

RESEARCH ARTICLE

GSMN-ML- a genome scale metabolic network reconstruction of the obligate human pathogen *Mycobacterium leprae*

Khushboo Borah¹, Jacque-Lucca Kearney¹, Ruma Banerjee², Pankaj Vats^{2aa}, Huihai Wu¹, Sonal Dahale^{2ab}, Sunitha Manjari Kasibhatla², Rajendra Joshi², Bhushan Bonde³, Olabisi Ojo^{4ac}, Ramanuj Lahiri⁴, Diana L. Williams^{4‡}, Johnjoe McFadden^{1*}

1 Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom, **2** HPC-Medical and Bioinformatics Applications Group, Centre for Development of Advanced Computing, C-DAC Innovation Park, Panchavati, Pashan, India, **3** Head of Innovation Development, IT-Early Solutions, UCB Pharma, Slough, United Kingdom, **4** United States Department of Health and Human Services, Health Resources and Services Administration, Healthcare Systems Bureau, National Hansen's Disease Program, Baton Rouge, Louisiana, United States of America

^{aa} Current address: Michigan Center for Translational Pathology (MCTP), University of Michigan, Ann Arbor, Michigan, United States of America

^{ab} Current address: Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom

^{ac} Current address: Department of Biological Sciences, Albany State University, Albany, Georgia, United States of America

‡ Retired

* j.mcfadden@surrey.ac.uk



OPEN ACCESS

Citation: Borah K, Kearney J-L, Banerjee R, Vats P, Wu H, Dahale S, et al. (2020) GSMN-ML- a genome scale metabolic network reconstruction of the obligate human pathogen *Mycobacterium leprae*. PLoS Negl Trop Dis 14(7): e0007871. <https://doi.org/10.1371/journal.pntd.0007871>

Editor: Bradley R. Borlee, Colorado State University, UNITED STATES

Received: October 22, 2019

Accepted: June 2, 2020

Published: July 6, 2020

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

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

Funding: We thank BBSRC grants: BB/L022869/1, BB/L004569/1, MRC/Newton Fund: MR/M026434/1, BBSRC: India Partnering Award IPA1825 and Bioinformatics Resources and Applications Facility (BRAf) of Government of India for funding. We thank National Institute of Allergy and Infectious Diseases (NIAID) Interagency Agreement IAA 15006-004-00000 for the provision of viable M.

Abstract

Leprosy, caused by *Mycobacterium leprae*, has plagued humanity for thousands of years and continues to cause morbidity, disability and stigmatization in two to three million people today. Although effective treatment is available, the disease incidence has remained approximately constant for decades so new approaches, such as vaccine or new drugs, are urgently needed for control. Research is however hampered by the pathogen's obligate intracellular lifestyle and the fact that it has never been grown *in vitro*. Consequently, despite the availability of its complete genome sequence, fundamental questions regarding the biology of the pathogen, such as its metabolism, remain largely unexplored. In order to explore the metabolism of the leprosy bacillus with a long-term aim of developing a medium to grow the pathogen *in vitro*, we reconstructed an *in silico* genome scale metabolic model of the bacillus, GSMN-ML. The model was used to explore the growth and biomass production capabilities of the pathogen with a range of nutrient sources, such as amino acids, glucose, glycerol and metabolic intermediates. We also used the model to analyze RNA-seq data from *M. leprae* grown in mouse foot pads, and performed Differential Producibility Analysis to identify metabolic pathways that appear to be active during intracellular growth of the pathogen, which included pathways for central carbon metabolism, co-factor, lipids, amino acids, nucleotides and cell wall synthesis. The GSMN-ML model is thereby a useful *in silico* tool that can be used to explore the metabolism of the leprosy bacillus, analyze functional genomic experimental data, generate predictions of nutrients required for growth of the bacillus *in vitro* and identify novel drug targets.

leprae for this study. The 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.

Author summary

Mycobacterium leprae, the obligate human pathogen is uncultivable in axenic growth medium, and this hinders research on this pathogen, and the pathogenesis of leprosy. The development of novel therapeutics relies on the understanding of growth, survival and metabolism of this bacterium in the host, the knowledge of which is currently very limited. Here we reconstructed a metabolic network of *M. leprae*- GSMN-ML, a powerful *in silico* tool to study growth and metabolism of the leprosy bacillus. We demonstrate the application of GSMN-ML to identify the metabolic pathways, and metabolite classes that *M. leprae* utilizes during intracellular growth.

Introduction

The mycobacterial genus includes two of the greatest scourges of humanity: *Mycobacterium tuberculosis* and *M. leprae*, responsible for tuberculosis (TB) and leprosy respectively. Leprosy is one of the oldest plagues of mankind yet remains prevalent in developing countries. In 1981, the World Health Organization (WHO) recommended multidrug treatment (MDT) of disease with dapsone, rifampicin and clofazimine. Implementation of MDT achieved a 98% cure rate leading to reduction of prevalence by a remarkable 45% within three years. Yet, although MDT reduced global prevalence of leprosy, annual new case detection has remained fairly constant over the last decade [1]. Problems with control include the extremely lengthy drug treatment complicated by reactions that can lead to nerve damage and worsening of symptoms. Relapse of infection can also occur after treatment if the MDT therapies fail to clear drug resistant bacteria [2]. Additionally, new cases occurring in children, reported at 9.2% in 2013, have not reduced, indicating that active transmission continues [1], [2]. According to the WHO reports, 210,758 new cases of leprosy were detected in 136 countries in 2015 with the highest number of cases recorded in the developing nations of South-east Asia [3]. New approaches are clearly needed for case-finding, diagnosis, treatment and tackling social stigma of this devastating disease [3], [4].

One of the most striking facts about leprosy is how different it is from the disease caused by its related pathogen, *M. tuberculosis*. Both *M. tuberculosis* and *M. leprae* are intracellular pathogens that replicate primarily in macrophages but whereas the TB bacillus appears to primarily infect macrophages (and possibly also dendritic cells), the leprosy bacillus replicates in many cell types but particularly Schwann cells and endothelial cells [5, 6]. Possibly as a consequence of their differing cell tropism, tuberculosis is predominantly a pulmonary inflammatory infection (although extra-pulmonary infections are common) characterized by strong cellular immune reactions; whereas leprosy causes both disseminated infections and localized infections [7], [8]. The replicative capabilities of the two pathogens are also different, with *M. tuberculosis* being able to grow both intracellularly and extracellularly; whereas *M. leprae* is an obligate intracellular pathogen. Additionally, *M. tuberculosis* is able to grow *in vitro*, in minimal medium containing a single carbon source such as glycerol, a nitrogen source, such as ammonium ions and a few essential mineral ions [9], whereas *M. leprae* has never been grown in axenic media, even with the addition of various growth supplements [10]. This latter feature severely hampers most biological studies of the pathogen. The replication rate of the pathogens also differs markedly. Although both are very slow growing organisms, *M. tuberculosis* replicates with a doubling time of about a day in optimal conditions whereas the doubling time of

M. leprae in infected mouse is 10–14 days [11], [12]. The reasons for these differences in growth, virulence characteristics and cell tropism still remain unknown.

A major advance in understanding the pathogenesis of both TB and leprosy was achieved with the publication of their complete genome sequences, first with *M. tuberculosis* in 1998 [13] and then *M. leprae* in 2001 [10]. Analysis of the *M. leprae* genome sequence revealed that the pathogen had undergone massive gene decay with only 1604 predicted open reading frames compared to 3959 in *M. tuberculosis*. The reduced genome is a consequence of multiple mutations in coding sequences leading to gene decay, and the accumulation of 1116 pseudogenes and thereby loss of their associated functions [12], [14]. The loss of so many genes raised the obvious possibility that the inability of *M. leprae* to grow in axenic culture is due to mutational loss of biosynthetic function and consequent reliance on importing complex nutrients from the host which are not available in artificial media, as has been found for several other intracellular parasites. However, analysis of the genome sequence of *M. leprae* demonstrated that most of the anabolic capability of the pathogen, relative to *M. tuberculosis*, seems to be intact. The pathogen appears to have retained complete enzyme systems for synthesis of all amino acids, except for methionine and lysine [10]. The biosynthetic pathways of purines, pyrimidines, nucleosides, nucleotides, vitamins and cofactors are mostly intact [10]. Gene deletion and decay does however appear to have eliminated many important catabolic systems and respiratory systems. For example, as previously reported by Cole *et al.*, 2001, the aerobic respiratory chain of *M. leprae* is truncated due to the loss of several genes in the NADH oxidase operon, *nuoA-N*, potentially eliminating the ability of the pathogen to produce ATP from the oxidation of NADH. Yet central carbon metabolism appears to be intact in the pathogen with an intact glycolytic pathway, pentose phosphate pathway (PPP), gluconeogenesis and the tricarboxylic acid (TCA) cycle. Thus, despite the availability of the complete genome sequence of *M. leprae*, the inability of the pathogen to grow in axenic culture remains a mystery. Also, the substrates utilized during growth of *M. leprae in vivo*, remain unclear. Recent research has also revealed intriguing capabilities of the leprosy bacillus to manipulate its host cell including its ability to dedifferentiate and reprogram its host Schwann cell to a stem-cell-like state in which the pathogen can proliferate more efficiently and which likely promotes its dissemination through the body [15].

