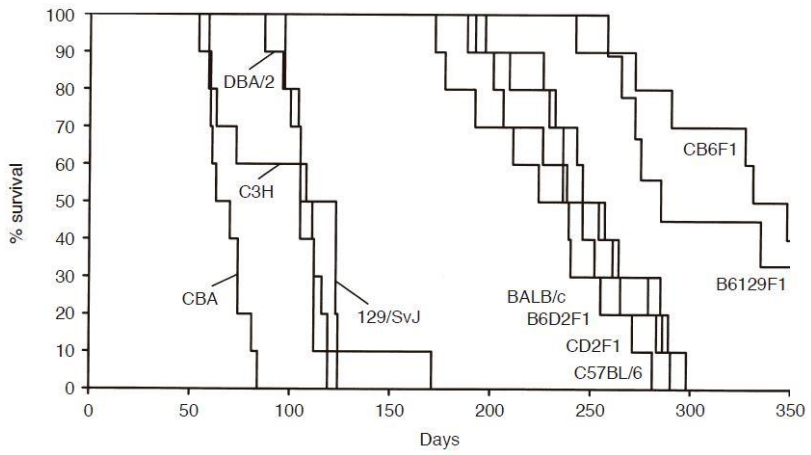


Figure 53. Survival curve of inbred strains and F_1 hybrids. Time of survival after inoculation of 10^5 CFU by intravenous route. CBA, C3H, DBA/2 and 129/SvJ are susceptible strains and C57BL/6 and BALB/c are resistant strains. All F_1 hybrids are resistant. From (8).

Kamath and colleagues observed that C3H (H-2k) mice recognize CFP-10 peptides in contrast to C57BL/6 (H-2b) and BALB/c (H-2d) (10). Extensive studies showed that C3H mice recognize ESAT-6 and CFP-10 but does not recognize Ag85B. In contrast, C57BL/6 recognizes ESAT-6 and Ag85B but no CFP-10. BALB/c does not recognize any of the three antigens. A reduction of half log was observed in MTBVAC compared to BCG Danish in C3H and similar protection of MTBVAC and BCG Danish was observed in C57BL/6 and BALB/c mice. When ESAT-6 and

CFP-10 were deleted in MTBVAC, same bacterial burden was observed compared with BCG Pasteur (Figure 54) (11).

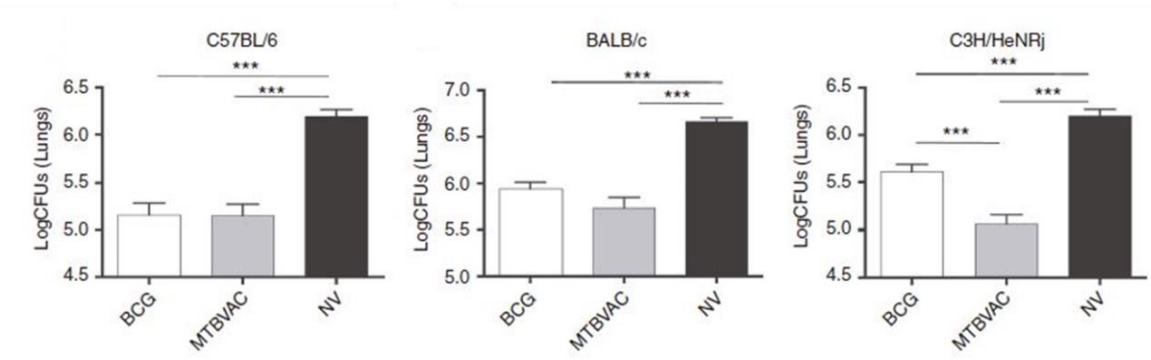


Figure 54. Protection of BCG Danish and MTBVAC against C57BL/6, BALB/c and C3H mice. Bacterial load in lungs in C57BL/6, BALB/c and C3H/HeNRj mice of BCG Danish and MTBVAC vaccinated mice and unvaccinated (NV) and challenged with H37Rv. MTBVAC conferred increased protection in C3H/HeNRj mice compared to BCG Danish vaccinated group. From (11).

Safety studies for TB research are usually performed in immunocompromised mice such as SCID, Nude or CD4 or interferon gamma knock-out (12).

SCID mice have severe combined immunodeficiency and these mice are defective in functional B and T cell due to a defect in the variable region gene recombination. Nude strains harbor a mutation that lead to absence of hair and lack of a functional thymus. These mice cannot produce functional T-cells (2).

In this chapter we evaluate attenuation of MTBVAC, MTBVAC-L2 and MTBVAC-L3 in immunocompromised SCID mice. We also study lineage-dependent protection of these vaccines against clinical isolates belonging to modern linages of *M. tuberculosis.*

Material and Methods

Quantification of bacterial glycerol stocks

Strains were cultured in 7H9-ADC-0.05% Tween 80 strain until log-phase was reached (Hyg was added if required). Cultures were centrifuged and pellets were resuspended in 0.1 volume of PBS-0.05% Tween 80. Suspensions were incubated at room temperature for 10 min to allow sedimentation of the clumps. Supernatants were centrifuged (1400 rpm for 10 min) to remove all clumps. 50% glycerol was added to the supernatants until final concentration of 5% glycerol and stored at -80 °C.

Duplicates of serial dilutions were plated onto 7H10 supplemented with ADC (or OADC for HMS13037 strain) to assess bacterial concentration in the stock. Hyg was added if needed.

Neutral red staining

Evaluation of fixation of neutral red was performed as described in Chapter 1, Materials and Methods section.

Ethics for animal procedures, facilities, mouse strains and anesthesia

The care and use of animals were performed accordingly with the Spanish Policy for Animal Protection RD53/2013 and the European Union Directive 2010/63 on the protection of animals used for experimental and other scientific purposes. All procedures were performed under Project License 50/14 approved by Ethic Committee for Animal Experiments of the University of Zaragoza. Experiments were performed in the facilities of *Centro de Investigación en Encefalopatías y Enfermedades Transmisibles Emergentes* (Zaragoza, Spain, Reference number ES 50 297 0012 009).

All animal procedures were performed by qualified researchers with the degree of "Carrying out procedures, Function C" (Orden ECC 566/2015) and taking into account the three R´s principles.

For protective and safety experiments, female C57BL/6JRj, C3H/HeNRj, and SCID (mice with Severe Combined Immunodeficiency) (Janvier Labs) were used for *in vivo* experiments. For virulence experiments, C3H, C57BL/6 and C57BL/6 $IgA^{-/-}$ mice were used.

Mice were allowed to acclimatized one week before the experiments started. Food and water were provided *ad libitum.* For animal procedures (vaccination and challenge), animals were anesthetized by inhalation route with 5% of isofluorane using a vaporizer (Isoba Vet®) and anesthesia was maintained with a 2% of isofluorane. At the end of the experiment for virulence and protective efficacy experiments, or when the 20% of lost weight was observed in the attenuation experiment, mice were euthanatized.

Virulence experiments

To confirm virulence of *M. tuberculosis* strains that will be used for the challenge in the mouse model to evaluate the protection efficacy of vaccine candidates, C3H or C57BL/6 $IgA^{-/-}$ mice were inoculated by intranasal route with different doses (200 CFU – 10^4 CFU). Four weeks post-challenge, bacterial load in lungs were quantified to evaluate the virulence of the strains in the mouse model.

Protective efficacy experiments

Groups of six female of C57BL/6JRj or C3H/HeNRj mice were used to evaluate protective efficacy. Mice were eight-week old when the experiment started. Mice were vaccinated subcutaneously with 10^6 bacteria per mouse contained in 100 µl of PBS. Vaccination was performed with BCG Pasteur, MTBVAC, MTBVAC-L2::*hyg* or MTBVAC-L3::*hyg*, and one unvaccinated group was used as control. Eight weeks later, mice were challenged by intranasal route with 200 CFU of a *M. tuberculosis* strain belonging to one of the modern lineages (lineages 2, 3 and 4). Challenge of 200 CFU was prepared in 40 µl of PBS. Preparations used for inoculation were plated to confirm correct dosage. Four weeks post-challenge, bacterial burden was evaluated in lungs and spleen. Serial dilutions were plated onto 7H10-ADC.

We are grateful to Dr. Gilla Kaplan for sharing HN878 and W4 strains and to Dr. Sofía Samper for sharing HCU3524 strain. These strains were used for the challenge in the protective efficacy experiments.

Attenuation experiment

Groups of six SCID mice were inoculated by intraperitoneal route with 10^6 bacteria in 100 µl PBS, equivalent to 2 times the dose of BCG recommended for humans. Mice were eight-weeks old when the experiment started. Mice were controlled and examined in case any abnormality of behavior was observed. Weight of mice was followed during the experiment. The euthanasia endpoint was defined at the point that the loss of weight was more than 20%.

Results

Protective efficacy of MTBVAC and MTBVAC-L2::*hyg* against H37Rv, a virulent strain from lineage 4

Group of six C57BL/6 mice were subcutaneously vaccinated with 10^6 CFU by subcutaneous route with MTBVAC, MTBVAC-L2::*hyg* or non-vaccinated (NV) as control. Eight weeks after vaccination, mice were challenged with 200 CFU by intranasal route with the virulent reference strain H37Rv. Four weeks post-challenge (endpoint experiment) bacterial burden were assessed in lung and spleen of the different groups (Figure 55A).

Results showed a reduction of one log in lungs in mice vaccinated with MTBVAC and 1.3 log reduction in vaccination with MTBVAC-L2::*hyg* in comparison with unvaccinated group control. In the spleen, a reduction of approximately 1.3 log was observed in both groups vaccinated with MTBVAC and MTBVAC-L2 compared to naive group (Figure 55B and C). These results demonstrate that both MTBVAC and MTBVAC-L2::*hyg* protect against H37Rv in C3H mice, with a slightly higher reduction of the bacillary load in MTBVAC-L2::*hyg* vaccinated mice.

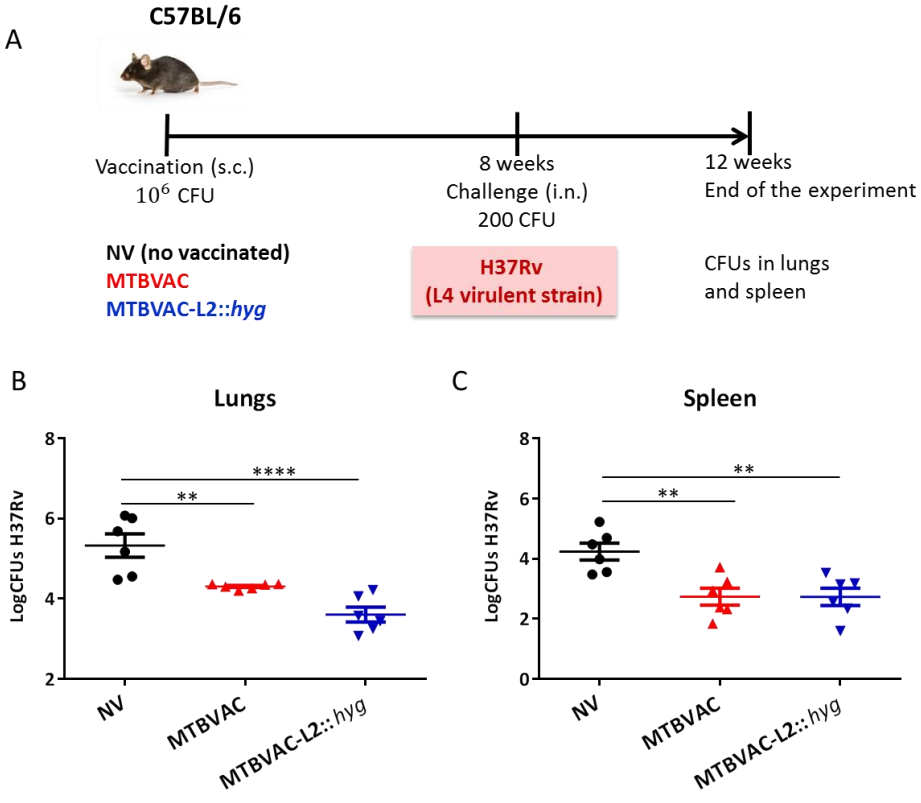


Figure 55. Protective efficacy of MTBVAC and MTBVAC-L2::*hyg* in C57BL/6 mice against H37Rv. (A) Schematic procedure. Bacterial burden in (B) lungs and (C) spleen of vaccinated and non-vaccinated mice (NV) four weeks post-challenge. Significant reduction in bacterial load was observed in mice vaccinated with MTBVAC and MTBVAC-L2::*hyg* in lungs and spleen compared to unvaccinated mice. All data are mean ± sem. Statistical analysis, by one-way ANOVA and *Bonferroni* post-test. **p<0.01, ****p<0.0001

Protective efficacy of MTBVAC and MTBVAC-L2::*hyg* against HN878, a strain from lineage 2

The same experiment as described before was performed to evaluate protective efficacy of MTBVAC and MTBVAC-L2::*hyg* against a strain from lineage 2 of *M. tuberculosis* (Figure 56A). The strain chosen for the challenge was a Beijing strain, HN878. This strain has been previously described as a hyper virulent strain (13).

C57BL/6 mice were vaccinated with MTBVAC, MTBVAC-L2::*hyg* or unvaccinated as control (six mice per group). Eight weeks post-vaccination, mice were challenged with 200 CFU of HN878 by intranasal route. Protective efficacy was evaluated by comparison of the number of bacilli in lungs and spleen.

Surprisingly, almost none bacteria was detected in lungs and spleens in none of the animals of all the groups (vaccinated and unvaccinated mice) (Figure 56B and C).

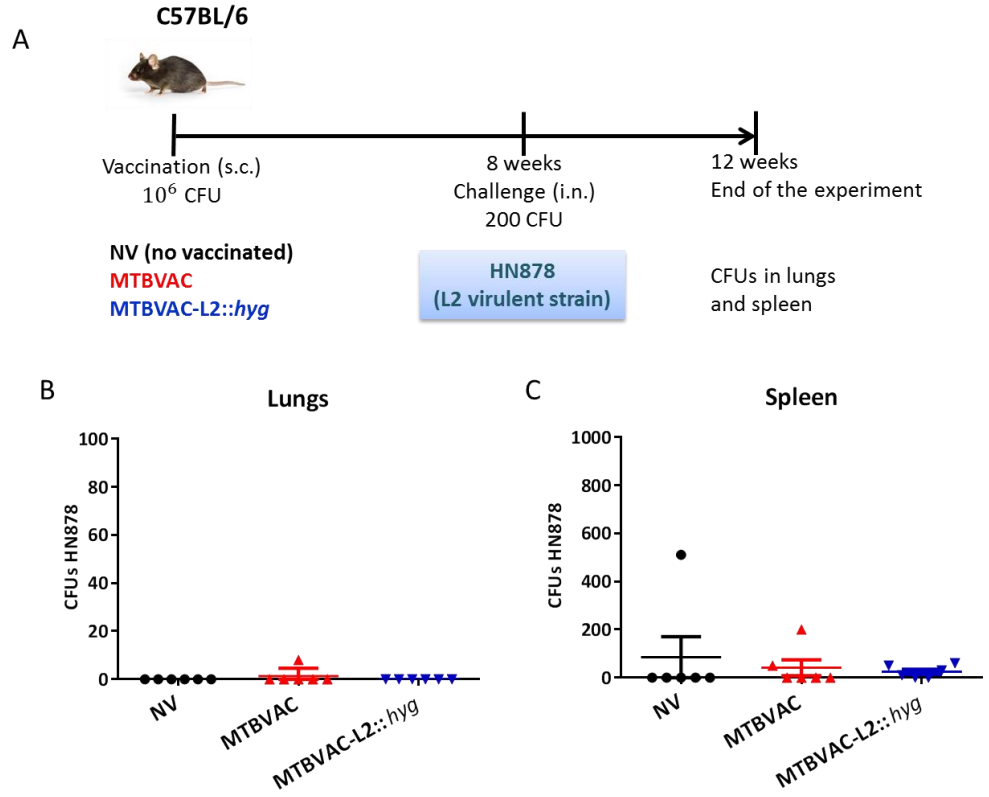


Figure 56. CFU detected in C57BL/6 mice after HN878 challenge. (A) Scheme of the **e**xperimental procedure. Homogenized of the organs were plated to evaluated the number or bacteria. (B) Total CFU in lungs. (C) Total CFU in spleen. All data are mean ± sem.

The results suggested that the strain of the challenge, HN878, had lost virulence. To further investigate this result, we performed a proof of virulence in which a high dose of bacteria was inoculated by intranasal route in a C57BL6 $IgA^{-/-}$ female mouse ($5x10^3$ bacilli). Four weeks later bacterial load in lungs was 10^4 (Figure 57A). This result was in agreement with the previous results and again suggested that this strain seemed to be no virulent in the mouse model.

Additionally, we performed a neutral red staining of the strain of the stock (which is quantified for the mice experiment) and the original received strain (HN878). Remarkably, the original

strain fixed the dye while the HNH878 strain used for the mice experiment has lost the ability to fix neutral red (Figure 57B). With all these findings we suggest that in the process of laboratory cultivation to obtain the stock quantified for mice experiments, the strain has suffered some changes in the external surface and probably has lost more than one type of methyl-branched lipids.

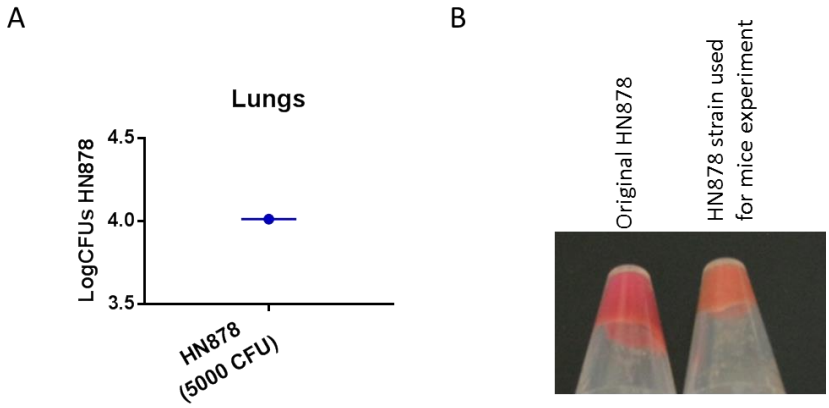


Figure 57. Bacterial load assessed in lungs and neutral red staining. (A) Bacterial burden in lungs four weeks after high dose challenge with HN878. (B) Variation of fixation of neutral red in the original HN878 and the strain HN878 used in the mice experiment.

A neutral red staining of six colonies obtained from the lungs were performed. We observed a reversion of neutral red phenotype in the six strains tested (Figure 58). All colonies were able to fix the dye after one month in the mouse whereas the strain was neutral red negative before being inoculated in the animal (Figure 57B). We hypothesized that during the infection in the mouse, some mutations could occur in the bacilli that could revert to the neutral red phenotype.

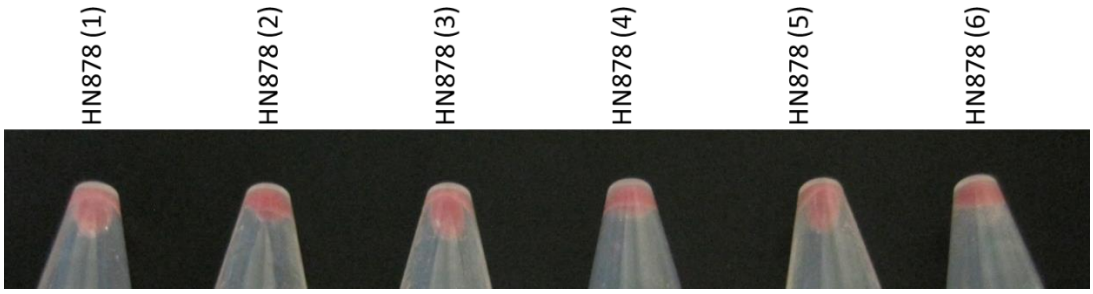


Figure 58. Fixation of neutral red dye. Fixation of neutral red was observed in the all colonies tested obtained from the lungs after 4 weeks in the mouse.

Virulence assays of lineages 2 and 3 strains of *M. tuberculosis* used in the mouse infection model

Taking into account previous results, two virulent strains from lineages 2 and 3 are required for the challenge to test protective efficacy of MTBVAC, MTBVAC-L2 and MTBVAC-L3 vaccine candidates against modern lineages of *M. tuberculosis.* For lineage 4, H37Rv will be used, the

reference laboratory strain of *M. tuberculosis*, widely used and a virulent phenotype confirmed in the mouse model.

Consequently, before protective efficacy assays, an experiment in mice to evaluate virulence of W4 and HCU3524 were performed. W4 and HCU3524 are two *M. tuberculosis* strains that belong to lineages 2-Beijing and lineage 3 respectively. HCU3524 was isolated at the Hospital Clínico Universitario Lozano Blesa (HCU).

Inoculation was performed by intranasal route and bacterial load was evaluated in lungs four weeks post-inoculation. 300 CFU of W4 (two C3H female mice and one C57BL/6 IgA$^{-/-}$ female mice) were inoculated and different doses of HCU3524 were evaluated (300, 1,5x10^3 and 2x10^4 CFU) in the mouse model. Two C3H female were used for the lowest dose, two C3H female for the intermediate dose and four C3H male were used for the highest dose (Figure 59A). At the end of the experiment, animals with the highest dose had started to lose weight and piloerection was observed. Both signs suggested that HCU3524 was virulent.

Results showed that in both challenges with 300 CFU the bacterial burden in lungs was 10^6 or increased in higher doses (Figure 59B). Therefore, once demonstrated the virulence of these strains in the mouse model, these two strains from lineages 2 and 3 of *M. tuberculosis* will be used for the challenge in protective efficacy experiments.

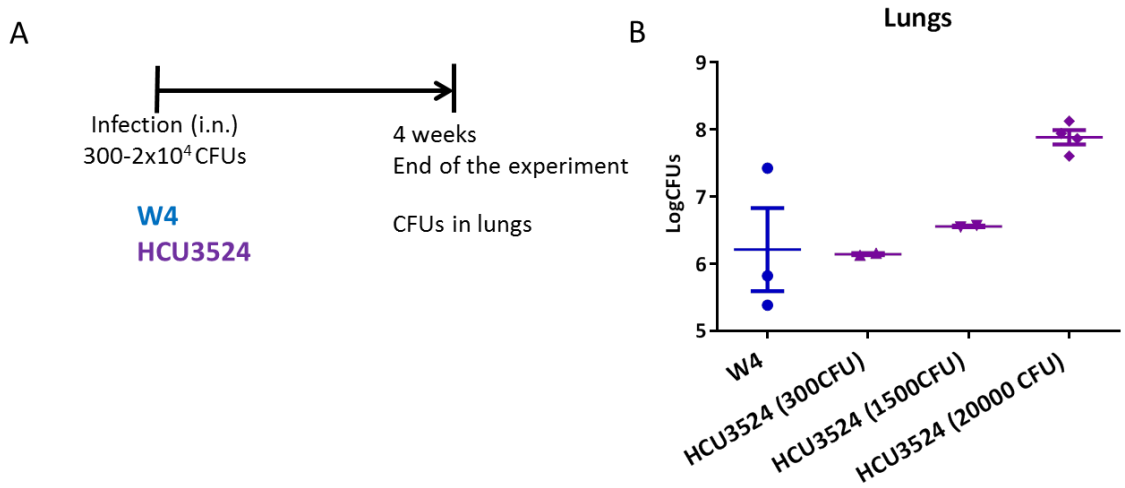


Figure 59. Log CFU of W4 or HCU3524 in lungs. (A) Schematic procedure to test virulence of the strains in the mouse model. (B) Log CFU of lineages 2 and 3 strains four week post-challenge. All data are mean ± sem

Protective efficacy of the three *phoP* and *fadD26* mutants (MTBVAC, MTBVAC-L2 and MTBVAC-L3) against modern lineages of *M. tuberculosis*

Once new double *phoP* and *fadD26* mutants in lineages 2 and 3 were obtained (MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg*), and virulence of strains from lineages 2, 3 and 4 of *M. tuberculosis* were confirmed to be used for the challenge, protective efficacy assays were performed.

Protective efficacy of the three MTBVAC vaccine candidates constructed in the three modern lineages of *M. tuberculosis* was evaluated against strains from the modern lineages. H37Rv from lineage 4, W4-Beijing strain from lineage 2 and HCU3524 from lineage 3 were used for the challenge. BCG Pasteur was used as a comparator for the vaccine candidates, as the current vaccine against TB.

The aim is to evaluate the protective efficacy of the new double mutants constructed and to study lineage-dependence protection. For these experiments, C3H/HeNRj mice were used.

The experiments of protection were performed as described below:

- **Protection of MTBVAC vaccine strains and BCG Pasteur against H37Rv (lineage 4 strain)**

C3H mice were subcutaneously vaccinated with 10^6 CFU of MTBVAC, MTBVAC-L2::*hyg*, MTBVAC-L3::*hyg*, BCG Pasteur or unvaccinated as control. After the resting period of eight weeks, mice were challenged by intranasal route with 200 CFU of H37Rv, a virulent strain from lineage 4. The end of experiment is defined four weeks after the challenge, and bacterial load was assessed in lungs and spleen (Figure 60A).

Results showed that all MTBVAC strains and BCG Pasteur had a comparably bacterial burden in lungs. An approximately 1.3 log reduction was observed in lungs all vaccinated groups compared to unvaccinated animals (Figure 60B). In the spleen, all vaccinated mice had diminished bacterial burden than unvaccinated group, and a reduction of 0.86 to 1.25 log was observed in vaccinated mice in comparison with unvaccinated group (Figure 60C). In conclusion, new constructions MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* conferred protection against H37Rv and this protection was comparable with MTBVAC and BCG Pasteur.

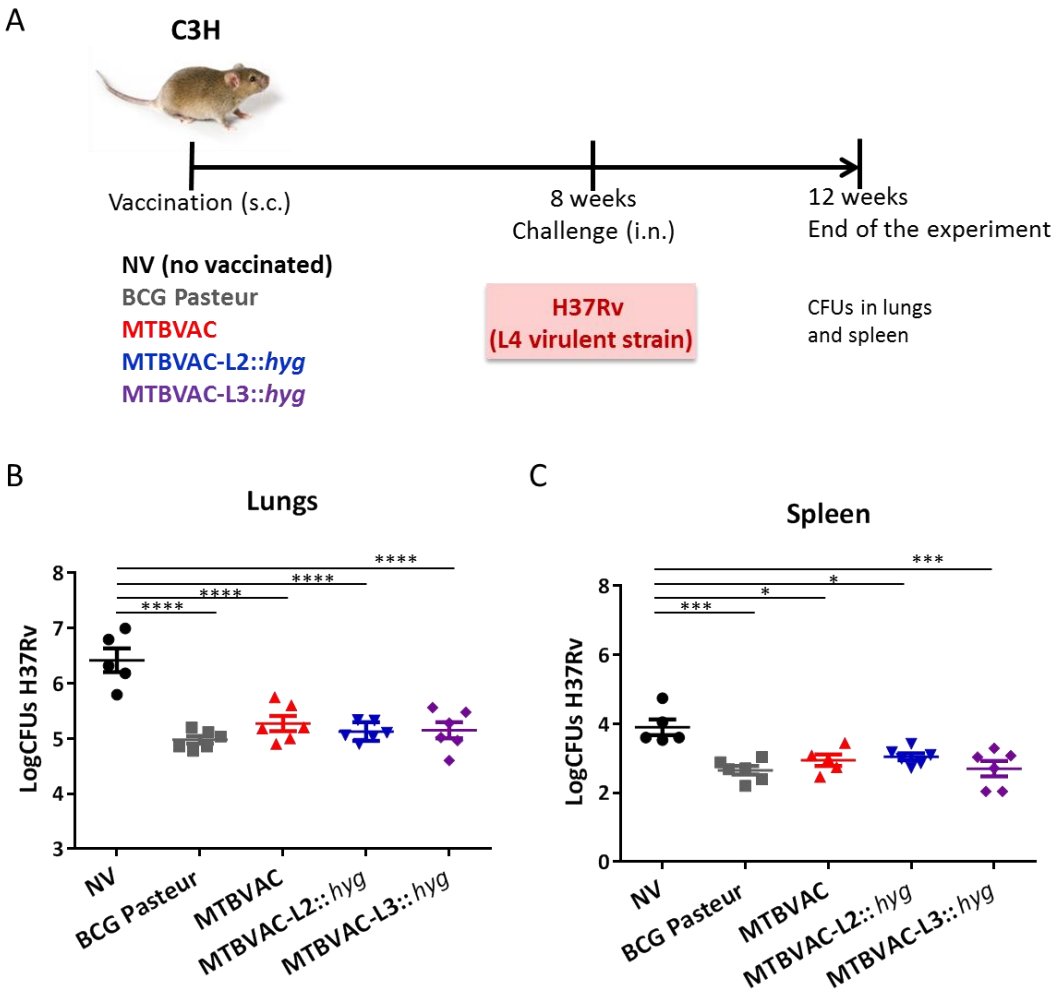


Figure 60. Protective efficacy of MTBVAC strains and BCG Pasteur in C3H mice against H37Rv. (A) Schematic representation of the experiment. (B) and (C), bacterial burden in lungs and spleen respectively. Significant and similar reduction in both organs were observed in all vaccinated groups compared to unvaccinated group. All data are mean ± sem. Statistical analysis by one-way ANOVA and *Bonferroni* post-test. *p<0.5, ***p<0.001, ****p<0.0001

- **Protection of MTBVAC vaccine strains and BCG Pasteur against W4 (linage 2, Beijing strain)**

Mice received a dose of 10^6 CFU of MTBVAC, MTBVAC-L2::*hyg*, MTBVAC-L3::*hyg* or BCG Pasteur by subcutaneous route. A group of unvaccinated mice was used as control. Two months later, mice were challenged by intranasal route with 200 CFU of W4 strain. Four weeks after the challenge, differences of bacterial load in lungs and spleen were assessed in all groups (Figure 61A).

All vaccinated mice showed a significant reduction in the bacterial burden in lungs compared to unvaccinated mice. The most notable decrease in bacillary load was observed in MTBVAC and MTBVAC-L2::*hyg* vaccinated groups (2.8 and 2.6 log respectively) in comparison with unvaccinated group. Mice vaccinated with MTBVAC-L3::*hyg* and BCG Pasteur also showed a significant decrease (2.3 and 2 log). Remarkably, a significant reduction in bacterial burden in

lungs was observed in MTBVAC and MTBVAC-L2::*hyg* vaccinated mice compared to BCG Pasteur vaccinated group (Figure 61B).

In the spleen, variable reduction of bacterial burden was observed (from 1.5 log until 2.2 log) compared to unvaccinated mice, but the reduction was significant in all vaccinated groups. MTBVAC vaccinated mice showed significantly lower bacterial burden compared to MTBVAC-L3::*hyg* vaccinated mice (Figure 61C).

These results suggest that all the vaccines tested protect against W4-Beijijng strain, but increased protection was conferred after vaccination with MTBVAC and MTBVAC-L2::*hyg* compared with the current vaccine BCG Pasteur.

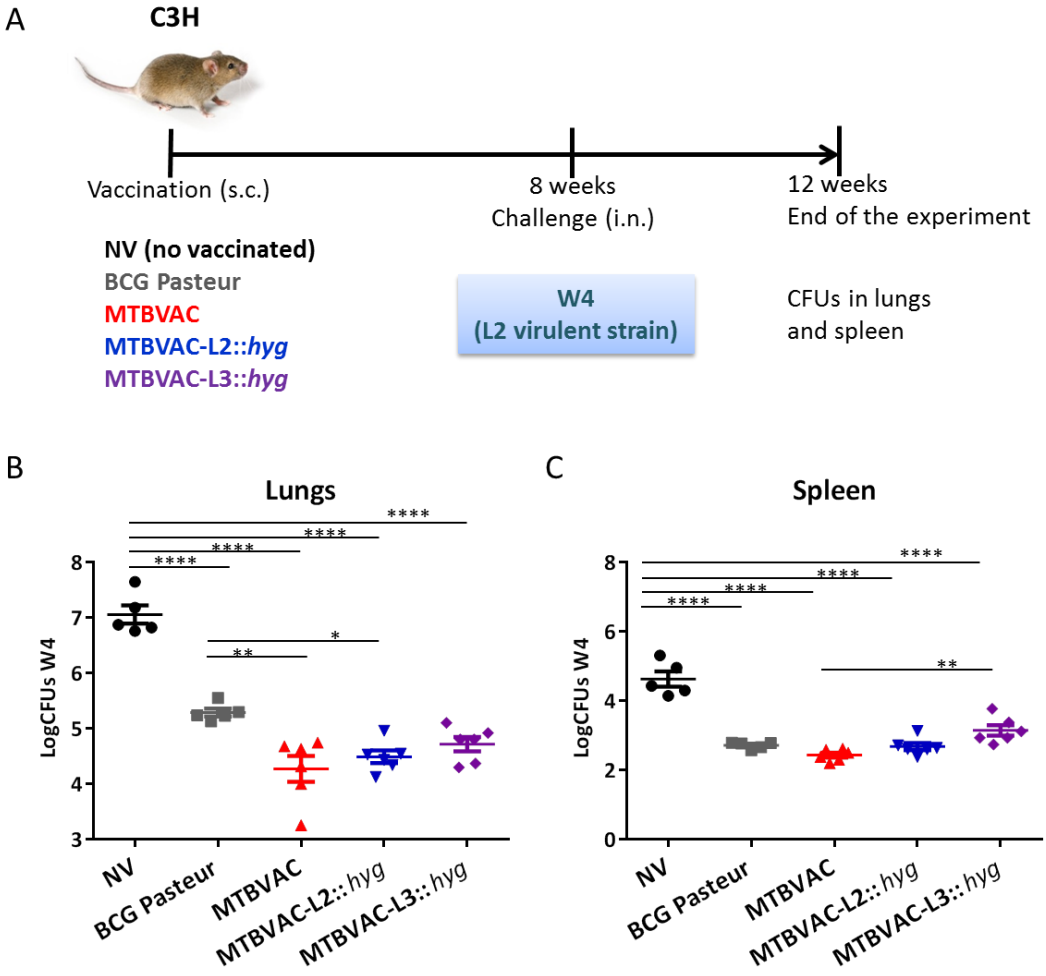


Figure 61. Protective efficacy of MTBVAC strains and BCG Pasteur in C3H mice against W4-Beijing strain. (A) Schematic representation of the experiment. (B) and (C), bacterial burden in lungs and spleen respectively. Significant reduction in both organs was observed in all vaccinated groups compared to unvaccinated mice. MTBVAC and MTBVAC-L2::*hyg* conferred significant increased protection compared to BCG Pasteur. All data are mean ± sem. Statistical analysis, by one-way ANOVA and *Bonferroni* post-test. *p<0.5, **p<0.01, ****p<0.0001

- **Protection of MTBVAC vaccine strains and BCG Pasteur against HCU3524 (lineage 3 strain)**

Experiment was performed as the two experiments described above. Briefly, mice were vaccinated subcutaneously (10^6 CFU) with MTBVAC, MTBVAC-L2::*hyg*, MTBVAC-L3::*hyg*, BCG Pasteur or non-vaccinated as control. Eight weeks post-vaccination, mice were challenged by intranasal route with 200 CFU of HCU3524, a virulent strain belonging to lineage 3 of *M. tuberculosis.* Four weeks later, bacterial burden was evaluated in all groups in lungs and spleen (Figure 62A).

Results showed variable protection of vaccinated animals compared to unvaccinated group. The highest reduction in bacterial burden was observed with MTBVAC and MTBVAC-L3::*hyg* with a reduction of 1.1 log in both vaccine candidates followed by vaccination with MTBVAC-L2::*hyg* with a reduction of 0.75 log and finally BCG Pasteur with a 0.6 log of reduction (Figure 62B).

In the spleen, the bacterial load in the saline group was relatively low, and no significant differences were observed in vaccinated groups compared to non-vaccinated with the exception of MTBVAC-L3::*hyg* (Figure 62C).

These results exhibit variable protection of the vaccines tested. MTBVAC and MTBVAC-L3::*hyg* conferred the best protection against the lineage 3 HCU3524 strain in C3H mice, followed by MTBVAC-L2::*hyg* and BCG Pasteur.

