

## Aberystwyth University

### *Defining the Genes Required for Survival of Mycobacterium bovis in the Bovine Host Offers Novel Insights into the Genetic Basis of Survival of Pathogenic Mycobacteria*

Gibson, AJ; Stiens, J; Passmore, IJ; Faulkner, V; Miculob, J; Willcocks, S; Coad, M; Berg, S; Werling, D; Wren, BW; Nobeli, I; Villarreal-Ramos, B; Kendall, Sharon L.

*Published in:*  
mBIO

*DOI:*  
[10.1128/mbio.00672-22](https://doi.org/10.1128/mbio.00672-22)  
[10.5281/zenodo.6576716](https://doi.org/10.5281/zenodo.6576716)

*Publication date:*  
2022

*Citation for published version (APA):*

Gibson, A.J., Stiens, J., Passmore, I.J., Faulkner, V., Miculob, J., Willcocks, S., Coad, M., Berg, S., Werling, D., Wren, B.W., Nobeli, I., Villarreal-Ramos, B., & Kendall, S. L. (2022). Defining the Genes Required for Survival of Mycobacterium bovis in the Bovine Host Offers Novel Insights into the Genetic Basis of Survival of Pathogenic Mycobacteria. *mBIO*, 13(4). <https://doi.org/10.1128/mbio.00672-22>, <https://doi.org/10.5281/zenodo.6576716>

#### **Document License** CC BY

#### **General rights**

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400  
email: [is@aber.ac.uk](mailto:is@aber.ac.uk)



# Defining the Genes Required for Survival of *Mycobacterium bovis* in the Bovine Host Offers Novel Insights into the Genetic Basis of Survival of Pathogenic Mycobacteria

Amanda J. Gibson,<sup>a\*</sup> Jennifer Stiens,<sup>b</sup> Ian J. Passmore,<sup>c</sup> Valwynne Faulkner,<sup>a§</sup> Josephous Miculob,<sup>a</sup> Sam Willcocks,<sup>c</sup> Michael Coad,<sup>d</sup> Stefan Berg,<sup>d</sup> Dirk Werling,<sup>a</sup>  Brendan W. Wren,<sup>c</sup> Irene Nobeli,<sup>b</sup> Bernardo Villarreal-Ramos,<sup>d</sup>  Sharon L. Kendall<sup>a</sup>

<sup>a</sup>Centre for Emerging, Endemic and Exotic Diseases, Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, United Kingdom

<sup>b</sup>Institute of Structural and Molecular Biology, Biological Sciences, Birkbeck, University of London, London, United Kingdom

<sup>c</sup>London School of Hygiene and Tropical Medicine, London, United Kingdom

<sup>d</sup>Animal and Plant Health Agency, Addlestone, Surrey, United Kingdom

Amanda Gibson and Jennifer Stiens contributed equally to this work. Author order was determined alphabetically.

**ABSTRACT** Tuberculosis has severe impacts on both humans and animals. Understanding the genetic basis of survival of both *Mycobacterium tuberculosis*, the human-adapted species, and *Mycobacterium bovis*, the animal-adapted species, is crucial to deciphering the biology of both pathogens. There are several studies that identify the genes required for survival of *M. tuberculosis in vivo* using mouse models; however, there are currently no studies probing the genetic basis of survival of *M. bovis in vivo*. In this study, we utilize transposon insertion sequencing in *M. bovis* AF2122/97 to determine the genes required for survival in cattle. We identify genes encoding established mycobacterial virulence functions such as the ESX-1 secretion system, phthiocerol dimycocerosate (PDIM) synthesis, mycobactin synthesis, and cholesterol catabolism that are required *in vivo*. We show that, as in *M. tuberculosis* H37Rv, *phoPR* is required by *M. bovis* AF2122/97 *in vivo* despite the known defect in signaling through this system. Comparison to studies performed in species that are able to use carbohydrates as an energy source, such as *M. bovis* BCG and *M. tuberculosis*, suggests that there are differences in the requirement for genes involved in cholesterol import (*mce4* operon) and oxidation (*hsd*). We report a good correlation with existing mycobacterial virulence functions but also find several novel virulence factors, including genes involved in protein mannosylation, aspartate metabolism, and glycerol-phosphate metabolism. These findings further extend our knowledge of the genetic basis of survival *in vivo* in bacteria that cause tuberculosis and provide insight for the development of novel diagnostics and therapeutics.