The availability of full genome sequences does however allow reconstruction of genome-scale metabolic reaction networks for micro-organisms. These not only instantiate current knowledge of the metabolism of particular organisms; but may be used to analyze and integrate data, particularly functional genomic data, to generate novel predictions that may be subjected to experimental validation. Genome-scale metabolic networks have been constructed for many microbes including *M. tuberculosis* and *Mycobacterium bovis* where they have been used to identify novel metabolic pathways and examine their growth and likely substrate utilization within their host cells [16], [17], [18], [19]. In this study we constructed the first genome-scale network of *M. leprae*, GSMN-ML. We used the network to examine the organism's catabolic and anabolic capabilities to gain insight into why the pathogen is unable to grow *in vitro*. We also used GSMN-ML to interrogate a RNA-seq dataset of *M. leprae* obtained from mouse foot pads and then maintained *ex vivo* to probe the pathogen's metabolic response to its host environment and analyzed the data using differential producibility analysis, a bioinformatic tool that translates transcriptomic data to predictions of metabolic responses [20].

Materials and methods

Ethics statement

All animal studies were performed under the scientific protocol number A-105 reviewed and approved by the National Hansen's Disease Programs (NHDP) Institutional Animal Care and

Use Committee, and were conducted in strict accordance with all state and federal laws in adherence with PHS policy and as outlined in The Guide to Care and Use of Laboratory Animals, Eighth Edition.

***M. leprae* culture and growth conditions**

Athymic nude mice (Envigo) were infected by inoculating 3×10^7 *M. leprae*, strain Thai-53, in each hind footpad. At 4–5 months post inoculation mice were euthanized and footpad tissues removed. These tissues were either fixed in 70% ethanol for at least 48 h or minced and homogenized immediately for purification of viable *M. leprae* [21]: *in vivo M. leprae*. Viable bacteria were then added to axenic medium 7H9 containing Caesitone (0.1% w/v), BSA (0.5% w/v), dextrose (0.75% w/v) and ampicillin (50 µg/ml), and incubated at 33°C, 5% O₂ for 48h. Following incubation, bacteria were pelleted and fixed in 70% ethanol: *in vitro M. leprae*. After fixation for at least 48h, ethanol was removed and *M. leprae* were treated with 0.1N NaOH to remove mouse tissues. Bacteria were adjusted to 2×10^9 /ml and 1 ml aliquots were pelleted and resuspended in 1 ml Trizol and stored at -20°C. *M. leprae* in ethanol-fixed tissues were also purified, treated with NaOH, resuspended in Trizol and stored at -20°C.

RNA purification and RNA-seq analysis

Bacterial lysis and RNA purification was performed using FastPrep-24 vertical homogenizer (MP BioMedicals), and FastPrep Lysing Matrix B tubes. After DNase treatment, an aliquot of sample RNA was converted to cDNA and analyzed for the presence of *M. leprae* *esxA* expression and DNA contamination [21]. Ribosomal RNA was depleted from total RNA preparations using Ribo-Zero rRNA Removal Kit according to manufacturer's recommendations, and RNA quantity and quality was determined using NanoDrop 2000 and Agilent 2100 BioAnalyzer. Sequencing samples were prepared using the SOLiD Total RNA-Seq Kit (Life Technologies, PN4445374) using 100 ng ribosome RNA-depleted and barcoded t-RNA (SOLiD RNA Barcoding Kit, Module 1–16 (Life Technologies, PN 4427046) according to the manufacturer's protocol. Emulsion PCR and SOLiD sequencing, 75 base pairs single direction, were performed for the SOLiD 5500 System (Life Technologies, Thermo Fisher Scientific). Sequencing analysis was done using Bacterial RNA-seq Analysis Kit 1.0 on Maverix Analytic Platform (Maverix Biomix Inc.). Raw sequencing reads from SOLiD sequencing platform that were converted into FASTQ file format were quality checked for potential sequencing issues and contaminants using FastQC (FastQC). Adapter sequences, primers, Ns, and reads with quality score below 13 were trimmed using fastq-mcf of ea-utils [22] and Trimmomatic [23]. Reads with a remaining length of < 20 bp after trimming were discarded. Pre-processed reads were mapped to the *Mycobacterium leprae* TN genome (RefSeq Accession Number: NC_002677) using EDGE-pro [24]. Read coverage on forward and reverse strands for genome browser [25] visualization was computed using BEDtools [26], SAMtools [27] and UCSC Genome Browser utilities [23]. Read counts for RefSeq genes generated by EDGE-pro [25] were normalized across all samples and then used for differential expression analysis using DESeq [28]. Significant differentially expressed genes were determined by adjusted P-value with a threshold of 0.05. Fold change (log₂) between samples were hierarchically clustered using Pearson correlation.

Construction of GSMN-ML

The updated *M. tuberculosis* genome scale metabolic network (GSMN-TB_2) [16] created for H37Rv strain (Supplementary S1 File) was the basis of reconstruction of *M. leprae* metabolic model GSMN-ML (Supplementary S3 File). The reconstruction of the network followed the

methods as described previously [9], [16]. *M. leprae* (RefSeq ID: NC_002677) functional genes were identified using the online GenBank tool [29], and compared to the same list used for the development of the GSMN-TB of *M. tuberculosis* using the bi-directional best hit criteria to search for orthologs [30]. Using *M. tuberculosis* H37Rv (RefSeq ID: NC_000962) as a template, the metabolic network was reconstructed using Pathway Tools (Version 24.0) [31]. Non-functional pathways in, and gene essentiality data for *M. leprae* were identified and the model refined to remove any non-homologous pathways, identified by literature search and metabolic pathway data from previous mutagenesis study [32]. Surrey-FBA was used in the study for the flux balance analysis to further refine the model by iteration [33]. Single genes, and the respective pathways were removed and the model's feasibility re-assessed. Only edits which did not render the model infeasible were used in the final product. If found to be essential, transport reactions were introduced for their product(s), so that the inactivation of these genes will be compensated by uptake of the predicted metabolites from the artificial media (presumed to be obtained from host *in vivo*). The full list of reactions, for GSMN-ML and GSMN-TB_2 models can be found in Supplementary S1 and S3 Files respectively.

Flux balance analysis (FBA)

FBA simulations in this study were also performed with the SurreyFBA software as described previously [9], [33]. Simulations were conducted under steady-state conditions, using biomass production as the objective function to obtain the steady state theoretical fluxes through the reactions in the network that maximised biomass production. For FBA, substrate inputs for carbon and nitrogen sources were specified in the test medium that was used to constrain the simulations. External metabolites, except in the cases of those identified as being rate-limiting, were considered to be freely available and disposable. Rate-limiting external nutrients are bound by upper and lower flux rate limits, as described in the model's set-up. FBA analysis were performed as described in Gevorgyan et al., 2011 to obtain the upper and lower flux limits through a particular reaction [33]. Gene essentiality predictions were tested using SurreyFBA software. The maximal theoretical growth rate of each *in silico* gene knock out was calculated by removing single genes from the network and performing FBA linear programming as described previously [9], [16], [33].

Differential producibility analysis (DPA)

DPA seeks to translate transcriptomic signatures into metabolic fluxes using FBA to identify metabolites or metabolic pathways most affected by the changes in gene transcription and is described in detail by Bonde et al., 2011 [20]. RNA-seq data from *M. leprae*, strain Thai-53 cultivated in mice foot pads were used for DPA. DPA utilized a FBA-based metabolite expression analysis using glucose or glycerol as the carbon source input to GSMN-ML to calculate the maximal theoretical flux towards each metabolite in the network. For every metabolite, reactions involved were then ranked based on the expression values, with those ≥ 2 being highly expressed and those < 2 lowly expressed. A median value was assessed, creating a matrix of values for each metabolite. The same process is repeated with down-regulated reactions and the result is two lists of the most differentially produced metabolites for a given experimental environment. Statistical significance of the DPA is assessed by the following:

$$\begin{aligned} &\geq 0, \quad \text{or} \\ &\leq 0, \quad th = 1 \\ &oth = 0 \\ &E = S/N \end{aligned}$$

where RUm = median expected value of up-regulated metabolite

DUm = median expected value of down-regulated metabolite

O = observed value in real dataset

S = arbitrary score value

N = number of datasets

E = E-value

A significance threshold of $E < 0.15$ was decided for further analysis

Results

Construction of GSMN-ML genome scale metabolic network

Our approach was to start with our previously-constructed metabolic model of *M. tuberculosis*, GSMN-TB [16] and step-by-step, replace all *M. tuberculosis* genes with orthologues from *M. leprae*. We used an updated version of GSMN-TB_2, which includes reactions for cholesterol catabolism (Supplementary S1 File). We adjusted the biomass composition to reflect what is known about the macromolecular composition of *M. leprae*, for example, removing *M. tuberculosis*-specific components and adding the phenolic glycolipid (PGL1) which is a major antigen in leprosy. This entailed adding additional reactions for PGL1 synthesis. Also, *M. leprae* lacks N-glycosylated muramic acid in its peptidoglycan; and glycine is substituted in place of L-alanine [34], so this was reflected in the *in silico* peptidoglycan composition. Mannosyl β -1 phosphodolichol (MPD) of *M. tuberculosis* is a polyketide known to invoke strong T-cell response and pks12 [35], which is required for formation of MPD a pseudogene in *M. leprae* (ML1437c) and therefore, the corresponding biosynthesis reactions were removed from the model. A small number of additional metabolic reactions that are found in *M. leprae* but absent in *M. tuberculosis*, were added to the network. These included ML2177, a putative uridine phosphorylase involved in nucleotide salvage, highlighting the previously described difference between nucleotide salvage in *M. leprae* compared to *M. tuberculosis* [36]; and ML0247, a putative arsenate reductase arsC. Using the template of reactions available for *M. tuberculosis* H37Rv in GSMN-TB_2 (856 reactions catalysed by 729 enzymes), we were able to transfer 713 reactions to construct the *M. leprae* model using gene orthology. However, only 396 out of the total orthologues appeared to be functional in *M. leprae*, the remainders were pseudogenes. 21 additional reactions were added to the *M. leprae* model based on manual literature curation, and the resulting *M. leprae* network encoded a total of 417 genes encoding metabolic enzymes. This reduced the number of functional enzymes (nearly half the number as compared to *M. tuberculosis*), which is a direct consequence of the massive genome reduction as well as pseudogene formation in *M. leprae*.