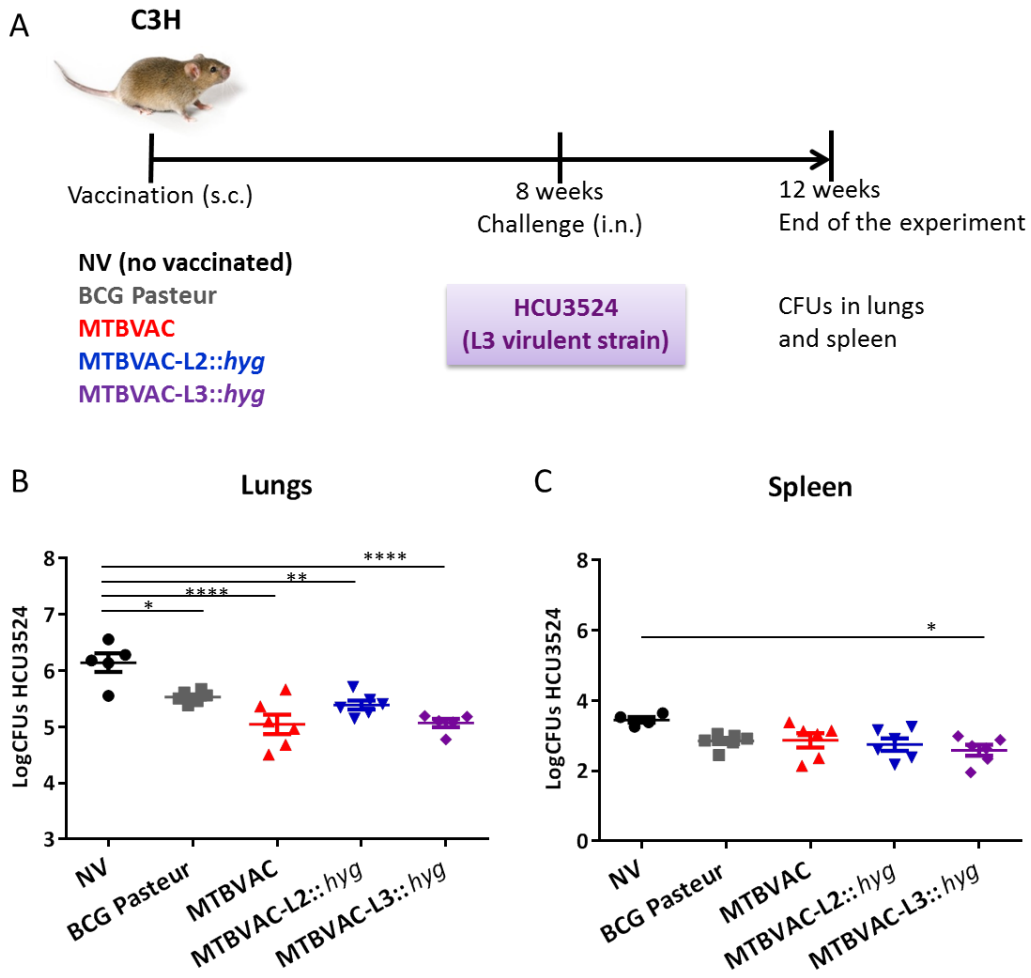


Figure 62. Protective efficacy of MTBVAC strains and BCG Pasteur in C3H mice against HCU3524 strain. (A) Schematic representation of the experiment. (B) and (C), bacterial burden in lungs and spleen respectively. Reduction of bacterial burden in lung was observed in all vaccinated groups compared to unvaccinated animals. Significant difference was observed in MTBVAC-L3 vaccinated mice in the spleen. Data are represented mean with SEM. Statistical analysis, by one-way ANOVA and *Bonferroni* post-test. *p<0.5, **p<0.01, ****p<0.0001

In summary, the three MTBVAC strains conferred protection against the modern strains of *M. tuberculosis* tested (H37Rv, W4 and HCU3524) without significant differences among them. This result suggests the lack of lineage dependent protection is congruent with hyperconservation of T cell epitopes (14). Even though the lack of lineage dependent protection, this work in consistent with the use of MTBVAC as a prototype of new vaccine against tuberculosis, because the double mutants conferred protection, independently of the modern lineage in which it is constructed.

Protective efficacy experiments were performed in C3H mice, which have a MHC haplotype that recognize ESAT-6 and CFP-10. Superior protection was observed in MTBVAC compared to BCG Danish in previous data published. When ESAT-6 and CFP-10 where deleted in MTBVAC,

these superior protection disappeared and equivalent protection were observed comparing BCG Danish and MTBVAC ΔESAT-6 ΔCFP-10 (11).

The expected superior protection in MTBVAC strains compared to BCG, because of recognition of ESAT-6 and CFP-10 was not observed with the exception of the experiment challenged with W4-lineage 2 strain. Further studies to understand this disagreement result are required, but one possibility could be the different BCG strains used in the previous study, BCG Danish (11), and the present work, BCG Pasteur. Both BCG Pasteur and Danish contain PDIM. Differences in the TCS *phoPR* have been described. Both contain the mutation Gly71Ile in PhoR (15), but BCG Danish additionally contains a 10 bp deletion in codon 91 in *phoR* (16). Additional polymorphisms could be involved in different phenotypes of BCG substrains that can influence in protective efficacy.

Several studies have tried to reveal differences of the diverse BCG strains and protection in different animal models. In mice, Zhang and colleagues demonstrated that BCG strains showed different levels of attenuation in SCID mice. Protective efficacy of BCG strains against H37Rv in BALB/c mice exhibited a trend that the more virulent the BCG strain was, the better protection conferred. In BCG Danish and BCG Pasteur, similar protection efficacy was observed (0.07 log reduction in BCG Danish compared to BCG Pasteur) although BCG Danish is more attenuated than BCG Pasteur (17). In other study, protective efficacy of 10 BCG substrains was compared in BALB/c mice. Comparing BCG Danish and Pasteur, increased reduction in bacterial burden was observed in BCG Danish (0.14 log reduction between BCG Danish and BCG Pasteur) (18). By contrast, in other mice study, a reduced bacterial burden was observed in BCG Pasteur vaccinated mice compared to BCG Danish after H37Rv challenge (19).

An outstanding result is the superior protection of MTBVAC and MTBVAC-L2::*hyg* against W4 Beijing strain in comparison with BCG Pasteur. Previous studies proposed the inefficacy of BCG in conferring protection against Beijing strains (20), although this is not widely accepted because of the controversial results obtained in protection experiments using BCG vaccination against Beijing strains (21, 22) (more detailed in Introduction section, Chapter 1, Lineage 2 *of M. tuberculosis* – East Asian Lineage).

Even though different protection was not observed using the mouse model, it is worth to remember that intralineage heterogenicity exist in *M. tuberculosis.* Future studies with a representative collection of clinical isolates would provide further insights into the lineage-dependent protection of the vaccines studied in this Thesis.

Evaluation of current and future clinical data of MTBVAC-vaccinated individuals would provide information about the protective efficacy of this vaccine against different lineages of *M. tuberculosis*. In this context, the new MTBVAC-L2 and MTBVAC-L3 constructed in this Thesis would pave the way to progress *phoP*, *fadD26*-based vaccines.

Safety experiment in SCID mice

Attenuation of the three MTBVAC strains and BCG Pasteur was evaluated by survival of immunocompromised SCID mice after inoculation of 10^6 CFU by intraperitoneal route (Figure 63A). In SCID mice, survival after and inoculation of 10^6 CFU of Mt103 by intraperitoneal route was 24 days (23). In this study, MTBVAC, MTBVAC-L2::*hyg*, MTBVAC-L3::*hyg* and BCG Pasteur showed attenuation in SCID mice. However, differences in the level of attenuation were observed. MTBVAC was the most attenuated strain of all strains assayed while MTBVAC-L2::*hyg* was the least attenuated strain of the strains tested. BCG Pasteur and MTBVAC-L3::*hyg* exhibited an intermediate attenuation between MTBVAC-L2::*hyg* and MTBVAC (Figure 63B).

The median of survival for MTBVAC-L2::*hyg* was 85.5 days, followed by MTBVAC-L3::*hyg* with 93.5 days, 108.5 days in BCG Pasteur and the most attenuated strain was MTBVAC with a median of survival of 125.5 days.

In conclusion, new constructions MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* showed attenuation in SCID mice, albeit both were less attenuated than MTBVAC. Additionally, MTBVAC was the only MTBVAC strain that exhibited higher attenuation than BCG Pasteur.

Lack of PE_PGRS proteins in MTBVAC-L2::*hyg* can contribute to the inferior attenuation observed, because lack of secretion of these proteins has been correlated with increased virulence. Other phenotype that could be involved in this inferior attenuation is the presumable presence of PGL in GC1237, because of an intact *pks1/15* region (24), required for PGL synthesis. Disruption of PGL synthesis in HN878 strain was correlated with loss of its hypervirulent phenotype in the mouse model (13).

Lineage 3 has been less characterized than lineage 2-Beijing strains and lineage 4 strains, and the possible reasons for its inferior attenuated phenotype in comparison with MTBVAC are still unknown.

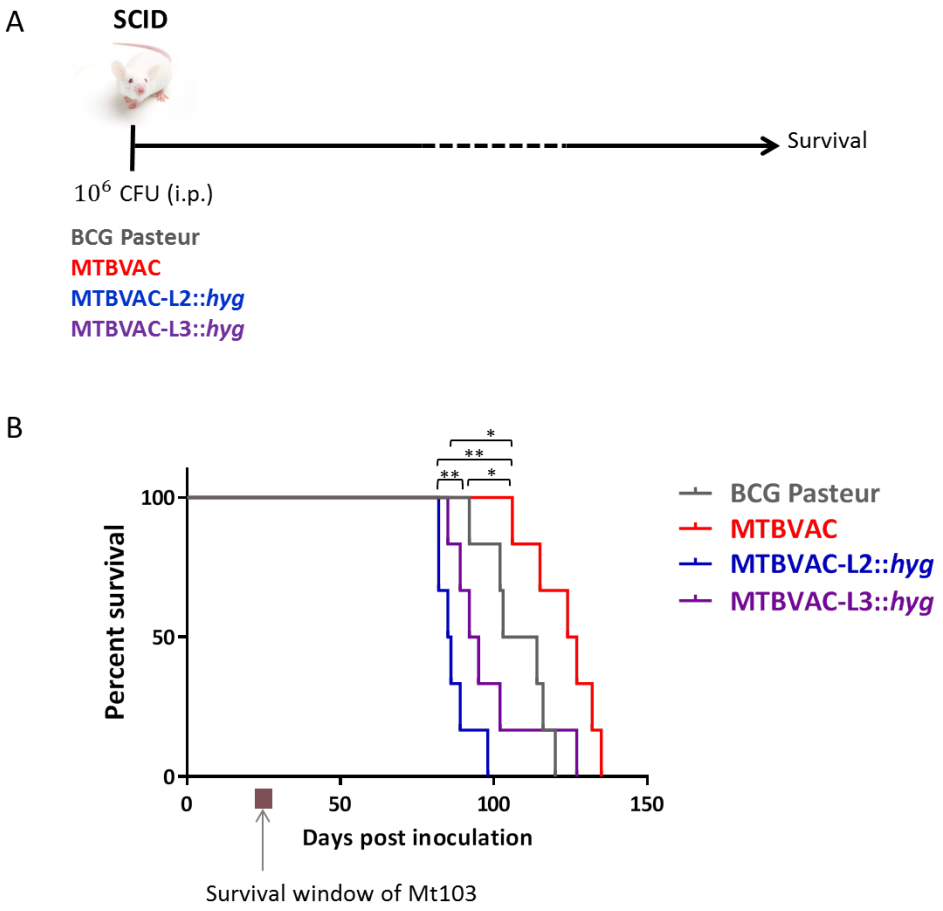


Figure 63. Attenuation in SCID mouse. (A) Scheme of the experimental procedure. (B) Survival curves of SCID mice after inoculation of 106 bacilli by intraperitoneal route. Statistical analysis, Gehan-Breslow-Wilcoxon with a confidence interval or 95%. Rectangle showed survival of SCID mice after intraperitoneal route with 10^6 CFU of Mt103, the parental strain of MTBVAC (23).

Conclusions

After construction and characterization of new live attenuated vaccine candidates MTBVAC-L2 and MTBVAC-L3, we performed preclinical characterization of these new double mutants in the mouse model.

A first experiment of protective efficacy was performed in mice to compare the protection of MTBVAC and MTBVAC-L2::*hyg* vaccine candidates against H37Rv and HN878 (lineage 4 and lineage 2-Beijing strains respectively). Results showed that both MTBVAC and MTBVAC-L2::*hyg* conferred similar protection against H37Rv, but unexpected results were obtained in the challenge of HN878 because the strain seemed to have suffered some changes as a consequence of cultivation in the laboratory that lead the loss of virulence in mice, presumably by lack of methyl-branched derived lipids involved in virulence.

Consequently, we previously confirmed virulence of lineages 2 and 3 strains used as challenge to study the protection of the three MTBVAC strains against lineages 2, 3 and 4 of *M. tuberculosis*. The strains tested, W4 from lineage 2, HCU3524 from lineage 3 and H37Rv from lineage 4 were virulent in the mouse model.

Protective efficacy experiments were performed to evaluate the protection of MTBVAC, MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* against virulent strains representative of each modern lineages (H37Rv, W4 and HCU3524). BCG Pasteur, the current vaccine against TB, was used as comparator and unvaccinated group as control.

Significant and similar protection efficacy was observed in MTBVAC, MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* and BCG Pasteur against the lineage 4 strain, H37Rv. In the challenge against the lineage 2-Beijing strain, W4, all vaccine candidates tested conferred protection. Significant superior protection of MTBVAC and MTBVAC-L2::*hyg* was observed compared to BCG Pasteur against W4 strain. After challenge with the lineage 3 strain, significant and comparable protection was observed in MTBVAC and MTBVAC-L3::*hyg*, followed by MTBVAC-L2::*hyg* and the lowest bacterial decrease was observed in the group of BCG Pasteur vaccinated mice.

In safety studies performed in immunocompromised SCID mice, significant differences were observed in attenuation although all vaccine candidates (MTBVAC, MTBVAC-L2::*hyg*, MTBVAC-L3::*hyg* and BCG Pasteur) were attenuated. MTBVAC was the most attenuated of all strains tested in the experiment and resulted the only one that was more attenuated than BCG Pasteur. MTBVAC-L2::*hyg* was the least attenuated strain, and MTBVAC-L3::*hyg* exhibited similar survival curve than BCG Pasteur, and their attenuated profile was intermediate between MTBVAC-L2::*hyg* and MTBVAC.

Results suggested that *phoP*, *fadD26* deleted strains succeed in protection against TB in the mouse model. MTBVAC, MTBVAC-L2::*hyg* and MTBVAC-L3::*hyg* conferred protection against *M. tuberculosis* strains belonging to modern lineages. Additionally, MTBVAC protective efficacy was tested against lineages 2 and 3 strains for the first time and results showed that MTBVAC succeed in protection against the three modern lineages of *M. tuberculosis.*

References

1. **Fonseca KL, Rodrigues PNS, Olsson IAS, Saraiva M.** 2017. Experimental study of tuberculosis: From animal models to complex cell systems and organoids. *PLoS Pathog* **13:**e1006421.
2. **Gupta UD, Katoch VM.** 2005. Animal models of tuberculosis. *Tuberculosis (Edinb)* **85:**277-293.
3. **Myllymaki H, Niskanen M, Oksanen KE, Ramet M.** 2015. Animal models in tuberculosis research - where is the beef? *Expert Opin Drug Discov* **10:**871-883.
4. **Cooper AM, Magram J, Ferrante J, Orme IM.** 1997. Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with mycobacterium tuberculosis. *J Exp Med* **186:**39-45.
5. **Flynn JL, Chan J, Triebold KJ, Dalton DK, Stewart TA, Bloom BR.** 1993. An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection. *J Exp Med* **178:**2249-2254.
6. **Flynn JL, Goldstein MM, Chan J, Triebold KJ, Pfeffer K, Lowenstein CJ, Schreiber R, Mak TW, Bloom BR.** 1995. Tumor necrosis factor-alpha is required in the protective immune response against Mycobacterium tuberculosis in mice. *Immunity* **2:**561-572.
7. **Singh AK, Gupta UD.** 2018. Animal models of tuberculosis: Lesson learnt. *Indian J Med Res* **147:**456-463.
8. **Medina E, North RJ.** 1998. Resistance ranking of some common inbred mouse strains to Mycobacterium tuberculosis and relationship to major histocompatibility complex haplotype and Nramp1 genotype. *Immunology* **93:**270-274.
9. **Medina E, North RJ.** 1999. Genetically susceptible mice remain proportionally more susceptible to tuberculosis after vaccination. *Immunology* **96:**16-21.
10. **Kamath AB, Woodworth J, Xiong X, Taylor C, Weng Y, Behar SM.** 2004. Cytolytic CD8+ T cells recognizing CFP10 are recruited to the lung after Mycobacterium tuberculosis infection. *J Exp Med* **200:**1479-1489.
11. **Aguilo N, Gonzalo-Asensio J, Alvarez-Arguedas S, Marinova D, Gomez AB, Uranga S, Spallek R, Singh M, Audran R, Spertini F, Martin C.** 2017. Reactogenicity to major tuberculosis antigens absent in BCG is linked to improved protection against Mycobacterium tuberculosis. *Nat Commun* **8:**16085.
12. **Walker KB, Brennan MJ, Ho MM, Eskola J, Thiry G, Sadoff J, Dobbelaer R, Grode L, Liu MA, Fruth U, Lambert PH.** 2010. The second Geneva Consensus: Recommendations for novel live TB vaccines. *Vaccine* **28:**2259-2270.
13. **Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, Kaplan G, Barry CE, 3rd.** 2004. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431:**84-87.
14. **Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, Ernst JD, Gagneux S.** 2010. Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. *Nat Genet* **42:**498-503.
15. **Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, Laval F, Daffe M, Martin C, Brosch R, Guilhot C.** 2014. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. *Proc Natl Acad Sci U S A* **111:**11491-11496.
16. **Broset E, Martin C, Gonzalo-Asensio J.** 2015. Evolutionary landscape of the Mycobacterium tuberculosis complex from the viewpoint of PhoPR: implications for virulence regulation and application to vaccine development. *MBio* **6:**e01289-01215.

17. **Zhang L, Ru HW, Chen FZ, Jin CY, Sun RF, Fan XY, Guo M, Mai JT, Xu WX, Lin QX, Liu J.** 2016. Variable Virulence and Efficacy of BCG Vaccine Strains in Mice and Correlation With Genome Polymorphisms. *Mol Ther* **24:**398-405.
18. **Castillo-Rodal AI, Castanon-Arreola M, Hernandez-Pando R, Calva JJ, Sada-Diaz E, Lopez-Vidal Y.** 2006. Mycobacterium bovis BCG substrains confer different levels of protection against Mycobacterium tuberculosis infection in a BALB/c model of progressive pulmonary tuberculosis. *Infect Immun* **74:**1718-1724.
19. **Gheorghiu M, Lagrange PH.** 1983. Viability, heat stability and immunogenicity of four BCG vaccines prepared from four different BCG strains. *Ann Immunol (Paris)* **134C:**125-147.
20. **Lopez B, Aguilar D, Orozco H, Burger M, Espitia C, Ritacco V, Barrera L, Kremer K, Hernandez-Pando R, Huygen K, van Soolingen D.** 2003. A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. *Clin Exp Immunol* **133:**30-37.
21. **Jeon BY, Derrick SC, Lim J, Kolibab K, Dheenadhayalan V, Yang AL, Kreiswirth B, Morris SL.** 2008. Mycobacterium bovis BCG immunization induces protective immunity against nine different Mycobacterium tuberculosis strains in mice. *Infect Immun* **76:**5173-5180.
22. **Ordway DJ, Shang S, Henao-Tamayo M, Obregon-Henao A, Nold L, Caraway M, Shanley CA, Basaraba RJ, Duncan CG, Orme IM.** 2011. Mycobacterium bovis BCG-mediated protection against W-Beijing strains of Mycobacterium tuberculosis is diminished concomitant with the emergence of regulatory T cells. *Clin Vaccine Immunol* **18:**1527-1535.
23. **Solans L, Uranga S, Aguilo N, Arnal C, Gomez AB, Monzon M, Badiola JJ, Gicquel B, Martin C.** 2014. Hyper-attenuated MTBVAC erp mutant protects against tuberculosis in mice. *Vaccine* **32:**5192-5197.
24. **Alonso H, Aguilo JI, Samper S, Caminero JA, Campos-Herrero MI, Gicquel B, Brosch R, Martin C, Otal I.** 2011. Deciphering the role of IS6110 in a highly transmissible Mycobacterium tuberculosis Beijing strain, GC1237. *Tuberculosis (Edinb)* **91:**117-126.

Chapter 3

Implications of the metabolite c-di-AMP in MTBVAC in the host-pathogen interaction

Abstract

Nucleotides derived metabolites have been described as important second messenger signaling molecules in bacteria, archaea and eukaryotic cells. During the last years, enzymes involved in synthesis and degradation of the cyclic di-nucleotide, cyclic-di-AMP (c-di-AMP), with diadenylyl cyclase and phosphodiesterase activities respectively, have been described in different bacteria and this molecule has been proposed to be involved in diverse processes.

One of the pathways in which this molecule is involved in *M. tuberculosis* is to induce type I IFN response in the infected host cell. Bacterial c-di-AMP can reach the cytosol and triggers IFN-β response in the eukaryotic cell. This response is enhanced when phosphodiesterase is deleted or cyclase is overexpressed, as consequence of increased levels of c-di-AMP in the bacilli. However, beneficial or detrimental effect of IFN-β in bacterial infections is still under debate.

Evaluation of c-di-AMP in *M. tuberculosis phoPR* mutants revealed higher production of this metabolite in the *phoPR* deleted strain compared to WT strain. The same phenotype was described in MTBVAC in comparison with Mt103. These results demonstrated that c-di-AMP production is a new PhoPR-regulated phenotype. Evaluation of the secreted fractions also confirmed the role of *phoPR* in c-di-AMP export.

ESAT-6 secreted protein (secreted through ESX-1) is involved in phagosome rupture and the escape of the bacilli into the cytosol. It has been described that c-di-AMP of *M. tuberculosis*, that can reach the cytosol, triggers host IFN-β response via cytosolic sensor STING.

To decipher the possible involvement of bacterial c-di-AMP in the host immune response, knock-outs in the cyclase (Δ*disA*) and phosphodiesterase (Δ*cnpB*) involved in c-di-AMP metabolism were obtained in MTBVAC. Even though the different levels of c-di-AMP in MTBVAC and knock-outs, none of the strains triggers IFN-β response in THP-1 cells, suggesting that bacterial c-di-AMP would not reach the cytosol. By contrast, MTBVAC and knock-outs exhibited response of the proinflammatory cytokine IL-1β despite the impaired ESAT-6 secretion.

Introduction

The cyclic di-nucleotide, c-di-AMP

Nucleotide derived molecules have been described as signaling metabolites in bacteria, archaea and eukaryotic cells. One of the nucleotide metabolite identified in a wide range of bacteria is cyclic-di-AMP (c-di-AMP) (1). This cyclic dinucleotide is synthesized from two molecules of ATP by a diadenylyl cyclase enzyme and degraded to either phosphoadenylyl adenoseine (pApA) or AMP by a phosphodiesterase enzyme (2-8) (Figure 64).

c-di-AMP has been described as important second messenger to be involved in diverse cellular processes: growth of *Staphylococcus aureus* in low-potassium conditions (9), regulation of fatty acid synthesis in *Mycobacterium smegmatis* by binding to the transcription factor *DarR* (a TeTR-like regulator) (10), control of sporulation checkpoint in *Bacillus subtilis* by sensing DNA integrity (11) or cell wall homeostasis in *S. aureus, B. subtilis* and *L. monocytogenes* suggesting that the phosphodiesterase deletion leads to increase resistance to β-lactam antibiotics and conversely, increased susceptibility to cell-wall antibiotics is observed when the phosphodiesterase is overexpressed (4, 6, 12-14). In *L. monocytogenes*, which utilizes the is listeriolysin O (LLO) to reach the cytosol, c-di-AMP is described to be involved in modulation of host immune response via c-di-AMP secretion to the cytosol that triggers IFN-β response (15).

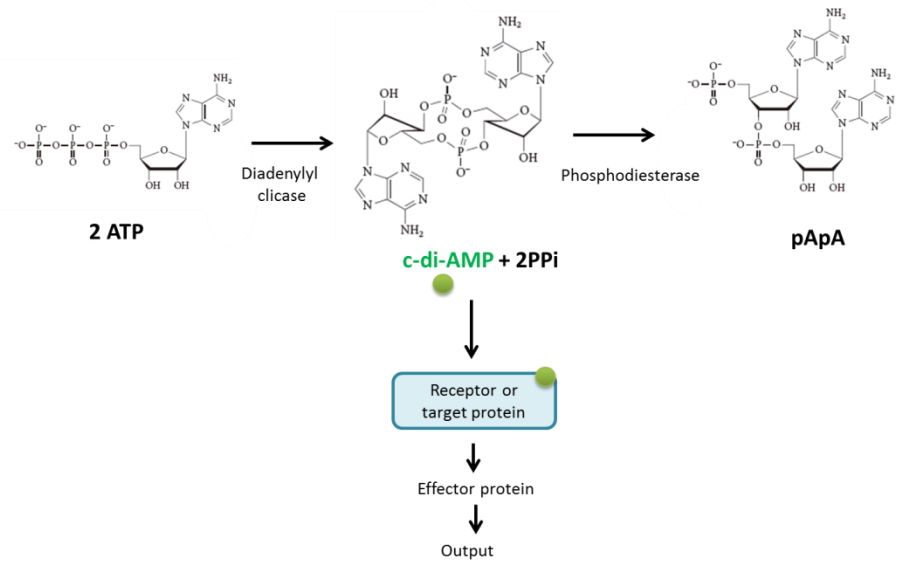


Figure 64. Synthesis and degradation of c-di-AMP. c-di-AMP is synthesized from two molecules of ATP by a diadenylyl cliclase enzyme and is cleavage by a phosphodiesterase enzyme. The metabolite can control diverse cellular pathways. Adapted from (1).

Synthesis and degradation of c-di-AMP in *M. tuberculosis*

Genes involved in synthesis and degradation of c-di-AMP in *M. tuberculosis* have been described during the last decade.

Rv3865 (*disA,* also named *dacA* from diadenylate cyclase), which is an ortholog of *Bacillus subtilis* DisA, codes for a diadenylate cyclase that converts ATP into c-di-AMP. In *B. subtilis,* the cyclase is named DisA for DNA integrity scanning protein A, which is involved in cell-cycle checkpoints. Characterization of DisA from *M. tuberculosis* shows that DisA can catalyze the c-di-AMP synthesis reaction from ATP, although it also catalyzes c-di-AMP synthesis from ADP, but at lower yields. Additional ATPase and ADPase activities were observed in DisA but no specific activity was observed with GTP. RHR motif in DisA is essential for interaction with ATP and therefore, to have the adenylate cyclase activity. DisA constitutes an octamer and divalent cations (Mg^{2+}, Mn^{2+} or Co^{2+}) are required for enzymatic activity and higher activity is observed in basic pH than in acid pH (3, 16).

The gene that codifies for the phosphodiesterase responsible of c-di-AMP degradation, Rv2837c (named *cnpB)*, was identified as ortholog of *S. penumoniae pd2* (2), being a soluble protein in contrast to *gdpP* and *yybT* phosphodiesterases from *S. aureus* and *B. subtilis* respectively (4, 5). CnpB contains a DHH domain and a DHHA1 domain, divalent cations Co^{2+} or Mg^{2+} are required for its activity and the pH 8.5 is the condition with the highest activity. CnpB cleavages c-di-AMP to AMP in two steps: c-di-AMP is first linearized into pApA that will be hydrolyzed to 5' AMP (8, 16).

In vitro studies exhibit that deletion of *disA* abolish the production of c-di-AMP in *M. tuberculosis* and BCG suggesting this is the only cyclase that synthesis c-di-AMP in these bacilli. On the other hand, higher levels of c-di-AMP are observed after deletion of *cnpB* in *M. tuberculosis* and BCG in comparison with WT strains. Surprisingly, in the secreted fraction, c-di-AMP is detected in a *cnpB* knock-out in *M. tuberculosis* in contrast to a BCG Δ*cnpB* mutant (8, 17, 18).

IFN-β and IL-1β host immune response in *M. tuberculosis* infection

Type I IFN (IFN α/β) is important for defense against viruses. However the role of type I IFN in bacterial infections remains elusive: protective effects have been described in some bacterial infections including *Chlamydia trachomatis, Chlamydia pneumoniae, Legionella pneumophila* and *Bacillus anthracis* in contrast to detrimental effects observed in *L. monocytogenes* and *M. tuberculosis* infections (19).

Detrimental effects of type I IFN in *M. tuberculosis* infection have been described in humans and mice. In humans, transcriptomics revealed increased transcriptomic profile of IFNα/β that correlates with reduced successful treatment and with the extent of radiographic disease. In other study devoid of signaling of type I IFN because of a functional mutation in the receptor was correlated with increased resistance to the disease. In mice, lack of IFNα/β led to a decreased bacterial load together with increased survival. Decreased IL-1α and IL-1β production, which are crucial for host defense, inhibited by IFNα/β was observed both in mice

and humans (19, 20). In addition, it has been described that the hypervirulent isolate HN878 showed increased induction of type I IFN and failed to induce Th1 immunity response (21). Despite the strong support of detrimental effects of type I IFN response, it has been also described some protective effects of type I IFN response under specific conditions. Comparable bacterial burden and survival have been observed in presence and absence of type I IFN response, protective effect of type I IFN has been described in absence of IFN-γ and even improved symptoms have been reported after IFN-α administration in combination with antimycobacterial drugs to patients that conventional treatment failed (20, 22). Furthermore, BCG::ESX-1 ,which triggers IFN-β response, exhibited superior protection in comparison with BCG. However, the contribution of this IFN-β production in the protection requires further studies (23).

IL-1 cytokines comprises IL-1α, IL-1β, IL-18 and IL-33 which are proinflammatory cytokines. IL-1β is involved in control of bacteria, demonstrated by diminished survival of IL-1β and IL-1R (IL-1α/β receptor) knockouts mice after *M. tuberculosis* infection. Knockout IL-1β mice exhibit increased susceptibility to *M. tuberculosis* infection and confirming that IL-1β has a non-redundant function that is not compensated with IL-1α (24, 25). Additionally, subcutaneous administration of anti-IL-1β antibodies to *M. tuberculosis* infected mice exhibited loss of body weight and lethality. IL-1 is suggested to help the maintenance of the balance between sufficient phagocytes to control the pathogen and inhibit the recruitment of permissive macrophages by inhibition of type IFN (26).

In addition, polymorphisms in the *IL1* gene cluster have been described to influence the susceptibility of individuals to TB (24). Recent studies associated upregulation of IL-1β with induction of trained immunity, defined as non-specific protection against infections after previous exposure to certain microbial components) (27). It has been demonstrated that BCG vaccination induces epigenetic reprogramming in human monocytes *in vivo* which induced protection against non-related viral infection and IL-1β played a key role for these trained immunity responses (28).

Type I IFN response and IL-1β pathways after *M. tuberculosis* infection

After aerosol infection, *M. tuberculosis* is engulfed by alveolar macrophages. *M. tuberculosis* has developed different strategies to prevent macrophage defenses, such as prevention the acidification of the phagosome which results in optimal activity of lysosomal digestive enzymes and production of oxygen species to control multiplication of the bacilli.

Additionally, *M. tuberculosis* has developed the ability to escape of the phagosome and this event depends on the correct function of ESX-1 secretion system (29, 30). Consequently BCG, with the RD1 deleted region that induces the ESX-1 system, is not able to reach the cytosol.

Type I IFN response is produced by different pathways: bacterial DNA is sensed by cGAS which converts one molecule of ATP and one molecule of GTP into cyclic-di-GAMP (c-di-GAMP) which can activate STING which triggers type I IFN response (IFN-β) (31-33). Additionally, mitochondrial DNA released into the cytosol because of mitochondrial stress

caused by bacterial infection and recognition of peptidoglycan fragments in the cytosol recognized by the cytosolic sensor NOD2 can also induce type I IFN responses (20).

Type I IFN response is suggested to depend on the functional ESX-1 protein secretion system (32, 34), although some controversial results have been described. IFN-β response is not observed after BCG infection in THP-1 cells and BMDMs (18, 23, 32) while IFN-β is observed in the macrophage cell line RAW264.7 from mouse and in BMDMs and bone marrow dendritic cells (BMDCs) after BCG infection (17). Groschel and colleagues observed that complementation of BCG with ESX-1 leads to induce IFN-β response in THP-1 cells (23) whereas Zhang and colleagues did not observe IFN-β in BCG RD1 complemented strain in BMDMs from C57BL/6 (18).

Production of IL-1β is tightly controlled and its synthesis leads to production of reactive oxygen and nitrogen intermediates and activation of other proinflammatory cytokines (25, 35). IL-1β is produced by the canonical pathway of inflammasome activation and the action of caspase-1 in the precursor pro-IL-1β (26) (Figure 65). In addition, mycobacterial DNA in the cytosol can be sensed by AIM-2 and can activate the NLRP3-inflammasone (20) (Figure 65). However, other inflammasome independent pathways can induce IL-1β production (25, 36).

It has been generally supposed that response of IL-1β via inflammasome is mediated by ESX-1 (37, 38). On the other hand, a recent study exhibited that inflammasome activation and IL-1β response is not affected when ESAT-6 secretion is impaired. Blocked ESAT-6 secretion downregulated IFN expression while IL-1β response is still observed (32).

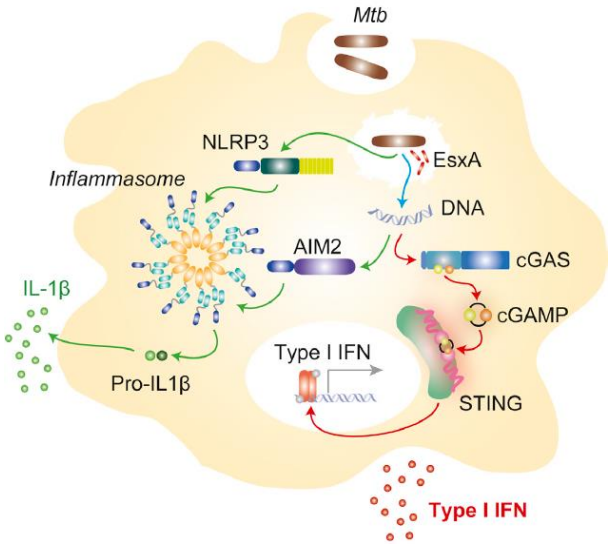


Figure 65. Type I IFN and IL-1β pathways after *M. tuberculosis* infection, from (32). Bacterial DNA can be sensed by cGAS and AIM2 that lead to type I IFN response and IL-1β responses respectively. These two phenotypes are suggested to be dependent on ESX-1 secretion system and ESAT-6 secretion. An independent mechanism of ESAT-6 secretion can produce IL-1β response.

- **Role of c-di-AMP from *M. tuberculosis* and BCG in IFN-β host immune response**

The ability of *M. tuberculosis* to produce c-di-AMP led to study the possibility that bacterial c-di-AMP contributes to enhance IFN-β response in the host cell. For these studies, several mutants in *cnpB* and *disA* were obtained in different strains from *M. tuberculosis*. *cnpB* mutant exhibited

increased levels of c-di-AMP compared to the WT whereas *disA* mutant abolished c-di-AMP production.

In vitro and *ex vivo* infections using these mutants in *M. tuberculosis* revealed that bacterial c-di-AMP modulates host immune response by triggering IFN-β response (Figure 66). This activation was STING-dependent and independent of cGAS. Studies in different cells (J774.1 cells, BMDMs and BMDCs from C57BL/6) showed that Δ*cnpB* or *disA* overexpression strains of *M. tuberculosis* or BCG overexpressing *disA* (higher levels of c-di-AMP compared to WT strains) enhanced IFN-β response, and this phenotype was reverted in complemented strains (8, 17, 18, 39). Additionally, Dey and colleagues observed an increased production of TNF, IL-6 and IL-1α in BMDMs and BMDCs obtained from C57BL/6 mice when the phosphodiesterase was deleted (39) and Dey and colleagues observed an increase of TNF-α when *disA* is overexpressed (17). These phenotypes suggest that increased levels of bacterial c-di-AMP led to enhanced type I IFN and proinflammatory cytokine response. Increased pro-inflammatory response was also observed in BCG overexpressing *disA* (40).

Disagreement results were observed in BCG derivatives regarding IFN-β response. Overexpression of *disA* enhanced IFN-β response in RAW264.7 murine cells, BMDMs and BMDCs from C57BL/6 and this response was significantly decreased in STING-KO cells (17). By contrast, neither BCG Δ*cnpB* nor BCG Δ*cnpB* complemented with RD1 produce IFN-β response in BMDMs from C57BL/6 mice (18).

In vivo studies in BALB/c and C57BL/6 mice exhibited that levels of c-di-AMP are involved in virulence. *cnpB* mutants or *disA* overexpressing strains, showed attenuation by enhance survival in mice, lower bacterial burden in lung and smaller and fewer lesions compared to WT strains. In agreement with these results, increased virulence was observed in *disA* mutants, which were unable to synthesize c-di-AMP (mice survival was diminished) (8, 17, 39).

- **Role of cyclic dinucleotides in vaccines**

Cyclic dinucleotides have been described as mediators of innate host defenses and have been tested as new adjuvants in mice (41).

In TB, administration of cyclic dinucleotides has been described to conferred protection. STING was required for protection whereas type I IFN signaling was not needed for protection (42).

One disadvantage of cyclic dinucleotides as adjuvants vaccines is their rapid elimination. To overcome this problem, a BCG strain overexpressing *disA* was obtained by Dey and colleagues. In protective efficacy studies in guinea pig model, BCG overexpressing *disA* conferred superior protection after 18 weeks post-challenge with H37Rv compared to BCG Pasteur. Additionally, increased proinflammatory cytokine response have been observed after infection of bone-marrow derived macrophages (BMDMs) (40).