**IMPORTANCE** This is the first report of the genetic requirements of an animal-adapted member of the *Mycobacterium tuberculosis* complex (MTBC) in a natural host. *M. bovis* has devastating impacts on cattle, and bovine tuberculosis is a considerable economic, animal welfare, and public health concern. The data highlight the importance of mycobacterial cholesterol catabolism and identify several new virulence factors. Additionally, the work informs the development of novel differential diagnostics and therapeutics for TB in both human and animal populations.

**KEYWORDS** Tn-seq, gene essentiality, One Health, tuberculosis, virulence factors

**B**acteria belonging to the *Mycobacterium tuberculosis* complex (MTBC) have devastating impacts on both animal and human populations. *Mycobacterium bovis*, an animal-adapted member of the MTBC and one of the main causative agents of bovine tuberculosis (bTB), remains endemic in some high-income settings despite the implementation of a test-and-slaughter policy. In low- and middle-income settings, the presence of bTB in livestock combined with the absence of rigorous control measures contributes to

**Editor** Sabine Ehrh, Weill Cornell Medical College

**Copyright** © 2022 Gibson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Sharon L. Kendall, skendall@rvc.ac.uk.

\*Present address: Amanda J. Gibson, Centre of Excellence for Bovine Tuberculosis, IBERS, Aberystwyth University, Penllys, Aberystwyth, Ceredigion, United Kingdom.

§Present address: Valwynne Faulkner, Systems Chemical Biology of Infection and Resistance Laboratory, The Francis Crick Institute, London, United Kingdom.

◇Present address: Bernardo Villarreal-Ramos, Centre of Excellence for Bovine Tuberculosis, IBERS, Aberystwyth University, Penllys, Aberystwyth, Ceredigion, United Kingdom.

The authors declare no conflict of interest.

**Received** 5 April 2022

**Accepted** 22 June 2022

**Published** 14 July 2022

the risk of zoonotic transmission (1, 2). Control measures based on cattle vaccination utilize the live attenuated vaccine *M. bovis* BCG, but the efficacy of this vaccine remains low in field situations (3, 4). In addition to vaccines, the development of diagnostic tools for the identification of infected individuals is crucial for the management of transmission. Vaccination with *M. bovis* BCG sensitizes animals to the diagnostic tuberculin skin test; therefore, sensitive and specific differentiating diagnostic strategies are a current imperative (5, 6).

The increased accessibility of whole-genome fitness screens has allowed the assessment of the impacts of the loss of gene function on bacterial survival (7). Such screens have been invaluable in identifying novel drug targets or candidates for the generation of new live attenuated vaccines in a number of bacterial pathogens, including *M. tuberculosis* (8–13). Studies utilizing whole-genome transposon mutagenesis to examine gene fitness *in vivo* in *M. tuberculosis* H37Rv have been limited to mouse models (8–10). These models do not faithfully replicate the granulomatous pathology associated with TB, nor do mice contain the same repertoire of CD1 molecules expressed by bovine T cells required to present mycobacterial lipid antigens (14). Whole-genome transposon mutagenesis screens utilizing non-human primates are limited because screening is restricted to smaller mutant pools (15). To date, transposon insertion sequencing (Tn-seq)-based studies in the context of bTB in cattle have only been performed using *M. bovis* BCG strains (16, 17).