Incorporation of pseudogenes and orphan reactions facilitated *in silico* growth of GSMN-ML

We applied FBA to investigate the ability of the *in silico* *M. leprae* model to synthesize biomass from basic nutrients, starting from the minimal nutrient requirements (Supplementary S2 File) needed to grow *M. tuberculosis in vitro*. GSMN-ML was however unable to synthesize biomass from the minimal requirements that are needed by GSMN-TB. To address this problem we reintroduced reactions that were removed in constructing GSMN-ML from GSMN-TB (because genes encoding these reactions were pseudogenes in *M. leprae*) until we obtained a feasible network. We started with 260 orphan reactions that appeared to be essential in *M. leprae*, but whose function is not encoded by any known functional *M. leprae* gene (Supplementary S3 File). We then adopted two approaches to minimize the number of these GSMN-ML orphan reactions. First, we performed bioinformatic analysis to identify any

hypothetical gene that could complement a function provided by an essential orphan reaction. If such a gene was found, it was added to the network. However, in many cases, no putative gene could be identified to potentially encode the function of the orphan gene. We then investigated whether the orphan reaction could be made non-essential by opening a transport gate for a metabolite that was potentially imported from the intracellular environment of *M. leprae*. If that was the case then we opened that transport gate to allow import and removed that orphan reaction from the network. Lastly, essential orphan reactions that could not be complemented by a hypothetical gene whose product could not be feasibly obtained from the host were left as functional orphan reactions. This operation led to the generation of a feasible network GSMN-ML with 872 reactions that included eight orphan transport reactions (Table 1). A full list of network reactions and the medium components used for this analysis is included in Supplementary S2 and S3 Files respectively.

The following eight potentially imported nutrients that were identified by this procedure were formate, L,L-2,6-diaminopimelate, lysine, methionine, tetrahydrofolate, coenzyme-B (COB-I), pantothenate and protoporphyrinogen (Table 1). The requirement for methionine and lysine biosynthesis is caused by the genes *metC* encoding O-acetyl-L-homoserine sulfhydrylase (ML0683c) and *dapB* that encodes dihydrodipicolinate reductase (ML1527) being pseudogenes in *M. leprae*. The requirement for L,L-2,6-diaminopimelate (DAPIM), an intermediate for lysine synthesis and one of the principle substrates for peptidoglycan biosynthesis, is due to the absence of diaminopimelate decarboxylase (EC: 4.1.1.20) in *M. leprae* based on the annotation used for GSMN-ML (Table 1). The requirement for tetrahydrofolate, formate and COB-I is due to many genes involved in their synthesis being pseudogenes in *M. leprae*. The requirement for pantothenate, the substrate for coenzyme-A (CoA) synthesis, is due to *M. leprae* lacking 2-dehydropantoate 2-reductase (EC: 1.1.1.169) that is needed to synthesize CoA from glycolysis. The requirement for protoporphyrinogen is a consequence of *M. leprae*'s inability to synthesise porphyrin, and thereby any hemoproteins, from L-glutamate due to the absence of the *hemN* gene, coding for oxygen-independent coproporphyrinogen III oxidase. The import flux of these nutrients was fixed to the minimum needed for maximum biomass production (see Supplementary S3 File and Fig 1).

Substrate utilization by GSMN-ML

A major finding of the genome sequencing project by Cole *et al.*, 2001 was that *M. leprae* has lost many catabolic genes and pathways [10]. To investigate this in more detail we investigated the ability GSMN-ML to utilize various carbon and nitrogen substrates for biomass production and compared it to GSMN-TB_2 (Supplementary S4 File). The results, (Fig 2A) shows

Table 1. Essential *in silico* medium supplements and their transport gates in GSMN-ML.

<u>Metabolite</u>	<u>Transport gate</u>
L, L-diaminopimelate	R827
Lysine	R834
Methionine	R835
Formate	R806
Tetrahydrofolate	R936
Pantothenate	R939
COB-I	R937
Protoporphyrinogen	R938

Medium supplements for growth of GSMN-ML. Transport gates for the metabolites shown are included in S3 File.

<https://doi.org/10.1371/journal.pntd.0007871.t001>

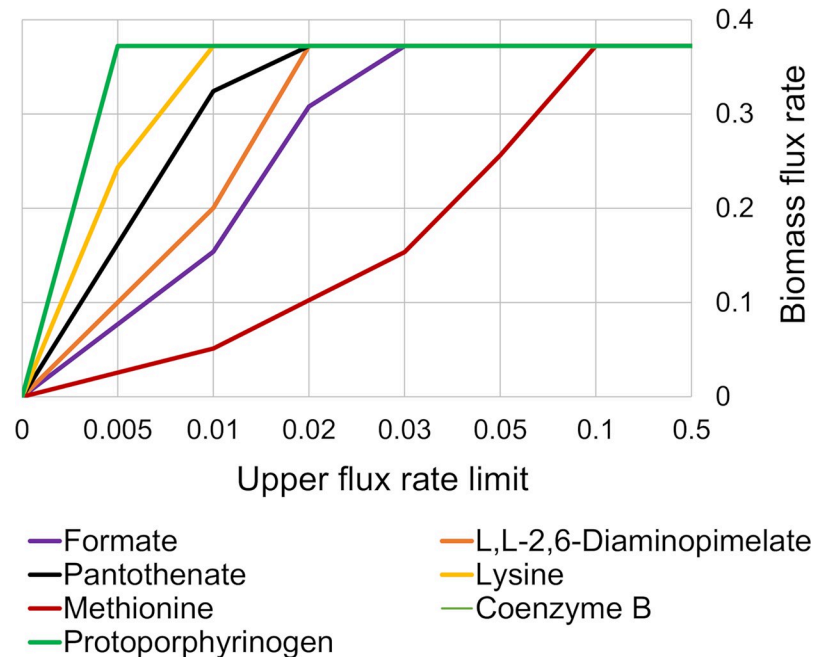


Fig 1. Influence of opening transport gates for essential *in silico* medium supplements on predicted biomass production for the *M. leprae* GSMN model. Media component uptake was limited to maximal value of 1 mmol g⁻¹ dryweight h⁻¹. The biomass flux rate for GSMN-ML was calculated with respect to the media components.

<https://doi.org/10.1371/journal.pntd.0007871.g001>

that, perhaps surprisingly, *M. leprae* has retained the ability to utilize many carbon sources; yet, in comparison with *M. tuberculosis*, has lost the ability to utilize acetate and glycolate due to pseudogenization of acetate kinase, phosphate acetate transferase and acetyl-coA synthase. *M. leprae* has also lost almost the entire pathway involved in cholesterol utilization and cannot thereby utilize host-derived cholesterol [37], which appears to be a growth substrate for *M. tuberculosis in vivo*. Catabolism of amino acids, such as valine, threonine and isoleucine as carbon sources was also impaired in *M. leprae* (Fig 2B). To explore the nitrogen requirements of GSMN-ML, FBA was performed with five different single nitrogen sources: ammonia, nitrate, urea, glutamate and glutamine (Fig 2C) and with glucose as the primary carbon source, comparing to GSMN-TB_2. Both models were able to utilize glutamate and glutamine to generate approximately equal biomass. The GSMN-TB_2 model was able to generate 31.3% more BIOMASS using ammonia as compared to GSMN-ML, and unlike *M. leprae*, was able to grow on urea as a single nitrogen source. Neither of the models could generate biomass using nitrate as nitrogen source. These data suggest that glutamate or glutamine is the most likely nitrogen source for *M. leprae*.

DPA to predict intracellular metabolism of *M. leprae* from transcriptome data

Directly measuring metabolism is very difficult, particularly for intracellular bacteria. It is however relatively straightforward to obtain gene expression profiles of intracellular bacteria and, from that, discover whether key metabolic genes are up- or down-regulated in a certain condition. However, single genes may contribute to several different metabolic pathways; and most metabolic pathways involve many genes, which may show contradictory expression patterns. To provide a systems-wide insight into intracellular metabolism, we previously developed DPA that uses a genome scale model to associate metabolites with system-wide analysis