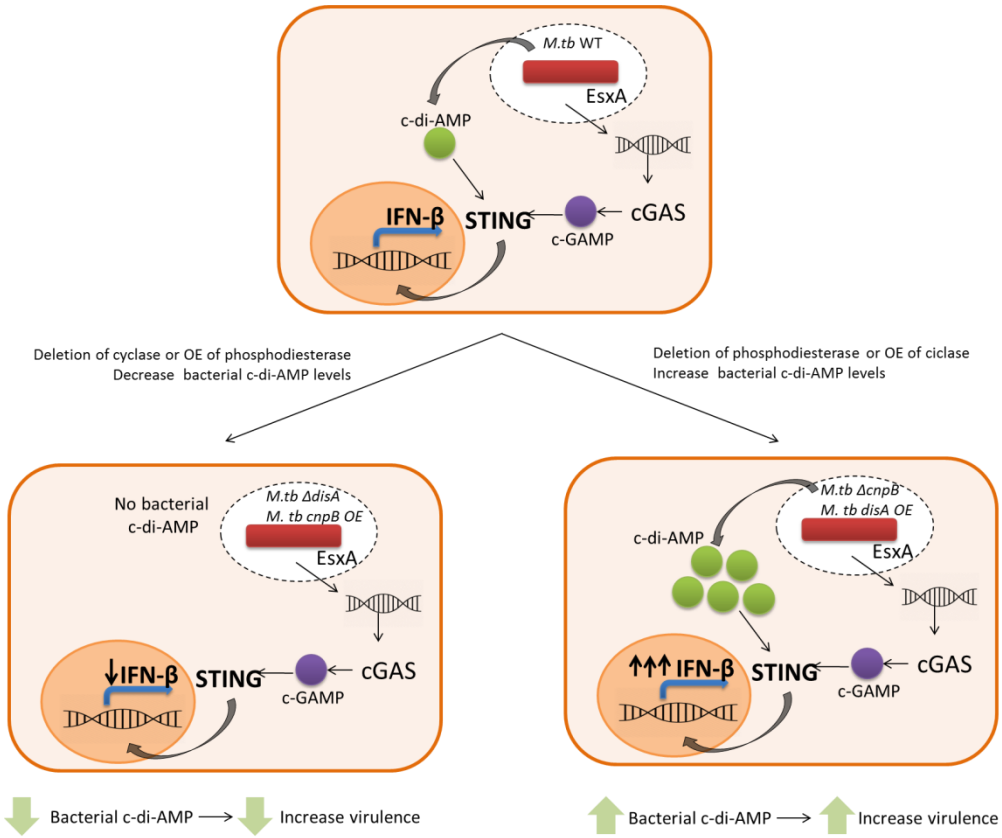


Figure 66. Modulation of IFN-β host immune response by bacterial c-di-AMP after *M. tuberculosis* infection. Deletion of *disA*, that codes for the cyclase to synthesize c-di-AMP or overexpression of *cnpB* gene which codes for the phosphodiesterase that cleavages c-di-AMP leads to no production of the metabolite in the bacilli and lower IFN-β response in comparison with WT strain. By contrast, *cnpB* deletion or *disA* overexpression drive to increased levels of c-di-AMP and consequently increased IFN-β response in the host cell. These results conclude that bacterial c-di-AMP can reach the cytosol and activate IFN-β response via STING.

In this chapter we evaluate the production and secretion of c-di-AMP in H37Rv, ΔphoPR, complemented strain, Mt03 and MTBVAC. We obtain knock-outs in MTBVAC of the cyclase and phosphodiesterase involved in the synthesis of c-di-AMP and evaluate IFN-β and IL-1β response after infection of THP-1 cells with these strains.

Material and methods

Bacterial strains, media and growth conditions

E. coli DH10B strains were grown at 37°C in Luria-Bertani (LB) broth. When required, media were supplemented with ampicillin (Amp, 100 μg/ml), kanamycin (Km, 20 μg/ml), hygromycin (Hyg, 50μg/ml) or chloramphenicol (Cm, 12.5 μg/ml); 15% arabinose.

E.coli DH10B containing a BAC library of H37Rv chromosome cloned in pBeloBAC were grown at 30°C when contained the thermosensitive pKD46 plasmid.

Mycobacteria strains (Table 6) were grown at 37 °C in Middlebrook 7H9 liquid medium containing 0.05% Tween 80 and 10% (vol/vol) ADC (0.2 % dextrose, 0.5% bovine serum albumin, 0.085% NaCl and 0.0003% beef catalase). For solid media Middlebrook 7H10 broth supplemented with ADC was used. Hyg 20 μg/ml, Km 20 μg/ml were added to the media if required. 0.2% acetamide was used to induce the expression of gp60/61 in pJV53H plasmid.

Eukaryotic cells and culture conditions

THP-1 human monocytic cells line was grown in RPMI1640 supplemented with 10% (vol/vol) fetal bovine serum (FBS). Culture conditions were grown at 37 °C, 5% CO_2 and 95% humidity. THP-1 cells were differentiated with 10 ng/ml of PMA for 72 h.

Knock-out construction in MTBVAC

Deletion of *cnpB* and *disA* were obtained by BAC-recombineering strategy (described in chapter 1).

Plasmid extraction (mini-prep), preparation of electrocompetent cells in *M. tuberculosis* for recombineering, polymerase chain reaction (PCR) and plasmids used for gene deletion by BAC-recombineering (pKD46, pKD4 and pJV53H) are described in detail in Materials and Methods section of Chapter 1. For PCR amplification to obtain AES for MTBVAC, 2 U of MyTaq DNA polymerase (BIOLINE) and 0,06 U of Pwo DNA polymerase (Roche) were used in a total volume of 50 μl with MyTaq Buffer 5x.

Primers used for PCR are detailed in Table 7.

Extraction of c-di-AMP

Mycobacterial cultures were grown in 7H9-0.05% Tween-80 supplemented with Dextrose/NaCl to avoid presence of contaminating albumin in the supernatant.

Bacteria cultures were grown until log-phase (OD_{600} ~ 0.6) and pelleted by centrifugation (4000 rpm 5min at 4 °C). Supernatant fraction was filtered through 0.22 μm-pore-size to evaluate c-di-AMP secretion of the strains.

For total c-di-AMP extraction, pellets were resuspended in 300 μL of extraction solution (acetronitrile/methanol/water, 2/2/1, vol/vol/vol, HPLC grade) and transferred to tubes containing glass beads (MP Biomedals). Mycobacterial suspensions were disrupted by Fast Prep (6.5 m/s, 45 s), incubated on ice for 15 min and heated at 95 °C for 10 min. Suspensions were cooled on ice and supernatants were transferred into a new vial after centrifugation at 14000 rpm, at 4 °C for 10 min. Two additional extractions were performed by addition of 200 μL to the tubes containing glass beads. Suspensions were disrupted by Fast Prep (6.5 m/s, 45 s), incubated on ice for 15 min and supernatants were combined into the vial after centrifugation (14000 rpm, at 4 °C for 10 min). The vials with the combination of the supernatants (approximately 700 μL) were stored at -20 ° overnight. Supernatants were centrifuged at 14000 rpm for 10 min at 4 °C and supernatants were filtered through 0.22 μm-pore-size.

Quantification of total c-di-AMP and c-di-GMP of H37Rv, H37Rv Δ*phoPR*, H37Rv Δ*phoPR* complemented, Mt103 and MTBVAC strains were performed in collaboration with Institute of Pharmacology, Hannover Medical School, Hannover (Germany). These measurements were performed using an isotope labeled $^{13}C^{15}N$ of c-di-AMP. Same extractions were also quantified in Proteomics facility of Servicios Científico Técnicos del CIBA (IACS-Universidad de Zaragoza) using c-di-GMP as internal control, because of our previous finding demonstrated lack of production of this metabolite in *M. tuberculosis*. Comparable results were obtained between both institutions.

All secreted fractions were evaluated in collaboration with Fundación Medina, Granada (Spain). Detection limit is 14.7 ng/ml.

Table 6. Mycobacteria strains used in this chapter

Strain	Despcrition	Reference
Mt103	Parental clinical isolate of *M. tuberculosis* lineage 4	(43)
H37Rv	Reference laboratory strain of *M. tuberculosis*	(44)
H37Rv Δ*phoPR*	Mutant of *phoP* and *phoR* genes in H37Rv	(45)
H37Rv Δ*phoPR* complemented	Complemented *phoPR* genes in H37Rv Δ*phoPR* strain	(45)
H37Rv ΔRD1	H37Rv deleted RD1 region	(46)
BCG Pasteur	Current vaccine against TB, attenuated from *M. bovis*	(47)
BCG Danish	Current vaccine against TB, attenuated from *M. bovis*	(47)
MTBVAC::*hyg* (Mt103 Δ*fadD26* Δ*phoP*::*hyg*)	Double mutant marked in *phoP* with res-Ωhyg-res cassette	(48)
MTBVAC	Double unmarked mutant of *phoP* and *fadD26* in Mt103 strain	(49)
MTBVAC Δ*cnpB*::*km*	Marked mutant of *cnpB* gene in MTBVAC	This study
MTBVAC Δ*disA*::*km*	Marked mutant of *disA* gene in MTBVAC	This study

Table 7. Primers used in this chapter for PCR

Primers	Sequence 5´→3´
cnpB-P1-Fw	GGGTGCGGCAAGCGGGTAGAGGTCAGCTTTGCCGCGCCGGGTGTAGGCTGGAGCTGCTTC
cnpB-P2-Rv	GCGTCGTCGATCGAGCCGGTGGTCGTATACCCCGCGGCCACATATGAATATCCTCCTTAGT
KO-cnpB-Fw	GGGCGTTGTTCCGGTCTTCG
KO-cnpB-Rv	CCTAGCCCTAACGCCGCTG
Confirm-KO-cnpB-Fw	GCGGTGATCATCGGTTTCAATGTGCG
Confirm-KO-cnpB-Rv	CAACGCACTCGCCACCGTC
cnpB-I1	GGGCGATCTAACTGATTCCGGGC
cnpB-I2	CGACCTCCTTGAACACCGCCG
disA-P1-Fw	GCACGCTGTGACTCGTCCGACCCTGCGTGAGGCTGTCGCCGTGTAGGCTGGAGCTGCTTC
disA-P2-Rv	CGTCCGACAGGGCGTCCAGTTCGTCCAGGGTGGCATTGATCATATGAATATCCTCCTTAGT
KO-disA-Fw	CTCGCCGCACAGTCCGAC
KO-disA-Rv	GATCGCGGTGGCCATCGTCATC
Confirm-KO-disA-Fw	GCGCGCTCTATGTCTCTGGTGAG
Confirm-KO-disA-Rv	CGTTTGCGTGACCTGGCCCTAC
disA-I1	GATGGTGGCTTCTCCCTCGATGTC
disA-I2	CACCACCGTCATCACATCGCG
pKD46-gam-fw	CCTGTTTTCCTAATCAGCCCGGC
pKD46-bet-rv	AAATGCCGTCTGGCGAAGAGTG
P1-inv	GAAGCAGCTCCAGCCTACAC
P2-inv-long	CTTCGGAATAGGAACTAAGGAGGATATTCATATG
gp60fr	ATCCGGCTCTACGCCGAC
gp61rv	CGGCAAATGACTCTTGCGT

THP-1 infection

THP-1 cells were cultured in RPMI1640 supplemented with 10% of FBS and cells were differentiated with 10 ng/ml of PMA. Infections were performed in 24 well plates and $5*10^5$ cells were distributed per well at multiplicity of infection (MOI) of 5:1. For bacteria preparation, liquid cultures of bacteria were centrifuged at 4000 rpm for 10 min at room temperature. Supernatants were discarded and bacteria were resuspended in 5 ml of DPBS 1x, and $OD_{600\ nm}$ were measured to determine the volume required for infection. Bacteria were added to cells for infection experiment and incubated at 37 °C and 5 % CO_2. Supernatants were collected after 24 and filtered through 0.22 μm-pore-size. IL-1β and IFN-β released were detected by ELISA.

Results

Production and secretion of c-di-AMP is a novel PhoPR-dependent phenotype in *M. tuberculosis*

To study the potential regulation of c-di-AMP by PhoPR, c-di-AMP was detected in *phoPR* mutant in H37Rv *M. tuberculosis* strain. Total c-di-AMP and c-di-GMP were measured in H37Rv, H37Rv Δ*phoPR* and complemented strain (grown until log-phase) using radiolabelled isotopes. Results exhibited that total c-di-AMP were higher in *phoPR* mutant compared to WT and complemented strains (Figure 67A). Complemented strain with *phoPR* partially recovered WT phenotype. Remarkably, no c-di-GMP production was observed in none of the samples.

The above extractions were quantified using c-di-GMP as internal control, because of the previous finding of no production of this metabolite. Comparable results were obtained; total c-di-AMP measured was higher in *phoPR* mutant compared to the WT strains and lower levels of c-di-AMP were observed after complementation in comparison with Δ*phoPR* strain (Figure 67B).

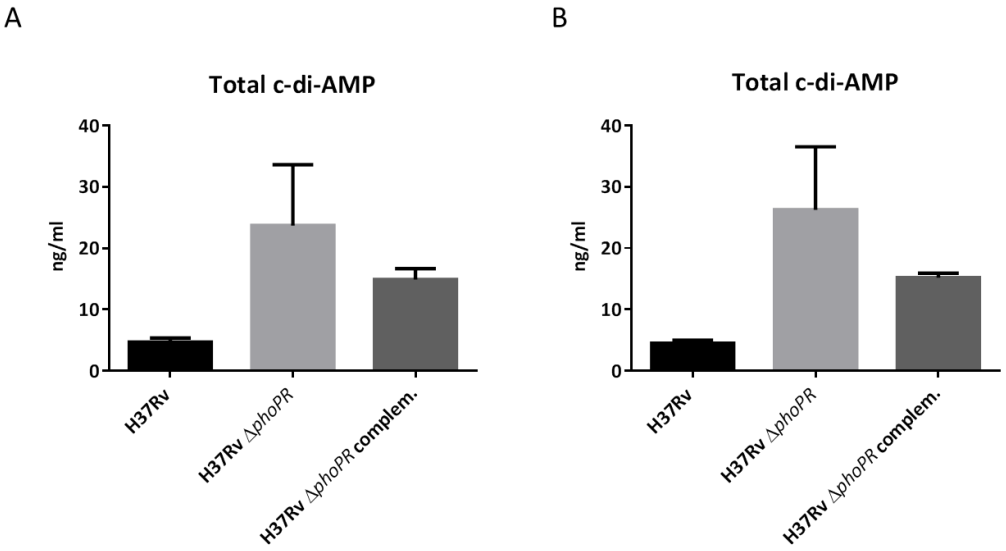


Figure 67. Comparable results of total c-di-AMP in WT strain, Δ*phoPR* mutant and complemented strain in 7H9-ADC-0.05% Tween 80. Quantification obtained (A) using labelled isotopes and (B) using c-di-GMP as internal control. Data are represented mean ± sd.

To further support the *phoPR* dependent phenotype, total c-di-AMP was quantified in Mt103 and MTBVAC (Mt103 Δ*phoP* Δ*fadD26*) using two different methods with comparable results (again no detection of c-di-GMP was obtained) (Figure 68A and B). MTBVAC vaccine candidate produced higher levels than the WT strain observing a 20 fold increase in comparison with Mt103.

Both experiments confirmed a new PhoPR-dependent phenotype. However, the reasons of the increased levels of c-di-AMP in MTBVAC compared to H37Rv Δ*phoPR* are still unknown.

Further studies are required to decipher the precise mechanism by which PhoPR might control c-di-AMP production or can inhibit its degradation.

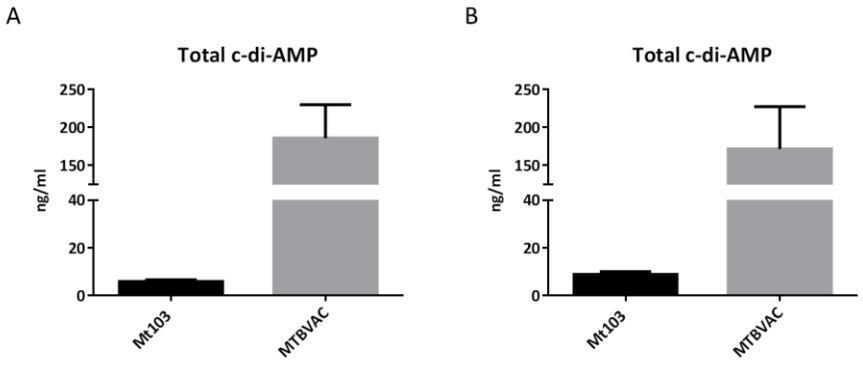


Figure 68. Comparable results of total c-di-AMP in Mt103 and the double mutant *phoP, fadD26* MTBVAC in 7H9-ADC-0.05% Tween 80. Quantification obtained (A) using labelled isotopes and (B) using c-di-GMP as internal control. Data are represented mean ± sd.

After total metabolite detection in the lysate fraction, c-di-AMP was quantified.in the supernatant fraction.

Results exhibited presence of c-di-AMP exclusively in H37Rv Δ*phoPR* in contrast to H37Rv and complemented strain (Figure 69A). Similarly, in Mt103 and MTBVAC secreted fractions, c-di-AMP was only observed in MTBVAC while no c-di-AMP detection was obtained in the parental Mt103 strain (Figure 69B).

In conclusion, c-di-AMP was only detected in the secreted fraction in *phoPR* mutant strains. This phenotype could be as a consequence of the increased levels of c-di-AMP in the total fraction in *phoPR* mutants that led to the bacteria to secrete the metabolite to counter the high levels of the metabolite or regulation of a transporter system by PhoPR.

Mechanisms by which c-di-AMP could be exported are still unknown. It has been reported that Rv3877 and Rv1877 efflux pumps are not involved in the secretion (18).

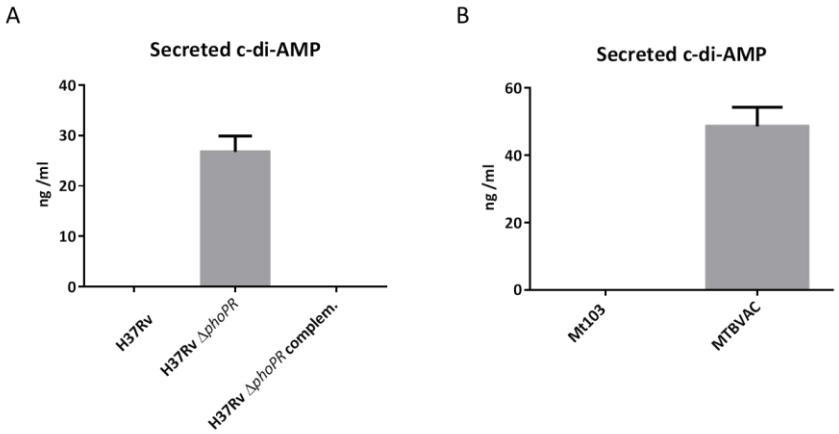


Figure 69. Secreted c-di-AMP. (A) Secretion of c-di-AMP was observed in H37Rv Δ*phoPR* in contrast to H37Rv and complemented strain. (B) Secretion of c-di-AMP was observed in MTBVAC and no detection was observed in Mt103. Data are represented mean ± sd.

BCG, the current vaccine against TB, produced comparable levels to WT *M. tuberculosis* strains and no secretion was detected

Once observed higher levels of c-di-AMP in total and secreted fraction in the MTBVAC vaccine candidate, we aimed at measuring production of this metabolite in the current vaccine against TB, BCG.

Total production of c-di-AMP in BCG Pasteur and BCG Danish were evaluated in total and secreted fraction. Comparable results were obtained in the two strains in the total production of c-di-AMP (9 and 5 ng/ml respectively) (Figure 70) and no secretion of the metabolite was detected. Higher c-di-AMP production levels were expected because of the SNP in *phoR* (codon 71 GGT→ATT) in all animal-adapted mycobacteria and L6 from *M. africanum* (45) responsible of downregulation of *phoPR* operon. Taking into account this evidence, we would expect comparable levels of production and secretion of c-di-AMP in BCG and *M. tuberculosis phoPR* mutants. By contrast, c-di-AMP production in BCG showed similar levels than *M. tuberculosis* WT strains, suggesting that other SNPs or mutations might have an impact on c-di-AMP production or degradation. Further studies are required to decipher other genetic polymorphisms involved in production and cleavage of c-di-AMP.

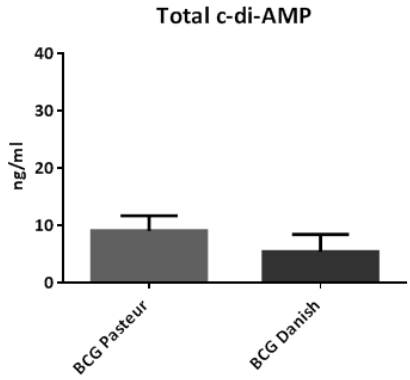


Figure 70. Total c-di-AMP in BCG strains. Similar results in production of c-di-AMP were detected in BCG Pasteur and BCG Danish strains. Data are represented mean ± sd.

Role of c-di-AMP in the vaccine phenotype of MTBVAC

In order to study whether bacterial c-di-AMP from MTBVAC has a role in modulating IFN-β and IL-1β response, deletion of the cyclase and the phosphodiesterase involved in c-di-AMP synthesis and degradation (*disA* and *cnpB* genes respectively) were obtained in MTBVAC.

For evaluation of c-di-AMP in immune host response, THP-1 cells were infected with MTBVAC and *cnpB* and *disA* knock-outs and IFN-β and IL-1β responses were evaluated.

- **Construction of *disA* (Rv3586) and *cnpB* (Rv2837c) knock-outs in MTBVAC**

A bacterial artificial chromosome from H37Rv (BAC) library was constructed in pBeloBAC11 vector and transformed in *E. coli* DH10B. DH10B harboring BAC-Rv222 and BAC-Rv404 containing *disA* and *cnpB* respectively were transformed with pKD46 thermosensitive plasmid

and plated onto LB-Amp and incubated at 30 °C overnight (Figure 71A). pKD46 contains γ, β and *exo* genes from λ Red system and their expression is controlled by the presence of arabinose.

DH10B BAC Rv222 and DH10B BAC Rv404 both containing pKD46 plasmid and induced with arabinose were transformed with PCRs obtained with disA-P1-Fw/disA-P2-rv and cnpB-P1-Fw/cnpB-P2-rv primers using pKD4 as template (Figure 71B). PCR products contained FRT-*km*-FRT cassette flanked with 40 bp of identity arms for site specific recombination in *cnpB* and *disA* genes respectively (Figure 71C). Transformation were plated onto LB-Km and incubated at 37 °C overnight. Recombinants were confirmed using two pairs of primers, one 5´upstream of the recombination site and the other inside the Km cassette and the other pair of primers one of them inside the Km cassette and the other 3´downstream recombination site (Figure 72A-D).

DH10B BAC Rv222 Δ*disA*::*Km* and DH10B BAC Rv404 Δ*cnpB*::*Km* were grown overnight and DH10B BAC Rv222 Δ*disA*::*Km* and BAC Rv404 Δ*cnpB*::*Km* were purified. PCR products containing 1 kb of identity arms were obtained using DH10B BAC Rv222 Δ*disA*::*Km* and BAC Rv404 Δ*cnpB*::*Km* as templates. These PCR products will be used as allelic exchange substrate (AES) to obtain the knock-outs in MTBVAC (Figure 73A and B).

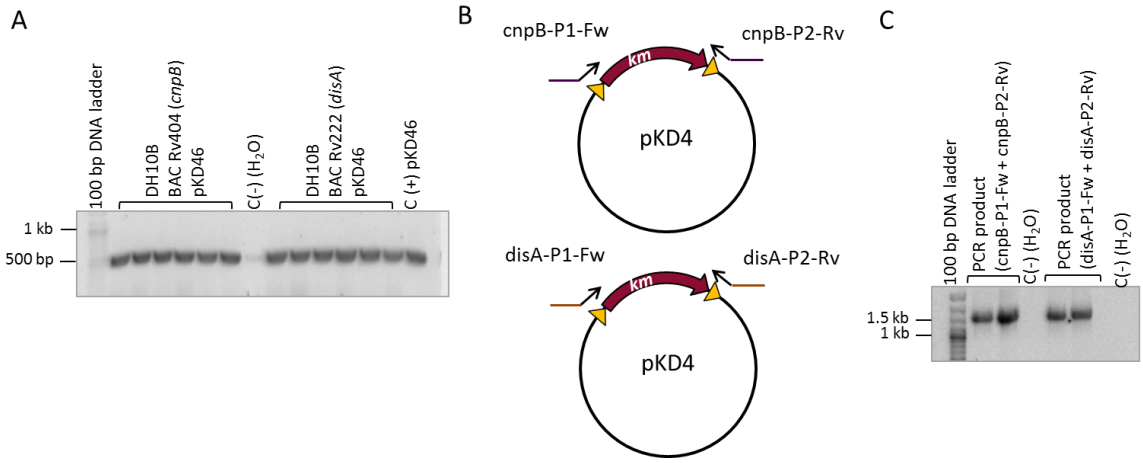


Figure 71. Confirmation of pKD46 plasmid and PCR product for transformation in *E. coli*. (A) Transformation of pKD46 plasmid in DH10B Rv404 and DH10B Rv222 was confirmed using pKD46-gam-fw and pKD46-bet-rv primers. All colonies tested were positive. pKD46 was used as control positive. (B) Schematic representation of primers used for PCR amplification of FRT-*km*-FRT flanked with 40 bp identity arms (added to the primers) of *cnpB* and *disA* using pKD4 as template. (C) PCR product obtained using cnpB-P1-Fw/cnpB-P2-Rv and disA-P1-Fw/disA-P2-Rv.

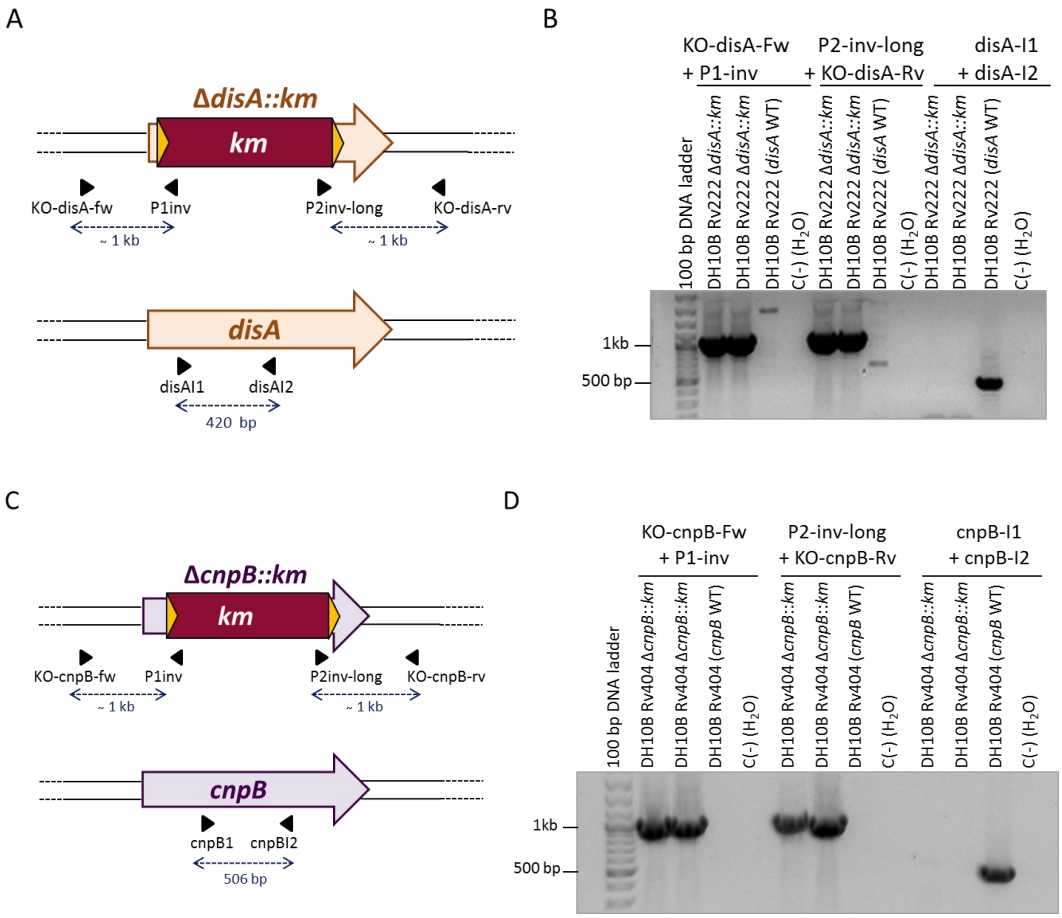


Figure 72. Confirmation of deletion in *disA* and *cnpB* in BAC Rv404 and BAC Rv222 respectively in DH10B. (A and C) Genetic structure of disrupted *disA* and *cnpB* genes by FRT-*km*-FRT cassette. (B and D) PCR amplification using different combination of primers represented in panels A or C to confirm deletion of *disA* and *cnpB* gene in BACRv222 and BACRv404. WT BAC Rv222 or BAC Rv404 were used as control.

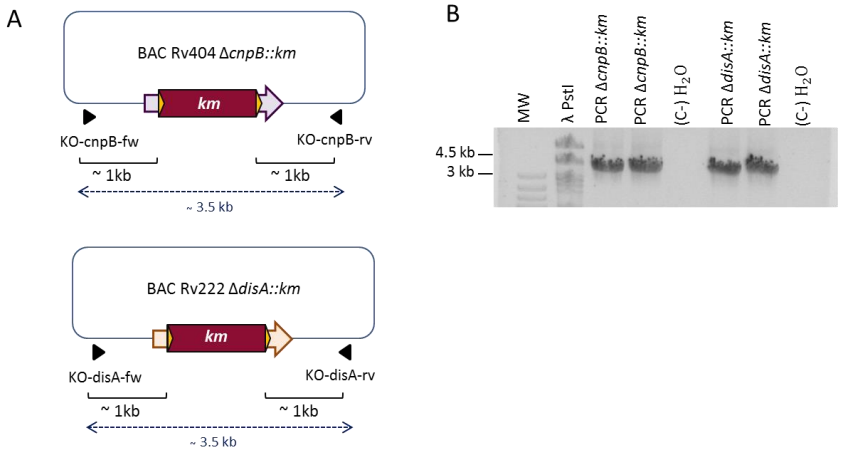


Figure 73. AES for recombination in mycobacteria. (A) Schematic representation to obtain PCR product of FRT-*km*-FRT cassette flanked with 1 kb of identity arms for site specific recombination to obtain knock-outs in *cnpB* and *disA* genes (named PCR product Δ*cnpB::km* and PCR product Δ*disA::km* respectively) using BAC Rv404 Δ*cnpB::km* and BAC Rv222 Δ*disA::km* respectively. (B) PCR products obtained after amplification with primers detailed in A.

MTBVAC was transformed with pJV53H plasmid, which contains Che9c gene products 60 and 61 that allow recombination in mycobacteria. PCR products (AES) were transformed in MTBVAC harboring pJV53H plasmid, induced with acetamide, and plated onto 7H10-Km. Recombinants were confirmed using different combinations of primers that hybridize 5´upstream or 3´downstream of recombination site, inside the FRT-*km*-FRT cassette or in the deleted region of *disA* and *cnpB* genes (Figure 74A-D).

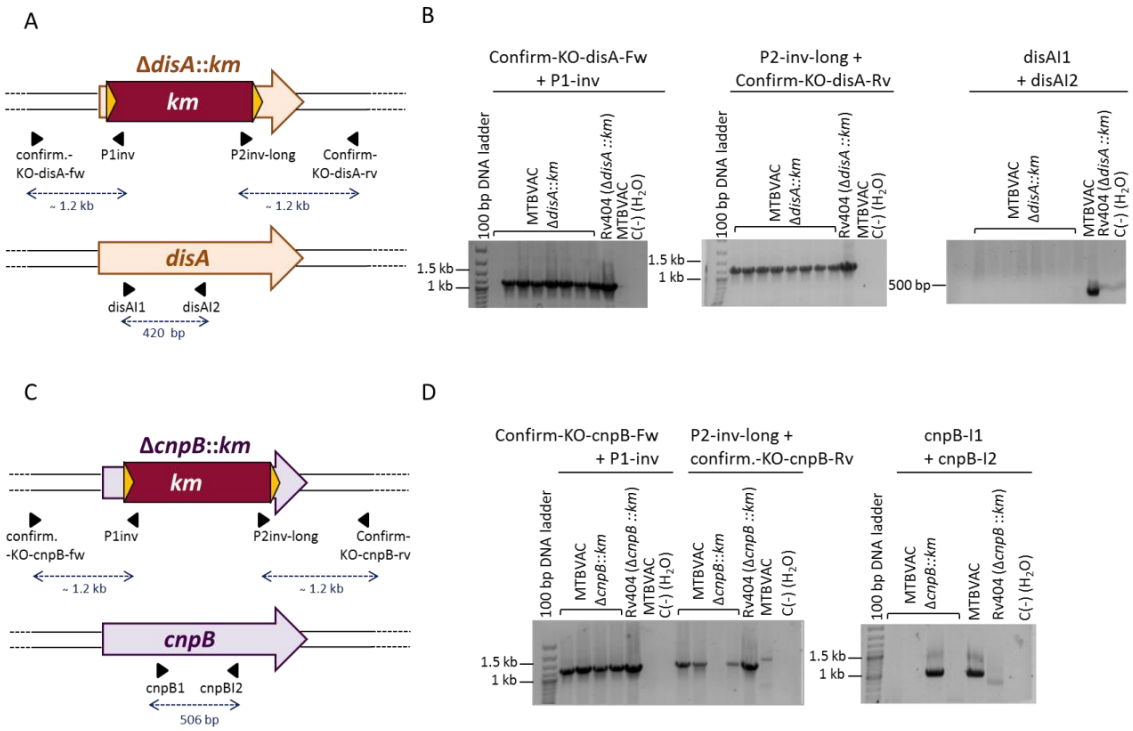


Figure 74. Confirmation of deletion of *disA* and *cnpB* genes in MTBVAC. (A and C) Genetic structures of *disA* and *cnpB* disrupted with FRT-*km*-FRT cassette or WT genes. (B and D) PCR amplification using the primers represented in panels A and C. Amplification should be observed using confirm-Fw/P1-inv and confirm-Rv/P2-inv-long primers and absence of amplification with internal primers which hybridize in the deleted region (disAI1/disAI2 and cnpBI/cnpBI2 primers). MTBVAC was used as control.

A schematic representation of the mutant constructions in MTBVAC is detailed in Figure 75.

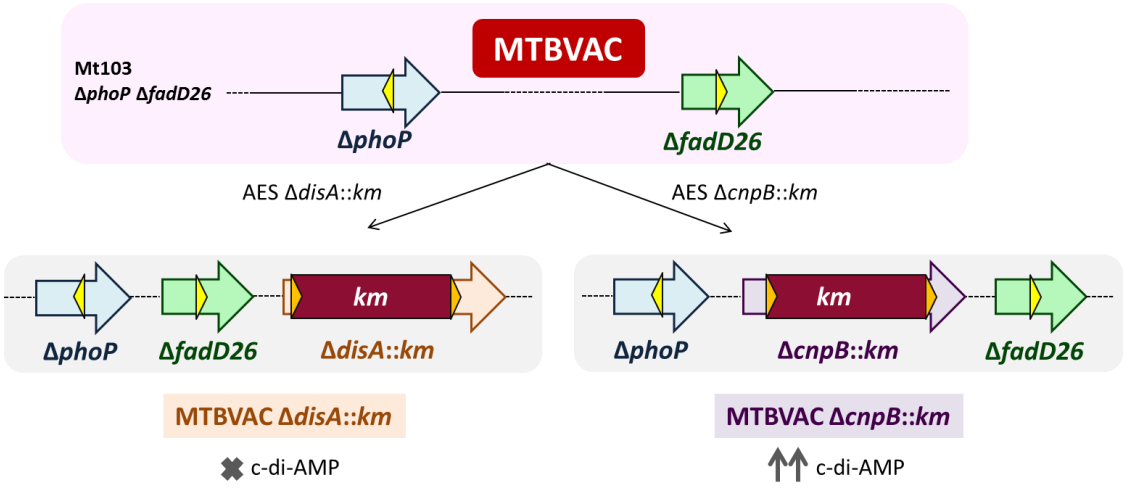


Figure 75. Scheme of genomic structure of mutants obtained in MTBVAC. Km marked mutants were obtained in *disA* and *cnpB* genes involved in c-di-AMP metabolism

- **Production and secretion of c-di-AMP in MTBVAC and *disA* and *cnpB* knock-outs**

MTBVAC, MTBVAC Δ*disA::km* and MTBVAC Δ*cnpB::km* were grown until log-phase and evaluation of total and secreted c-di-AMP were obtained. When *disA* gene was deleted in MTBVAC, which codes for the cyclase involved in c-di-AMP synthesis, no detection of c-di-AMP was observed in total and secreted fraction. By contrast, when the phosphodiesterase *cnpB* was deleted in MTBVAC a 70-fold increase of total c-di-AMP and almost 20-fold increase in secreted fraction were observed (Figure 76A and B). Therefore, three different strains in the MTBVAC genetic background (MTBVAC, MTBVAC Δ*cnpB::km* and MTBVAC Δ*disA::km*) were obtained. This strains displayed different levels of total production and secretion c-di-AMP representing useful genetic tools to study its role in the host immune response in the vaccine candidate MTBVAC.