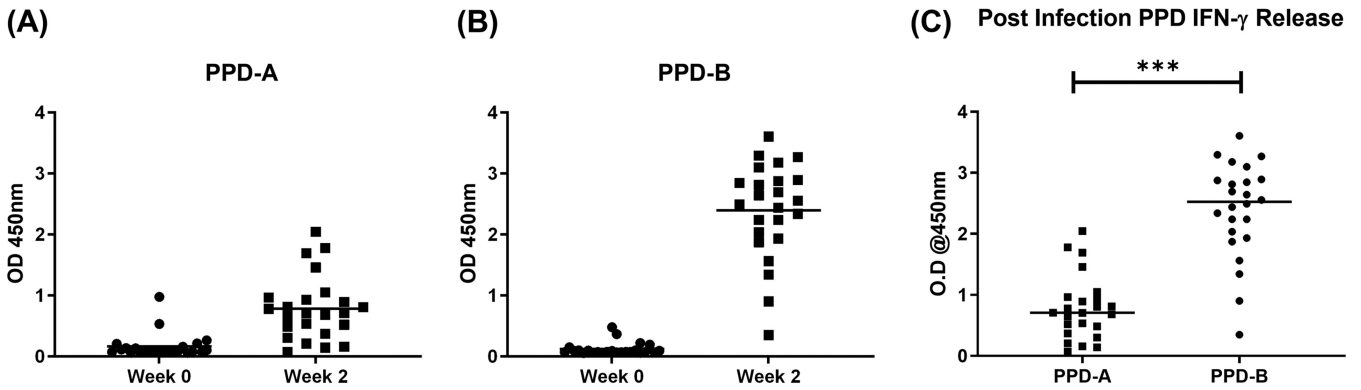
In this study, we use Tn-seq to determine the genes required for survival of *M. bovis* AF2122/97 directly in cattle. We show that genes involved in the biosynthesis of phthiocerol dimycocerosates (PDIMs), the ESX-1 secretion system, cholesterol catabolism, and mycobactin biosynthesis are essential for survival in cattle, corroborating current knowledge of gene essentiality in members of the MTBC (8–10, 16, 17). We identify differences in the requirement for genes involved in cholesterol transport and oxidation in the fully virulent *M. bovis* AF2122/97 strain. We also identify several novel genes required for survival *in vivo* that have not been previously described in members of the MTBC.

## RESULTS AND DISCUSSION

**Generation and sequencing of the input library.** We generated a transposon library in *M. bovis* AF2122/97 using the phagemid system as previously described (18, 19). Sequencing of the input library showed that transposon insertions were evenly distributed around the genome, and 27,419 of the permissible 66,931 thymine-adenine dinucleotide (TA) sites contained an insertion representing an insertion density of ~41% (see input library Fig. S1 and Table S1 at <https://doi.org/10.5281/zenodo.6598446>). The *M. bovis* AF2122/97 genome has 3,989 coding sequences, and insertions were obtained in 3,319 of these; therefore, the input library contained insertions in 83% of the total coding sequences.

***Mycobacterium bovis*-specific immune responses were observed in cattle.** Twenty-four clinically healthy calves of approximately 6 months of age were inoculated with the library through the endobronchial route. Infection was monitored by interferon gamma (IFN- $\gamma$ ) release assay (IGRA) at the time of inoculation and 2 weeks postinfection. *M. bovis*-specific immune responses were observed for all study animals at 2 weeks postinfection (Fig. 1A and B). Each animal presented a very low background of circulating IFN- $\gamma$  together with a statistically significant (\*\*\*) ( $P \leq 0.001$ ) increase in IFN- $\gamma$  release in response to purified protein derivative from *M. bovis* (PPD-B) compared to PPD from *M. avium* (PPD-A) antigens (Fig. 1C). This indicates that infection with the library was successfully established in the cattle.