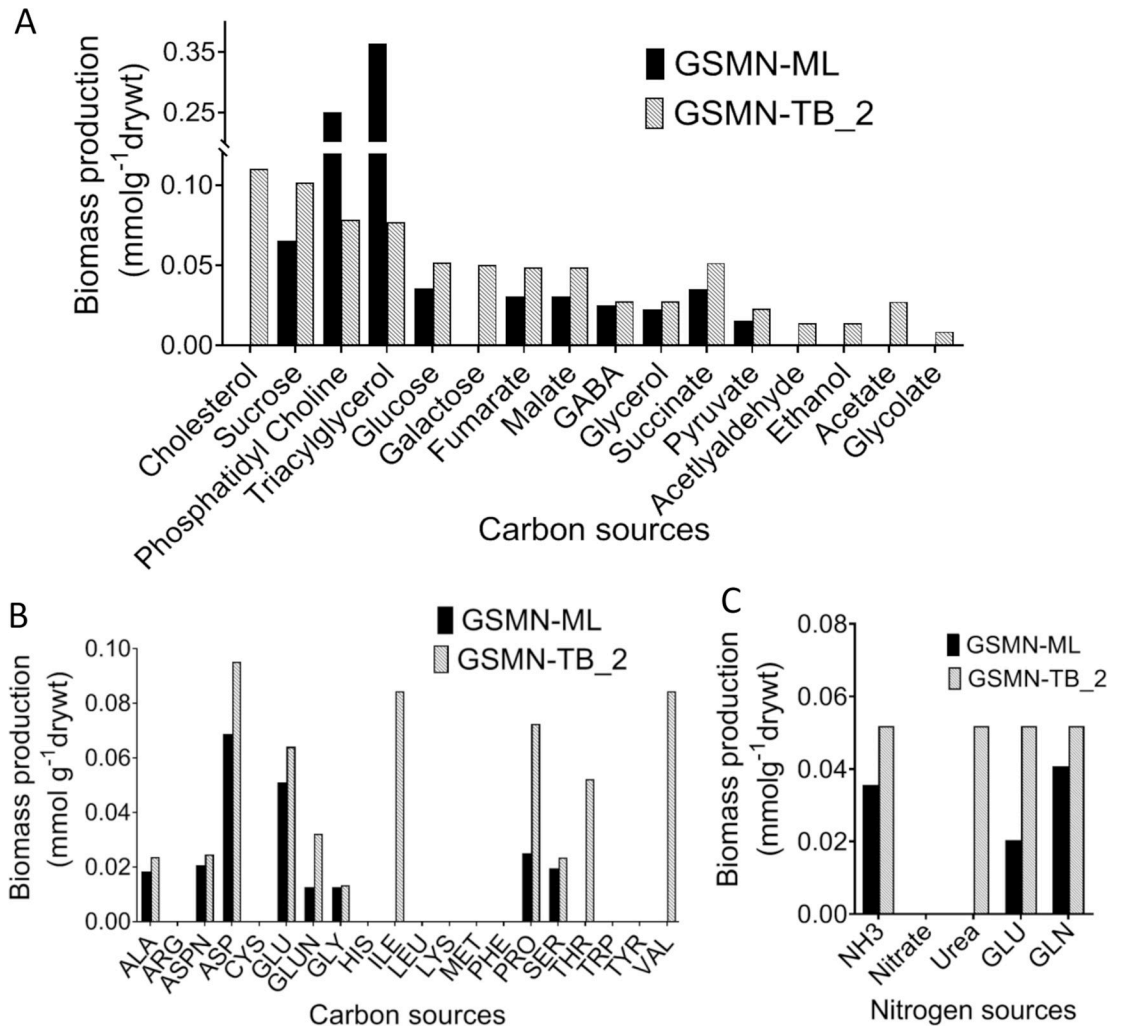


Fig 2. In silico prediction of biomass production for various substrates of *M. leprae* compared to the *M. tuberculosis* GSMN models. Carbon substrate utilization profiles of *M. leprae* GSMN-ML vs. *M. tuberculosis* GSMN-TB_2. Biomass production is shown for both the models on (A) various carbon substrates and (B) amino acids as carbon sources (C) various nitrogen sources. Biomass production flux was calculated on the various substrates using FBA. Abbreviations for amino acids are alanine (ALA), arginine (ARG), asparagine (ASPN) aspartic acid (ASP), cysteine (CYS), glutamic acid (GLU), glutamine (GLUN), histidine (HIS), isoleucine (ILE), leucine (LEU), lysine (LYS), methionine (MET), phenylalanine (PHE), proline (PRO), serine (SER), threonine (THR), tryptophan (TRP), tyrosine (TYR) and valine (VAL).

<https://doi.org/10.1371/journal.pntd.0007871.g002>

of gene expression patterns [20]. The first step in DPA is to identify all genes (the metabolite producibility gene list) that are required (in the genome-scale model) for production of every metabolite, given a set of nutrients inputs. The next step in DPA is to interrogate gene expression data to identify whether the genes that contribute to the production of a particular metabolite tend to be associated with either up-regulated or down-regulated genes in the test condition. Note that metabolites may appear in both lists if, for example, one set of genes associated with their production is found to be up-regulated and another set is found to be down-regulated. In our previous study, we applied DPA to gene expression data obtained from both *E. coli* and *M. tuberculosis*, including *M. tuberculosis* replicating within macrophages [20]. In this study, we applied the same approach to investigate metabolic changes associated with *in vivo* growth of *M. leprae*.

The nutrients used by *M. leprae* *in vivo* still remains unknown. Recently we showed that intracellular *M. leprae* in host schwann cells used glucose or glucose derivatives as the primary carbon sources [38]. Enzymes for catabolism of hexoses and glycerol have been detected in *M. leprae* isolated from armadillos [39, 40], prompting us to perform the DPA growth simulations with both (independently) glucose and glycerol. RNA was extracted and immediately processed from *in vivo*-grown viable *M. leprae* harvested and from mouse footpads [21] (see methodology for details). The control condition was *M. leprae* from mouse footpads, as above, but then incubated in axenic culture for 48 hours under conditions (see methodology) where the bacilli remain viable but do not replicate. *In vivo* up-regulated genes were defined as being the product of an enzyme transcribed from a gene with a fold change >2 in the RNA-Seq data. *In vivo* down-regulated genes are defined as genes with a fold change <0.5 in the RNA-Seq data (Supplementary S5 File). These expression lists were then cross-checked against the metabolite genes to identify metabolite whose production (genes encoding the enzymes involved in their production) is significantly associated with the up-regulated gene list, or is significantly associated with the down-regulated gene list. The top 100 metabolites associated with up- or down-regulated genes for both carbon sources is presented in Supplementary S6 File. For up-regulated genes, 67 and 54 metabolites were shared between GSMN-ML using glucose and glycerol as carbon sources (Fig 3A and 3B). Pathways and areas of metabolism likely to be affected by up and down-regulated genes in *M. leprae* during growth on the two carbon sources are illustrated in Fig 3C–3F. Metabolites required for the synthesis of amino acids comprised the major category of metabolites predominantly associated with up-regulated genes in both glucose and glycerol; whereas metabolites involved in cofactor synthesis were mostly associated with up-regulated genes on glucose simulations (Fig 3C and 3D). Metabolites involved in lipid synthesis were the major family of metabolites identified as being associated with down-regulated genes for growth on glucose (Fig 3E). Biomass production of GSMN-ML on glucose was higher than that on glycerol, which implies that the carbon and energy resource allocations on two carbon sources may be different (Fig 2A). Down-regulation of lipid biosynthesis may be a consequence or a requirement to up-regulate amino acid synthesis and promote bacterial replication. On glycerol, metabolites for co-factor biosynthesis were down-regulated on glycerol (Fig 3F). Several metabolites involved in cell wall synthesis of *M. leprae*'s cell wall were associated with up-regulated genes only during glycerol utilisation simulations. This is likely because cell wall components, such as arabinogalactan-decaprenyl phosphate [41], and d-alanyl-d-alanine, a peptidoglycan precursor [42], fatty-acid derived mycolates and arabinogalactan-peptidoglycan molecules are derivatives from glycerol metabolism.

This pattern of gene regulation of *M. leprae* in response to its *in vivo* environment appears to be very different from that of intracellular *M. tuberculosis* [20]. For example, in *M. tuberculosis* 25% of metabolites associated with phospholipid synthesis were associated with up-regulated genes compared to only 1% associated with down-regulated genes [20]; whereas, in *M. leprae*, for glycerol and glucose-based simulations, a similar fraction of around 7% of metabolites involved in phospholipid synthesis were found to be associated with both up and down-regulated genes. Also, in *M. tuberculosis*, there was a clear tendency for metabolites involved in lipid metabolism (other than phospholipids) to be associated with up-regulated genes, whereas, for *M. leprae*, they were mostly associated with down-regulated genes. This pattern was reversed for amino acid metabolism, with a clear association with down-regulated genes in *M. tuberculosis* (25% compared with 15% associated with up-regulated genes) [20], whereas, in *M. leprae*, 18% of metabolites involved in amino acid synthesis were associated with up-regulated genes and only 9% associated with down-regulated genes. Comparing the two pathogens, the adaptation of *M. tuberculosis* to its intracellular environment appears to be associated