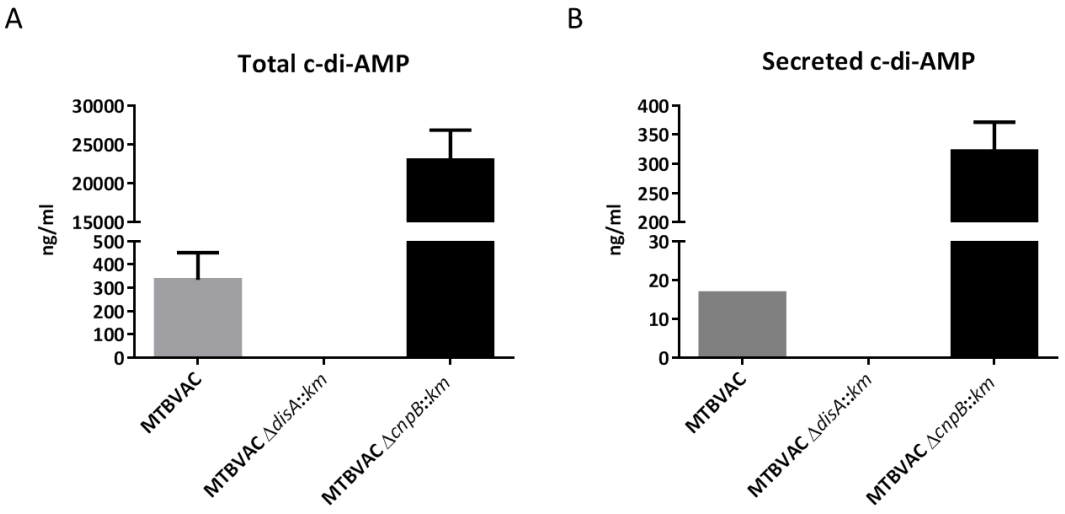


Figure 76. Total and secreted c-di-AMP of MTBVAC and *disA* and *cnpB* mutants. Deletion of *disA* gene abolished the production of c-di-AMP and therefore no secretion of c-di-AMP was observed in these mutants, while *cnpB* deletion led to an increase of c-di-AMP in total and secreted fraction compared to MTBVAC. Data are represented mean ± sd.

- **Evaluation of host-immune response in THP-1 infection with MTBVAC, *disA* and *cnpB* mutants**

Previous observations of increased c-di-AMP levels in MTBVAC compared to Mt103 led to consider whether this bacterial c-di-AMP from MTBVAC could reach the cytosol and triggers IFN-β expression in the host cell. To solve this question THP-1 cells were infected with MTBVAC and derivatives and IFN-β production were measured by ELISA. For the experimental infection, we also used BCG Pasteur, H37Rv ΔRD1 and *M. tuberculosis* strains Mt103 and H37Rv.

Results showed no production of IFN-β in MTBVAC, *disA* and *cnpB* mutants. This observation suggested that despite the high levels of c-di-AMP in MTBVAC, and especially in MTBVAC Δ*cnpB::km*, c-di-AMP did not trigger type I IFN response. We hypothesize that MTBVAC resides in the phagosome, and as consequence, bacterial c-di-AMP cannot reach the cytosol of the cell (because of inability to secrete ESAT-6) and activate INF type response via STING. IFN-β was detected in the virulent strains of *M. tuberculosis* H37Rv and Mt103, used as positive control for the infection. No detection in BCG Pasteur and H37ΔRD1 were observed (Figure 77A).

In other studies, IFN-β response was observed after infection of BCG and BCG overexpressing *disA* in RAW264.7 murine cells, BMDMs and BMDCs from C57BL/6 (17) whereas no IFN-β response was detected after BCG, BCG Δ*cnpB* or BCG::RD1 infection in BMDMs from C57BL/6 (18).

These controversial results in BCG open the question whether bacterial c-di-AMP could reach the host cytosol even in an ESX-1 defective strain. Our results suggested that a functional ESX-1 system were required for induction IFN-β response, in agreement with several studies (18, 23, 32). However, Dey and colleagues observed IFN-β response after BCG infection, although different cell lines were used (17).

IL-1β response was also measured by ELISA (Figure 77B). Results showed no significant variability among MTBVAC, *cnpB* and *disA* mutants suggesting no relation of c-di-AMP in triggering IL-1β response. This response was comparable to Mt103 and H37Rv IL-1β levels.

IL-1β response was observed in MTBVAC in agreement with Wasserman and colleagues (32), observing a robust IL-1β response even though ESAT-6 secretion was blocked. Diminished IL-1β response was observed in BCG Pasteur and H37Rv ΔRD1, suggesting that RD1 is partially involved in IL-1β response. These data are in agreement with several studies (29, 32, 37, 38). No significant differences were observed in MTBVAC, MTBVAC Δ*cnpB*::km and MTBVAC Δ*disA*::km. However, Dey and colleagues observed increased IL-1β response in BCG overexpressing *disA* compared to BCG WT in mouse bone marrow-derived macrophages(40).

In summary, MTBVAC and derivatives induced proinflammatory cytokine IL-1β response. Conversely, no IFN-β response was observed after infection with MTBVAC and derived strains. Despite increased bacterial c-di-AMP in MTBVAC Δ*cnpB::km* no IFN-β response was observed. We hypothesize that even though MTBVAC produces higher levels of c-di-AMP in comparison to Mt103, this c-di-AMP was not involved in host immune signaling.

Our results suggested that a functional ESX-1 is required for IFN-β response. For IL-1β, results suggested that RD1 trigerred less IL-1β response compared to WT strains. In MTBVAC, IL-1β response was observed despite impaired ESAT-6 secretion.

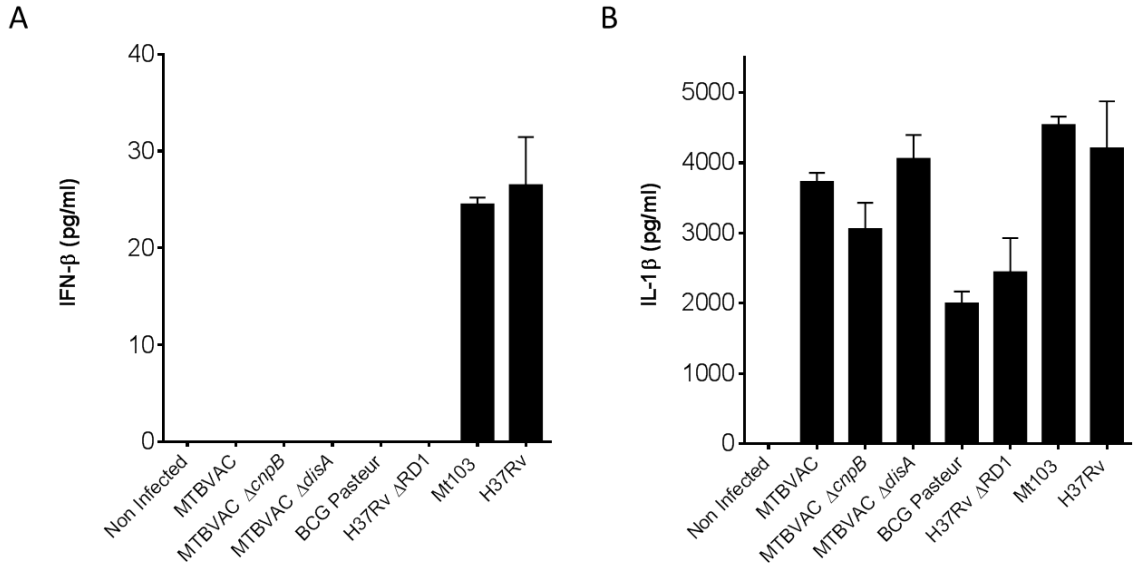


Figure 77. IFN-β and IL-1β response after 24 h of infection in THP-1 cells. (A) No IFN-β response was observed in MTBVAC and derivatives confirmed that bacterial c-di-AMP of MTBVAC could not trigger IFN-β response. (B) No significant differences were observed in IL-1β in MTBVAC and derivatives.

Further studies are required to define the influence of c-di-AMP in the immune responses conferred by MTBVAC. In a recent study, BCG overexpressing *disA* in guinea pig model conferred higher protection in comparison with BCG (40). In this context, future experiments of attenuation and protective efficacy of MTBVAC, MTBVAC Δ*cnpB* and MTBVAC Δ*disA* should be performed.

Conclusions

Several PhoP-dependent phenotypes have been extensively characterized in *M. tuberculosis:* control of SL, DAT and PAT synthesis, post-transcripcional regulation of *tatC mRNA* by ncRNA *mcr7*, ESAT-6 secretion and secretion of Ag85.

In this chapter, a new PhoPR-dependent mechanism has been described. Production of c-di-AMP metabolite, which acts as a second messenger, was increased in H37Rv Δ*phoPR* mutant compared to parental strain and partially restored after the complemetation. Presence of the metabolite in the supernatant was only observed in *phoPR* mutant. Therefore, levels of c-di-AMP were higher in MTBVAC compared to Mt103 in the total fraction and the metabolite was also detected in supernatants MTBVAC but no in Mt103.

To further investigate the role of this metabolite in MTBVAC, in the context of vaccines, knock-outs in *cnpB* and *disA* genes, involved in degradation and synthesis of bacterial c-di-AMP were obtained. Accordingly, three strains with different levels of c-di-AMP production and secretion in MTBVAC genetic background were obtained. Production of c-di-AMP was abolished in *disA* knock-out and increased levels of c-di-AMP were observed in MTBVAC Δ*cnpB* knock-out in comparison to MTBVAC.

Neither MTBVAC nor MTBVAC Δ*cnpB::km* triggered IFN-β response in THP-1 cells, suggesting that even though increased c-di-AMP was observed in MTBVAC compared to Mt103, bacterial c-di-AMP from MTBVAC cannot activate the IFN-β response. These results indicate the requirement of a functional ESX-1 for IFN-β response. The biological significance in the context of protection needs to be studied.

By contrast, a robust IL-1β was observed after THP-1 infection with MTBVAC, *cnpB* and *disA* mutants. Despite ESAT-6 impaired secretion, MTBVAC and derivatives exhibited comparable levels of IL-1β response with *M. tuberculosis* WT strains. However, BCG Pasteur and H37Rv ΔRD1 exhibited lower levels of IL-1β, suggesting that RD1 is involved in this process. The implication of IL-1β in protection needs to be studied.

References

1. **Corrigan RM, Grundling A.** 2013. Cyclic di-AMP: another second messenger enters the fray. *Nat Rev Microbiol* **11:**513-524.
2. **Bai Y, Yang J, Eisele LE, Underwood AJ, Koestler BJ, Waters CM, Metzger DW, Bai G.** 2013. Two DHH subfamily 1 proteins in Streptococcus pneumoniae possess cyclic di-AMP phosphodiesterase activity and affect bacterial growth and virulence. *J Bacteriol* **195:**5123-5132.
3. **Bai Y, Yang J, Zhou X, Ding X, Eisele LE, Bai G.** 2012. Mycobacterium tuberculosis Rv3586 (DacA) is a diadenylate cyclase that converts ATP or ADP into c-di-AMP. *PLoS One* **7:**e35206.
4. **Corrigan RM, Abbott JC, Burhenne H, Kaever V, Grundling A.** 2011. c-di-AMP is a new second messenger in Staphylococcus aureus with a role in controlling cell size and envelope stress. *PLoS Pathog* **7:**e1002217.
5. **Rao F, See RY, Zhang D, Toh DC, Ji Q, Liang ZX.** 2010. YybT is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. *J Biol Chem* **285:**473-482.
6. **Witte CE, Whiteley AT, Burke TP, Sauer JD, Portnoy DA, Woodward JJ.** 2013. Cyclic di-AMP is critical for Listeria monocytogenes growth, cell wall homeostasis, and establishment of infection. *MBio* **4:**e00282-00213.
7. **Witte G, Hartung S, Buttner K, Hopfner KP.** 2008. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol Cell* **30:**167-178.
8. **Yang J, Bai Y, Zhang Y, Gabrielle VD, Jin L, Bai G.** 2014. Deletion of the cyclic di-AMP phosphodiesterase gene (cnpB) in Mycobacterium tuberculosis leads to reduced virulence in a mouse model of infection. *Mol Microbiol* **93:**65-79.
9. **Corrigan RM, Campeotto I, Jeganathan T, Roelofs KG, Lee VT, Grundling A.** 2013. Systematic identification of conserved bacterial c-di-AMP receptor proteins. *Proc Natl Acad Sci U S A* **110:**9084-9089.
10. **Zhang L, Li W, He ZG.** 2013. DarR, a TetR-like transcriptional factor, is a cyclic di-AMP-responsive repressor in Mycobacterium smegmatis. *J Biol Chem* **288:**3085-3096.
11. **Bejerano-Sagie M, Oppenheimer-Shaanan Y, Berlatzky I, Rouvinski A, Meyerovich M, Ben-Yehuda S.** 2006. A checkpoint protein that scans the chromosome for damage at the start of sporulation in Bacillus subtilis. *Cell* **125:**679-690.
12. **Banerjee R, Gretes M, Harlem C, Basuino L, Chambers HF.** 2010. A mecA-negative strain of methicillin-resistant Staphylococcus aureus with high-level beta-lactam resistance contains mutations in three genes. *Antimicrob Agents Chemother* **54:**4900-4902.
13. **Griffiths JM, O'Neill AJ.** 2012. Loss of function of the gdpP protein leads to joint beta-lactam/glycopeptide tolerance in Staphylococcus aureus. *Antimicrob Agents Chemother* **56:**579-581.
14. **Luo Y, Helmann JD.** 2012. Analysis of the role of Bacillus subtilis sigma(M) in beta-lactam resistance reveals an essential role for c-di-AMP in peptidoglycan homeostasis. *Mol Microbiol* **83:**623-639.
15. **Woodward JJ, Iavarone AT, Portnoy DA.** 2010. c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response. *Science* **328:**1703-1705.
16. **Manikandan K, Sabareesh V, Singh N, Saigal K, Mechold U, Sinha KM.** 2014. Two-step synthesis and hydrolysis of cyclic di-AMP in Mycobacterium tuberculosis. *PLoS One* **9:**e86096.

17. **Dey B, Dey RJ, Cheung LS, Pokkali S, Guo H, Lee JH, Bishai WR.** 2015. A bacterial cyclic dinucleotide activates the cytosolic surveillance pathway and mediates innate resistance to tuberculosis. *Nat Med* **21:**401-406.
18. **Zhang Y, Yang J, Bai G.** 2018. Cyclic di-AMP-mediated interaction between Mycobacterium tuberculosis DeltacnpB and macrophages implicates a novel strategy for improving BCG vaccination. *Pathog Dis* **76**.
19. **McNab F, Mayer-Barber K, Sher A, Wack A, O'Garra A.** 2015. Type I interferons in infectious disease. *Nat Rev Immunol* **15:**87-103.
20. **Moreira-Teixeira L, Mayer-Barber K, Sher A, O'Garra A.** 2018. Type I interferons in tuberculosis: Foe and occasionally friend. *J Exp Med* **215:**1273-1285.
21. **Manca C, Tsenova L, Bergtold A, Freeman S, Tovey M, Musser JM, Barry CE, 3rd, Freedman VH, Kaplan G.** 2001. Virulence of a Mycobacterium tuberculosis clinical isolate in mice is determined by failure to induce Th1 type immunity and is associated with induction of IFN-alpha /beta. *Proc Natl Acad Sci U S A* **98:**5752-5757.
22. **Donovan ML, Schultz TE, Duke TJ, Blumenthal A.** 2017. Type I Interferons in the Pathogenesis of Tuberculosis: Molecular Drivers and Immunological Consequences. *Front Immunol* **8:**1633.
23. **Groschel MI, Sayes F, Shin SJ, Frigui W, Pawlik A, Orgeur M, Canetti R, Honore N, Simeone R, van der Werf TS, Bitter W, Cho SN, Majlessi L, Brosch R.** 2017. Recombinant BCG Expressing ESX-1 of Mycobacterium marinum Combines Low Virulence with Cytosolic Immune Signaling and Improved TB Protection. *Cell Rep* **18:**2752-2765.
24. **Carvalho NB, de Lourdes Bastos M, Souza AS, Netto EM, Arruda S, Santos SB, Carvalho EM.** 2018. Impaired TNF, IL-1beta, and IL-17 production and increased susceptibility to Mycobacterium tuberculosis infection in HTLV-1 infected individuals. *Tuberculosis (Edinb)* **108:**35-40.
25. **Cooper AM, Mayer-Barber KD, Sher A.** 2011. Role of innate cytokines in mycobacterial infection. *Mucosal Immunol* **4:**252-260.
26. **Domingo-Gonzalez R, Prince O, Cooper A, Khader SA.** 2016. Cytokines and Chemokines in Mycobacterium tuberculosis Infection. *Microbiol Spectr* **4**.
27. **Kleinnijenhuis J, Quintin J, Preijers F, Joosten LA, Ifrim DC, Saeed S, Jacobs C, van Loenhout J, de Jong D, Stunnenberg HG, Xavier RJ, van der Meer JW, van Crevel R, Netea MG.** 2012. Bacille Calmette-Guerin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. *Proc Natl Acad Sci U S A* **109:**17537-17542.
28. **Arts RJW, Moorlag S, Novakovic B, Li Y, Wang SY, Oosting M, Kumar V, Xavier RJ, Wijmenga C, Joosten LAB, Reusken C, Benn CS, Aaby P, Koopmans MP, Stunnenberg HG, van Crevel R, Netea MG.** 2018. BCG Vaccination Protects against Experimental Viral Infection in Humans through the Induction of Cytokines Associated with Trained Immunity. *Cell Host Microbe* **23:**89-100 e105.
29. **Houben D, Demangel C, van Ingen J, Perez J, Baldeon L, Abdallah AM, Caleechurn L, Bottai D, van Zon M, de Punder K, van der Laan T, Kant A, Bossers-de Vries R, Willemsen P, Bitter W, van Soolingen D, Brosch R, van der Wel N, Peters PJ.** 2012. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol* **14:**1287-1298.
30. **van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, Brenner M, Peters PJ.** 2007. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129:**1287-1298.
31. **Collins AC, Cai H, Li T, Franco LH, Li XD, Nair VR, Scharn CR, Stamm CE, Levine B, Chen ZJ, Shiloh MU.** 2015. Cyclic GMP-AMP Synthase Is an Innate Immune DNA Sensor for Mycobacterium tuberculosis. *Cell Host Microbe* **17:**820-828.

32. **Wassermann R, Gulen MF, Sala C, Perin SG, Lou Y, Rybniker J, Schmid-Burgk JL, Schmidt T, Hornung V, Cole ST, Ablasser A.** 2015. Mycobacterium tuberculosis Differentially Activates cGAS- and Inflammasome-Dependent Intracellular Immune Responses through ESX-1. *Cell Host Microbe* **17:**799-810.
33. **Watson RO, Bell SL, MacDuff DA, Kimmey JM, Diner EJ, Olivas J, Vance RE, Stallings CL, Virgin HW, Cox JS.** 2015. The Cytosolic Sensor cGAS Detects Mycobacterium tuberculosis DNA to Induce Type I Interferons and Activate Autophagy. *Cell Host Microbe* **17:**811-819.
34. **Stanley SA, Johndrow JE, Manzanillo P, Cox JS.** 2007. The Type I IFN response to infection with Mycobacterium tuberculosis requires ESX-1-mediated secretion and contributes to pathogenesis. *J Immunol* **178:**3143-3152.
35. **Master SS, Rampini SK, Davis AS, Keller C, Ehlers S, Springer B, Timmins GS, Sander P, Deretic V.** 2008. Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3:**224-232.
36. **Mayer-Barber KD, Barber DL, Shenderov K, White SD, Wilson MS, Cheever A, Kugler D, Hieny S, Caspar P, Nunez G, Schlueter D, Flavell RA, Sutterwala FS, Sher A.** 2010. Caspase-1 independent IL-1beta production is critical for host resistance to mycobacterium tuberculosis and does not require TLR signaling in vivo. *J Immunol* **184:**3326-3330.
37. **Koo IC, Wang C, Raghavan S, Morisaki JH, Cox JS, Brown EJ.** 2008. ESX-1-dependent cytolysis in lysosome secretion and inflammasome activation during mycobacterial infection. *Cell Microbiol* **10:**1866-1878.
38. **Kurenuma T, Kawamura I, Hara H, Uchiyama R, Daim S, Dewamitta SR, Sakai S, Tsuchiya K, Nomura T, Mitsuyama M.** 2009. The RD1 locus in the Mycobacterium tuberculosis genome contributes to activation of caspase-1 via induction of potassium ion efflux in infected macrophages. *Infect Immun* **77:**3992-4001.
39. **Dey RJ, Dey B, Zheng Y, Cheung LS, Zhou J, Sayre D, Kumar P, Guo H, Lamichhane G, Sintim HO, Bishai WR.** 2017. Inhibition of innate immune cytosolic surveillance by an M. tuberculosis phosphodiesterase. *Nat Chem Biol* **13:**210-217.
40. **Dey RJ, Dey B, Singh AK, Praharaj M, Bishai W.** 2019. BCG overexpressing an endogenous STING agonist provides enhanced protection against pulmonary tuberculosis. *J Infect Dis* doi:10.1093/infdis/jiz116.
41. **Dubensky TW, Jr., Kanne DB, Leong ML.** 2013. Rationale, progress and development of vaccines utilizing STING-activating cyclic dinucleotide adjuvants. *Ther Adv Vaccines* **1:**131-143.
42. **Van Dis E, Sogi KM, Rae CS, Sivick KE, Surh NH, Leong ML, Kanne DB, Metchette K, Leong JJ, Bruml JR, Chen V, Heydari K, Cadieux N, Evans T, McWhirter SM, Dubensky TW, Jr., Portnoy DA, Stanley SA.** 2018. STING-Activating Adjuvants Elicit a Th17 Immune Response and Protect against Mycobacterium tuberculosis Infection. *Cell Rep* **23:**1435-1447.
43. **Jackson M, Raynaud C, Laneelle MA, Guilhot C, Laurent-Winter C, Ensergueix D, Gicquel B, Daffe M.** 1999. Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the Mycobacterium tuberculosis cell envelope. *Mol Microbiol* **31:**1573-1587.
44. **Cole ST, Barrell BG.** 1998. Analysis of the genome of Mycobacterium tuberculosis H37Rv. *Novartis Found Symp* **217:**160-172; discussion 172-167.
45. **Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, Laval F, Daffe M, Martin C, Brosch R, Guilhot C.** 2014. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. *Proc Natl Acad Sci U S A* **111:**11491-11496.

46. **Pym AS, Brodin P, Brosch R, Huerre M, Cole ST.** 2002. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol Microbiol* **46:**709-717.
47. **Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, Inwald JK, Golby P, Garcia JN, Hewinson RG, Behr MA, Quail MA, Churcher C, Barrell BG, Parkhill J, Cole ST.** 2007. Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* **104:**5596-5601.
48. Ainhoa Arbués, PhD Thesis, University of Zaragoza
49. **Arbues A, Aguilo JI, Gonzalo-Asensio J, Marinova D, Uranga S, Puentes E, Fernandez C, Parra A, Cardona PJ, Vilaplana C, Ausina V, Williams A, Clark S, Malaga W, Guilhot C, Gicquel B, Martin C.** 2013. Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated M. tuberculosis-based vaccine to enter clinical trials. *Vaccine* **31:**4867-4873.

General conclusions

- MTBVAC-L2 and MTBVAC-L3 have been constructed in two clinical isolated from lineages 2 and 3 *M. tuberculosis*. These new live attenuated vaccine candidates are based on *phoP* and *fadD26* deletions as MTBVAC.

- PhoP-dependent phenotypes previously described have been characterized in MTBVAC-L2 and MTBVAC-L3 (inability to secrete ESAT-6 and downregulation of PhoP-regulated genes). Moreover, it has been demonstrated that MTBVAC and MTBVAC-L3 but not MTBVAC-L2 secrete PE_PGRS proteins.

- MTBVAC, MTBVAC-L2 and MTBVAC-L3 are attenuated in SCID mice. MTBVAC is the only vaccine candidate more attenuated than BCG Pasteur.

- MTBVAC, MTBVAC-L2 and MTBVAC-L3 confer protection in C3H mice against *M. tuberculosis* strains from modern lineages.

- Production and secretion of c-di-AMP is increased in *M. tuberculosis phoPR* mutants. MTBVAC and mutants in c-di-AMP cyclase and phosphodiesterase fail to trigger IFN-β response in THP-1 cells. In contrast, all these strains exhibit a robust IL-1β response in macrophages.

Conclusiones generales

- MTBVAC-L2 y MTBVAC-L3 han sido construidas en dos aislados clínicos de los lineages 2 y 3 de *M. tuberculosis*. Estas nuevas candidatas a vacunas están basadas en las deleciones de *phoP* y *fadD26* como MTBVAC.

- Fenotipos dependientes de PhoP previamente descritos fueron caracterizados en MTBVAC-L2 y MTBVAC-L3 (no secreción de ESAT-6 y disminución de la expresión de genes regulados por PhoP). Además, se demostró que MTBVAC y MTBVAC-L3, pero no MTBVAC-L2, secretan las proteínas PE_PGRS.

- MTBVAC, MTBVAC-L2 y MTBVAC-L3 son cepas atenuadas en ratones SCID. MTBVAC es la una vacuna candidata más atenuada que BCG Pasteur.

- MTBVAC, MTBVAC-L2 y MTBVAC-L3 confieren protección en ratones C3H frente a cepas de *M. tuberculosis* de los linajes modernos.

- La producción y secreción de c-di-AMP está aumentada en mutantes *phoPR* de *M. tuberculosis.* MTBVAC y los mutantes en la ciclasa y fosfodiesterasa de c-di-AMP no desencadenan respuesta de IFN-β en células THP-1. Por el contrario, todas estas cepas exhiben una respuesta robusta de IL-1β en macrófagos.

APPENDICES

During the Thesis I have collaborated in two papers published on April 2018 and May 2019.

Appendix 1

New insights into the transposition mechanisms of IS*6110* and its dynamic distribution between *Mycobacterium tuberculosis* Complex lineages.

PLOS GENETICS

RESEARCH ARTICLE

New insights into the transposition mechanisms of IS*6110* and its dynamic distribution between *Mycobacterium tuberculosis* Complex lineages

Jesús Gonzalo-Asensio[1,2,3☯]*, Irene Pérez[1,2☯], Nacho Aguiló[1,2☯], Santiago Uranga[1,2], Ana Picó[1,2], Carlos Lampreave[1,2], Alberto Cebollada[1,2], Isabel Otal[1,2], Sofía Samper[1,2,4], Carlos Martín[1,2,5]*

1 Grupo de Genética de Micobacterias, Departamento de Microbiología y Medicina Preventiva. Facultad de Medicina, Universidad de Zaragoza, IIS Aragón, Zaragoza, Spain, **2** CIBER Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain, **3** Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Zaragoza, Spain, **4** Unidad de Investigación Translacional, Hospital Universitario Miguel Servet, Instituto de Investigación Sanitaria Aragón. Zaragoza, Spain, **5** Servicio de Microbiología, Hospital Universitario Miguel Servet, Zaragoza, Spain

☯ These authors contributed equally to this work.
* jagonzal@unizar.es (JGA); carlos@unizar.es (CM)

Citation: Gonzalo-Asensio J, Pérez I, Aguiló N, Uranga S, Picó A, Lampreave C, et al. (2018) New insights into the transposition mechanisms of IS*6110* and its dynamic distribution between *Mycobacterium tuberculosis* Complex lineages. PLoS Genet 14(4): e1007282. https://doi.org/10.1371/journal.pgen.1007282

Editor: Carmen Buchrieser, Institut Pasteur, CNRS UMR 3525, FRANCE

Received: January 4, 2018

Accepted: February 28, 2018

Published: April 12, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the European Commission Horizon 2020 (TBVAC2020, H2020-PHC-643381), Instituto de Salud Carlos III (FIS 15/0317), the Spanish Ministry of Science and Competitiveness (BIO2014-52580P) and Gobierno de Aragón/Fondo Social Europeo. IP was recipient of a "DGA-Fondo Social Europeo" grant. The

Abstract

The insertion Sequence IS*6110*, only present in the pathogens of the *Mycobacterium tuberculosis* Complex (MTBC), has been the gold-standard epidemiological marker for TB for more than 25 years, but biological implications of IS*6110* transposition during MTBC adaptation to humans remain elusive. By studying 2,236 clinical isolates typed by IS*6110*-RFLP and covering the MTBC, we remarked a lineage-specific content of IS*6110* being higher in modern globally distributed strains. Once observed the IS*6110* distribution in the MTBC, we selected representative isolates and found a correlation between the normalized expression of IS*6110* and its abundance in MTBC chromosomes. We also studied the molecular regulation of IS*6110* transposition and we found a synergistic action of two post-transcriptional mechanisms: a -1 ribosomal frameshift and a RNA pseudoknot which interferes translation. The construction of a transcriptionally active transposase resulted in 20-fold increase of the transposition frequency. Finally, we examined transposition in *M. bovis* and *M. tuberculosis* during laboratory starvation and in a mouse infection model of TB. Our results shown a higher transposition in *M. tuberculosis*, that preferably happens during TB infection in mice and after one year of laboratory culture, suggesting that IS*6110* transposition is dynamically adapted to the host and to adverse growth conditions.

Author summary

Since the pioneering discovery of transposition by Barbara McClintock in eukaryotes and later in prokaryotes by Robert W. Hedges and Alan E. Jacob, it has become clear the key

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

role of mobile genetics elements in chromosome remodelling, microbial evolution and host adaptation. The insertion sequence IS*6110* is widely recognized for its utility in TB diagnosis and epidemiology because it is only present in the *M. tuberculosis* Complex (MTBC) and its transposition provides an excellent chromosomal polymorphic variability allowing the study of recent TB transmission. This inherent feature of IS*6110* leads us to hypothesize that IS*6110* plays a crucial role during the TB infectious cycle. However, the biological significance of IS*6110* has been hindered by its almost exclusive use as an epidemiological marker. Here, we study the regulatory mechanisms and the distribution of IS*6110* in the different MTBC lineages. We discuss the potential biological implications of IS*6110*, that is much more than an excellent TB epidemiological tool. Since IS*6110* could play an important role in the adaptation of MTBC to the host, this study opens new avenues to decipher the biological roles of IS*6110* in TB pathogenesis.

Introduction

Tuberculosis (TB) is the largest infectious cause of death in history having claimed more deaths than smallpox, malaria, plague, influenza and AIDS together [1]. In addition to the alarming 1.7 million deaths and 10,4 million of new TB cases in 2016, the emergence of multidrug resistant strains is an increasing threat which makes TB treatment difficult or occasionally impossible [2]. Thus, early diagnostics and identification of transmission chains greatly contribute to control the TB epidemic.

The adaptation of *M. tuberculosis* to the host is extremely complex. Most of the infected individuals are chronically infected in the form of latent TB infection (LTBI) and only one of 10 will develop clinical TB disease. The essential, yet unanswered question, on the natural history of TB is when *M. tuberculosis* decides to establish either LTBI in the host, resembling the lysogenic cycle of lambda phage, or to cause pulmonary TB disease, like the lytic cycle of lambda phage. In this latter case, *M. tuberculosis* decide to kill the host with the aim of achieving transmission to new hosts [3].

Seminal studies by Barbara McClintock deciphered the key role of mobile genetic elements in chromosome remodelling of maize in 1950 [4]. In the late 60's insertion sequences were described by the groups of Shapiro, Malamy, Sybalsky and Starlinger and in 1974 Robert W. Hedges and Alan E. Jacob coined the term "transposition" in bacteria [5]. The insertion sequence IS*6110* is a mobile genetic element exclusively found in the *M. tuberculosis* Complex (MTBC) [6], the causative agent of TB in humans and other mammals including farm animals responsible for zoonotic TB transmission. This feature makes IS*6110* a valuable tool in the diagnosis of MTBC in biological samples [7, 8]. In addition, IS*6110* is present in multiple copies in the chromosome of *M. tuberculosis* and IS*6110* restriction fragment length polymorphism (RFLP) analysis of strains isolated from patients who developed TB showed identical patterns over years [9]. On the other side a high degree of polymorphism was observed between strains of the MTBC isolated from different patients due to IS*6110* transposition [10]. Standardized IS*6110* RFLP typing has been the gold standard for more than 25 years, being the most reliable TB epidemiological marker [11]. IS*6110* typing allows the detection of TB outbreaks as well as to identify transmission chains using conventional and molecular methods [12]. To date tens of thousands of MTBC stains all around the world have been typed by this method but the biological role, if any, of IS*6110* remains elusive. In the last 5–10 years IS*6110* typing is being replaced by less time-consuming methods based in PCR amplification of

mycobacterial interspersed repetitive units (MIRU) [13, 14], or more recently by whole genome sequencing (WGS) [15, 16].

The MTBC comprises eight defined phylogenetic lineages. *M. tuberculosis sensu-stricto* includes lineages L1–L4 and L7. These human-adapted lineages are responsible for the vast majority of global human TB cases, whereas *M. africanum* lineages (L5, L6) are mainly restricted to humans from West Africa. The L8 comprises animal-adapted strains with ecotypes adapted to different mammals, such as *M. caprae* and *M. bovis*, which branched from the *M. africanum* lineage [17]. All these lineages are classified into sub-lineage / clonal complexes or families on the basis of different spoligotyping profiles [18] or on specific genomic signatures [19, 20]. The more distantly related *M. canettii* is outside the clonal population of the MTBC and it is considered the most ancestral progenitor from which the above mentioned MTBC members emerged [21].

According to the IS*6110* content, MTBC members are classified into high (>6) and low (<7) IS*6110* copy number strains [22]. It is not clear whether differences in IS*6110* content account for biological phenotypic consequences in bacterial physiology and pathogenesis. However, it is well known that the *M. tuberculosis* Beijing/W lineage (L2), with a remarkably high content of IS*6110* [23], is associated with higher virulence and massive spread of drug resistant strains, being possibly better adapted to high density populations [19]. Beijing/W lineage was originally described in the 1990's as a predominant genotype found in countries of East Asia designated Beijing-family [24] and after observing an interstate spread from New York of the multidrug-resistant *M. tuberculosis* clone family named W [25].

During its transposition, the IS*6110* promotes a number of important genetic modifications in MTBC strains. This confers plasticity to the MTBC genomes and could have significant biological implications. As for other IS, insertion of IS*6110* into a coding region frequently renders the gene inactive, the basis of transposon mutagenesis, or the recombination between two IS*6110* copies can lead to either inversion or deletion of the chromosomal fragment between them [26–28]. Furthermore, it has been demonstrated that IS*6110* acts as a mobile promoter and this phenotype is selectively activated during *in vitro* infection of monocytes/macrophages [29, 30]. This latter finding has extraordinary consequences in the host-pathogen evolution of the MTBC, as will be discussed below.

It has been suggested that a moderate number of IS*6110* might translate into strain-specific phenotypes that provide selective advantages during the course of the infection [26]. Conversely, it has been demonstrated that excessive accumulation of IS*6110* copies could result in inactivation or deletion of essential genetic regions, being deleterious to the bacterium [31]. This later finding implies that transposition rates of IS*6110* should be finely regulated and maintained at relatively low levels (7.9×10^{-5} events per site per generation) [32]. Considering the clonal evolution of the MTBC, the rate of point mutations is estimated at 10^{-9} events per site per generation and comparatively the mutation rate of IS*6110* is orders of magnitude higher. This reinforces the notion that IS*6110* transposition is under positive selection when infecting or causing disease to the host [32] and accordingly it constitutes an excellent TB epidemiological marker.

At the genetic level the IS*6110* belongs to the IS*3* family and it is annotated as two open reading frames: ORF1 (327 bp) and ORF2 (987 bp) which overlap in 52 bp and are flanked by 28 bp imperfect Inverted Repeats (IR). The 3–4 bp boundaries of IS*6110* are duplicated upon transposition [10]. Despite the massive use of WGS, the repetitive nature of IS*6110* makes difficult to finely map their localizations in the MTBC chromosomes. Although some studies have attempted to localize IS*6110* in *M. tuberculosis* genomes [27, 33–35], little is known about its involvement in other MTBC members including *M. canettii*, *M. africanum* and ecotypes responsible for animal TB (i.e. *M. bovis* and *M. caprae*) which possess a zoonotic risk.

Similar to other members of the IS*3* family, it is thought that transposition of IS*6110* occurs when its two constituent ORFs are translationally fused producing an active transposase [36]. Former studies using *M. smegmatis* (a non-pathogen fast growing mycobacteria) as surrogate host to demonstrate that IS*6110* transposition occurs more readily when this element is located in transcriptionally active locations and also upon exposure to a microaerobic environment [37, 38]. However, there is a definite lack of evidence about the precise mechanisms leading to the production of an active IS*6110* transposase and the physiological conditions that promote transposition in the MTBC. In the present study, we analyse biological data from more than two-thousand clinical isolates covering the MTBC to dissect the molecular mechanism of IS*6110* transposition and its dynamic distribution between the different MTBC lineages. We discuss its biological significance in the tubercle bacillus and also in the clinical presentations of TB.