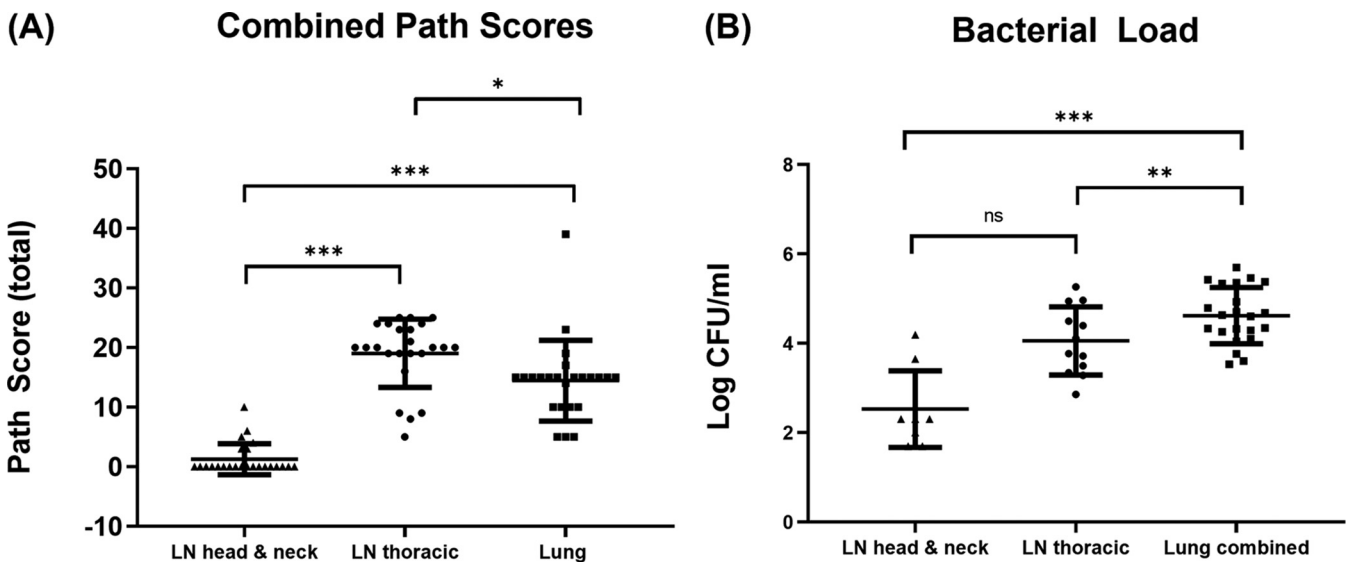
**Pathology associated with infection was greater in the lung and thoracic lymph nodes.** Animals were culled at 6 weeks postinfection. Lung sections and upper (head and neck) and lower (thoracic) respiratory tract-associated lymph nodes were examined for gross lesions. Lesions typical of *M. bovis* infection were observed in the tissues examined (see Table S2). Pathology scores are shown in Fig. 2A. Greater pathology was observed in the lung and thoracic lymph nodes than in the head and neck lymph nodes.



**FIG 1** bTB-specific IFN gamma release in cattle infected with the *M. bovis* Tn library. (A and B) Blood was collected from all 24 animals on the day of infection and 2 weeks later. No response was detected to either PPD-A or PPD-B antigen stimulation prior to infection (week 0). (C) All animals presented a significant and specific response to PPD-B compared to PPD-A as determined by a paired *t* test using GraphPad Prism. \*\*\*, *P* ≤ 0.001.

**Higher bacterial loads were associated with the lung and thoracic lymph nodes.** Bacterial counts were highest in lesions derived from the lung compared to those from the thoracic lymph nodes and head and neck lymph nodes (Fig. 2B). The lowest bacterial counts were observed within the head and neck lymph nodes. However, this was not significant compared to thoracic lymph nodes. The volume of each macerate varied depending on lesion size. Considering macerate volume, average bacterial loads of 10<sup>7</sup>, 10<sup>6</sup>, and 10<sup>5</sup> were recovered from lesions from samples of the lungs, thoracic lymph nodes, and head and neck lymph nodes, respectively.