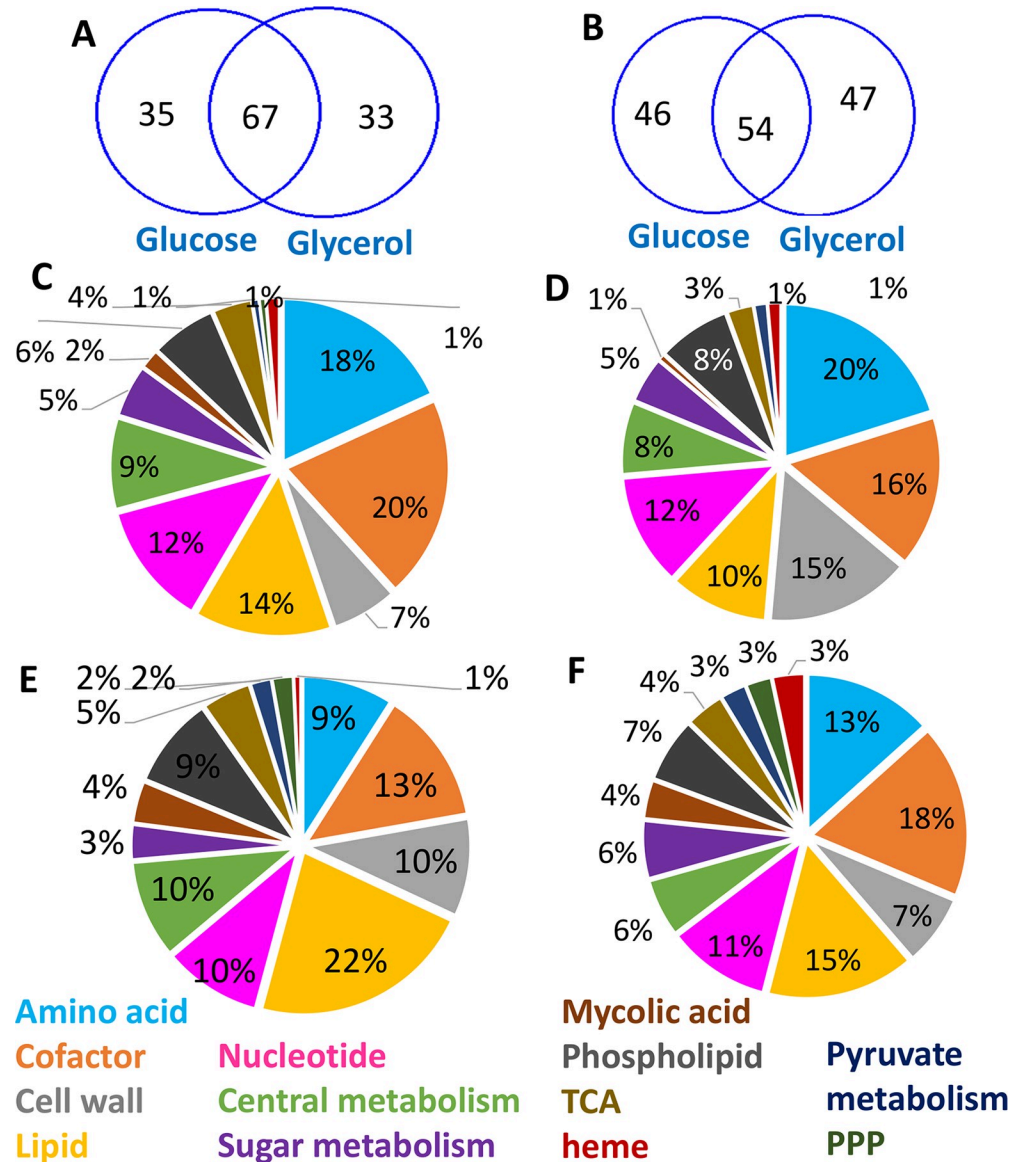


Fig 3. DPA analysis of gene expression patterns to identify metabolic processes associated with up-regulated (A) and down-regulated *M. leprae* genes (B) *in vivo*. Venn diagram compares metabolites identified as associated with up-regulated (C and D) or down-regulated (E and F) genes using either glucose (C and E) or glycerol (D and F) as carbon source in DPA.

<https://doi.org/10.1371/journal.pntd.0007871.g003>

with an up-regulation of lipid synthesis and down-regulation of amino acid synthesis; whereas the reverse appears to be true for *in vivo M. leprae*.

To assess the importance of *M. leprae* genes, i.e., whether a gene is essential or non-essential for growth on glycerol and glucose, we used GSMN-ML for gene knockout predictions. We calculated the growth rate when a gene in the GSMN-ML network was removed. We calculated gene essentiality (GS) ratios- the ratio of the growth rate of the knock out to that of wild-type for each gene in the network (Supplementary S7 and S8 Files). Figs 4 and 5 shows the central carbon network genes and the GS ratios. A gene was considered essential if the GS ratio was < 0.01. We found 21 genes involved in glycolysis, gluconeogenesis, TCA cycle and PPP to be essential central metabolic genes for growth of *M. leprae* on both glucose and

glycerol. We compared the GS ratios for the central metabolic network genes with that of *M. tuberculosis*, by performing GS analysis of GSMN-TB_2 using glucose or glycerol as the sole carbon source. Around 64–66% genes of the total genes in the network were predicted to be essential in GSMN-ML compared to ~28% in GSMN-TB_2. The higher percentage of essential metabolic genes in *M. leprae* is likely a consequence of the loss of non-essential genes due to deletion or pseudogenization during evolution of *M. leprae* from its progenitor.

Discussion

In this study, we constructed GSMN-ML, a genome scale metabolic model of *M. leprae*. Interestingly, the model predicted that the pathogen would be unable to grow without the addition of supplements (Table 1) that the pathogen probably obtains from the host *in vivo*. Although we predict that these supplements are required for *in vitro* growth of *M. leprae*, they are likely to be present in many of the media that have been tested in attempts to grow *M. leprae in vitro*. So it seems likely that additional, as yet unidentified, vital nutrients are needed for *in vitro* growth of the pathogen. Nonetheless, Table 1 does at least provide a minimal nutrient requirement that could be further investigated. Host-derived lipids are a well-established carbon source in mycobacterial infection such as *M. tuberculosis* [43–45]. Although there is no direct evidence of *M. leprae* using host cell lipids as a nutrient source, *M. leprae* infection did induce changes in lipid metabolism in the host Schwann cells and macrophages, making them lipid-rich foamy cells [46–48]. Also, lipid droplet biogenesis in the host provides growth advantage to *M. leprae* by providing lipid nutrient pools to the pathogen [47–49]. Our FBA-based simulations indicated that carbon sources triacylglycerol and phosphatidyl choline provided the highest biomass production for *in silico* growth on GSMN-ML, suggesting these, or similar lipids, as potential carbon sources for *M. leprae*. In comparison to *M. tuberculosis*, GSMN-ML failed to use either amino acids or a range of other lipid-derived metabolites as carbon sources. Similarly, and in contrast to *M. tuberculosis*, GSMN-ML was incapable of *in silico* growth with cholesterol as a carbon source (Fig 2A). This is in agreement with previous studies that indicate that, despite retaining an ability to oxidize cholesterol, *M. leprae* is unable to use cholesterol as an energy or carbon source [37]. This is consistent with the genome studies, which indicate that evolutionary gene-decay in *M. leprae* has resulted in the loss of the *mce4* operon and that encodes for active transport system for cholesterol [10], [37], as well as many genes required for cholesterol catabolism. Despite being part of a major component to mycobacterial cell walls [48], GSMN-ML also failed to utilize galactose; indicating that galactose-cell wall components are likely synthesised by the Leloir pathway in the leprosy bacillus from UDP-glucose, rather than from galactose directly [10], [11]. Amino acid utilisation in *M. leprae* and *M. tuberculosis* were also significantly different. While *M. leprae* retains the ability to biosynthesize amino acids, catabolic pathways are degraded, accounting for why the pathogen cannot utilize amino acids as carbon sources (Fig 2B) [10]. This deficit is largely explained by the absence of the urease operon in *M. leprae*; also accounting for the inability of GSMN-ML to utilize urea as the nitrogen source (Fig 2C).

Our FBA analysis of GSMN-ML indicates that *M. leprae* is able to utilize both glucose and glycerol as carbon sources, consistent with evidence that glucose and/or glucose-derived metabolites were used as nutrients by *M. leprae* during intracellular growth [38], [50]. DPA analysis of gene expression data from *M. leprae* harvested from mouse foot pads showed down-regulation of lipid metabolic genes of on glucose. This was unexpected as growth on glucose requires mycobacteria to synthesize lipids. *M. leprae* produced different amounts of biomass on glucose (six carbon molecule) and glycerol (three carbon molecule). The down-regulation of lipid metabolic genes on glucose may be required to balance carbon resource