Results

IS*6110* copy number in the MTBC is lineage-specific

Different members from the MTBC have evolved by accumulation of genomic deletions and specific polymorphisms [39, 40]. Accordingly, the MTBC phylogeny is the result of a genomic decay after an evolutionary bottleneck which led to speciation [41]. Upon examination of fully sequenced and assembled MTBC genomes, we observe that the *M. bovis* AF2122/97 reference strain contains a single IS*6110* while *M. africanum* and *M. tuberculosis* have higher copy numbers of this element (an average of 6 and 17 respectively) (**Fig 1A**). When interrogating *M. canettii*, considered as the most ancestral linage known from which all MTBC members emerged, we only found potentially functional IS*6110* sequences in subgroups STB-A, -D, and–L. Those subgroups that show greater phylogenetic distances (STB–J and–K) have no traces of IS*6110* [21]. Only STB-L carries identical IS*6110* sequences to the MTBC (**S1 Fig**). Supporting this finding, another study demonstrated the presence of IS*6110* in evolutionarily closer *M. canettii* isolates [42].

The IS*6110* content in MTBC genomes suggested to us that the copy number of this transposon might be lineage-specific. However, since a limited number of genomes have been fully sequenced and assembled, we decided to investigate this hypothesis in a representative collection of TB causing strains. We systematically genotyped clinical isolates from TB patients during the last 25 years and subdivided them in families according to spoligotyping profiles. A total of 2,236 clinical samples from our data base covering the MTBC were analysed by standardised Restriction Fragment Length Polymorphism (RFLP) of IS*6110*. Results confirmed that the average IS*6110* content is lineage-specific, ranging from low copy number (*M. bovis*, L1, the L4 sub-lineage X and *M. africanum* L5 and L6) to high copy number in modern *M. tuberculosis* lineages (LAM, CAS and Beijing from L4, L3 and L2 respectively) (**Fig 1B**). Among the *M. tuberculosis* human-adapted species, these high copy number families are globally distributed and accordingly they could be considered as generalists capable of infecting and causing disease in many different human populations [43]. Of these, the LAM family is distributed in America, Africa, Europe, Oceania and East Asia [20], the CAS family affects the Indian continent and East Africa [44] and the Beijing family is amply distributed in East Asia and East Europe [19].

Normalised IS*6110* expression data indicates an exponential transposition dynamics in the MTBC

Once we established that MTBC phylogenetic clades have different IS*6110* content, we interrogated the molecular mechanisms underlying this observation. First, we selected representative

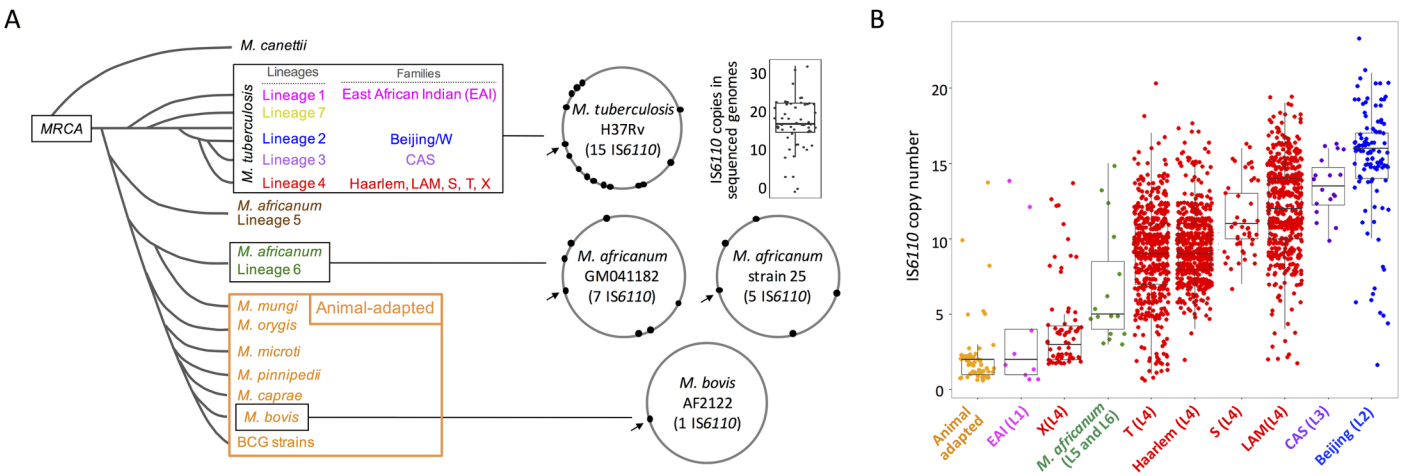


Fig 1. IS*6110* in the MTBC. (**a**) Schematic phylogenetic relationships of MTBC members arisen from a most recent common ancestor (MRCA) after an evolutionary bottleneck. For *M. tuberculosis* different lineages and families are indicated. The position of IS*6110* sequences in fully assembled genomes in indicated by black dots. The arrow indicates the position of IS*6110* in the Direct Repeat region of the CRISPR-Cas locus, which is common to most MTBC strains. For the remaining 17 *M. tuberculosis* strains different from H37Rv, the number of IS*6110* sequences is indicated by a box plot (median = 17). (**b**) Box plots showing the IS*6110* copies in MTBC families. For each family, the lineage according to panel (a) is provided in parenthesis in the X-axis. For clarity, L4 have been subdivided into 5 different families according to spoligotyping.

https://doi.org/10.1371/journal.pgen.1007282.g001

isolates of the MTBC (BCG, *M. bovis*, *M. caprae*, *M. africanum* and *M. tuberculosis*). Then we analysed their global IS*6110* mRNA levels and found that animal-adapted and *M. africanum* species have lower levels of IS*6110* mRNA than *M. tuberculosis* L2 and L4 (Fig 2A). We also found that ORF1 and ORF2 were similarly expressed (S2 Fig), with is compatible with the presence of a single RNA molecule with two out-of-phase reading frames translated into a single ORF by way of a translational frameshift. This result resembles other IS3 family members [36, 45]. Altogether these results indicated a proportional relationship between the copy number content and IS*6110* mRNA expression, which led us to quantitate the "normalised mRNA expression" by calculating expression ratios relative to the IS*6110* content in every MTBC strain (IS*6110* mRNA / IS*6110* copy number). First, the IS*6110* copy number in the above-mentioned strains was checked by RFLP (Fig 2B) and our previous results were reanalysed considering this IS*6110* content. Our results demonstrated that expression per IS*6110* copy is lower in animal adapted strains and in *M. africanum* than in *M. tuberculosis* (Fig 2C).

IS*6110* distribution in representative MTBC members indicates a proportional relationship between the copy number content and the normalised expression of this element and this relation follows an exponential trend ($r^2 = 0.80$) (Fig 2D). To gain further insight into the transposition dynamics of IS*6110*, we analysed *M. bovis* and *M. tuberculosis* isolates showing an uncommon copy number of this element. We selected *M. tuberculosis* clinical isolates from X and T families of lineage 4 containing 1, 2, 3 and 11 IS*6110* copies. Beijing strains from lineage 2 known to possess the highest copy number of IS*6110* were also included [34]. We also selected *M. bovis* strains causative of human TB with 3, 4 and 5 copies of *IS6110* [46], which represent an unusually high copy number for the animal adapted lineage (S3 Fig). Results demonstrated that *M. tuberculosis* strains having a single IS*6110* expresses this mRNA similarly to *M. bovis* BCG. Accumulation of additional copies of IS*6110* resulted in higher mRNA expression of the transposase to reach expression levels comparable to *M. tuberculosis* H37Rv (S3 Fig). On the other hand, accumulation of more than one IS*6110* in *M. bovis* resulted in exacerbated expression of its coding gene. This expression was 5-fold higher than that

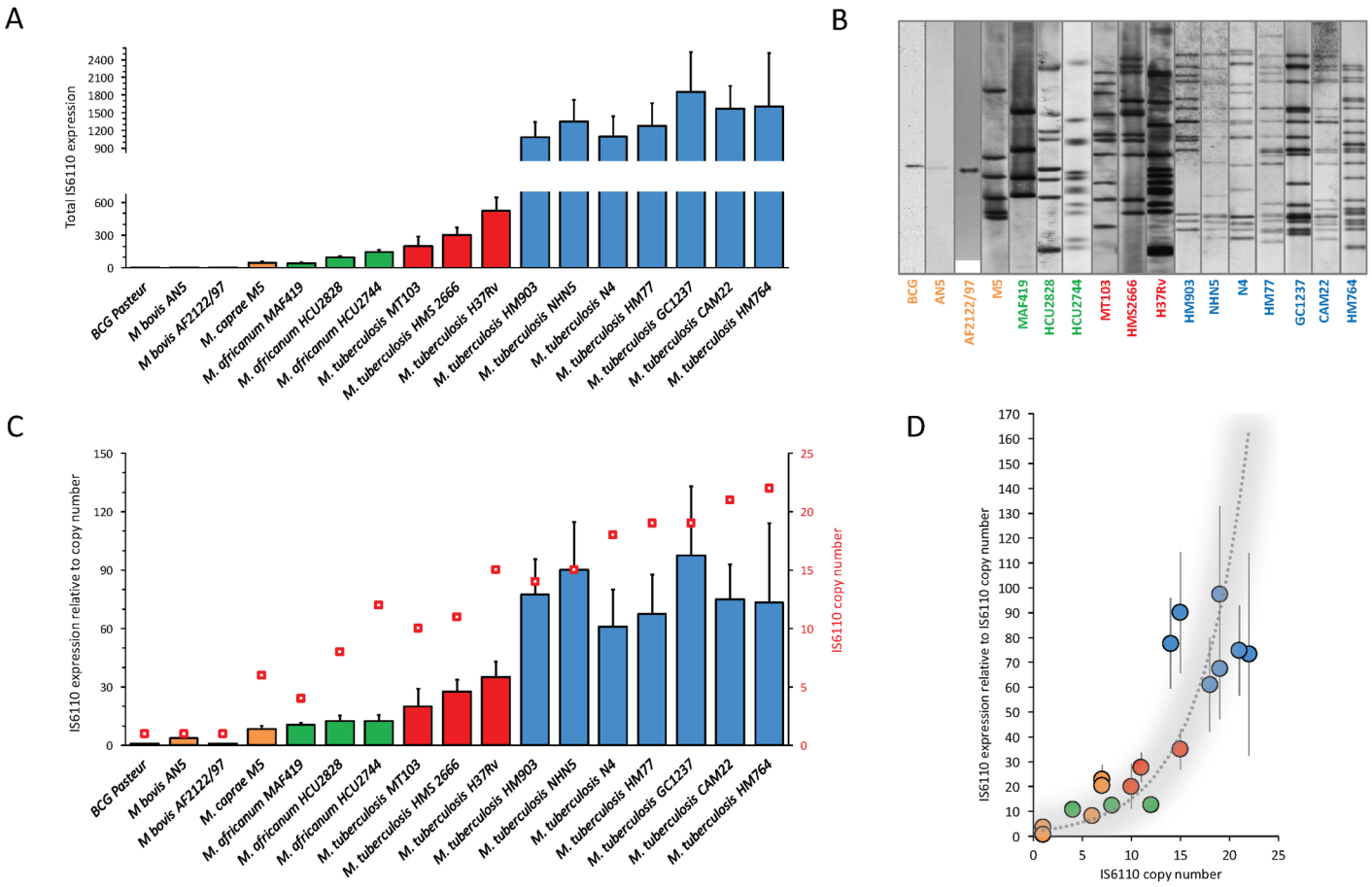


Fig 2. IS*6110* gene expression and determination of transposition dynamics in the MTBC. (**a**) Total IS*6110* expression in representative strains from the MTBC. Data are relative to BCG Pasteur. Columns and error bars are the average and standard deviation from three independent cultures. (**b**) IS*6110* RFLP from MTBC strains analysed in panel (a). (**c**) IS*6110* expression values normalised to the copy number content of this element. Columns represent normalised expression of IS*6110* according to the left Y-axis. Red squares show the IS*6110* copy number in each strain indicated in the right Y-axis. Normalised expression of BCG Pasteur is used as reference. (**d**) Expression per *IS6110* copy relative to the copy number content in MTBC strains. Data fit with an exponential curve ($r^2 = 0.80$) indicated by a grey shadowed line.

https://doi.org/10.1371/journal.pgen.1007282.g002

observed in *M. tuberculosis* H37Rv even if the latter contains 15 IS*6110* compared to the 3–5 copies in these atypical *M. bovis* isolates (S3 Fig).

Translation of an active IS*6110* transposase is post-transcriptionally regulated by a ribosomal frameshift and a RNA pseudoknot

The use of IS*6110* as molecular epidemiological marker is useful due to its relatively low frequency of transposition which allows investigators to distinguish between currently circulating strains (transmission) and older episodes of TB (reactivation) in individual patients. Since transcription per IS*6110* copy is within the range of other genes producing physiological phenotypes in *M. tuberculosis* (S4 Fig), it is predictable that low transposition rates must be subjected to some type of post-transcriptional regulation. Our results show that both ORFs are similarly transcribed (S2 Fig). The transposase is composed of a DNA binding domain (N-term) and a catalytic integrase domain (C-term) which contains the residues forming the putative active site (D310, D350, E379) (S5A and S5B Fig). By analysing the IS*6110* genetic sequence we found that the intergenic region of the constituent ORFs contained a putative

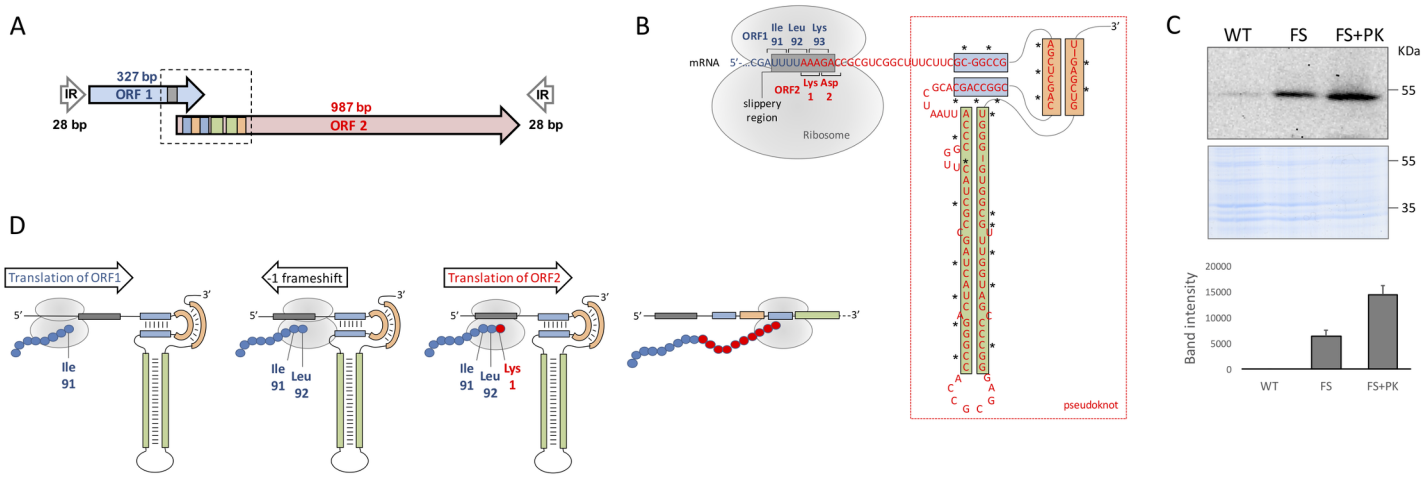


Fig 3. Post-transcriptional regulatory mechanisms of IS*6110* translation. (**a**) Genetic organization of IS*6110*. Overlapping ORF1 and ORF2 and the sense of transcription are indicated as blue and red arrows respectively. The scheme also shows the 28bp inverted repeats (IR) flanking both overlapping ORFs. (**b**) Mechanisms of post-transcriptional regulation of IS*6110*. The image shows an enlarged view of the region indicated with a dotted box from panel (a). The UUUUAAAG slippery sequence is indicated by a grey box. ORF1 and ORF2 as well as their coding triplets are indicated by blue and red letters according to panel (a). The RNA pseudoknot is included within a red rectangle and those regions involved in base pairing formation of secondary structures are indicated by blue, green and orange boxes. The position of the ribosome and the translated codons are also indicated. Asterisks in the pseudoknot indicate positions carrying mutations that disrupt this structure. (**c**) Expression of 3xFLAG variants of IS*6110*-WT, the transcriptionally active transposase IS*6110*-FS and the latter variant carrying mutations to disrupt pseudoknot formation IS*6110*-FS+PK. The upper and lower parts of the panel show a western-blot using and anti-FLAG antibody and a Coomassie staining which serves as loading control respectively. The right side of the panel shows the band intensity average from three independent experiments. (**d**) Post-transcriptional regulation of IS*6110* to produce a biologically active transposase. The image shows translation steps indicating the sense of ribosomal advance and the mRNA structure indicated in panel (b). Translation of the ORF1 produces the aminoacids from the N-terminus of IS*6110* (blue spheres) until it translates Ile91 and Leu92 coded by AUU and UUA triplets in the slippery region (grey box). At this position ribosome stalls probably because the presence of the downstream pseudoknot presenting a tight secondary structure. Stalling favours a -1 frameshift in the slippery region. Translation continues in the AAA codon coding for the Lys1 position of ORF2 (red sphere) until the ribosome reaches the C-terminus of IS*6110* coded in this latter ORF.

https://doi.org/10.1371/journal.pgen.1007282.g003

translational frameshift that could produce an active transposase as described for other members of the IS*3* family [47]. Since the precise translational frameshift has not been documented for IS*6110*, we searched for heptanucleotide U/A-rich sequences defined by the motif XXX-YYY-Z [47] in the overlapping region of ORF1 and ORF2 since these sequences are prone to ribosomal slippage. A auUUU-AAA-Gac sequence was located in the appropriate location (**Fig 3A and 3B**). This sequence codes for Ile91 (AUU), Leu92 (UUA) and Lys93 (AAG) codons of ORF1 and upon translational slippage it codes for Lys1 (AAA) and Asp2 (GAC) of ORF2 (**Fig 3B**). Additionally, we found a tight RNA secondary structure known as pseudoknot immediately downstream of the slippage sequence (**Fig 3A and 3B**). Pseudoknots are very complex and stable RNA structures with diverse biological functions, which include self-catalytic activity or the induction of ribosomal frameshifting [48]

To validate these mechanisms, we constructed three genetic variants of IS*6110* fused to a 3xFLAG epitope in order to detect the functional transposase by western blot. These variants were: the wild type (WT) sequence containing the UUU-AAA-G slippage region (IS*6110*-WT-FLAG), a construct with an A insertion in the slippage sequence (UUU-AAAA-G) to produce a complete transposase in the absence of ribosomal frameshift (IS*6110*-FS-FLAG) and a third construct including the previous A insertion and several mutations to disrupt pseudoknot formation without affecting the coding sequence (IS*6110*-FS+PK-FLAG) (**S6 Fig**). These variants were introduced in *Escherichia coli* to detect IS*6110* protein expression. We barely detected the IS*6110* using the WT sequence. In contrast, by introducing an A insertion a transcriptionally active transposase was detected as a discrete band (**Fig 3C**). Further, introduction of mutations in the pseudoknot sequence resulted in even more increased translation of the

functional transposase (Fig 3C). Based on these findings, we infer that post-transcriptional regulation of IS*6110* occurs by the combined action of two genetic mechanisms inherent to its coding sequence. The presence of a slippage sequence and a downstream pseudoknot would favour ribosome stalling at the appropriate location and the subsequent -1 translational frameshift (Fig 3D). In addition, the 5' end of the IS*6110* transcript is predicted to form a hairpin structure which occludes the ribosome binding sequence (S5C Fig) and possibly interferes with translation.

Transcriptionally active IS*6110* results in higher transposition frequencies during laboratory growth

Our next step was to demonstrate that the IS*6110* transposase produced after translational frameshift is biologically active when mycobacteria are grown under laboratory conditions. To avoid homologous recombination or other potential confusing effects that could be produced from orthologue IS*6110* sequences, we decided to study transposition in *M. smegmatis* mc^2155, a fast growing, non pathogen mycobacterial surrogate host in which the IS*6110* is not present [49]. It is important to remark that IS*6110* is exclusive of the MTBC and albeit a related IS*6110* (67% aminoacid identity) has been found in the MKD8 strain of *M. smegmatis*, this copy is non functional [49]. We cloned in a mycobacterial integrative plasmid either the wild type (pIS*6110*-WT) or a variant carrying the A insertion in the slippage region (pIS*6110*-FS) expected to be transcriptionally active (Fig 4A). Active transposases recognize the ends flanking the transposon, which in the case of IS*6110* are the IR, and catalyse "copy-out-paste-in" transposition [50, 51]. Accordingly, we constructed a third plasmid to act as a transposition reporter. A kanamycin resistance cassette flanked by the IR regions of the IS*6110* was cloned in a conditionally replicating plasmid with thermosensitive origin and *sacB* counter-selectable marker and named pIR-Km (Fig 4A). Plasmid pIR-Km was introduced in *M. smegmatis* mc^2155 carrying either pIS*6110*-WT or pIS*6110*-FS and maintained at 30˚C. We confirmed that both strains grew at comparable rates (Fig 4B). To measure transposition frequency, aliquots were plated on 7H10 medium to enumerate total CFU or on 7H10 medium containing kanamycin and sucrose and incubated at 42˚C. Under these latter conditions pIR-Km does not replicate and consequently kanamycin and sucrose resistant colonies arise from transposition of the IR-Km-IR construct into the chromosome (Fig 4C). Our results revealed that the transcriptionally active transposase in pIS*6110*-FS exhibited 20-fold higher transposition rates than the wild type IS*6110* (Fig 4D). Differences in transposition frequencies between both transposase variants were notably significant during the exponential and early stationary growth with a higher proportion of colonies resulting from transposition in this latter phase (Fig 4D). This result opens the door to hypothesize whether transposition *in vitro* is phase-dependent or conversely it results from accumulation of transposition events during mycobacterial growth. In order to confirm that transposition occurs randomly across *M. smegmatis* chromosome, we used a similar RFLP-IS*6110* analysis to that used in MTBC clinical isolates. Several kanamycin and sucrose resistant colonies were chosen at random and their restriction fragments were hybridised with a probe against the IR-Km-IR fragment. The RFLP showed loss of signal from the pIR-Km indicative of the appropriate plasmid loss. A polymorphic RFLP pattern was observed, indicative that IS*6110* transposition occurred at random locations in the chromosome (Fig 4E).

Transposition in *M. bovis* and *M. tuberculosis* preferentially occurs during laboratory starvation and in a mouse infection model of TB

Once studied the IS*6110* distribution in more than 2.000 strains covering various MTBC lineages and after we have experimentally demonstrated that low transposition frequencies of

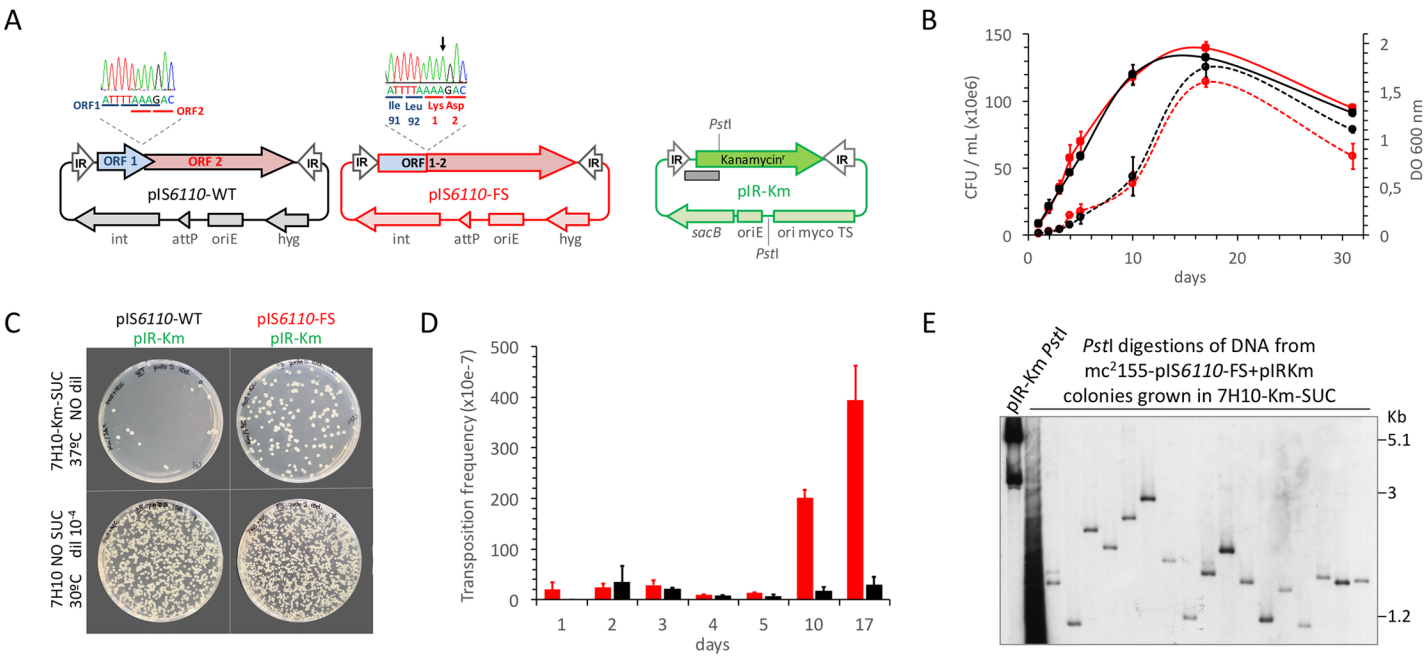


Fig 4. Construction of a transcriptionally active IS*6110* and measure of transposition frequencies during laboratory growth. (**a**) Plasmids used in the transposition reporter system. pIS*6110*-WT and pIS*6110*-FS are mycobacterial integrative plasmids carrying either the wild type IS*6110* or a mutated variant producing a transcriptionally active transposase respectively. The upper side of this panel shows Sanger sequencing histograms indicating the position of the A insertion in pIS*6110*-FS. pIR-Km is a conditionally replicating plasmid with thermosensitive origin of replication and the *sacB* gene conferring sucrose sensitivity. This plasmid contains a Kanamycin resistance cassette (km) flanked by the IR regions of IS*6110*. Positions of *Pst*I sites and probe (grey rectangle) used in RFLP shown in panel (e) are indicated. (**b**) Growth rates of liquid cultures at 30˚C of *M. smegmatis* transformed with either pIS*6110*-WT+pIR-Km (black lines) or pIS*6110*-FS+pIR-Km (red lines). Growth curves measured by OD_{600} and enumeration of CFU/mL are represented by continuous or dotted lines respectively. Error bars represent the standard deviation from three independent cultures. (**c**) CFU from *M. smegmatis* cotransformed with pIS*6110*-WT+pIR-Km or pIS*6110*-FS+pIR-Km and plated on 7H10 media supplemented with or without kanamycin and sucrose. Dilution used and incubation temperature are indicated. Note the increase in the number of CFU grown in kanamycin and sucrose medium for the pIS*6110*-FS variant relative to the pIS*6110*-WT. (**d**) Determination of transposition frequencies in *M. smegmatis* cotransformed with pIS*6110*-WT+pIR-Km (black columns) or pIS*6110*-FS+pIR-Km (red columns). Error bars indicate the standard deviation from three independent experiments. Note that the transcriptionally active transposase in pIS*6110*-FS increases up to 20-fold its transposition frequency relative to the wild type transposase in exponential and stationary periods. (**e**) RFLP analysis of DNA from colonies grown in kanamycin and sucrose plates resulting from transposition events. Note the loss of signal for pIR-Km indicative of the appropriate plasmid loss and the concomitant presence of an aleatory band pattern indicative of randomised transposition in the *M. smegmatis* chromosome.

https://doi.org/10.1371/journal.pgen.1007282.g004

IS*6110* are due to a post-transcriptional mechanism in *M. smegmatis*, we pursued our investigations in analysing potential biological conditions" promoting IS*6110* transposition in slow growing MTBC. We chose as reference strains *M. tuberculosis* H37Rv (15 IS*6110* copies) belonging to L4 and *M. bovis* AF2122/97 (1 IS*6110* copy) as representative of the animal-adapted L8. Each strain was transformed with the IS*6110* transposition reporter pIR-Km plasmid to measure transposition during growth on laboratory media or in a mouse infection model of TB. Aliquots of the culture or from organ homogenates at different time points were plated on conventional 7H10 medium to enumerate total CFU or on 7H10 supplemented with sucrose and kanamycin to recover colonies resulting from IS*6110* transposition (Fig 5A). For *in vitro* transposition experiments, we first confirmed that both strains carrying pIR-Km grew at comparable rates at 30˚C, a permissive temperature for this plasmid (S7 Fig). Then, we selected 1, 4 and 12 months' time points as representative for exponential, stationary and starvation periods in *in vitro* cultures according to growth curves at 30˚C (S7 Fig). Our results for *M. tuberculosis* H37Rv indicate that transposition rates were 10- and 60-fold higher in stationary and starvation periods respectively relative to exponential growth (Fig 5B). When examining *M. bovis* AF2122/97, similar transposition frequencies were observed under exponential

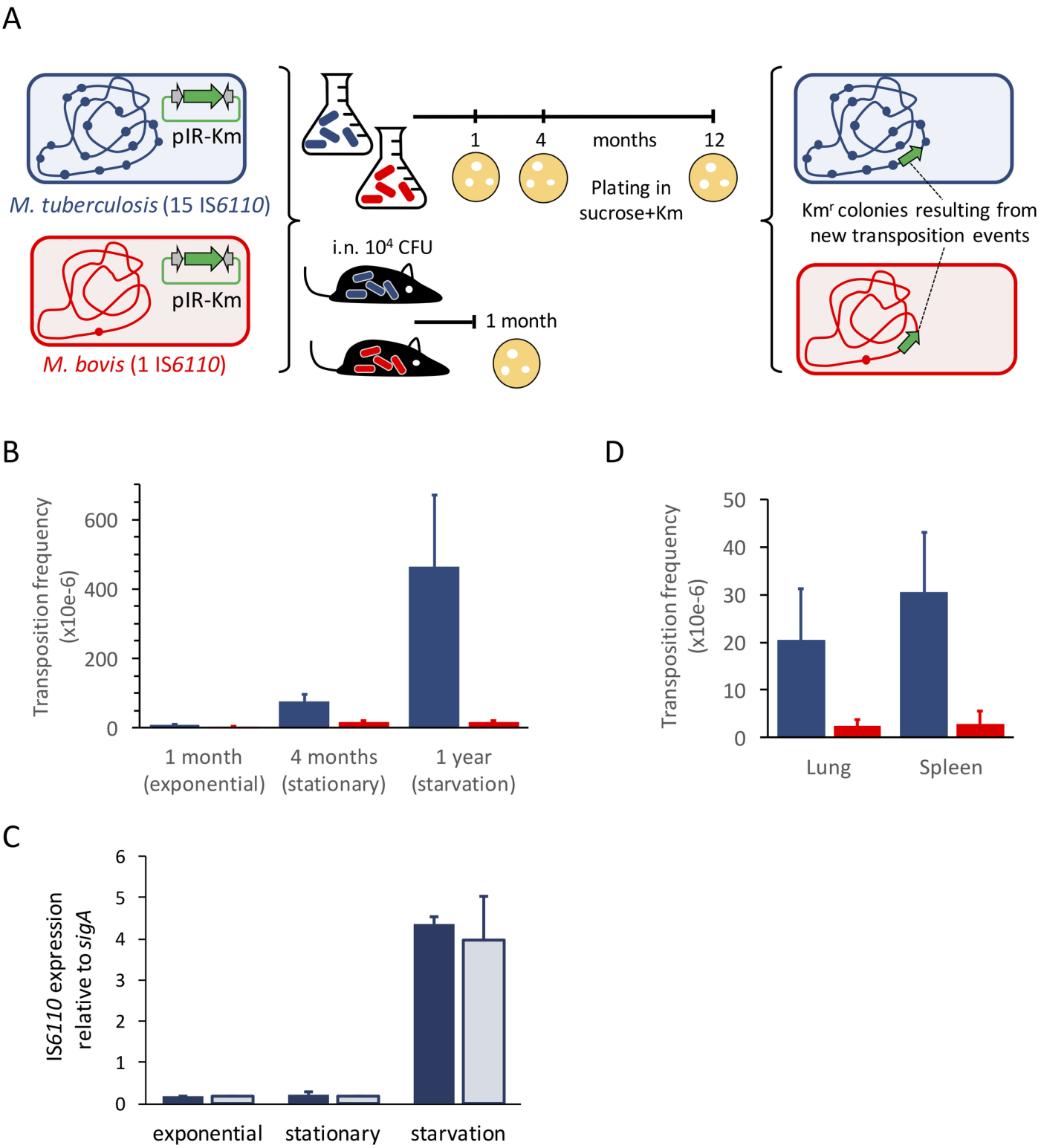


Fig 5. Transposition in the laboratory and in a mouse infection model using reference *M. bovis* and *M. tuberculosis* strains. (**a**) Experimental model to measure transposition rates in *M. tuberculosis* H37Rv (15 IS*6110* copies) and *M. bovis* AF2122/97 (1 IS*6110* copy). Both strains are transformed with pIR-Km and used to inoculate liquid media or to intranasally infect C57BL/6 mice. After the indicated time points aliquots are plated in kanamycin and sucrose containing plates to ensure pIR-km loss and to recover colonies resulting from transposition. (**b**) Transposition frequencies in laboratory medium in *M. bovis* and *M. tuberculosis* are indicated by red and blue columns respectively. Error bars indicate the standard deviation of the mean value from three independent cultures. Transposition

preferentially occurs in *M. tuberculosis* after the stationary phase reaching it maximum in the starvation period. (**c**) Expression of *M. tuberculosis* IS*6110* during exponential, stationary and starvation periods *in vitro*. Expression of ORF1 and ORF2 are indicated by dark and light blue columns respectively. Results are from three independent cultures. (**d**) Transposition frequencies during mouse infection with *M. bovis* or *M. tuberculosis* are indicated by red and blue columns respectively. Data from lung and spleen are shown and error bars indicate the standard deviation of the mean value from three independent mice. *M. bovis* does not exhibit increased transposition rates *in vivo* relative to liquid culture. Conversely *M. tuberculosis* show 10-fold higher transposition rates compared to exponential growth *in vitro*. In all cases, transposition frequencies were calculated relative to the total number of CFU in either cultures or mouse organs.

https://doi.org/10.1371/journal.pgen.1007282.g005

growth with respect to *M. tuberculosis*. However, although transposition in *M. bovis* strain was 5-fold higher in stationary and starvation periods, this was noticeably lower than that observed for *M. tuberculosis* (Fig 5B). We also quantitated *M. tuberculosis* IS*6110* expression during laboratory growth and we found higher mRNA transcription in the starvation period (Fig 5C). This result indicates that even if high mRNA expression does not necessarily imply high translation rates, there is a remarkable correlation between the transposase expression and the transposition frequencies. These results indicate that transposition increases starting from the stationary growth and similarly to that observed in *M. smegmatis* we cannot rule out the possibility that transposition events accumulate during growth *in vitro*. Further, the comparison of both strains allows us to establish lineage-defined transposition frequency. These results are remarkably comparable with our previous findings indicating that normalised expression of IS*6110* is lineage-specific, being 35-fold higher in *M. tuberculosis* than in *M. bovis* (Fig 2C).

Our transposition experiments in mice correlate with our findings during laboratory growth with an interesting exception: transposition rates for *M. bovis* AF2122/97 did not differ in the mouse model relative to exponential growth in laboratory medium, both being in the order 10^{-6} (Fig 5B and 5D). This result agrees with current biological and clinical data indicating that the single IS*6110* copy of *M. bovis* strains has been maintained during evolution with rare cases of transposition in this lineage. Further supporting these observations, our previous results demonstrate marginal levels of normalised IS*6110* mRNA expression in *M. bovis* isolates (Fig 2C). Conversely, for *M. tuberculosis* H37Rv, we observed a 10-fold increase in the transposition rates during mouse infection relative to exponential growth (Fig 5D). This increase was observed not only in the lung -the primary site of infection-, but also in the spleen of infected animals (Fig 5D). Finally, we also demonstrate that normalised expression of IS*6110* increases upon infection of murine alveolar macrophages (S8 Fig) and this result supports our transposition experiments in the mouse model of TB.