**Recovery and sequencing of *in vivo*-selected transposon libraries.** In order to recover the Tn library from harvested tissue, ~10<sup>5</sup> to 10<sup>6</sup> CFU from lungs and thoracic lymph nodes were plated. Samples from 4 cattle were lost due to fungal contamination; therefore, samples processed represent 20 cattle. Lung samples were plated from all 20 animals, and thoracic lymph nodes samples were plated from 6 cattle. Bacteria were grown for 4 to 6 weeks before harvesting for genomic DNA extraction and sequencing (see Table S1 for assignment of sequencing files). The insertion densities of the output libraries were



**FIG 2** Tissue pathology and bacterial load in tissue sites. (A) Six weeks after infection, animals were subjected to postmortem examination. Gross pathology and evidence of TB-like granulomas lesions were scored. Data presented are the mean across animals of the total scores for each tissue group from 24 animals ± the standard deviation. Lung and thoracic lymph nodes were observed to contain the highest pathology compared to head and neck lymph nodes. (B) For bacterial load estimation, aliquots of macerates were plated onto modified 7H11 agar containing kanamycin. Colonies were counted after 3 to 4 weeks growth. Data are presented as mean log<sub>10</sub> CFU/mL per collected tissue group ± standard deviation. Lung tissue contained the highest bacterial burden compared to thoracic and head and neck lymph nodes as determined by one-way analysis of variance (ANOVA) using GraphPad Prism. \*, *P* = 0.01; \*\*, *P* = 0.002; \*\*\*, *P* ≤ 0.001.

compared to the input library for each sample (Fig. S2). Libraries recovered from lung lesions from 20 different cattle contained an average of 14,456 unique mutants, and those recovered from the thoracic lymph nodes contained an average of 16,210 unique mutants. Given that the input library contained 27,419 unique mutants, this meant that there was an ~40 to 50% reduction in insertion density in the output libraries compared with the input. Good coverage of coding sequences (CDSs) was maintained, as the output libraries still contained insertions in (on average) 68 to 70% of the open reading frames. Given the loss of diversity of the individual output libraries, we pooled samples from the lungs and separately from the lymph nodes. The insertion densities of the pooled samples from the lungs were 33,039/66,931 permissible TA sites, and from the nodes were 25,072/66,931 permissible TA sites. This represented ~50% and ~38% saturation. Therefore, using this approach, the diversity of the input pools was maintained.

Calculation of the  $\log_2$  fold change in the read counts between the input and output libraries allowed a measurement of the impact of the insertions on the survival of mutants in cattle. In order to determine statistically significant changes in the representation of mutants between the input and output libraries, we analyzed pooled samples from the lungs (20 cattle) and pooled samples from the lymph nodes (6 cattle). However, in recognition that cattle are genetically more heterogenous than standard laboratory mice, we have included data where we have calculated the  $\log_2$  fold change between the output library and the input library for individual samples in addition to the pooled data set. The entire data set is shown in Table S3, and a volcano plot is shown in Fig. S3.