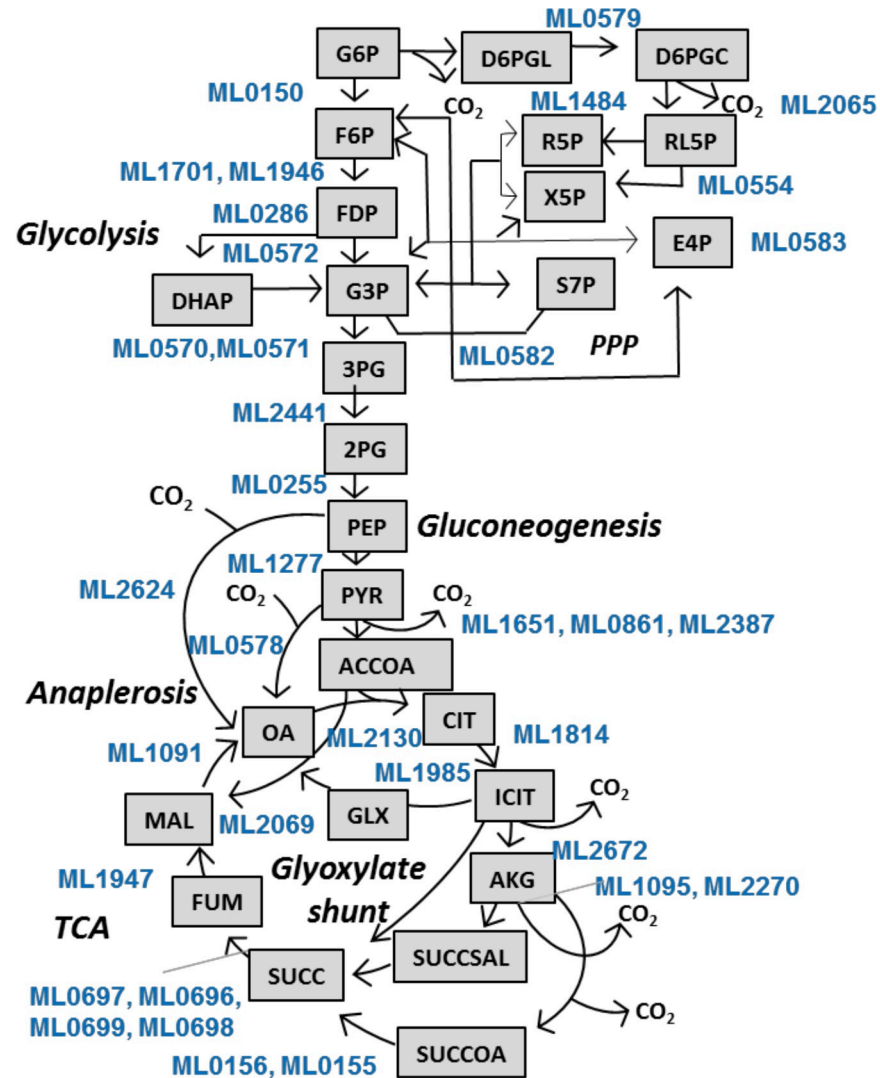


Fig 4. Central metabolic network of GSMN-ML. Reactions and their respective annotated genes (highlighted in blue) are shown for Glycolysis, Gluconeogenesis, PPP, TCA cycle, Glyoxylate shunt and anaplerotic CO₂ fixation. Abbreviations for metabolites are: G6P (glucose-6-phosphate), F6P (fructose-6-phosphate), FDP (fructose-1,6-bisphosphate), G3P (D-glyceraldehyde-3-phosphate), DHAP (dihydroxyacetone-phosphate), 3PG (3-phosphoglycerate), 2PG (2-phosphoglycerate), PEP (phosphoenolpyruvate), PYR (pyruvate), ACCOA (acetyl-CoA), OA (oxaloacetate), MAL (malate), FUM (fumarate), SUCC (succinate), SUCCOA (succinyl-CoA), SUCCSAL (succinate semialdehyde), AKG (α -ketoglutarate), ICIT (isocitrate), CIT (citrate), GLX (glyoxylate), D6PGL (D-gluconolactone-6-phosphate), D6PGC (6-phospho-D-gluconate), R5P (D-ribose-5-phosphate), RL5P (D-ribulose-5-phosphate), X5P (D-xylulose-5-phosphate), E4P (D-erythrose-4-phosphate) and S7P (D-sedoheptulose-7-phosphate).

<https://doi.org/10.1371/journal.pntd.0007871.g004>

allocation towards the synthesis of other biomass components in order to maintain higher proteome demand and amino acid synthesis for rapid growth [51], [52]. DPA analysis identified clear differences in intracellular metabolism of *M. leprae* compared to intracellular *M. tuberculosis*, particularly in lipid and amino acid metabolism. These differences may contribute to the very different virulence properties of these two related pathogens.

Genome scale metabolic reconstructions and flux balance analysis aims to find a state that maximizes growth of an organism in the given media. To validate these metabolic models, it is important to test experimentally if the bacteria grows optimally in the given media.

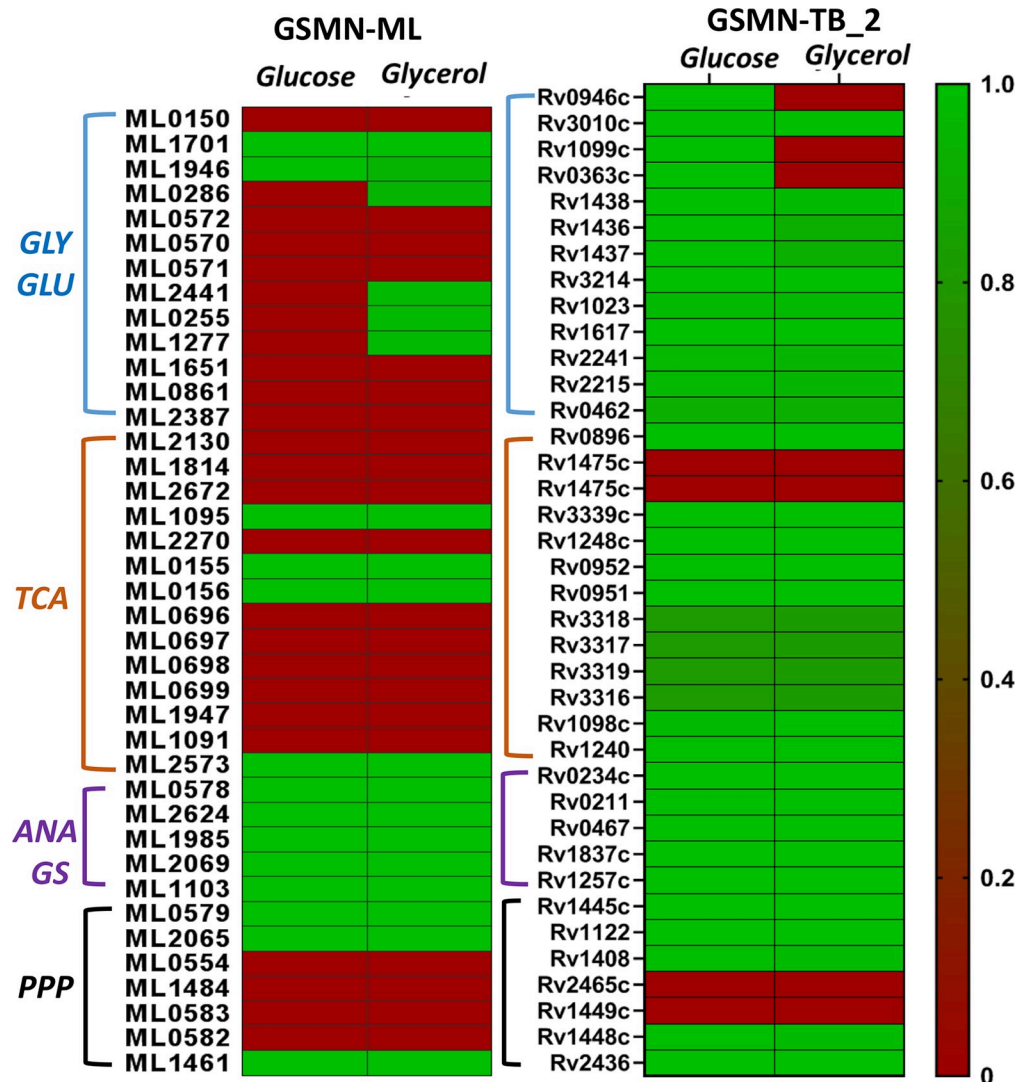


Fig 5. Gene essentiality analysis of GSMN-ML and GSMN-TB_2 utilizing glucose and glycerol as carbon sources. Heat map showing the gene essentiality (GS) predictions of central carbon metabolic genes of the two models using glucose and glycerol as the sole carbon sources (see Supplementary S7 and S8 Files for details and the calculated values). Simulations were performed in SurreyFBA platform to calculate GS which is the ratio of growth rate of the knockout to that of the wild type. A gene was considered essential if the GS ratio < 0.01 and the GS ratio of the essential genes were 0. Essentiality is indicated by single colour gradient with the maximum GS ratio of 1 (green) and minimum GS ratio (red).

<https://doi.org/10.1371/journal.pntd.0007871.g005>

Identification of discrepancies between model predictions and experimental data are informative in identifying errors in the model and gene annotation. The predictive capacities of our *M. tuberculosis* GSMN-TB at two different growth rates was validated with experimental data from *M. tuberculosis* grown in the chemostat at the two growth rates, highlighting the importance of the model to predict the metabolic state of *M. tuberculosis*, in experimentally challenging situation such as during infection [16]. But, the experimental validations of GSMN-ML predictions is challenging because of the current limitations to grow *M. leprae* in cultures. Our analysis did predict diaminopimelate, lysine, methionine, formate, tetrahydrofolate, pantothenate, cobalamin and protoporphyrinogen as nutrient supplements for growth and biomass production of GSMN-ML that could be further investigated.

GSMN-ML is applicable for interrogation and interpretation of RNA-seq data using DPA and for exploration, for testing scenarios, for scanning conditions, and for eliminating impossibilities in experimental design. A limitation of DPA analysis is that the predictions rely on the carbon substrate used for the analysis. In this study we chose glucose and glycerol based on our previous studies and evidences of these substrates available to *M. leprae* intracellularly [38]. As more studies and validations are performed in the future, DPA studies can be extended to simulate *M. leprae*'s metabolism for additional nutrients that are identified.

A limitation of GSMN-ML model is that it will test the metabolic abilities of the bacteria alone. Considering that metabolism of *M. leprae* relies primarily on the nutrient provision from the host, and that the pathogen probably does not maximize growth rate during infection, the use of an objective function based exclusively of maximization of growth rate is unlikely to correctly simulate the *M. leprae*'s *in vivo* metabolic state. In the future, it will be useful to construct a host-GSMN-ML dual model and validate with experimental datasets to explore the *M. leprae*'s metabolism in an infected cell.

With new research being available in the future, GSMN-ML will be developed by continuous curation of the metabolic model, improved gene annotation and integration of biochemical and high throughput datasets. The predictive capacities of the GSMN-ML confirmed the expected higher degree of gene essentiality in *M. leprae* compared to *M. tuberculosis*. This is likely to be associated with a lower degree of metabolic robustness in the leprosy pathogen, compared to the TB pathogen; a vulnerability that might provide novel opportunities for new drugs. We believe that GSMN-ML will be a useful model that will be made freely available to researchers aiming to help researchers perform *in silico* DPA analysis of RNA-seq data for metabolic predictions and pre-experimental gene essentiality analysis and growth tests aiding experimental design.

Conclusion

Despite intensive efforts, *M. leprae* has never been grown in the laboratory, severely hampering research on this important pathogen. In this study, we describe the construction of an *in silico* genome scale metabolic model of *M. leprae* GSMN-ML that is able to simulate growth. The model was used to devise a base medium for efforts to grow the pathogen in the laboratory. We also used the model to interrogate gene expression data from *M. leprae* grown in mouse foot pads and identified the active metabolic pathways during intracellular growth of the pathogen. We also found evidence for major differences in the intracellular metabolism of *M. leprae* compared to the related pathogen, *M. tuberculosis*. The availability of the GSMN-ML model provides a new tool for the further exploration of the biology of this important pathogen that could lead to development of new approaches for control of leprosy.