Discussion

The IS*6110* belongs to the IS*3* family whose more representative member is IS*911*. In this work, we first demonstrate that similarly to IS*911*, the IS*6110* is subjected to -1 ribosomal frameshifting [36, 52], contains a RNA pseudoknot [47] and its transposition occurs by a copy-out-paste-in mechanism [53–55]. Then, we go a step ahead to understand the biological role of IS*6110* transposition in the MTBC biology. The reliability of IS*6110* as a clinical epidemiological marker is unquestionable. In 1993 DNA fingerprinting using IS*6110* was standardized and became the gold standard for epidemiological studies of TB in the last 25 years. Since then tens of thousands MTBC strains have been studied by IS*6110* RFLP [11]. IS*6110* RFLP requires extraction of DNA from pure cultures/sputum samples which is then used in a Southern-Blot hybridization. Consequently, this is a laborious and time-consuming technique that in the last 5–10 years is being replaced by PCR methods based on amplification of MIRU [13, 14] or even more recently, by WGS [15, 16]. However, MIRU does not allow to know the number and position of IS*6110* insertions in the MTBC strains and most WGS studies fail to determine the number and localizing repeated sequences in the genome, such as the insertion sites of IS*6110*.

Hopefully, new PacBio and Oxford Nanopore sequencing technologies will improve the resolution of WGS.

After an in depth systematic analysis of 2,236 clinical isolates typed by IS*6110*-RFLP our findings show the different distribution of IS*6110* between the various MTBC lineages. Our results reveal that modern lineages of the MTBC (L2, L3 and L4) have accumulated higher IS*6110* copy number than ancient lineages (L1, L5 and L6) (**Fig 1**). Since modern lineages are widely distributed and consequently they are more successfully adapted to high density populations they have been referred to as generalists [43]. Conversely, lineages geographically restricted to certain regions are considered specialists [43]. Given the role of mobile genetic elements in providing chromosomal variability, it is tempting to think that the higher IS*6110* number in generalists might represent a strategy of the MTBC to adapt to different populations. A potential limitation of our study is the predominance of strains corresponding to L4, more frequent in Europe, Africa and America. Similar studies in other places of the world using larger number of the remaining MTBC lineages would be important to confirm the results of the present study.

As with any mutational event, transposition could be deleterious, neutral or advantageous and these events might impact on the pathogen fitness. Accordingly, another limitation of our study is inherent to the use of clinical isolates since only advantageous phenotypes are selected and we might be observing only those IS*6110* transposition events providing benefits in terms of enhanced transmissibility or pathogenicity. In this context, those transposition events observed during our mouse infection experiments might be the result of enhanced fitness *in vivo*. Accordingly, serially infecting batch of mice with those bacteria resulting from transposition events would surely enrich the bacterial population for IS*6110* locations conferring selective phenotypes.

The transposition dynamics of IS*6110* imply an exponential relationship between the copy number content and mRNA expression per IS*6110* copy, (**Fig 2D**). Accordingly, the increased expression per IS*6110* copy observed in high copy number strains (**Fig 2C**) provide more messenger molecules and this probably results in increased probability of ribosomal frameshift and translation of a functional transposase, leading to accumulation of this mobile element across the chromosome. On the other hand, even if transposition generally occurs at random across the MTBC chromosomes, it remains to be answered why some genomic regions such as a 600Kb close to the origin of replication lack IS*6110*, pointing to the detrimental effect of insertions in this location [56], while other regions such as *plcD* are prone to accumulate IS*6110* insertions and result in IS*6110*-mediated deletions such as RvD2 [57].

Since only subgroups STB-A, -D and–L (but not–J and–K) of the MTBC progenitor *M. canettii* contain IS*6110*, we can hypothesize about the origin of this transposon prior to or during the evolution of the MTBC progenitor. Recent evidence has shown that *M. canettii* strains, in contrast to the MTBC, are not clonal and could exchange DNA [58]. *M. canettii* STB-D, -A and -L share adjacent C-term and N-term truncated regions of IS*6110* separated by 1,2 Kb (**S1 Fig**). DNA binding and integrase domains are located in the opposite ends of the IS*6110* coding sequence (**S5 Fig**) and we can hypothesize about the origin of IS*6110* by a recombination between these adjacent regions (**S1 Fig**). Reinforcing this hypothesis, similar recombination events leading to surface remodelling have been recently documented in *M. canettii* [59].

The low transposition frequencies observed during the natural infection support the remarkable value of IS*6110* as a molecular epidemiology marker. Transposition is probably maintained at low levels by the action of several mechanisms. Here, we found two regulatory pathways involving translational slippage or the formation of secondary RNA structures, such as pseudoknot, but we cannot discard other regulatory mechanisms. The putative ribosome binding sequence of the IS*6110* is occluded by a stem loop (**S5C Fig**) and formation of this

structure is expected to have some impact over translation of the transposase. An important question is whether the mRNA initiates within the own IS*6110* or whether it initiates upstream from an adjacent promoter in the MTBC chromosome. In this latter case, IS*6110* transcription might very well depend on the precise location of this transposon within the host chromosome. This assumption would justify why high copy number strains are associated with higher expression rates per IS*6110* copy and *vice versa* (**Figs 1 and 2**). The exploration of *M. bovis* RNA-seq data indicates negligible transcription of the unique IS*6110* copy in this species with no presence of neighbour transcription start sites [60, 61]. The position of the IS*6110* copy in *M. bovis* is shared by most members of the MTBC and it is located within the Direct Repeats (DR) region of the CRISPR-Cas locus. Since this region is subjected to a complex post-transcriptional regulation involving RNA processing steps, this might explain the low expression of this IS*6110* copy. This is congruent with our expression data (**Fig 2**) and reinforces the hypothesis that lower transcription is likely associated with decreased probability of translational frameshift and consequently with low transposition rates in *M. bovis* (**Fig 5B and 5C**).

Our results with *M. tuberculosis* H37Rv indicate that transposition of IS*6110* is not limited to the natural infection and also occurs during growth *in vitro* (**Fig 5B**). Supporting this finding, the examination of H37Rv reference strains across multiple laboratories worldwide indicate different transposition events of IS*6110* [62]. Other example of changes in the IS*6110* pattern is the presence of 19 and 15 IS*6110* copies in H37Ra and H37Rv respectively. Since these strains arose during laboratory subculture of the original H37R strain in the 1930's, differential IS*6110* are likely the result of separate and individual transposition events during *in vitro* passage. It is interesting to observe higher transposition frequencies in long-term cultures (**Fig 5B**), but it remains to be answered whether this is the result of cumulative transposition events during the growth curve or transposition increases as a consequence of starvation signals. Examination of the IS*6110* mRNA expression indicates a strong upregulation during starvation (**Fig 5C**) which could indicate the presence of yet unknown stimulating signals triggering IS*6110* mobilization. To rule out that differences in the mutation rate of starved bacteria influence the transposition frequency, we should have performed a fluctuation test or similar. Nevertheless, a previous work demonstrated similar mutation rates during latency and during active disease or in a logarithmic growing culture [63], which agrees with the low mutation rates observed for MTBC chromosomes [40].

Several lines of evidence support a possible role for *IS6110* during adaptation to different hosts. Diverse epidemiological studies have demonstrated that IS*6110* RFLP presents distinct profiles in *M. tuberculosis* transmission clusters [28, 64–67]. Since these studies involve isolates from different patients isolated during prolonged periods of time, it is plausible to think that IS*6110*-mediated adaptive mechanisms might be involved in the patient-to-patient transmission of *M. tuberculosis*. Supporting this idea, our results indicate higher transposition rates during infection (**Fig 5D**). Another evidence comes from the observation that *M. bovis* are able to infect humans but rarely transmits between this population. However, a specific *M. bovis* strain was responsible of a deathly human MDR TB outbreak [68, 69] and this phenotype is largely related to transposition of a second IS*6110* [70]. This second IS*6110* is located upstream the *phoPR* virulence genes and acts as a mobile exogenous promoter increasing virulence phenotypes in *M. bovis* [71]. A very recent study demonstrates that IS*6110*-mediated deletions in the *ppe38-ppe71* genes are widespread in "modern" Beijing strains. This genotype result in lack of secretion of PE_PGRS and PPE-MPTR proteins and lead to increased virulence. Accordingly, this specific deletion mediated by IS*6110* may have contributed to the success and global distribution of this Beijing sublineage [72]. A previous study confirmed that Beijing (L2) strains have higher mutation rates than L4 strains, which result in increased acquisition of drug resistance in the former [73]. It is at present unknown whether varying

mutational rates across MTBC lineages can impact on transposition rates. Further work is needed to confirm whether higher mutation rates provide the driving force for increased transposition or *viceversa*.

In conclusion, our findings indicate that the lineage-specific number of IS*6110* results from differential transcriptional and posttranscriptional mechanisms inherent to the MTBC chromosomes in order to control the copy number of this transposon. Our results show that IS*6110* transposition increases during mouse infection and during growth in starvation suggesting the potential role of IS*6110* transposition during the MTBC adaptation to the host. In the future, many MTBC strains are being massively sequenced, this opportunity should be taken into consideration to locate the IS*6110* insertion sites, which would lead us to a better understanding of its biological role in TB pathogenesis and life cycle.

Material and methods

Ethics statement

All procedures were carried out under Project Licence PI14/14 approved by the Ethic Committee for Animal Experiments from the University of Zaragoza. The care and use of animals were performed accordingly with the Spanish Policy for Animal Protection RD53/2013, which meets the European Union Directive 2010/63 on the protection of animals used for experimental and other scientific purposes.

MTBC strains, growth conditions and chemicals

Strains from the MTBC and *M. smegmatis* mc^2155 were routinely grown at 37˚C in 7H9 medium (Difco) supplemented with 0.05% Tween 80 and 10% albumin-dextrose-catalase (ADC, Middlebrook) or on 7H10 plates supplemented with 10% ADC. For MTBC strains different from *M. tuberculosis*, 40 mM sodium pyruvate was added to the medium. *E. coli* DH5α used for cloning procedures was grown at 37˚C in LB broth or on LB agar plates. Ampicillin (100 μg/ml), kanamycin (20 μg/ml) and hygromycin (20 μg/ml) were used as appropriate. For transposition experiments, cultures were incubated at 30˚C or 37˚C and sucrose was added to 7H10 plates at a final concentration of 2% for *M. tuberculosis* and *M. bovis* and 10% for *M. smegmatis* if appropriate. All chemicals were purchased from Sigma-Aldrich, unless otherwise stated.

Plasmid construction

IS*6110* containing IR was PCR amplified from *M. tuberculosis* H37Rv DNA using primers *Nhe*I-IS*6110*-fw (GCTAGCTGAACCGCCCCGGCATG) and *Nhe*I-IS*6110*-rv (GCTAGCTG AACCGCCCCGGTGAGT). The PCR product was digested with *Nhe*I and cloned into *Nhe*I cut pMV361 to yield pIS*6110*-WT. To construct IS*6110* carrying the -1 translational frameshift, a two-step overlapping PCR strategy was used. ORF1 was amplified with primers *Nhe*I-IS*6110*-fw and *IS6110*-FS-rv (CGACGCGGTCTTTTAAAATCGCGT) and ORF2 with *Nhe*I-IS*6110*-rv and *IS6110*-FS-fw (ACGCGATTTTAAAAGACCGCGTCG). Both PCR products overlap in 24 nucleotides and carry the extra nucleotide required for the translational frameshifting (underlined nucleotides). These products were used as self-templates in a PCR reaction that was amplified using the flanking primers *Nhe*I-IS*6110*-fw and *Nhe*I-IS*6110*-rv, digested with *Nhe*I and introduced in the *Nhe*I site of pMV361 to yield pIS*6110*-FS. The resulting constructs were confirmed by Sanger sequencing, introduced in *M. smegmatis* mc^2155 by electroporation and colonies carrying a chromosome-integrated vector were checked by PCR.

To construct the transposition reporter, a kanamycin resistance cassette from pKD4 was amplified with primers *Bam*HI-IR1-P1 (cgcggatccgcgTGAACCGCCCCGGCATGTCCGGAG ACTCgtgtaggctggagctgcttc) and *Bam*HI-IR2-P2 (cgcggatccgcgTGAACCGCCCCGGTGAGT CCGGAGACTCcatatgaatatcctccttag), which include the IR from IS*6110* indicated in capital letters. The PCR product was confirmed by Sanger sequencing, digested with *Bam*HI and introduced into pPR27 cut with the same enzyme. The final plasmid was named pIR-Km and was introduced into *M. smegmatis* mc^2155, *M. tuberculosis* H37Rv and *M. bovis* AF2122/97 by electroporation. Transformants were selected with kanamycin at 30˚C and cultured at this permissive temperature to allow plasmid replication.

Tagged variants of IS*6110* were obtained by gene synthesis (Genescript) as follows: a 3xFLAG epitope (DYKDHDGDYKDHDIDYKDDDDK) with codons optimized for *M. tuberculosis* was placed in frame immediately after the IS*6110* coding sequence. To construct the transcriptionally active IS*6110*-FS, an A insertion was placed after the Leu92 codon. To construct the IS*6110*-FS+PK variant carrying mutations disrupting the pseudoknot, the original sequence (cgcggccgagctcgaccggccagcacgctaattacccggttcatcgccgatcatcagggccaccgcgagggcccc-gatggtttgcggtggggtgtcgag) was replaced by (cgggggcgtgcacgtcccgcgagtacgctaattacgcggtttattgcc-gaccaccaagggcaccgcgaggggcccgacggcttaaggtggggagtggaa) to maintain the aminoacid sequence. The final constructs were flanked by *Xmn*I and *Eco*RI sites at the 5' and 3' ends respectively and cloned between these sites in pMV361. These plasmids were introduced in *E. coli* DH5α for subsequent experiments.

Restriction fragment length polymorphism of IS*6110*

DNA from MTBC strains or *M. smegmatis* mc^2155 were extracted by the CTAB/NaCl procedure. DNA integrity was confirmed by agarose gel electrophoresis. For standard IS*6110* RFLP, DNA was digested with *Pvu*II and separated overnight in 0.8% agarose gels. DNA was transferred from the gel to a positively charged nylon membrane (Hybond N^+, Amersham) by using a vacuum transfer device. The membrane was hybridized with a probe amplified with primers INS-1 (CGTGAGGGCATCGAGGTGGC) and INS-2 (GCGTAGGCGTCGGTGAC AAA). After hybridization with labeled DNA probes, the bound probes were detected with an enhanced chemiluminescence direct nucleic acid detection system (Amersham) according to the manufacturer's recommendations.

For RFLP of colonies resulting from transposition of the IR-Km-IR cassette from the pIR-Km transposition reporter, these modifications were introduced in the RFLP protocol: DNA was digested with *Pst*I and hybridized with a probe amplified with P1 (GTGTAGGCTG-GAGCTGCTTC) and Km-pKD4-out1 (CCACGATAGCCGCGCTGCCTCG) primers using pIR-Km as template.

Bioinformatic analysis

Genome sequences were retrieved from NCBI GenBank (http://www.ncbi.nlm.nih.gov/). The copy number content and genomic polymorphisms in IS*6110* were calculated using nucleotide BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Secondary RNA structures were predicted using the RNA fold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Pseudoknot structures and their estimated free energy were located and computed using DotKnot (http://dotknot.csse.uwa.edu.au/).

RNA extraction and normalized expression of IS*6110*

Mycobacterial cultures were grown to exponential phase (OD_{600} = 0.5–0.6) and pelleted by centrifugation. To minimize RNA degradation bacteria were resuspended in 1 ml RNA Protect

Bacteria Reagent (Qiagen), incubated for 5 min at room temperature and then centrifuged. Bacterial pellets were resuspended in 0.4 ml lysis buffer (0.5% SDS, 20 mM NaAc, 0.1 mM EDTA) and 1 ml phenol:chloroform (pH = 4.5) 1:1. Suspensions were transferred to tubes containing glass beads (Qbiogene) and lysed using a Fast-prep instrument with a three-cycle program (15 sec at speed 6.5 m) including cooling the samples on ice for 5 min between pulses. Samples were then centrifuged and the homogenate was removed from the beads and transferred to a tube containing chloroform:isoamylalcohol 24:1. Tubes were inverted carefully before centrifugation and the upper (aqueous) phase was then transferred to a fresh tube containing 0.3 M Na-acetate (pH = 5.5) and isopropanol. Precipitated nucleic acids were collected by centrifugation. The pellets were rinsed with 70% ethanol and air dried before being re-dissolved in RNase-free water. DNA was removed from RNA samples using Turbo DNA free (Ambion) by incubation at 37˚C for 1 h. RNA integrity was assessed by agarose gel electrophoresis and absence of contaminating DNA was checked by lack of amplification products after 30 PCR cycles.

One microgram of MTBC RNA was converted to cDNA using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's recommendations. The 10 μl PCR reaction consisted of 1X SYBR Green PCR Master Mix (Applied Biosystems), 0.25 μM of each primer and 1 μl of 1:10 diluted cDNA or IP DNA from immunoprecipitation reactions. Reactions were carried out in triplicate in an Applied Biosystems StepOnePlus^{TM} Sequence Detection System (Applied Biosystems) according to the manufacturer's instructions. Melting curves were constructed to ensure that only one amplification product was obtained. Normalization was obtained to the number of *sigA* molecules in each sample. To obtain normalized expression values per IS*6110* copy number, data normalized with respect to *sigA* were subsequently divided by the total number of IS*6110* for every strain used. All qRT-PCR primers were designed using Primer Express software (Applied Biosystems) and sequences are as follows: RT-IS*6110*-1-fw (TCAGCACGATTCGGAGTGG), RT-IS*6110*-1-rv (CCAAGTAGACG GGCGACCT), RT-IS*6110*-2-fw (CGCAAAGTGTGGCTAACCCT), RT-IS*6110*-2-rv (GCAT CTGGCCACCTCGAT), RT-sigA-fw (CCGATGACGACGAGGAGATC) and sigA-rv (CGGAGGCCTTGTCCTTTTC).

Western blot

The pelleted fraction of bacterial cultures was resuspended in PBS containing 1% triton X100 and a cocktail of protease inhibitors (Roche) and disrupted using a Fast-Prep during three pulses, 1 minute each, cooling on ice between pulses. Samples were then centrifuged and the upper phase containing whole-cell lysate was quantitated using the RC DC protein assay (BioRad). Equal amounts of protein preparations were loaded per well. Proteins were separated on SDS-PAGE 12–15% gels and transferred onto PVDF membranes using a semidry electrophoresis transfer apparatus (Bio-Rad). Membranes were incubated in TBS-T blocking buffer (25 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween 20) with 5% w/v skimmed milk powder for 30 min prior to overnight incubation with primary antibodies at the dilution indicated below. Membranes were washed in TBS-T three times, and then incubated with secondary antibodies for 1 h before washing. Anti-FLAG (M2 clone, Sigma) antibody was used at 1:10,000 dilution and horseradish peroxidase (HRP) conjugated IgG secondary antibody (Sigma-Aldrich) was used at a 1:20,000 dilution. Signals were detected using chemiluminescent substrates (GE Healthcare).

Mouse infection procedures

All mice were kept under controlled conditions and observed for any sign of disease. Experimental work was conducted in agreement with European and national directives for

protection of experimental animals and with approval from the competent local ethics committees (approved protocol PI14/14). We performed a single biological replicate using 3 mice per group. Female C57BL/6 mice (Janvier Biolabs) were intranasally inoculated with 10^4 CFU of *M. tuberculosis* H37Rv or *M. bovis* AF2122/97 (both carrying the transposition reporter). Infection was left to progress for 4 weeks and bacterial burden was determined by plating homogenized lungs and spleen on solid medium. Transposition events were enumerated as described in the "transposition experiments" section.

Transposition experiments

Liquid cultures grown at 30˚C or organ homogenates from infected mice were serially diluted and plated on 7H10 medium without sucrose at 30˚C to enumerate viable bacteria. In parallel, appropriate dilutions were plated on 7H10 medium containing kanamycin and sucrose at 37˚C to obtain colonies resulting from transposition of the IR-Km-IR cassette in the mycobacterial chromosome. The transposition frequency was calculated as the number of bacteria resulting from a transposition events divided by the number of total viable bacteria.

Supporting information

S1 Fig. IS*6110* in *M. canettii*. Network phylogeny of *M. canettii* sequenced strains adapted from [21]. The position of IS*6110* sequences in fully assembled genomes of *M. canettii* is indicated. Black dots indicate a wild type sequence, red dots indicate a mutated protein and red squares show positions of truncated IS*6110*. Only STB-D, STB-A and STB-L subgroups of *M. canettii* containing IS*6110* are indicated. The lower part of this figure indicate the positional arrangement of 5' and 3' terminus of the IS*6110* shared by STB-A, STB-D and STB-L and that of a functional IS*6110* in the remaining members of the MTBC. This organization suggest a recombination between both fragment in *M. canettii* to produce an active transposase.
(TIFF)

S2 Fig. Normalised expression of ORF1 and ORF2. Blue and red bars indicate expression per IS*6110* copy for ORF1 and ORF2 respectively according to the left Y-axis. Red squares indicate the IS*6110* copy number n the analysed strains according to right Y-axis. Note that both constituent ORF from the transposase are equally expressed.
(TIFF)

S3 Fig. IS*6110* gene expression in atypical copy number strains. (**a**) Normalised IS*6110* expression in low copy number *M. tuberculosis* strains. Note that *M. tuberculosis* containing a single IS*6110* expresses this gene at comparable levels to BCG Pasteur used as reference. Accumulation of further IS*6110* copies results in higher normalised expression values. (**b**) Normalised IS*6110* expression in "high copy" number *M. bovis* strains. Note that the presence of >1 IS*6110* copies in *M. bovis* results in high expression values compared to either BCG Pasteur used as reference or *M. tuberculosis* H37Rv. (**c**) Normalised IS*6110* expression in *M. tuberculosis* Lineage 2 (Beijing) strains. Note that normalised expression values are noticeable higher than those observed in *M. tuberculosis* H37Rv from lineage 4. (**d, e, f**) RFLP from MTBC strains analysed in panels a, b and c. Columns and error bars from panels (a), (b) and (c) are the standard deviation of the mean value from three independent cultures according to the left Y-axis. Red squares in panels (a), (b) and (c) indicate the IS*6110* copy number according to the right Y-axis.
(TIFF)

S4 Fig. Expression of IS*6110* in the context of gene expression from diverse genes in *M. tuberculosis*. Each gene is measured relative to the *sigA* expression levels and columns indicate

log10 values of normalised expression values. Note that IS*6110* expression per copy is within the range of genes producing physiological phenotypes such as *tatC* involved in protein secretion or *pks3* involved in acyltrehalose containing lipids.
(TIFF)

S5 Fig. Genetic features and domain organization of the IS*6110* protein. (**a**) The two constituent ORF ara indicated by blue and red arrows. Position of transposase, integrase and helix-turn-helix domains are shown. The lower part of the panel include a description of the indicated domains. (**b**) The putative content of alpha-helices and beta-strands in the IS*6110* aminoacidic sequence is indicated by cylinders and arrows respectively (**c**) RNA secondary structure of the N-terminus. The RBS and the start codon are indicated by bold and underlined characters respectively. Note the presence of the stem loop occluding the RBS.
(TIFF)

S6 Fig. Pseudoknot structure and mutational analysis. (**a**) Structure of the IS*6110* pseudoknot indicating the positions selected for mutation (asterisks). (**b**) Alignment of wild type and mutated variants of the pseudoknot. (**c**) Formation of secondary structures in the wild type and mutated variant indicating the ΔG values. Note the formation of a stable pseudoknot in the wild type but not in the mutated variant.
(TIFF)

S7 Fig. Growth rates of liquid cultures at 30˚C of *M. tuberculosis* or *M. bovis* transformed with pIR-Km are indicated blue and red lines respectively. Enumeration of CFU/mL represent the average and standard deviation from three independent cultures.
(TIFF)

S8 Fig. IS*6110* expression during macrophage infection. Bars indicate normalised expression per IS*6110* copy after 4 and 24 hours of MHS macrophage infection (dark grey columns) relative to expression under laboratory growth (light grey columns). Data from two *M. tuberculosis* clinical isolates are provided. Results represent average and standard deviation from three independent infections.
(TIFF)

Author Contributions

Conceptualization: Jesús Gonzalo-Asensio, Irene Pérez, Nacho Aguiló, Carlos Martín.

Data curation: Jesús Gonzalo-Asensio, Alberto Cebollada, Sofía Samper.

Formal analysis: Jesús Gonzalo-Asensio, Nacho Aguiló, Carlos Martín.

Funding acquisition: Sofía Samper, Carlos Martín.

Investigation: Jesús Gonzalo-Asensio, Irene Pérez, Nacho Aguiló, Santiago Uranga, Ana Picó, Carlos Lampreave.

Methodology: Jesús Gonzalo-Asensio, Irene Pérez, Nacho Aguiló, Santiago Uranga, Ana Picó, Carlos Lampreave, Alberto Cebollada.

Project administration: Sofía Samper, Carlos Martín.

Resources: Jesús Gonzalo-Asensio, Irene Pérez, Nacho Aguiló, Santiago Uranga, Ana Picó, Isabel Otal.

Software: Jesús Gonzalo-Asensio, Alberto Cebollada.

Supervision: Jesús Gonzalo-Asensio.

Validation: Jesús Gonzalo-Asensio, Irene Pérez, Santiago Uranga, Ana Picó, Carlos Lampreave.

Visualization: Jesús Gonzalo-Asensio, Alberto Cebollada.

Writing – original draft: Jesús Gonzalo-Asensio, Carlos Martín.

Writing – review & editing: Jesús Gonzalo-Asensio, Carlos Martín.

References

1. Paulson T. Epidemiology: A mortal foe. Nature. 2013; 502(7470):S2–3. https://doi.org/10.1038/502S2a PMID: 24108078.
2. WHO. Global Tuberculosis Report 2017. 2017.
3. Gonzalo-Asensio J, Aguilo N, Marinova D, Martin C. Breaking Transmission with Vaccines: The Case of Tuberculosis. Microbiol Spectr. 2017; 5(4). https://doi.org/10.1128/microbiolspec.MTBP-0001-2016 PMID: 28710848.
4. Mc CB. The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A. 1950; 36(6):344–55. PMID: 15430309; PubMed Central PMCID: PMCPMC1063197.
5. Hedges RW, Jacob AE. Transposition of ampicillin resistance from RP4 to other replicons. Mol Gen Genet. 1974; 132(1):31–40. PMID: 4609125.
6. Thierry D, Cave MD, Eisenach KD, Crawford JT, Bates JH, Gicquel B, et al. IS6110, an IS-like element of Mycobacterium tuberculosis complex. Nucleic Acids Res. 1990; 18(1):188. PMID: 2155396; PubMed Central PMCID: PMCPMC330226.
7. Thierry D, Brisson-Noel A, Vincent-Levy-Frebault V, Nguyen S, Guesdon JL, Gicquel B. Characterization of a Mycobacterium tuberculosis insertion sequence, IS6110, and its application in diagnosis. J Clin Microbiol. 1990; 28(12):2668–73. PMID: 2177747; PubMed Central PMCID: PMCPMC268253.
8. Brisson-Noel A, Aznar C, Chureau C, Nguyen S, Pierre C, Bartoli M, et al. Diagnosis of tuberculosis by DNA amplification in clinical practice evaluation. Lancet. 1991; 338(8763):364–6. PMID: 1677709.
9. Otal I, Martin C, Vincent-Levy-Frebault V, Thierry D, Gicquel B. Restriction fragment length polymorphism analysis using IS6110 as an epidemiological marker in tuberculosis. J Clin Microbiol. 1991; 29 (6):1252–4. PMID: 1677943; PubMed Central PMCID: PMCPMC269979.
10. Mendiola MV, Martin C, Otal I, Gicquel B. Analysis of the regions responsible for IS6110 RFLP in a single Mycobacterium tuberculosis strain. Res Microbiol. 1992; 143(8):767–72. PMID: 1363676.
11. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol. 1993; 31(2):406–9. PMID: 8381814; PubMed Central PMCID: PMCPMC262774.
12. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, et al. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. N Engl J Med. 1994; 330(24):1703–9. https://doi.org/10.1056/NEJM199406163302402 PMID: 7910661.
13. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. J Clin Microbiol. 2006; 44(12):4498–510. https://doi.org/10.1128/JCM.01392-06 PMID: 17005759; PubMed Central PMCID: PMCPMC1698431.
14. de Beer JL, van Ingen J, de Vries G, Erkens C, Sebek M, Mulder A, et al. Comparative study of IS6110 restriction fragment length polymorphism and variable-number tandem-repeat typing of Mycobacterium tuberculosis isolates in the Netherlands, based on a 5-year nationwide survey. J Clin Microbiol. 2013; 51(4):1193–8. https://doi.org/10.1128/JCM.03061-12 PMID: 23363841; PubMed Central PMCID: PMCPMC3666783.
15. Kohl TA, Diel R, Harmsen D, Rothganger J, Walter KM, Merker M, et al. Whole-genome-based Mycobacterium tuberculosis surveillance: a standardized, portable, and expandable approach. J Clin Microbiol. 2014; 52(7):2479–86. https://doi.org/10.1128/JCM.00567-14 PMID: 24789177; PubMed Central PMCID: PMCPMC4097744.
16. Takiff HE, Feo O. Clinical value of whole-genome sequencing of Mycobacterium tuberculosis. Lancet Infect Dis. 2015; 15(9):1077–90. https://doi.org/10.1016/S1473-3099(15)00071-7 PMID: 26277037.
17. Smith NH, Hewinson RG, Kremer K, Brosch R, Gordon SV. Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis. Nat Rev Microbiol. 2009; 7(7):537–44. https://doi.org/10.1038/nrmicro2165 PMID: 19483712.

18. Streicher EM, Victor TC, van der Spuy G, Sola C, Rastogi N, van Helden PD, et al. Spoligotype signatures in the Mycobacterium tuberculosis complex. J Clin Microbiol. 2007; 45(1):237–40. https://doi.org/10.1128/JCM.01429-06 PMID: 17065260; PubMed Central PMCID: PMCPMC1828946.

19. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. Nat Genet. 2015; 47(3):242–9. https://doi.org/10.1038/ng.3195 PMID: 25599400.

20. Stucki D, Brites D, Jeljeli L, Coscolla M, Liu Q, Trauner A, et al. Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages. Nat Genet. 2016. https://doi.org/10.1038/ng.3704 PMID: 27798628.

21. Supply P, Marceau M, Mangenot S, Roche D, Rouanet C, Khanna V, et al. Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. Nat Genet. 2013; 45(2):172–9. Epub 2013/01/08. https://doi.org/10.1038/ng.2517 [pii]. PMID: 23291586; PubMed Central PMCID: PMC3856870.

22. Fomukong N, Beggs M, el Hajj H, Templeton G, Eisenach K, Cave MD. Differences in the prevalence of IS6110 insertion sites in Mycobacterium tuberculosis strains: low and high copy number of IS6110. Tuber Lung Dis. 1997; 78(2):109–16. PMID: 9692179.

23. Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NE, Kreiswirth BN, et al. Definition of the Beijing/W lineage of Mycobacterium tuberculosis on the basis of genetic markers. J Clin Microbiol. 2004; 42(9):4040–9. https://doi.org/10.1128/JCM.42.9.4040-4049.2004 PMID: 15364987; PubMed Central PMCID: PMCPMC516354.

24. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, et al. Predominance of a single genotype of Mycobacterium tuberculosis in countries of east Asia. J Clin Microbiol. 1995; 33(12):3234–8. PMID: 8586708; PubMed Central PMCID: PMCPMC228679.

25. Bifani PJ, Plikaytis BB, Kapur V, Stockbauer K, Pan X, Lutfey ML, et al. Origin and interstate spread of a New York City multidrug-resistant Mycobacterium tuberculosis clone family. JAMA. 1996; 275(6):452–7. PMID: 8627966.

26. McEvoy CR, Falmer AA, Gey van Pittius NC, Victor TC, van Helden PD, Warren RM. The role of IS6110 in the evolution of Mycobacterium tuberculosis. Tuberculosis (Edinb). 2007; 87(5):393–404. https://doi.org/10.1016/j.tube.2007.05.010 PMID: 17627889.

27. Roychowdhury T, Mandal S, Bhattacharya A. Analysis of IS6110 insertion sites provide a glimpse into genome evolution of Mycobacterium tuberculosis. Sci Rep. 2015; 5:12567. https://doi.org/10.1038/srep12567 PMID: 26215170; PubMed Central PMCID: PMCPMC4517164.

28. Millan-Lou MI, Otal I, Monforte ML, Vitoria MA, Revillo MJ, Martin C, et al. In Vivo IS6110 Profile Changes in a Mycobacterium tuberculosis Strain as Determined by Tracking over 14 Years. J Clin Microbiol. 2015; 53(7):2359–61. https://doi.org/10.1128/JCM.00607-15 PMID: 25948604; PubMed Central PMCID: PMCPMC4473185.

29. Safi H, Barnes PF, Lakey DL, Shams H, Samten B, Vankayalapati R, et al. IS6110 functions as a mobile, monocyte-activated promoter in Mycobacterium tuberculosis. Mol Microbiol. 2004; 52(4):999–1012. https://doi.org/10.1111/j.1365-2958.2004.04037.x PMID: 15130120.

30. Alonso H, Aguilo JI, Samper S, Caminero JA, Campos-Herrero MI, Gicquel B, et al. Deciphering the role of IS6110 in a highly transmissible Mycobacterium tuberculosis Beijing strain, GC1237. Tuberculosis (Edinb). 2011; 91(2):117–26. Epub 2011/01/25. https://doi.org/10.1016/j.tube.2010.12.007 S1472-9792(10)00146-0 [pii]. PMID: 21256084.

31. Tanaka MM, Rosenberg NA, Small PM. The control of copy number of IS6110 in Mycobacterium tuberculosis. Mol Biol Evol. 2004; 21(12):2195–201. https://doi.org/10.1093/molbev/msh234 PMID: 15317877.

32. Tanaka MM. Evidence for positive selection on Mycobacterium tuberculosis within patients. BMC Evol Biol. 2004; 4:31. https://doi.org/10.1186/1471-2148-4-31 PMID: 15355550; PubMed Central PMCID: PMCPMC518962.

33. Reyes A, Sandoval A, Cubillos-Ruiz A, Varley KE, Hernandez-Neuta I, Samper S, et al. IS-seq: a novel high throughput survey of in vivo IS6110 transposition in multiple Mycobacterium tuberculosis genomes. BMC Genomics. 2012; 13:249. https://doi.org/10.1186/1471-2164-13-249 PMID: 22703188; PubMed Central PMCID: PMCPMC3443423.

34. Alonso H, Samper S, Martin C, Otal I. Mapping IS6110 in high-copy number Mycobacterium tuberculosis strains shows specific insertion points in the Beijing genotype. BMC Genomics. 2013; 14:422. https://doi.org/10.1186/1471-2164-14-422 PMID: 23800083; PubMed Central PMCID: PMCPMC3701491.

35. Millan-Lou MI, Lopez-Calleja AI, Colmenarejo C, Lezcano MA, Vitoria MA, del Portillo P, et al. Global study of IS6110 in a successful Mycobacterium tuberculosis strain: clues for deciphering its behavior

and for its rapid detection. J Clin Microbiol. 2013; 51(11):3631–7. https://doi.org/10.1128/JCM.00970-13 PMID: 23985924; PubMed Central PMCID: PMCPMC3889744.

36. Sekine Y, Eisaki N, Ohtsubo E. Translational control in production of transposase and in transposition of insertion sequence IS3. J Mol Biol. 1994; 235(5):1406–20. https://doi.org/10.1006/jmbi.1994.1097 PMID: 8107082.

37. Wall S, Ghanekar K, McFadden J, Dale JW. Context-sensitive transposition of IS6110 in mycobacteria. Microbiology. 1999; 145 (Pt 11):3169–76. https://doi.org/10.1099/00221287-145-11-3169 PMID: 10589725.

38. Ghanekar K, McBride A, Dellagostin O, Thorne S, Mooney R, McFadden J. Stimulation of transposition of the Mycobacterium tuberculosis insertion sequence IS6110 by exposure to a microaerobic environment. Mol Microbiol. 1999; 33(5):982–93. PMID: 10476032.

39. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, et al. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proc Natl Acad Sci U S A. 2002; 99(6):3684–9. Epub 2002/03/14. https://doi.org/10.1073/pnas.052548299 052548299 [pii]. PMID: 11891304; PubMed Central PMCID: PMC122584.

40. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat Genet. 2013; 45(10):1176–82. Epub 2013/09/03. https://doi.org/10.1038/ng.2744 [pii]. PMID: 23995134; PubMed Central PMCID: PMC3800747.

41. Galagan JE. Genomic insights into tuberculosis. Nat Rev Genet. 2014; 15(5):307–20. https://doi.org/10.1038/nrg3664 PMID: 24662221.

42. Gutierrez MC, Brisse S, Brosch R, Fabre M, Omais B, Marmiesse M, et al. Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. PLoS Pathog. 2005; 1(1):e5. https://doi.org/10.1371/journal.ppat.0010005 PMID: 16201017; PubMed Central PMCID: PMCPMC1238740.

43. Brites D, Gagneux S. Co-evolution of Mycobacterium tuberculosis and Homo sapiens. Immunol Rev. 2015; 264(1):6–24. https://doi.org/10.1111/imr.12264 PMID: 25703549; PubMed Central PMCID: PMCPMC4339235.

44. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 2006; 103(8):2869–73. https://doi.org/10.1073/pnas.0511240103 PMID: 16477032; PubMed Central PMCID: PMCPMC1413851.