Supporting information

S1 File. Metabolic network of GSMN-TB_2. The network file includes metabolite lists, pathways and reactions for *M. tuberculosis* and the problem file (p-file) including the list of medium supplements and the upper and lower limits of each medium component uptake used for nutrient utilization analysis. The network was constructed using SurreyFBA software. (XLSX)

S2 File. *In silico* medium supplement composition. This composition was used to test and incorporate pseudogenes and orphan reactions that facilitated the growth of GSMN-ML. The p-file includes the list of medium supplements and the upper and lower limits of each medium

component uptake used for nutrient utilization analysis.
(XLSX)

S3 File. Metabolic network of GSMN-ML. The network file includes metabolite lists, pathways and reactions for *M. leprae* and the problem file (p-file) including the list of medium supplements and the upper and lower limits of each medium component uptake used for nutrient utilization analysis. The network was constructed using SurreyFBA software.
(XLSX)

S4 File. P-files used in SurreyFBA to generate growth and biomass production with GSMN-ML and GSMN-TB_2 model. The p-files include the list of nutrients and their uptake rates upper and lower limits.
(XLSX)

S5 File. RNA-seq data for *M. leprae* isolated from mouse foot pads. The file includes expression data and fold changes in expression of *M. leprae* freshly harvested vs. 48h incubation in axenic medium.
(XLSX)

S6 File. List of up-regulated and down-regulated metabolites identified using DPA analysis. Metabolites are listed for GSMN-ML and RNA-seq analysis using glucose and glycerol as sole carbon sources. The p-files include the list of nutrients and their uptake rates- upper and lower limits.
(XLSX)

S7 File. Flux balance analysis of GSMN-ML on glucose and glycerol as carbon sources and gene essentiality predictions. File shows flux variability analysis and gene essentiality analysis for the model. Two media p-files was set for this test. One is to use glycerol as only carbon source and the other is to use glycerol as only carbon source and the other necessary molecules. Table 'FVA' shows results for all reactions with maximized biomass. Table 'media' shows media settings, those substrates whose lower-bound < 0 are allowed to be uptaken. Table 'metabolite_name' shows metabolite IDs and their names. Table 'KO_gene' shows gene essentiality analysis results on the media set in 'Media'. GRateKO: growth rate when the gene knocked out. Ratio: calculated as GRateKO/GRateWT(wildType). Genes are considered as essential If Ratio < 0.01 (marked red).
(XLSX)

S8 File. Flux balance analysis of GSMN-TB_2 on glucose and glycerol as carbon sources and gene essentiality predictions. File shows flux variability analysis and gene essentiality analysis for the model. Two media p-files was set for this test. One is to use glycerol as only carbon source and the other is to use glycerol as only carbon source and the other necessary molecules. Table 'FVA' shows results for all reactions with maximized biomass. Table 'media' shows media settings, those substrates whose lower-bound < 0 are allowed to be uptaken. Table 'metabolite_name' shows metabolite IDs and their names. Table 'KO_gene' shows gene essentiality analysis results on the media set in 'Media'. GRateKO: growth rate when the gene knocked out. Ratio: calculated as GRateKO/GRateWT(wildType). Genes are considered as essential If Ratio < 0.01 (marked red).
(XLSX)

S1 Fig. Viability of *M. leprae* used for RNA-seq analyses. Radiorespirometry (RR) was used to measure oxidation rate of *M. leprae* using palmitic acid as the only carbon source [52]. For the RNaseq experiment the RR data shows that there was no significant difference ($P = 0.209$)

in RR between fresh and following 48hrs incubation in the axenic medium, while after 96hrs of incubation there was a significant difference ($P = 0.018$).
(TIF)

Author Contributions

Conceptualization: Johnjoe McFadden.

Formal analysis: Khushboo Borah, Jacque-Lucca Kearney, Ruma Banerjee, Pankaj Vats, Huihai Wu, Sonal Dahale, Sunitha Manjari Kasibhatla, Rajendra Joshi, Bhushan Bonde, Olabisi Ojo, Ramanuj Lahiri.

Funding acquisition: Sunitha Manjari Kasibhatla, Rajendra Joshi, Diana L. Williams, Johnjoe McFadden.

Investigation: Khushboo Borah, Ramanuj Lahiri, Diana L. Williams, Johnjoe McFadden.

Methodology: Khushboo Borah, Jacque-Lucca Kearney, Ruma Banerjee, Pankaj Vats, Bhushan Bonde, Ramanuj Lahiri, Diana L. Williams, Johnjoe McFadden.

Resources: Sunitha Manjari Kasibhatla, Rajendra Joshi, Ramanuj Lahiri, Diana L. Williams, Johnjoe McFadden.

Software: Johnjoe McFadden.

Supervision: Sunitha Manjari Kasibhatla, Rajendra Joshi, Ramanuj Lahiri, Diana L. Williams, Johnjoe McFadden.

Visualization: Khushboo Borah, Johnjoe McFadden.

Writing – original draft: Khushboo Borah, Johnjoe McFadden.

Writing – review & editing: Khushboo Borah, Jacque-Lucca Kearney, Ruma Banerjee, Huihai Wu, Sonal Dahale, Sunitha Manjari Kasibhatla, Rajendra Joshi, Bhushan Bonde, Olabisi Ojo, Ramanuj Lahiri, Diana L. Williams, Johnjoe McFadden.

References

1. Chaptini C, Marshman G. Leprosy: a review on elimination, reducing the disease burden, and future research. *Leprosy Rev.* 2015; 86: 307–15. PMID: [26964426](#).
2. Malathi M, Thappa DM. Fixed-duration therapy in leprosy: limitations and opportunities. *Indian J of Dermatology* 2013; 58: 93–100. <https://doi.org/10.4103/0019-5154.108029> PMID: [23716796](#).
3. World Health Organisation. *Weekly epidemiological record* 206; 91: 405–420.
4. Salgado CG, Barreto JG, da Silva MB, Frade MAC, Spencer JS. What do we actually know about leprosy worldwide? *The Lancet Infectious Diseases* 2016; 16: 778. [https://doi.org/10.1016/S1473-3099\(16\)30090-1](https://doi.org/10.1016/S1473-3099(16)30090-1) PMID: [27352757](#).
5. Spierings E, De Boer T, Zulianello L, Ottenhoff TH. The role of Schwann cells, T cells and *Mycobacterium leprae* in the immunopathogenesis of nerve damage in leprosy. *Leprosy Review* 2000; 71: 121–9. PMID: [11201869](#).
6. Scollard DM, McCormick G, Allen JL. Localization of *Mycobacterium leprae* to Endothelial Cells of Epineurial and Perineurial Blood Vessels and Lymphatics. *The American Journal of Pathology* 1999; 154:1611–20. [https://doi.org/10.1016/S0002-9440\(10\)65414-4](https://doi.org/10.1016/S0002-9440(10)65414-4) PMID: [10329613](#).
7. Suzuki K, Akama T, Kawashima A, Yoshihara A, Yotsu RR et al. Current status of leprosy: epidemiology, basic science and clinical perspectives. *The Journal of Dermatology* 2012; 39: 121–9. <https://doi.org/10.1111/j.1346-8138.2011.01370.x> PMID: [21973237](#).
8. Guinto RS, Doull JA, de Guia I. Mortality of persons with leprosy prior to sulfone therapy, Cordova and Talisay, Cebu, Philippines. *International journal of leprosy* 1954; 22: 273–84. PMID: [13232775](#).

9. Lofthouse EK, Wheeler PR, Beste DJ, Khatri BL, Wu H, Mendum TA, Kierzek A, McFadden J. Systems-based approaches to probing metabolic variation within the *Mycobacterium tuberculosis* complex. *PLoS One* 2013; 8: e75913. <https://doi.org/10.1371/journal.pone.0075913> PMID: 24098743.
10. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR et al. Massive gene decay in the leprosy bacillus. *Nature* 2001; 409: 1007–11. <https://doi.org/10.1038/35059006> PMID: 11234002.
11. Wheeler P Leprosy—clues about the biochemistry of *Mycobacterium leprae* and its host-dependency from the genome. *World Journal of Microbiology and Biotechnology* 2003; 19: 1–16.
12. Pattyn SR. The problem of cultivation of *Mycobacterium leprae*. A review with criteria for evaluating recent experimental work. *Bulletin of World Health Organisation* 2003; 49: 403–410. PMID: 4212439.
13. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393: 537–44. <https://doi.org/10.1038/31159> PMID: 9634230
14. Gómez-Valero L, Rocha EP, Latorre A, Silva FJ. Reconstructing the ancestor of *Mycobacterium leprae*: The dynamics of gene loss and genome reduction. *Genome Research* 2007; 17: 1178–1185. <https://doi.org/10.1101/gr.6360207> PMID: 17623808.
15. Masaki T, Qu J, Cholewa-Waclaw J, Burr K, Raaum R, Rambukkana A Reprogramming adult schwann cells to stem cell-like cells by leprosy bacilli promotes dissemination of infection. *Cell* 2013; 152: 51–67. <https://doi.org/10.1016/j.cell.2012.12.014> PMID: 23332746
16. Beste DJ, Hooper T, Stewart G, Bonde B, Avignone-Rossa C, Bushell ME, Wheeler P, Klamt S, Kierzek AM, McFadden J. GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biology* 2007; 8: R89. <https://doi.org/10.1186/gb-2007-8-5-r89> PMID: 17521419
17. Jamshidi N, Palsson BØ Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC Systems Biology* 2007; 1: 26. <https://doi.org/10.1186/1752-0509-1-26> PMID: 17555602
18. Rienksma RA, Suarez-Diez M, Spina L, Schaap PJ, Martins dos Santos VA. Systems-level modeling of mycobacterial metabolism for the identification of new (multi-)drug targets. *Seminars in Immunology* 2014; 26: 610–22. <https://doi.org/10.1016/j.smim.2014.09.013> PMID: 25453232.
19. Kavvas ES, Seif Y, Yurkovich JT, Norsigian C, Poudel S, Greenwald WW, Ghatak S, Palsson BØ, Monk JM. Updated and standardized genome-scale reconstruction of *Mycobacterium tuberculosis* H37Rv, iEK1011, simulates flux states indicative of physiological conditions. *BMC Systems Biology* 2018; 12: 25. <https://doi.org/10.1186/s12918-018-0557-y> PMID: 29499714
20. Bonde BK, Beste DJ, Laing E, Kierzek AM, McFadden J. Differential Producibility Analysis (DPA) of Transcriptomic Data with Metabolic Networks: Deconstructing the Metabolic Response of *M. tuberculosis*. *PLOS Computational Biology* 2011; 7: e1002060. <https://doi.org/10.1371/journal.pcbi.1002060> PMID: 21738454
21. Davis GL, Ray NA, Lahiri R, Gillis TP, Krahenbuhl JL, Williams DL et al. Molecular assays for determining *Mycobacterium leprae* viability in tissues of experimentally infected mice. *PLoS Neglected Tropical Disease* 2013; 7: e2404. <https://doi.org/10.1371/journal.pntd.0002404> PMID: 24179562
22. Aronesty E. Comparison of sequencing utility programs. *The Open Bioinformatics Journal* 2013; 7: 1–8.
23. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30: 2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
24. Magoc T, Wood D, Salzberg SL. EDGE-pro: Estimated degree of gene expression in prokaryotic genomes. *Evolutionary Bioinformatics* 2013; 9: 127–36. <https://doi.org/10.4137/EBO.S11250> PMID: 23531787
25. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, et al. The UCSC genome browser database: extensions and updates. *Nucleic Acids Research* 2013; 41: D64–9. <https://doi.org/10.1093/nar/gks1048> PMID: 23155063
26. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26: 841–2. <https://doi.org/10.1093/bioinformatics/btq033> PMID: 20110278
27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 2009; 25: 2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
28. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biology* 2010; 11: R106. <https://doi.org/10.1186/gb-2010-11-10-r106> PMID: 20979621
29. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Research* 2011; 39: D32–7. <https://doi.org/10.1093/nar/gkq1079> PMID: 21071399
30. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. *Science* 1997; 278: 631–7. <https://doi.org/10.1126/science.278.5338.631> PMID: 9381173

31. Karp PD, Paley S, Romero P. The pathway tools software. *Bioinformatics* 2002; 18: S225–S232. https://doi.org/10.1093/bioinformatics/18.suppl_1.s225 PMID: 12169551
32. Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology* 2003; 48: 77–84. <https://doi.org/10.1046/j.1365-2958.2003.03425.x> PMID: 12657046
33. Gevorgyan A, Bushell ME, Avignone-Rossa C, Kierzek AM. SurreyFBA: a command line tool and graphics user interface for constraint-based modeling of genome-scale metabolic reaction networks. *Bioinformatics* 2011; 27: 433–434. <https://doi.org/10.1093/bioinformatics/btq679> PMID: 21148545
34. Bhamidi S, Scherman MS, Jones V, Crick DC, Belisle JT, Brennan PJ, McNeil MR. Detailed structural and quantitative analysis reveals the spatial organization of the cell walls of *in vivo* grown *Mycobacterium leprae* and *in vitro* grown *Mycobacterium tuberculosis*. *The Journal of Biological Chemistry* 2011; 286: 23168–23177. <https://doi.org/10.1074/jbc.M110.210534> PMID: 21555513
35. Matsunaga I, Bhatt A, Young DC, Cheng TY, Eyles SJ, Besra GS et al. *Mycobacterium tuberculosis* pks12 Produces a novel polyketide presented by CD1c to T Cells. *The journal of experimental medicine* 2004; 200: 1559–1569. <https://doi.org/10.1084/jem.20041429> PMID: 15611286
36. Dawes SS, Mizrahi V. DNA metabolism in *Mycobacterium leprae*. *Leprosy Review* 2001; 72: 408–14. PMID: 11826477
37. Marques MA, Berrêdo-Pinho M, Rosa TL, Pujari V, Lemes RM, Lery LM et al. The essential role of cholesterol metabolism in the intracellular survival of *Mycobacterium leprae* is not coupled to central carbon metabolism and energy production. *Journal of bacteriology* 2015; 197: 3698–707. <https://doi.org/10.1128/JB.00625-15> PMID: 26391209
38. Borah K, Girardi KDCDV, Mendum TA, Lery LMS, Beste DJV, Lara FA, et al. Intracellular *Mycobacterium leprae* utilizes host glucose as a carbon source in Schwann cells. *mBio* 2019; 10:e02351–19. <https://doi.org/10.1128/mBio.02351-19> PMID: 31848273.
39. Wheeler PR. Catabolic pathways for glucose, glycerol and 6-phosphogluconate in *Mycobacterium leprae* grown in armadillo tissues. *Journal of General Microbiology* 1983; 129: 1481–95. <https://doi.org/10.1099/00221287-129-5-1481> PMID: 6352857
40. Franzblau SG, Harris EB. Biophysical Optima for metabolism of *Mycobacterium leprae*. *Journal of Clinical Microbiology* 1988; 26: 1124–1129. PMID: 3290244
41. Kaur D, Brennan PJ, Crick DC. Decaprenyl diphosphate synthesis in *Mycobacterium tuberculosis*. *Journal of Bacteriology* 2004; 186: 7564–7570. <https://doi.org/10.1128/JB.186.22.7564-7570.2004> PMID: 15516568
42. Doan TTN, Kim JK, Tran HT, Cha SS, Chung KM, Huynh KH et al. Crystal structures of d-alanine-d-alanine ligase from *Xanthomonas oryzae* pv. *oryzae* alone and in complex with nucleotides. *Archives of Biochemistry and Biophysics* 2014; 545: 92–99. <https://doi.org/10.1016/j.abb.2014.01.009> PMID: 24440607
43. Norman E1, De Smet KA, Stoker NG, Ratledge C, Wheeler PR, Dale JW. Lipid synthesis in mycobacteria: characterization of the biotin carboxyl carrier protein genes from *Mycobacterium leprae* and *M. tuberculosis*. *Journal of Bacteriology* 1994; 176: 2525–31. <https://doi.org/10.1128/jb.176.9.2525-2531.1994> PMID: 7909542
44. Pandey AK, Sassetti CM. Mycobacterial persistence requires the utilization of host cholesterol. *Proceedings of the National Academy of Sciences of the United States of America* 2008; 105: 4376–80. <https://doi.org/10.1073/pnas.0711159105> PMID: 18334639
45. Beste DJ, Nöh K, Niedenführ S, Mendum TA, Hawkins ND, Ward JL, Beale MH, Wiechert W, McFadden J. ¹³C-Flux spectral analysis of host-pathogen metabolism reveals a mixed diet for intracellular *Mycobacterium tuberculosis*. *Chemistry and Biology* 2013; 20: 1012–1021. <https://doi.org/10.1016/j.chembiol.2013.06.012> PMID: 23911587
46. Cruz D, Watson AD, Miller CS, Montoya D, Ochoa MT, Sieling PA et al. Host-derived oxidized phospholipids and HDL regulate innate immunity in human leprosy. *The Journal of Clinical Investigation* 2008; 118: 2917–28. <https://doi.org/10.1172/JCI34189> PMID: 18636118
47. Mattos KA, Oliveira VG, D'Avila H, Rodrigues LS, Pinheiro RO, Sarno EN et al. TLR6-driven lipid droplets in *Mycobacterium leprae*-infected Schwann cells: immunoinflammatory platforms associated with bacterial persistence. *Journal of immunology* 2011; 187: 2548–58. <https://doi.org/10.4049/jimmunol.1101344> PMID: 21813774
48. Elamin AA, Stehr M, Singh M. Lipid droplets and *Mycobacterium leprae* infection. *Journal of Pathogens* 2012;361374. <https://doi.org/10.1155/2012/361374> PMID: 23209912
49. Mattos KA, D'Avila H, Rodrigues LS, Oliveira VG, Sarno EN, Atella GC et al Lipid droplet formation in leprosy: Toll-like receptor-regulated organelles involved in eicosanoid formation and *Mycobacterium leprae* pathogenesis. *Journal of Leukocyte Biology* 2010; 87: 371–84. <https://doi.org/10.1189/jlb.0609433> PMID: 19952355

50. Medeiros RC, Girardi KD, Cardoso FK, Mietto BS, Pinto TG, Gomez LS et al. Subversion of Schwann cell glucose metabolism by *Mycobacterium leprae*. *The journal of Biological Chemistry* 2016; 291: 21375–21387. <https://doi.org/10.1074/jbc.M116.725283> PMID: 27555322
51. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR et al. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature*. 2015; 528:99–104. <https://doi.org/10.1038/nature15765> PMID: 26632588
52. Lahiri R, Randhawa B, Krahenbuhl J. Application of a viability-staining method for *Mycobacterium leprae* derived from the athymic (nu/nu) mouse foot pad. *Journal of Medical Microbiology* 2005; 54:235–42. <https://doi.org/10.1099/jmm.0.45700-0> PMID: 15713606.