45. Chandler M, Fayet O. Translational frameshifting in the control of transposition in bacteria. Mol Microbiol. 1993; 7(4):497–503. PMID: 8384687.

46. Otal I, Gomez AB, Kremer K, de Haas P, Garcia MJ, Martin C, et al. Mapping of IS6110 insertion sites in Mycobacterium bovis isolates in relation to adaptation from the animal to human host. Vet Microbiol. 2008; 129(3–4):333–41. https://doi.org/10.1016/j.vetmic.2007.11.038 PMID: 18207337.

47. Mazauric MH, Licznar P, Prere MF, Canal I, Fayet O. Apical loop-internal loop RNA pseudoknots: a new type of stimulator of -1 translational frameshifting in bacteria. J Biol Chem. 2008; 283(29):20421–32. https://doi.org/10.1074/jbc.M802829200 PMID: 18474594.

48. Staple DW, Butcher SE. Pseudoknots: RNA structures with diverse functions. PLoS Biol. 2005; 3(6): e213. https://doi.org/10.1371/journal.pbio.0030213 PMID: 15941360; PubMed Central PMCID: PMCPMC1149493.

49. Coros A, DeConno E, Derbyshire KM. IS6110, a Mycobacterium tuberculosis complex-specific insertion sequence, is also present in the genome of Mycobacterium smegmatis, suggestive of lateral gene transfer among mycobacterial species. J Bacteriol. 2008; 190(9):3408–10. https://doi.org/10.1128/JB.00009-08 PMID: 18326566; PubMed Central PMCID: PMCPMC2347380.

50. Hickman AB, Dyda F. DNA Transposition at Work. Chem Rev. 2016. https://doi.org/10.1021/acs.chemrev.6b00003 PMID: 27187082.

51. Chandler M, Fayet O, Rousseau P, Ton Hoang B, Duval-Valentin G. Copy-out-Paste-in Transposition of IS911: A Major Transposition Pathway. Microbiol Spectr. 2015; 3(4). https://doi.org/10.1128/microbiolspec.MDNA3-0031-2014 PMID: 26350305.

52. Polard P, Prere MF, Chandler M, Fayet O. Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS911. J Mol Biol. 1991; 222(3):465–77. PMID: 1660923.

53. Duval-Valentin G, Marty-Cointin B, Chandler M. Requirement of IS911 replication before integration defines a new bacterial transposition pathway. EMBO J. 2004; 23(19):3897–906. https://doi.org/10.1038/sj.emboj.7600395 PMID: 15359283; PubMed Central PMCID: PMCPMC522794.

54. Ton-Hoang B, Polard P, Chandler M. Efficient transposition of IS911 circles in vitro. EMBO J. 1998; 17 (4):1169–81. https://doi.org/10.1093/emboj/17.4.1169 PMID: 9463394; PubMed Central PMCID: PMCPMC1170465.

55. Polard P, Chandler M. An in vivo transposase-catalyzed single-stranded DNA circularization reaction. Genes Dev. 1995; 9(22):2846–58. PMID: 7590258.

56. Gordon SV, Heym B, Parkhill J, Barrell B, Cole ST. New insertion sequences and a novel repeated sequence in the genome of Mycobacterium tuberculosis H37Rv. Microbiology. 1999; 145 (Pt 4):881–92. https://doi.org/10.1099/13500872-145-4-881 PMID: 10220167.

57. Ho TB, Robertson BD, Taylor GM, Shaw RJ, Young DB. Comparison of Mycobacterium tuberculosis genomes reveals frequent deletions in a 20 kb variable region in clinical isolates. Yeast. 2000; 17 (4):272–82. https://doi.org/10.1002/1097-0061(200012)17:4<272::AID-YEA48>3.0.CO;2-2 PMID: 11119304; PubMed Central PMCID: PMCPMC2448390.

58. Boritsch EC, Khanna V, Pawlik A, Honore N, Navas VH, Ma L, et al. Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. Proc Natl Acad Sci U S A. 2016; 113(35):9876–81. https://doi.org/10.1073/pnas.1604921113 PMID: 27528665; PubMed Central PMCID: PMCPMC5024641.

59. Boritsch EC, Frigui W, Cascioferro A, Malaga W, Etienne G, Laval F, et al. pks5-recombination-mediated surface remodelling in Mycobacterium tuberculosis emergence. Nature Microbiology. 2016. https://doi.org/10.1038/NMICROBIOL.2015.19 PMID: 27571976

60. Dinan AM, Tong P, Lohan AJ, Conlon KM, Miranda-CasoLuengo AA, Malone KM, et al. Relaxed selection drives a noisy noncoding transcriptome in members of the Mycobacterium tuberculosis complex. MBio. 2014; 5(4):e01169–14. https://doi.org/10.1128/mBio.01169-14 PMID: 25096875; PubMed Central PMCID: PMCPMC4128351.

61. Solans L, Gonzalo-Asensio J, Sala C, Benjak A, Uplekar S, Rougemont J, et al. The PhoP-Dependent ncRNA Mcr7 Modulates the TAT Secretion System in Mycobacterium tuberculosis. PLoS Pathog. 2014; 10(5):e1004183. Epub 2014/05/31. https://doi.org/10.1371/journal.ppat.1004183 [pii]. PMID: 24874799.

62. Ioerger TR, Feng Y, Ganesula K, Chen X, Dobos KM, Fortune S, et al. Variation among genome sequences of H37Rv strains of Mycobacterium tuberculosis from multiple laboratories. J Bacteriol. 2010; 192(14):3645–53. https://doi.org/10.1128/JB.00166-10 PMID: 20472797; PubMed Central PMCID: PMCPMC2897344.

63. Ford CB, Lin PL, Chase MR, Shah RR, Iartchouk O, Galagan J, et al. Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection. Nat Genet. 2011; 43 (5):482–6. https://doi.org/10.1038/ng.811 PMID: 21516081; PubMed Central PMCID: PMCPMC3101871.

64. de Boer AS, Borgdorff MW, de Haas PE, Nagelkerke NJ, van Embden JD, van Soolingen D. Analysis of rate of change of IS6110 RFLP patterns of Mycobacterium tuberculosis based on serial patient isolates. J Infect Dis. 1999; 180(4):1238–44. https://doi.org/10.1086/314979 PMID: 10479153.

65. Perez-Lago L, Herranz M, Bouza E, Garcia de Viedma D. Dynamic and complex Mycobacterium tuberculosis microevolution unrevealed by standard genotyping. Tuberculosis (Edinb). 2012; 92(3):232–5. https://doi.org/10.1016/j.tube.2012.01.003 PMID: 22342248.

66. Schurch AC, Kremer K, Kiers A, Daviena O, Boeree MJ, Siezen RJ, et al. The tempo and mode of molecular evolution of Mycobacterium tuberculosis at patient-to-patient scale. Infect Genet Evol. 2010; 10(1):108–14. https://doi.org/10.1016/j.meegid.2009.10.002 PMID: 19835997.

67. Small PM, Shafer RW, Hopewell PC, Singh SP, Murphy MJ, Desmond E, et al. Exogenous reinfection with multidrug-resistant Mycobacterium tuberculosis in patients with advanced HIV infection. N Engl J Med. 1993; 328(16):1137–44. https://doi.org/10.1056/NEJM199304223281601 PMID: 8096066.

68. Samper S, Martin C, Pinedo A, Rivero A, Blazquez J, Baquero F, et al. Transmission between HIV-infected patients of multidrug-resistant tuberculosis caused by Mycobacterium bovis. AIDS. 1997; 11 (10):1237–42. PMID: 9256941.

69. Samper S, Martin C. Spread of extensively drug-resistant tuberculosis. Emerg Infect Dis. 2007; 13 (4):647–8. https://doi.org/10.3201/eid1304.061329 PMID: 17561563; PubMed Central PMCID: PMCPMC2725978.

70. Soto CY, Menendez MC, Perez E, Samper S, Gomez AB, Garcia MJ, et al. IS6110 mediates increased transcription of the phoP virulence gene in a multidrug-resistant clinical isolate responsible for tuberculosis outbreaks. J Clin Microbiol. 2004; 42(1):212–9. Epub 2004/01/13. https://doi.org/10.1128/JCM.42.1.212-219.2004 PMID: 14715755; PubMed Central PMCID: PMC321672.

71. Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, et al. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. Proc Natl

Acad Sci U S A. 2014; 111(31):11491–6. Epub 2014/07/23. https://doi.org/10.1073/pnas.1406693111 [pii]. PMID: 25049399; PubMed Central PMCID: PMC4128152.

72. Ates LS, Dippenaar A, Ummels R, Piersma SR, van der Woude AD, van der Kuij K, et al. Mutations in ppe38 block PE_PGRS secretion and increase virulence of Mycobacterium tuberculosis. Nat Microbiol. 2018; 3(2):181–8. https://doi.org/10.1038/s41564-017-0090-6 PMID: 29335553.

73. Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, et al. Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat Genet. 2013; 45(7):784–90. https://doi.org/10.1038/ng.2656 PMID: 23749189; PubMed Central PMCID: PMCPMC3777616.

Appendix 2

Comparative metabolomics between *Mycobacterium tuberculosis* and the MTBVAC vaccine candidate.

ACS Infectious Diseases

Article

Cite This: ACS Infect. Dis. XXXX, XXX, XXX−XXX

pubs.acs.org/journal/aidcbc

Comparative Metabolomics between *Mycobacterium tuberculosis* and the MTBVAC Vaccine Candidate

Caridad Díaz,[†] José Pérez del Palacio,[†] Pedro Luis Valero-Guillén,[‡] Patricia Mena García,[†] Irene Pérez,[§,∥] Francisca Vicente,[†] Carlos Martín,[§,∥,⊥] Olga Genilloud,[†] Antonio Sánchez Pozo,[*,#] and Jesús Gonzalo-Asensio[*,§,∥,▽]

[†]Fundación MEDINA, Parque Tecnológico de Ciencias de la Salud, Avenida del Conocimiento 34, 18016 Granada, Spain

[‡]Departamento de Genética y Microbiología, Facultad de Medicina, Universidad de Murcia, Instituto Murciano de Investigación Biosanitaria (IMIB), Campus de Espinardo, 30100 Murcia, Spain

[§]Grupo de Genética de Micobacterias, Departamento de Microbiología y Medicina Preventiva, Facultad de Medicina, Universidad de Zaragoza, IIS Aragón, C/Domingo Miral s/n, 50019 Zaragoza, Spain

[∥]CIBER Enfermedades Respiratorias, Instituto de Salud Carlos III, 28029 Madrid, Spain

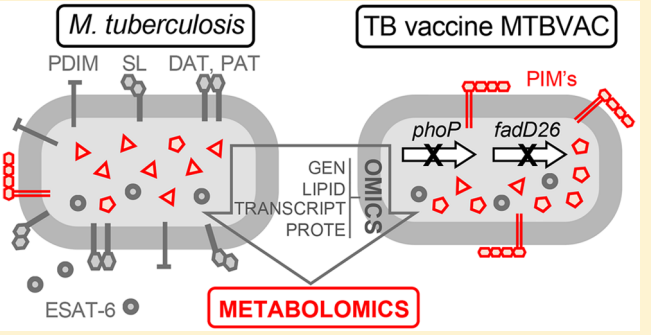
[⊥]Servicio de Microbiología, Hospital Universitario Miguel Servet, Paseo Isabel la Católica 1-3, 50009 Zaragoza, Spain

[#]Departamento de Bioquímica y Biología Molecular II, Facultad de Farmacia, Universidad de Granada, Campus Universitario de Cartuja, 18071 Granada, Spain

[▽]Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), C/Mariano Esquillor, Edificio I + D, Campus Río Ebro, 50018 Zaragoza, Spain

Supporting Information

ABSTRACT: MTBVAC is a live attenuated *M. tuberculosis* vaccine constructed by genetic deletions in the *phoP* and *fadD26* virulence genes. The MTBVAC vaccine is currently in phase 2 clinical trials with newborns and adults in South Africa, one of the countries with the highest incidence. Although MTBVAC has been extensively characterized by genomics, transcriptomics, lipidomics, and proteomics, its metabolomic profile is yet unknown. Accordingly, in this study we aim to identify differential metabolites between *M. tuberculosis* and MTBVAC. To this end, an untargeted metabolomics approach based on liquid chromatography coupled to high-resolution mass spectrometry was implemented in order to explore the main metabolic differences between *M. tuberculosis* and MTBVAC. As an outcome, we identified a set of 34 metabolites involved in diverse bacterial biosynthetic pathways. A consistent increase in the phosphatidylinositol species was observed in the vaccine candidate relative to its parental strain. This phenotype resulted in an increased production of phosphatidylinositol mannosides, a novel PhoP-regulated phenotype in the most widespread lineages of *M. tuberculosis*. This study represents a step ahead in our understanding of the MTBVAC vaccine, and some of the differential metabolites identified in this work might be used as potential vaccination biomarkers.



KEYWORDS: *live attenuated vaccine, phoP, virulence, phosphatidylinositol mannosides, vaccination biomarkers*

A live attenuated derivative of *M. tuberculosis* -MTBVAC- was constructed by the rational attenuation of a clinical strain belonging to the most widespread lineage (Europe−America−Africa; lineage 4) of the tubercle bacillus. Notably, MTBVAC is the first live attenuated *M. tuberculosis* vaccine satisfying the requirements of its clinical evaluation in humans.[1] The first in-human phase 1 clinical trials began in 2013 in adults in Switzerland (clinical trial registration number NCT02013245) and continued in 2016 in South African newborns (NCT02729571). The successful results allowed advancement into two phase 2 trials in 2019 in South Africa, both in newborn (NCT03536117) and in adult (NCT02933281) populations. In accordance with the Second Geneva Consensus document,[2] MTBVAC contains two unmarked genetic deletions in virulence genes *phoP* and *fadD26*. As a consequence, MTBVAC is devoid of virulence phenotypes described in the next paragraph but maintains the whole antigenic repertoire of *M. tuberculosis*.[3]

PhoP is a transcriptional regulator that controls at least 2% of the genome of *M. tuberculosis*.[4−6] Among those genes under PhoP regulation, we found some within the ESX-1 region

Received: January 10, 2019
Published: May 17, 2019

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

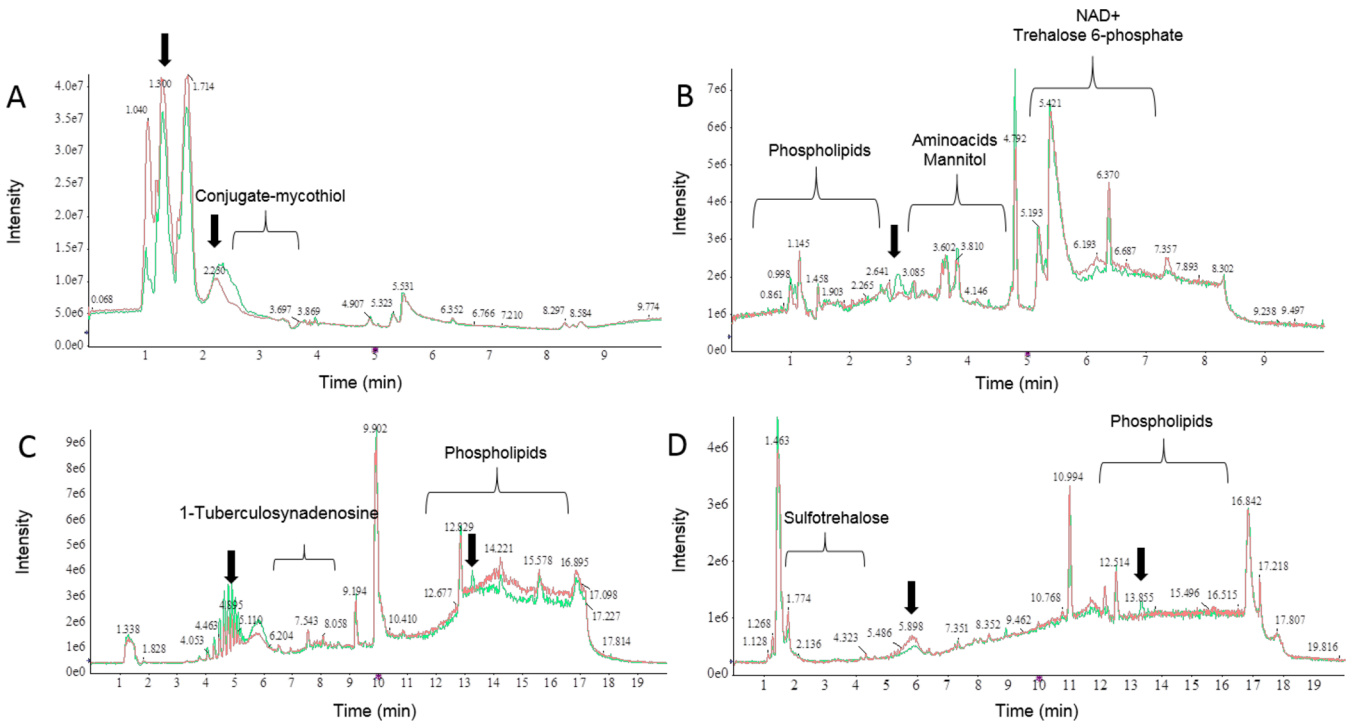


Figure 1. Total ion chromatograms showing the metabolite content of the *M. tuberculosis* MT103 parental strain (red lines) and its MTBVAC vaccine derivative (green lines). (A) HILIC ESI(+). (B) HILIC ESI(−). (C) Reverse-phase ESI(+). (D) Reverse-phase ESI(−). Arrows indicate those metabolites showing the most significant differences between *M. tuberculosis* and MTBVAC, and some differential compounds between both strains identified in this study are indicated above the chromatograms.

involved in the secretion of the main antigen and virulence factor ESAT-6. Consequently, although *M. tuberculosis phoP* mutants produce ESAT-6, they are unable to export this virulence factor.[6,7] ESAT-6 has been described to inhibit autophagy,[8] and it is also involved in cell-to-cell spread by promoting the apoptosis of infected host cells.[9] In addition, *phoP* mutants exhibit a smaller bacillary size and diminished cording formation compared to the wild type and also present an altered colony morphology.[10] These phenotypes, posibly related to the cell wall content, led to the study of the lipid composition in various *phoP* and *phoPR* knockouts, demonstrating that *phoP* controls the synthesis of polyketide-derived lipids, namely, sulfolipids (SL) and di- and polyacyltrehaloses (DAT and PAT).[11,12] On the other side, gene *fadD26* is involved in the biosynthesis of phticeroldimycocerosates (PDIMs), a major virulence lipid in the *M. tuberculosis* cell wall.[13] PDIMs are essential to the survival of *M. tuberculosis* in the host; accordingly, PDIM-deficient mutants are severely attenuated. A lack of PDIM renders *M. tuberculosis* more susceptible to attack from the early innate immune response[14] and also less able to arrest phagosomal maturation, leading to accumulation in acidified phagosomes and consequent bacterial killing.[15] Recently, the concerted role of PDIM and ESAT-6 secretion in promoting cell apoptosis and phagosomal rupture has been shown.[16]

Metabolomics is a unique top-down approach for studying complex systems.[17] Metabolomics information is representative of the cellular status and therefore can be used to link the phenotype of a cell to genetic alterations or environmental changes. Metabolomics complements traditional -omics techniques used to identify molecules that cannot be directly inferred by transcriptomics, proteomics, and genomics.[18] In this context, the integration of transcriptomics and metabolomics recently allowed us to elucidate the metabolic network of *M. tuberculosis* during infection.[19]

Previous metabolomics studies on *M. tuberculosis* were mainly focused on detection from culture and species/strain identification by the analysis of intracellular metabolites.[20] However, little is known about the differential metabolites between *M. tuberculosis* strains or the potential of *M. tuberculosis*-specific metabolites as potential biomarkers of infection.[21]

Our objective was to explore differences in intracellular metabolites between vaccine candidate MTBVAC and its parental strain of *M. tuberculosis*. We hypothesize that some differential metabolites might be specific to *M. tuberculosis* and may be related to the virulence of this pathogen; consequently, they could be used as biomarkers of tuberculosis in bodily fluids. Conversely, MTBVAC-specific metabolites could be used to differentiate vaccinated from infected individuals in clinical trials. For this purpose, four different modes of liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) analysis were implemented for a nontargeted metabolomics approach. In this regard, we used a combination of reverse-phase liquid chromatography (RPLC) and hydrophilic interaction liquid chromatography (HILIC). Intracellular metabolites are complex mixtures containing molecules over a wide polarity range.[22] Traditional techniques such as RPLC enable the separation of nonpolar metabolites. Alternatively, HILIC, because of its multimodal separation ability, yields an improvement in the selectivity and reproducibility of more polar compounds[23]

The choice of ionization mode in LC-HRMS analyses is of utmost importance in metabolomics because of the differentiation in the physicochemical properties of the molecules. In this work, we used electrospray ionization (ESI) in both positive

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

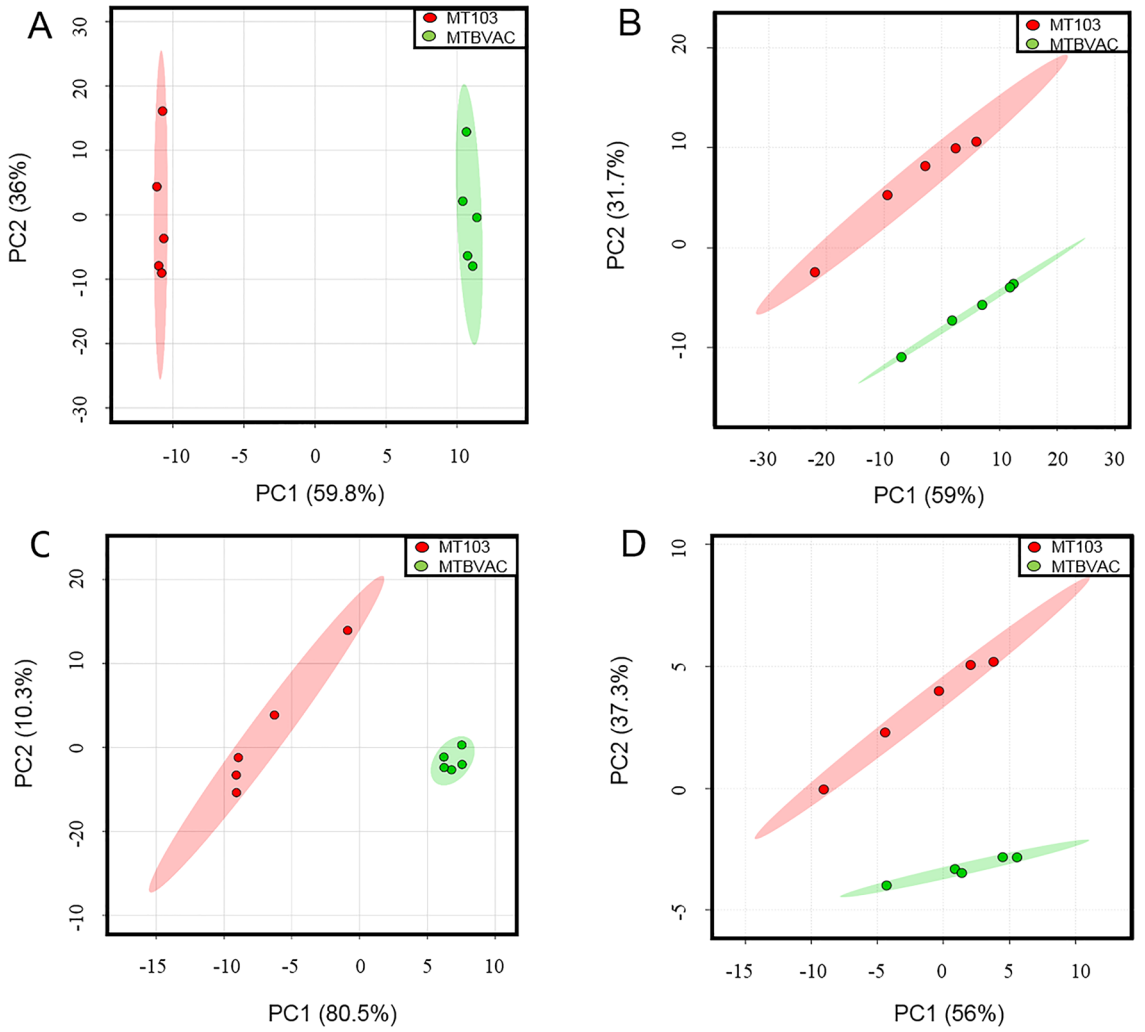


Figure 2. (A) PCA score plot generated using MetaboAnalyst for reverse-phase ESI (−) analysis. (B) PCA score plot for reverse-phase ESI (+) analysis. (C) PCA score plot for HILIC ESI (−). (D) PCA score plot for reverse-phase ESI (−). Green dots and clouds correspond to the MTBVAC sample, and red dots and clouds indicate the *M. tuberculosis* MT103 strain sample.

(+) and negative (−) ion modes as applied to the previously mentioned chromatographic approaches.

For data analysis, we used a combination of multi- and univariate statistical analysis: principal component analysis (PCA) is appropriate for detecting clusters and outliers,[18] and the Student's *t* test is used to select differential metabolites. This approach enabled us to observe significant differences in intracellular metabolites of *M. tuberculosis* and MTBVAC.

RESULTS AND DISCUSSION

Metabolomics Profiling of Virulent and Vaccine Strains of *M. tuberculosis*. After the extraction of intracellular metabolites, we compared the metabolite profiles from *M. tuberculosis* and MTBVAC strains for the four analyses, RPLC ESI(−), RPLC ESI(+), HILIC ESI(−), and HILIC ESI(+). The total ion chromatogram (TIC) representative of each HPLC/HRMS analysis is shown in Figure 1. As a result of the data extraction and alignment process, 4.872, 1.496, 1.781, and 1.523 *m*/*z* were obtained in HILIC ESI(+), ESI(−), reverse-phase ESI(+), and ESI(−), respectively. To diminish the dimensionality of the data matrix and lessen redundant signals, only peaks indicating monoisotopic ions were selected. In addition, there were signals that were reproducible in the five biological replicates and signals that did not come from the solvent. After data filtration, we reduced the number of metabolites to 271 for HILIC ESI(+), 142 for HILIC ESI(−), 214 for RPLC ESI(+), and 267 for RPLC ESI(−). The application of PCA to the data matrices explained more than 93% of the total variance using the first two principal components in all cases. The scores plot revealed close clustering indicative of a clear difference between the intracellular metabolites of *M. tuberculosis* and MTBVAC (Figure 2).

After applying the Student's *t* test with FDR correction, we obtained a total of 125 metabolites with $p < 0.05$ and at least a 2-fold change for the four LCMS analyses. A total of 26 metabolites for HILIC ESI(+), 26 for HILIC ESI(−), 48 for reverse-phase ESI(−), and 25 for reverse-phase ESI(+) were obtained. Metabolites present in more than one analysis were identified by comparing the mass spectra, fragment mass, and retention time, and only one of them was the selected differential in the study. Also, different ion adducts of the same metabolite were identified using mass spectra, the fragment mass, and the retention time, and only one of these forms was the selected differential. A total of 34 of the 125 metabolites were tentatively identified, 25 were increased in MTBVAC, and 9 metabolites were higher in *M. tuberculosis*. For the rest of the metabolites, it

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

Table 1. Differential Intracellular Metabolites of the *M. tuberculosis* MT103 Strain and MTBVAC[a]

tentative structure	*m/z*	adduct	chromatography	retention time	mass error (ppm)	*p* value (FDR)	fold change	Shymanski et al. classification
glutamine	145.0624	−H	HILIC	4.13	3.0	1.60×10^{-7}	2.86	level 2
glutamate	146.0466	−H	HILIC	3.83	2.0	7.89×10^{-8}	2.01	level 2
mannitol	181.0718	−H	HILIC	3.78	0.2	7.89×10^{-8}	12.34	level 1
2-sulfotrehalose	421.0654	−H	reverse	1.52	0.1	7.82×10^{-6}	2.08	level 2
trehalose 6-phosphate	421.0747	−H	HILIC	6.10	−1.3	1.60×10^{-7}	−2.63	level 2
conjugate-mycothiol	487.1588	+H	HILIC	2.61	−0.9	7.74×10^{-3}	>52.49	level 3
1-tuberculosynadenosine	540.3536	+H	reverse	8.56	−0.8	4.40×10^{-2}	−2.86	level 2
NAD	662.0997	−H	HILIC	5.85	−0.5	4.57×10^{-7}	2.05	level 2

[a]Positive fold change values indicate that more metabolites were produced in the *M. tuberculosis* MT103 strain, and negative values indicate that more metabolites were produced in MTBVAC.

Table 2. Differential Intracellular Lipids of the *M. tuberculosis* MT103 Strain and MTBVAC[a]

tentative structure	*m/z*	adduct	chromatography	retention time	mass error (ppm)	*p* value (FDR)	fold change	Shymanski et al. classification
PPA(18:1(9Z))/18:1((9Z))	413.2247	+2 Na	reverse	5.50	1.0	8.00×10^{-3}	12.8	level 2
PI(16:0/0:0)	571.2871	−H	HILIC	3.27	−2.3	2.73×10^{-7}	−2.00	level 2
PE(16:0/16:0)	690.5047	−H	HILIC	1.58	−1.6	4.96×10^{-3}	−2.94	level 2
PE-NMe(16:0/16:0)	704.5236	−H	HILIC	1.58	−1.3	3.13×10^{-3}	−4.55	level 2
PE(16:0/18:1 ((9Z))	716.5228	−H	HILIC	1.55	−1.1	4.14×10^{-3}	−3.57	level 2
PE(16:0/18:0)	718.5360	−H	HILIC	1.59	−2.1	8.35×10^{-3}	−3.03	level 2
PE(R1CO2H+R2CO2H = 35:1)	730.5395	−H	HILIC	1.55	0.1	9.08×10^{-3}	−2.67	level 3
PE(R1CO2H+R2CO2H = 36:1)	744.5553	−H	HILIC	1.54	−0.4	1.82×10^{-2}	−2.47	level 3
PE(R1CO2H+R2CO2H = 36:0)	746.5670	−H	HILIC	1.61	−3.1	1.22×10^{-2}	−2.94	level 3
PG(16:0/18:1(9Z))	747.5159	−H	HILIC	1.70	−3.0	1.16×10^{-5}	3.06	level 2
PG(R1CO2H+R2CO2H = 38:0)	805.5985	−H	reverse	14.47	1.2	1.74×10^{-5}	−24.39	level 3
PG(R1CO2H+R2CO2H = 37:0)	823.6067	+FA−H	reverse	12.15	0.4	3.02×10^{-5}	−58.80	level 3
PI(16:0/18:1 ((9Z))	835.5325	−H	HILIC	2.74	−2.0	2.89×10^{-4}	−3.33	level 2
PI(16:0/18:0)	837.5477	−H	HILIC	2.79	−2.6	1.19×10^{-3}	−3.23	level 2
PC(R1CO2H+R2CO2H = 44:7)	846.6026	+H	reverse	12.07	−1.0	5.33×10^{-5}	−16.67	level 3
PI(R1CO2H+R2CO2H = 37:0)	847.6066	+H−H_2O	reverse	13.24	0.0	8.00×10^{-3}	−11.63	level 3
PI(R1CO2H+R2CO2H = 38:0)	861.6226	+H−H_2O	reverse	13.85	1.2	8.00×10^{-3}	−10.00	level 3
PI(R1CO2H+R2CO2H = 40:5)	863.5814	+H−2H_2O	reverse	13.27	0.6	1.20×10^{-2}	−8.07	level 3
PI(R1CO2H+R2CO2H = 36:0)	865.5763	−H	HILIC	2.88	−1.1	4.61×10^{-3}	−3.57	level 3
PI(R1CO2H+R2CO2H = 36:4)	903.5199	+FA−H	HILIC	2.74	4.0	7.50×10^{-4}	−4.34	level 3
PI(R1CO2H+R2CO2H = 36:3)	905.5352	+FA−H	HILIC	2.78	4.0	1.27×10^{-3}	−3.70	level 3
Mbt+Fe (*R* = 20:1)	921.4509	−H	reverse	6.05	−2.6	1.38×10^{-7}	11.30	level 2
PI(R1CO2H+R2CO2H = 42:3)	925.6127	−H	reverse	11.01	5.0	1.94×10^{-6}	−10.10	level 3
PI(R1CO2H+R2CO2H = 44:2)	927.6259	−H_2O−H	reverse	13.34	−3.4	1.38×10^{-7}	−12.98	level 3
PI(R1CO2H+R2CO2H = 40:2)	937.5990	+FA−H	reverse	13.34	−3.7	1.11×10^{-7}	−18.52	level 3
PIP(R1CO2H+R2CO2H = 38:3)	1013.5347	+FA−H	reverse	5.66	0.6	1.21×10^{-5}	−2.95	level 3

[a]Positive fold change values indicate lipids of higher abundance in *M. tuberculosis* MT103, and negative values indicate lipids of higher abundance in MTBVAC.

was not possible to assign a molecular identity. This fact suggests that some of them might be novel compounds exclusive to *M. tuberculosis*. This is not unexpected because metabolomics applied to the study of bacteria is an emerging field, and in particular, the number of known metabolites from *M. tuberculosis* is very limited.[21]

Identification of Differential Metabolites between Virulent and Vaccine Strains of *M. tuberculosis*. The metabolites were identified at different confidences from the confirmed structure to a putative structure.[24,25] The structure of mannitol was confirmed using the standard compound, which was acquired and analyzed to check the compound identity of *m/z* 181.0718 with MS, MS/MS, and the retention time matching (Figure 1, Supporting Information). Sixteen of them were identified with a probable structure[25] {glutamine, glutamate, nicotinamide adenine dinucleotide (NAD), 2-sulfotrehalose, trehalose 6-phosphate, mycobactin (Mbt) + Fe (*R* = 20:1), 1-tuberculosynadenosine, 1-hexadecanoyl-*sn*-glycero-3-phospho-(1′-myo-inositol) PI(16:0/0:0), 1,2-dihexadecanoyl-*sn*-glycero-3-phospho-*N*-ethanolamine PE(16:0/16:0), 1,2-dihexadecanoyl-*sn*-glycero-3-phospho-*N*-methylethanolamine PNMe(16:0/16:0), 1-hexadecanoyl-2-(9Z-octadecenoyl)-*sn*-glycero-3-phosphoethanolamine PE(16:0/18:1), 1-hexadecanoyl-2-octadecenoyl-*sn*-glycero-3-phosphoethanolamine PE(16:0/18:0), 1-hexadecanoyl-2-(9Z-octavdecenoyl)-*sn*-glycero-3-phosphoglycerol PG(16:0/18:1), 1-hexadecanoyl-2-(9Z-octadecenoyl)-*sn*-glycero-3-phospho-(1′-myo-inositol) PI(16:0/18:0), 1-hexadecanoyl-2-octadecenoyl-*sn*-glycero-3-phospho-(1′-myo-inositol) PI(16:0/18:0), and 1,2-di(9Z-octadecenoyl)-*sn*-glycero-3-pyrophosphate (PPA(18:1/18:1)} using difference evidence matching literature[21] or library spectra, mainly NIST. Other phospholipids were putatively

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

identified as phosphatidylinositol (PI) {PI(R1CO2H +R2CO2H = 37:0), PI(R1CO2H+R2CO2H = 38:0), PI-(R1CO2H+R2CO2H = 40:5), PI(R1CO2H+R2CO2H = 36:0), PI(R1CO2H+R2CO2H = 36:4), PI(R1CO2H +R2CO2H = 36:3), PI(R1CO2H+R2CO2H = 42:3), PI-(R1CO2H+R2CO2H = 44:2), and PI(R1CO2H+R2CO2H = 40:2)}, phosphatidylethanolamine (PE) {PE(R1CO2H +R2CO2H = 35:1), PE(R1CO2H+R2CO2H = 36:0), and PE(R1CO2H+R2CO2H = 36:0)}, and phosphatidylglycerol (PG) {PG(R1CO2H+R2CO2H = 38:0) and PG(R1CO2H +R2CO2H = 37:0)} using the match with known compounds in different databases, the in-silico fragmentation of metabolites, and/or retention behavior. Also, a metabolite exclusively detected in *M. tuberculosis* samples was putatively identified as a conjugate of mycothiol using a formula finder for assigning molecular formulas comparing mass and mass/mass spectra. These differential metabolites are shown in Tables 1 and 2. The rest of metabolites could not be identified because they did not match the known compounds (Supporting Information Tables 1−4).

Targeted Metabolomics Analysis between *M. tuberculosis* and *phoP* Mutants. The consistent increase in PI species produced by the MTBVAC vaccine relative to *M. tuberculosis* led us to perform a targeted study of PI derived glycolipids as phosphatidyl-inositol mannosides (PIMs). Furthermore, to generalize this phenotype beyond the MTBVAC vaccine, we included in our study *M. tuberculosis phoP* mutants constructed in the H37Rv reference strain and the GC1237 clinical isolate belonging to lineages 4 and 2, respectively. PIMs are heterogeneous in structure and contain up to four fatty acids and six mannoses. The predominant variants of PIM are triacylated species termed AcPIM2 and AcPIM6, which contain two and six mannose residues, respectively.[26] PIM2 constitutes the anchor motif of a key component of the mycobacteria cell wall, named lipoarabinomannans (LAM), and both LAM and PIM are agonists of the Toll-like receptor involved in innate immunity.[27] The major monoisotopic ions detected and quantified for the different structural phospholipids of the mycobacteria studied are given in Supporting Information Table 5, and the ratio of PIM species in the mutants relative to the parent strains is shown in Table 3. It is noteworthy that the production of PIMs was higher in H37Rv Δ*phoP*, GC1237 Δ*phoP*, and MTBVAC than in their parent strains (Figure 3 and Table 3). Accordingly, the synthesis of PIM derivatives represents a novel PhoP-regulated phenotype generalizable to the most widespread lineages of *M. tuberculosis*.

Table 3. Comparative Abundance of PIMs in *M. tuberculosis phoP* Mutants Relative to Their Wild Type Strains

	PIM abundance		
strains compared	PIM2	monoacyl-PIM2	diacyl-PIM2
MTBVAC versus MT103	1.1	1.8	1.4
GC1237 Δ*PhoP* versus GC1237 wt	3.9	1.7	2.4
H37Rv Δ*PhoP* versus H37Rv wt	8.9	1.1	3.2

Biological Interpretation. Genomics, lipidomics, transcriptomics, and proteomics have been applied to understand the implications of *phoP* and *fadD26* mutations in the MTBVAC vaccine,[1,3,5,6,11,12,28] leading to a deep understanding of the mechanisms underlying its protective efficacy. This study represents a step ahead in MTBVAC characterization, and to our knowledge, this is the first metabolomics approach to compare a live attenuated vaccine with its parental strain. Because MTBVAC was derived from an *M. tuberculosis* clinical isolate by introducing well-known genetic modifications, it is expected that differential metabolites exist between both strains that could be otherwise overlooked or that would require complex techniques for their discovery.

The ability of *M. tuberculosis* to infect host cells results from multiple phenotypes, some of them related to the mycobacterial cell envelope.[29] The outer membrane is mainly composed of mycolates and a set of noncovalently bound lipids that include PDIM, acyltrehaloses, trehalose mycolates, sulfatides, and phospatidyl-myo-inositol mannosides, and it is surrounded by a capsule.[30,31] MTBVAC carries *phoP* and *fad2D6* gene deletions, with both genes being involved in lipid metabolism. In this study, we have identified metabolites involved in the synthesis of the cell envelope such as phospholipids, and in particular, we have found changes in PI, PG, and PE. Other metabolites also implicated in the synthesis of the cell envelope such as trehalose 6-phosphate and 2-sulfotrehalose were differential between strains.

PI plays an important structural, physiological, and immunological role in *M. tuberculosis*. For example, the mycobacteria cell wall carries complex lipoglycans such as lipoarabinomannan and lipomannan, anchored to the membrane by a common PI anchor that constitutes their first building block. In *M. tuberculosis*, these lipids act as modulators of the host immune response and virulence factors.[32] We demonstrate a differential composition in PIMs between MTBVAC and *M. tuberculosis*. Notably, this phenotype is not exclusive to the vaccine, and our results demonstrate that PIMs are more abundant in *M. tuberculosis phoP* mutants of lineages 2 and 4, which are the most widespread lineages of this pathogen. Accordingly, the synthesis of PIMs represents an as-yet unknown PhoP-regulated phenotype in *Mycobacterium*. This increased concentration of free PI and PIMs might be related to the attenuation of *phoP* mutants[1,10,12] as a result of a reduction in the number of anchored lipids such as lipoglycans.

In mycobacteria, PG is an intermediate in the biosynthesis of PI, cardiolipins, and lisinylated PG.[33] We have found higher values of PG(16:0/18:1) in *M. tuberculosis*, while other lipoforms were identified with an increased signal in MTBVAC. On the one hand, these differences might support the immunity of the vaccine and its decreased synthesis of cardiolipins and lisinylated PG, and on the other hand, PGs appear to be a molecular signature that allows immune recognition of bacterial infection.[34,35]

Differences in the detected concentrations of trehalose phosphate suggest a difference in the production between *M. tuberculosis* and MTBVAC. Nevertheless, we have not identified significant differences for trehalose in our analysis. The OtsA-OtsB pathway, wherein OtsA catalyzes the condensation of UDP-glucose and glucose-6-phosphate to form trehalose-6-phosphate and OtsB is a dephosphorylase releasing free trehalose from trehalose-6-phosphate, represents a key pathway required for *M. tuberculosis* growth in laboratory culture and for virulence in mice.[36] We suggest that a higher concentration of trehalose 6-phosphate in MTBVAC can be linked to the altered activity of OtsA-OtsB. Also, this sugar serves as a metabolic building block for some glycolipids, namely, trehalose monomycolate (TMM) and trehalose dimycolate (TDM), also known as the cord factor. TDM prevents *M. tuberculosis* killing by macrophages, acts as a potent modulator of the activation of macrophages, and increases the antibiotic

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

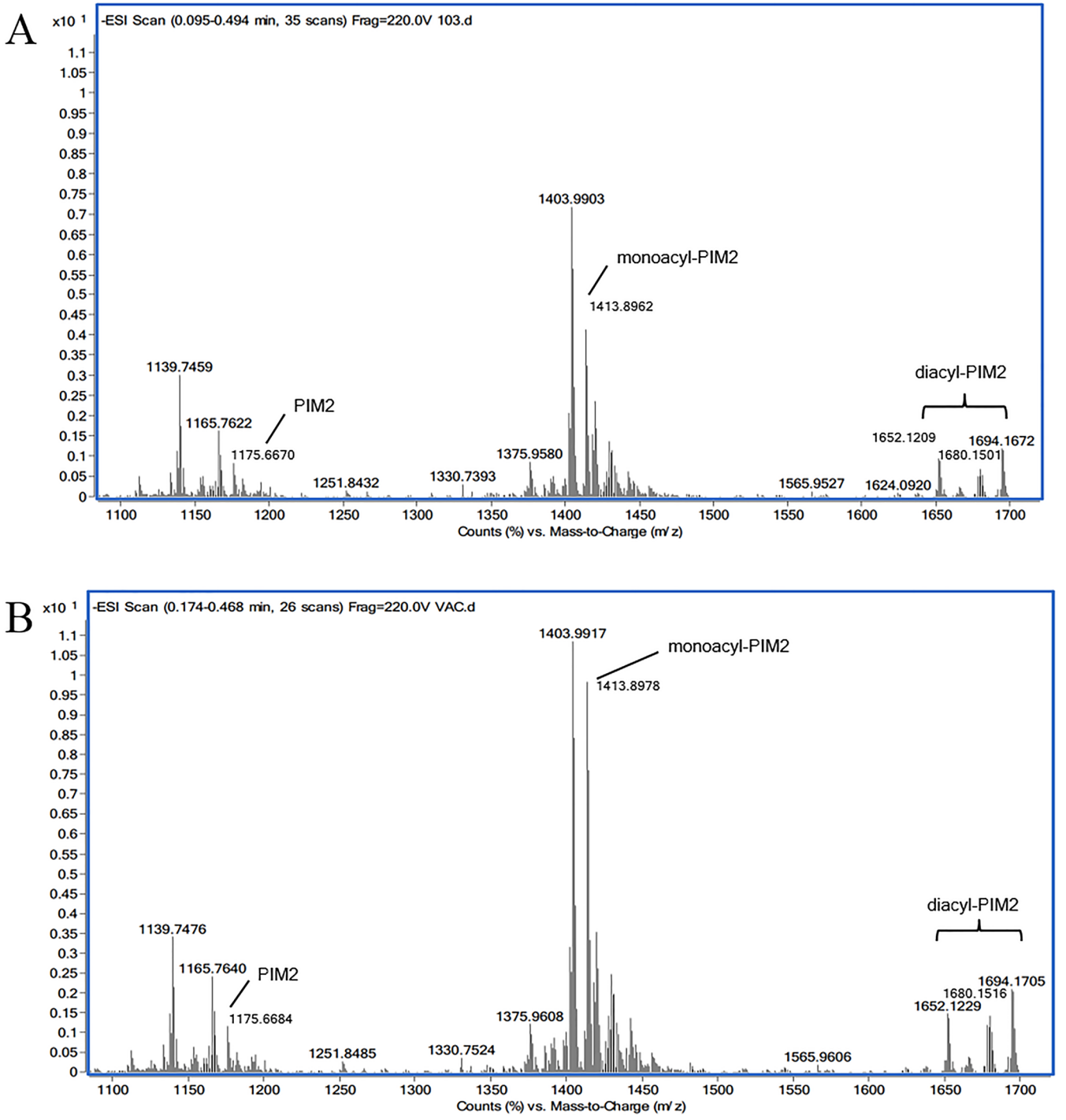


Figure 3. ESI−TOF-MS analysis of PIM2 in (A) *Mycobacterium tuberculosis* MT103 and (B) MTBVAC. The ratios of PIMs between these strains are calculated in Table 3, and the identification of the different species is detailed in Supporting Information Table 5. Ions used to detect each PIM subclass are as follows: for PIMan2 (phosphatidylinositol dimannoside), the ion considered was m/z 1175.6694 (TBS + 16:0); for monoacyl-PIMan2, m/z 1413.8999 (TBS + 16:0 + 16:0); and for diacyl-PIMan2, m/z 1652.1295 (TBS + 16:0 + 16:0 + 16:0), m/z 1680.1580 (TBS + 16:0 + 16:0 + 18:0), and m/z 1694.1798 (TBS + 16:0 + 16:0 + TBS).[27,48] TBS, tuberculostearic acid (10-methyl octadecanoic acid); 16:0, hexadecanoic acid; and 18:0, octadenoic acid.

resistance in mycobacteria.[37] A reduction in trehalose phosphate synthesis could also be related to the decreased production of these glycolipids and consequently to the absence of cord formation in *M. tuberculosis phoP* mutants.[11]

SL constitutes a family of polyacylated trehalose-2-sulfate glycolipids esterified with two to four acyl chains. PhoP is required for the production of SL,[4,11,12] and a Ser219Leu substitution in the *phoP* gene of the reference H37Ra strain is responsible for the absence of SL, DAT, and PAT in this strain.[28] Decreased production of trehalose 2-sulfate in MTBVAC in our metabolomics study contributes to the explanation of the lack of SL in the attenuated strain as a consequence of the *phoP* mutation.

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

Mycothiol (MSH) is the functional equivalent of glutathione (GSH) in mycobacteria[38] and protects *M. tuberculosis* from external factors such as antibiotics and oxidative stress.[39] We have tentatively identified a component conjugated with MSH exclusively in *M. tuberculosis* (Table 1). This result might indicate that the antioxidant capacity is diminished in MTBVAC, and we hypothesize that some of the MSH biosynthetic enzymes, and probably the MSH-S transferase, could have decreased activity in MTBVAC, lowering the production of the conjugate mycothiol.

We have found a reduction in glutamine and glutamate in MTBVAC. These amino acids represent nitrogen donors for the synthesis of cell metabolites.[40] They are also the two main metabolites of a metabolic feedback loop that integrates information from carbon and nitrogen metabolism in the control of the carbon−nitrogen ratio in the cell.[41] In mycobacteria, the assimilation of inorganic nitrogen and its conversion to glutamine and glutamate is carried out by glutamine synthetase (GS) and glutamate synthetase.[42] GS has traditionally been related to *M. tuberculosis* virulence and pathogenicity. Virulent mycobacteria secrete the GS enzyme, and this extracellular GS might play a role in the production of poly-L-glutamine-glutamate (PLG), a polymer found only in pathogenic mycobacterial cell walls and/or the extracellular GS activity may modulate phagosomal pH and thereby avoid phagosome−lysosome fusion.[43] We hypothesize that an elevated concentration of these amino acids could be related to a higher activity of GS and therefore to increased virulence. Also, we have found higher levels of NAD in *M. tuberculosis*. NAD acts widely in glutamate catabolism to form α-ketoglutarate through the action of nicotinamide adenine dinucleotide (NAD)-dependent glutamate dehydrogenase and can be reduced by GDH to form glutamate. NAD and glutamine/glutamate showed a clear increase in *M. tuberculosis* compared to MTBVAC, which might explain the attenuated phenotype of the vaccine in the context of energetically disfavored metabolism.

Altogether, these results expand our knowledge of the MTBVAC vaccine, particularly in the context of the lipid metabolism, and lay foundations to identify metabolites exclusively produced by virulent *M. tuberculosis*. Additional experiments using different media, growth conditions, or even *in vivo* models will help to identify either potential biomarkers of tuberculosis infection or alternatively MTBVAC-specific biomarkers, which will be enormously useful in future MTBVAC efficacy trials.

METHODS

Cultivation of Strains and Whole-Cell Metabolite Extraction. *M. tuberculosis* MT103 clinical isolate and MTBVAC were cultured at 37 °C in 7H9-ADC medium. Five independent biological replicates were carried out per strain. Once the cultures reached the appropriate growth (1 unit of optical density at 600 nm), they were centrifuged at 4000*g* for 10 min at 4 °C. Bacterial pellets were washed with 10 mL of ice-cold Tris-HCl, pH 7.4 and centrifuged at 4000*g* for 10 min at 4 °C. This step was repeated three times. Methanol (1 mL) at −80 °C was added to each sample to resuspend the bacterial pellet. Bacteria were lysed in the fast prep for 45 s, lysates were centrifuged at 12 000*g* for 5 min at 4 °C, and the supernatant was transferred to a tube kept on dry ice. This step was repeated two times. Methanol at −80 °C (7 mL) was added to the lysed bacteria. To further extract metabolites in the pellet, 1 mL of 50% methanol/water (v/v) at −20 °C was added to the tubes containin a bacterial pellet, and bacteria were lysed again in the fast prep. Lysates were centrifuged at 12 000*g* for 5 min at 4 °C. The total supernatant was filtered through 0.22 μm syringe filters. Quality control (QC) was conducted by pooling equal volumes of all extracted samples.

Liquid Chromatography−High Resolution Mass Spectrometric (LC−HRMS). Supernatants were transferred to the analytical vials, stored in the autosampler at 4 °C, and analyzed with LC−HRMS. For liquid chromatography, the separation was performed with an Agilent series 1290 LC system (Agilent Technologies, Santa Clara, CA, USA) in HILIC mode using a Waters XBridge BEH amide column (2.1 mm × 150 mm, 2.5 μm) (Waters Corporation, Milford, MA, U.S.) with an XBridge BEH amide guard column (2.1 mm × 5 mm, 2.5 μm). The injected sample volume was 3 μL. The mobile phase (MP) consisted of 0.1% formic acid−[water:AcN] [90:10] (MP A) and 0.1% formic acid−[AcN:water] [90:10] (MP B). The gradient program was applied as follows: t = 0−0.1 min 99% MP B; t = 7 min 30% MP B, t = 7.10 min 99% MP B. The stop time was 10 min. The flow rate was 0.4 mL/min.

The column temperature was maintained at 25 °C in ESI(+) and at 45 °C in ESI(−). For RPLC, the chromatographic separation and mass detection were performed using an Atlantis T3 column (C18:2.1 mm × 150 mm, 3 μm) from Waters, kept at 30 °C in ESI(+) and ESI(−). The injected sample volume was 5 μL. The MP was the same as that used for HILIC. The gradient elution was performed as follows: t = 0−0.5 min 1% MP B; t = 11 min 99% MP B; t = 15.50 min 99% MP B; t = 15.60 min 1% MP B. The stop run was 20 min. The flow rate was 0.3 mL/min. Mass detection was performed using an AB SCIEX Triple TOF 5600 quadrupole time-of-flight mass spectrometer (Q-TOF) (SCIEX, Concord, ON, Canada). The exact mass calibration was automatically achieved every 10 and 7 injections in HILIC and RPLC, respectively.

Data Set Creation. First, we evaluated the retention time and m/z variability of three peaks positioned at different points on the chromatogram for each analysis using Peak View software (version 1.1.2, AB SCIEX, Concord, ON). This information was used to correct the alignment of the peaks. All mass spectra data were processed using Marker View software (version 1.2.1, AB SCIEX, Concord, ON) to perform peak detection, alignment, and data filtering. Data mining was accomplished with an automated algorithm in the retention time between 1.20 and 8.70 min in HILIC and between 1.50 and 15.50 min in reverse-phase analysis. Collection parameters were set as follows: mass window 20.0 ppm, retention time window 0.12 min, and minimum intensity 100 counts/s in all cases. Only variables that are present in at least one of the groups were selected for the four analyses. Next, only monoisotopic peaks were considered in order to decrease mass redundancy and improve the selection of the true molecular feature.

Analytical Method Validation. The QC samples were used in order to check the precision and repeatability of our sample analysis. Precision was evaluated by using random injections of the QC during the study sample sequence. Also, QC samples were included to validate the quality of the analytical system's performance[44] using Marker View software. Our study was evaluated by performing PCA for all analyses. PCA is the most common unsupervised method used in metabolomics. It is a common first step in data analysis and entails an algorithm that reduces a set of high-dimensional data to a small number of dimensions that explain as much of the

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

variation in the data as possible.[19] We reviewed the relative standard deviation (RSD) for the peak intensity of all ions after data set creation. Variables with unacceptable reproducibility (RSD > 30%) were rejected from the data matrix. Organic solvent samples were run alongside QC samples to identify impurities in the solvents or extraction procedures and to check carryover contamination from intense analytes.[45] All signals from the mobile phases were excluded from the analysis.

Statistical Analysis. Metaboanalyst 2.0.was used for statistical analysis. First, we normalized the intensity of each ion to the summed total ion intensity of the chromatogram, and then the data was transformed using the Pareto scaling method to transform the data matrix into a more Gaussian-type distribution.[46] Multivariate statistical analysis, PCA was used to detect clusters and possible outliers. For the univariate analysis of metabolites in the culture supernatant, statistical significance was determined using the Student's *t* test with two experimental groups, *M. tuberculosis* and MTBVAC. The Benjamini–Hochberg false discovery rate (FDR) correction for multiple comparisons was performed afterward to reduce the expected proportion of false positives (type I errors). A p value <0.05 was considered to be statistically significant, and differences in the 2-fold change were considered for the selection of differential metabolites.

Identification of Differential Metabolites. Peak View software (version 1.0 with Formula Finder plug-in version 1.0, AB SCIEX, Concord, ON, Canada) was used to determine a molecular formula according to the exact mass and isotope pattern. The elemental formula approximation was completed from both single HRMS and MS/MS accurate mass spectra obtaining tentative formulas, reflecting the differences between calculated and measured m/z values for both parent and fragment ions. We conducted additional MS/MS experiments to identify metabolites in cases that were necessary. The identification of molecular components was achieved through comparative searches of available mass spectral using the CEU mediator as tool for searching metabolites in several databases such as Metlin, the Human Metabolome Data Base, and Lipid Maps in a first analysis. In the next steps, the identification was confirmed by comparing experimental fragmentation mass spectra in a database (NIST 2012 mass spectral library and mass bank). In cases in which it was not possible to assign an identification another database was consulted (Mtb LipidDB, PubChem, ChemSpider, Metacyc encyclopedia of metabolic pathways, the *Escherichia coli* Metabolome, and the Biocyc database collection) and/or a literature search was performed. Also, we used parent compound information and the experimental conditions, ionization behavior, and/or retention time under the different chromatographic conditions. Metabolite identities were searched for using a mass tolerance of <5 ppm as the mass window.

Targeted Metabolomics of PIMs. Strains MT103, MTBVAC, H37Rv, H37Rv Δ*phoP*::*hyg*, GC1237, and GC1237 Δ*phoP*::*hyg*[11] were grown in 7H9-ADC medium as previously described. The pelleted cells were successively extracted with chloroform–methanol (1:2 v/v, 1:1 v/v, and 2:1 v/v) overnight at room temperature. The extracts were combined, partitioned in chloroform–methanol–water (4:2:1 v/v/v), evaporated to dryness under N_2 (40 °C), adjusted to approximately 0.5 mg/mL, and dissolved in chloroform–methanol (2:1 v/v). Mass spectrometry analyses were performed in an ESI–TOF mass spectrometer 6220 (Agilent) connected to an HPLC 1200 series (Agilent) operated in negative mode (m/z 100–2000) and direct insertion; the mobile phase (0.4 mL/min) consisted of methanol/water (75:25 v/v) containing 0.1% formic acid and 0.5 mM ammonium formate. The spectra were acquired under standard conditions in centroid mode.[47] Phospholipids were identified according to their exact masses (Table 3) (calculated at www.chemcalc.org; error below 5 ppm). Phospholipids data from *Mycobacterium bovis* BCG[27,48] was employed as a reference. The pairs of strains were analyzed once the same day within a period of 10 min, employing the same reference calibration of the system. Ratios of lipids between pairs of strains (*phoP* mutants versus parent strain) were calculated by taking into account the TIC area and the intensities of the different monoisotopic ions of a given phospholipid.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.9b00008.

MS analysis information, MS and MS/MS data, and data tables (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail sanchezp@go.ugr.es.
*E-mail jagonzal@unizar.es.

ORCID

Jesús Gonzalo-Asensio: 0000-0001-8841-6593

Author Contributions

C.D., C.M., J.G.-A., and F.V. designed the study. I.P. and J.G.-A. carried out strain cultures and metabolite extraction. C.D. and P.M.G. performed analysis LC–MS. P.L.V-.G. performed targeted analysis. C.D. performed statistical analysis. C.D. J.P.d.P. completed the identification of metabolites. C.D., A.S.P., and J.G.-A. performed biological interpretation. F.V., A.S.P., and J.G.-A. supervised the work. C.D. wrote the manuscript. C.M., O.G., A.S.P., and J.G.-A. critically reviewed the manuscript. All authors contributed to and approved the final manuscript.

Notes

The authors declare the following competing financial interest(s): C.M. and J.G.A. are co-inventors in patent applications entitled TB vaccine filed by the University of Zaragoza (application number PCT/ES 2007/070051) and Compositions for use as a prophylactic agent to those at risk of infection of TB , or as secondary agents for treating infected TB patients (application number: 218382097.6-1112) filed by University of Zaragoza/Biofabri.

■ ACKNOWLEDGMENTS

This work was supported by the European Commission Horizon 2020 (TBVAC2020, H2020-PHC-643381), the Spanish Ministry of Science and Competitiveness (BFU2015-72190-EXP), and Gobierno de Aragón-Fondo Europeo de Desarrollo Regional (FEDER) 2014-2020: Construyendo Europa Desde Aragon. Part of this work will be included in a doctoral thesis in the University of Granada Pharmacy Program. The MEDINA authors disclosed the receipt of financial support from Fundación MEDINA, a public–private partnership of Merck Sharp & Dohme de España S.A./Universidad de Granada/Junta

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX–XXX

de Andalucía. I.P. was a recipient of a DGA-Fondo Social Europeo grant.

REFERENCES

(1) Arbues, A., Aguilo, J. I., Gonzalo-Asensio, J., Marinova, D., Uranga, S., Puentes, E., Fernandez, C., Parra, A., Cardona, P. J., Vilaplana, C., and Ausina, V. (2013) Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated *M. tuberculosis*-based vaccine to enter clinical trials. *Vaccine 31* (42), 4867−4873.

(2) Walker, K. B., Brennan, M. J., Ho, M. M., Eskola, J., Thiry, G., Sadoff, J., Dobbelaer, R., Grode, L., Liu, M. A., Fruth, U., and Lambert, P. H. (2010) The second Geneva Consensus: Recommendations for novel live TB vaccines. *Vaccine 28* (11), 2259−2270.

(3) Gonzalo-Asensio, J., Marinova, D., Martin, C., and Aguilo, N. (2017) MTBVAC: Attenuating the Human Pathogen of Tuberculosis (TB) Toward a Promising vaccine against the TB epidemic. *Front. Immunol. 8*, 1803.

(4) Walters, S. B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M., and Smith, I. (2006) The *Mycobacterium tuberculosis* PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol. Microbiol. 60* (2), 312−330.

(5) Gonzalo-Asensio, J., Mostowy, S., Harders-Westerveen, J., Huygen, K., Hernández-Pando, R., Thole, J., Behr, M., Gicquel, B., and Martín, C. (2008) PhoP: a missing piece in the intricate puzzle of *Mycobacterium tuberculosis* virulence. *PLoS One 3* (10), No. e3496.

(6) Solans, L., Aguiló, N., Samper, S., Pawlik, A., Frigui, W., Martín, C., Brosch, R., and Gonzalo-Asensio, J. (2014) A specific polymorphism in *Mycobacterium tuberculosis* H37Rv causes differential ESAT-6 expression and identifies WhiB6 as a novel ESX-1 component. *Infect. Immun. 82*, 3446.

(7) Frigui, W., Bottai, D., Majlessi, L., Monot, M., Josselin, E., Brodin, P., Garnier, T., Gicquel, B., Martin, C., Leclerc, C., and Cole, S. T. (2008) Control of *M. tuberculosis* ESAT-6 secretion and specific T cell recognition by PhoP. *PLoS Pathog. 4* (2), No. e33.

(8) Romagnoli, A., Etna, M. P., Giacomini, E., Pardini, M., Remoli, M. E., Corazzari, M., Falasca, L., Goletti, D., Gafa, V., Simeone, R., and Delogu, G. (2012) ESX-1 dependent impairment of autophagic flux by *Mycobacterium tuberculosis* in human dendritic cells. *Autophagy 8* (9), 1357−1370.

(9) Aguilo, J. I., Alonso, H., Uranga, S., Marinova, D., Arbués, A., de Martino, A., Anel, A., Monzon, M., Badiola, J., Pardo, J., Brosch, R., and Martin, C. (2013) ESX-1-induced apoptosis is involved in cell-to-cell spread of *Mycobacterium tuberculosis*. *Cell. Microbiol. 15* (12), 1994−2005.

(10) Pérez, E., Samper, S., Bordas, Y., Guilhot, C., Gicquel, B., and Martín, C. (2001) An essential role for *phoP* in *Mycobacterium tuberculosis* virulence. *Mol. Microbiol. 41* (1), 179−187.

(11) Gonzalo Asensio, J. G., Maia, C., Ferrer, N. L., Barilone, N., Laval, F., Soto, C. Y., Winter, N., Daffé, M., Gicquel, B., Martín, C., and Jackson, M. (2006) The virulence-associated two-component PhoP-PhoR system controls the biosynthesis of polyketide-derived lipids in *Mycobacterium tuberculosis*. *J. Biol. Chem. 281* (3), 1313−1316.

(12) Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C., Moreau, F., Laval, F., Daffé, M., Martin, C., Brosch, R., and Guilhot, C. (2014) Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. *Proc. Natl. Acad. Sci. U. S. A. 111* (31), 11491−11496.

(13) Camacho, L. R., Ensergueix, D., Perez, E., Gicquel, B., and Guilhot, C. (1999) Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol. Microbiol. 34* (2), 257−267.

(14) Day, T. A., Mittler, J. E., Nixon, M. R., Thompson, C., Miner, M. D., Hickey, M. J., Liao, R. P., Pang, J., Shayakhmetov, D., and Sherman, D. R. *Mycobacterium tuberculosis* strains lacking the surface lipid phthiocerol dimycocerosate are susceptible to killing by an early innate host response. *Infect. Immun.* 2014, *82*, DOI: 10.1128/IAI.01340-13.

(15) Passemar, C., Arbués, A., Malaga, W., Mercier, I., Moreau, F., Lepourry, L., Neyrolles, O., Guilhot, C., and Astarie-Dequeker, C. (2014) Multiple deletions in the polyketide synthase gene repertoire of *Mycobacterium tuberculosis* reveal functional overlap of cell envelope lipids in host−pathogen interactions. *Cell. Microbiol. 16* (2), 195−213.

(16) Augenstreich, J., Arbues, A., Simeone, R., Haanappel, E., Wegener, A., Sayes, F., Le Chevalier, F., Chalut, C., Malaga, W., Guilhot, C., and Brosch, R. (2017) ESX-1 and phthiocerol dimycocerosates of *Mycobacterium tuberculosis* act in concert to cause phagosomal rupture and host cell apoptosis. *Cell. Microbiol. 19* (7), No. e12726.

(17) Nicholson, J. K., and Lindon, J. C. (2008) Systems biology: metabonomics. *Nature 455* (7216), 1054.

(18) Kosmides, A. K., Kamisoglu, K., Calvano, S. E., Corbett, S. A., and Androulakis, I. P. (2013) Metabolomic fingerprinting: challenges and opportunities. *Crit. Rev. Biomed. Eng. 41* (3), 205.

(19) Zimmermann, M., Kogadeeva, M., Gengenbacher, M., McEwen, G., Mollenkopf, H. J., Zamboni, N., Kaufmann, S. H. E., and Sauer, U. (2017) Integration of metabolomics and transcriptomics reveals a complex diet of *Mycobacterium tuberculosis* during early macrophage infection. *MSystems 2* (4), e00057−17.

(20) du Preez, I., Luies, L., and Loots, D. T. (2019) The application of metabolomics toward pulmonary tuberculosis research. *Tuberculosis 115*, 126.

(21) Lau, S. K., Lam, C. W., Curreem, S. O., Lee, K. C., Lau, C. C., Chow, W. N., Ngan, A. H., To, K. K., Chan, J. F., Hung, I. F., and Yam, W. C. Identification of specific metabolites in culture supernatant of *Mycobacterium tuberculosis* using metabolomics: exploration of potential biomarkers. *Emerging Microbes Infect.* 2015, *4*, DOI: 10.1038/emi.2015.6.

(22) Creek, D. J., Jankevics, A., Breitling, R., Watson, D. G., Barrett, M. P., and Burgess, K. E. (2011) Toward global metabolomics analysis with hydrophilic interaction liquid chromatography−mass spectrometry: improved metabolite identification by retention time prediction. *Anal. Chem. 83* (22), 8703−8710.

(23) Buszewski, B., and Noga, S. (2012) Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Anal. Bioanal. Chem. 402* (1), 231−247.

(24) Brown, M., Wedge, D. C., Goodacre, R., Kell, D. B., Baker, P. N., Kenny, L. C., Mamas, M. A., Neyses, L., and Dunn, W. B. (2011) Automated workflows for accurate mass-based putative metabolite identification in LC/MS-derived metabolomic datasets. *Bioinformatics 27* (8), 1108−1112.

(25) Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., and Hollender, J. (2014) Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol. 48* (4), 2097−2098.

(26) Fukuda, T., Matsumura, T., Ato, M., Hamasaki, M., Nishiuchi, Y., Murakami, Y., and Kinoshita, T. Critical roles for lipomannan and lipoarabinomannan in cell wall integrity of mycobacteria and pathogenesis of tuberculosis. *mBio* 2013, *4*, DOI: 10.1128/mBio.00472-12.

(27) Gilleron, M., Quesniaux, V. F., and Puzo, G. (2003) Acylation state of the phosphatidylinositol hexamannosides from *Mycobacterium bovis* bacillus Calmette Guerin and *Mycobacterium tuberculosis* H37Rv and its implication in Toll-like receptor response. *J. Biol. Chem. 278* (32), 29880−29889.

(28) Chesne-Seck, M. L., Barilone, N., Boudou, F., Asensio, J. G., Kolattukudy, P. E., Martín, C., Cole, S. T., Gicquel, B., Gopaul, D. N., and Jackson, M. (2008) A point mutation in the two-component regulator PhoP-PhoR accounts for the absence of polyketide-derived acyltrehaloses but not that of phthiocerol dimycocerosates in *Mycobacterium tuberculosis* H37Ra. *Journal of bacteriology 190* (4), 1329−1334.

(29) Sartain, M. J., Dick, D. L., Rithner, C. D., Crick, D. C., and Belisle, J. T. (2011) Lipidomic analyses of *Mycobacterium tuberculosis* based on accurate mass measurements and the novel MtbLipidDB. *J. Lipid Res. 52*, 861.

(30) Kaur, D., Guerin, M. E., Škovierová, H., Brennan, P. J., and Jackson, M. (2009) Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis*. *Adv. Appl. Microbiol. 69*, 23−78.

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX

(31) Daffé, M., Crick, D. C., and Jackson, M. Genetics of capsular polysaccharides and cell envelope (glyco) lipids. *Microbiol. Spectrum* 2014, *2*, DOI: 10.1128/microbiolspec.MGM2-0021-2013.
(32) Clarke, O. B., Tomasek, D., Jorge, C. D., Dufrisne, M. B., Kim, M., Banerjee, S., Rajashankar, K. R., Shapiro, L., Hendrickson, W. A., Santos, H., and Mancia, F. (2015) Structural basis for phosphatidylinositol-phosphate biosynthesis. *Nat. Commun. 6*, 8505.
(33) Maloney, E., Stankowska, D., Zhang, J., Fol, M., Cheng, Q. J., Lun, S., and Madiraju, M. V. (2009) The two-domain LysX protein of *Mycobacterium tuberculosis* is required for production of lysinylated phosphatidylglycerol and resistance to cationic antimicrobial peptides. *PLoS Pathog. 5* (7), No. e1000534.
(34) Wolf, B. J., Tatituri, R. V., Almeida, C. F., Le Nours, J., Bhowruth, V., Johnson, D., and Rossjohn, J. (2015) Identification of a potent microbial lipid antigen for diverse NKT cells. *J. Immunol. 195* (6), 2540−2551.
(35) Van Rhijn, I., Van Berlo, T., Hilmenyuk, T., Cheng, T. Y., Wolf, B. J., Tatituri, R. V., and Willemsen, P. (2016) Human autoreactive T cells recognize CD1b and phospholipids. *Proc. Natl. Acad. Sci. U. S. A. 113* (2), 380−385.
(36) Murphy, H. N., Stewart, G. R., Mischenko, V. V., Apt, A. S., Harris, R., McAlister, M. S., Driscoll, P. C., Young, D. B., and Robertson, B. D. (2005) The OtsAB pathway is essential for trehalose biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem. 280*, 14524.
(37) Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H., and Jackson, M. (2014) The cell envelope glycoconjugates of *Mycobacterium tuberculosis*. *Crit. Rev. Biochem. Mol. Biol. 49* (5), 361−399.
(38) Rawat, M., and Av-Gay, Y. (2007) Mycothiol-dependent proteins in actinomycetes. *FEMS microbiology reviews 31* (3), 278−292.
(39) Buchmeier, N. A., Newton, G. L., Koledin, T., and Fahey, R. C. (2003) Association of mycothiol with protection of *Mycobacterium tuberculosis* from toxic oxidants and antibiotics. *Mol. Microbiol. 47* (6), 1723−1732.
(40) Newsholme, P., Procopio, J., Lima, M. M. R., Pithon-Curi, T. C., and Curi, R. (2003) Glutamine and glutamate—their central role in cell metabolism and function. *Cell Biochem. Funct. 21* (1), 1−9.
(41) Berney, M., and Berney-Meyer, L. *Mycobacterium tuberculosis* in the Face of Host-Imposed Nutrient Limitation. *Microbiol. Spectrum* 2017, *5*, DOI: 10.1128/microbiolspec.TBTB2-0030-2016.
(42) Harper, C. J., Hayward, D., Kidd, M., Wiid, I., and Van Helden, P. (2010) Glutamate dehydrogenase and glutamine synthetase are regulated in response to nitrogen availability in *Mycobacterium smegmatis*. *BMC Microbiol. 10* (1), 138.
(43) Tullius, M. V., Harth, G., and Horwitz, M. A. (2001) High extracellular levels of *Mycobacterium tuberculosis* glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism. *Infection and immunity 69* (10), 6348−6363.
(44) Ríos Peces, S., Díaz Navarro, C., Márquez López, C., Caba, O., Jiménez-Luna, C., Melguizo, C., Prados, J. C., Genilloud, O., Vicente Pérez, F., and Pérez del Palacio, J. (2017) Untargeted LC-HRMS-based metabolomics for searching new biomarkers of pancreatic ductal adenocarcinoma: A pilot study. *SLAS DISCOVERY: Advancing Life Sciences R&D 22* (4), 348−359.
(45) García-Fontana, B., Morales-Santana, S., Navarro, C. D., Rozas-Moreno, P., Genilloud, O., Pérez, F. V., del Palacio, J. P., and Muñoz-Torres, M. (2016) Metabolomic profile related to cardiovascular disease in patients with type 2 diabetes mellitus: A pilot study. *Talanta 148*, 135−143.
(46) Di Guida, R., Engel, J., Allwood, J. W., Weber, R. J., Jones, M. R., Sommer, U., Viant, M. R., and Dunn, W. B. (2016) Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics 12* (5), 93.
(47) Valero-Guillén, P. L., Fernández-Natal, I., Marrodán-Ciordia, T., Tauch, A., and Soriano, F. (2016) Ether-linked lipids of *Dermabacter hominis*, a human skin actinobacterium. *Chem. Phys. Lipids 196*, 24−32.
(48) Hsu, F. F., Turk, J., Owens, R. M., Rhoades, E. R., and Russell, D. G. (2007) Structural characterization of phosphatidyl-myo-inositol mannosides from *Mycobacterium bovis* Bacillus Calmette Guerin by multiple-stage quadrupole ion-trap mass spectrometry with electrospray ionization. II. Monoacyl-and diacyl-PIMs. *J. Am. Soc. Mass Spectrom. 18* (3), 479−492.

DOI: 10.1021/acsinfecdis.9b00008
ACS Infect. Dis. XXXX, XXX, XXX−XXX