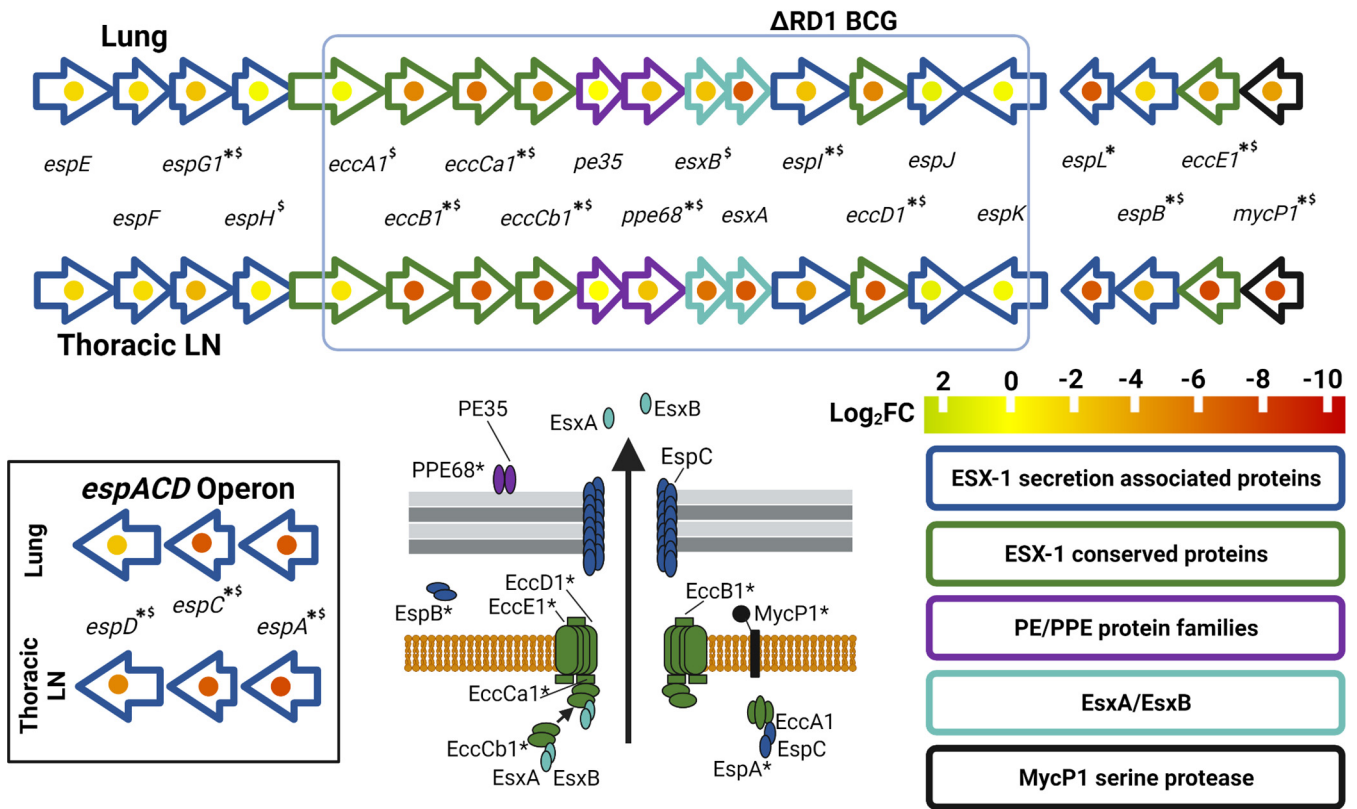
A comparison of the mean  $\log_2$  fold change between lung and lymph node samples showed good correlation (Spearman's rho, 0.88;  $P < 2.2e-16$ ) (Fig. S4). TRANSIT resampling was performed to compare the composition of the mutant population in the lungs and thoracic lymph nodes of paired cattle; it was also applied to compare all the thoracic lymph nodes with the lungs of all cattle samples. TRANSIT analysis did not find any statistically significant differences, indicating that there were no differences in mutant composition between the tissue sites.

Using an adjusted  $P$  value cutoff of  $\leq 0.05$  and a  $\log_2$  fold change of  $-1.5$  in either lungs or lymph node, insertions in 300 genes caused significant attenuation in cattle. Of these genes, 220 had been previously described as being required *in vivo* in *M. tuberculosis* H37Rv in standard mouse models through the use of whole-genome Tn screens representing ~73% overlap with the previous literature (8–10). These genes are given in Table S3 ("Significant genes" tab). No insertion mutants were significantly over-represented in the library. Although *Mb0025* was over-represented in both lungs and nodes ( $\log_2$  fold change, 7 to 8 in the pooled analyses), significant cutoffs were not reached, and this may be reflective of a lower number of TA sites in this gene, which limits statistical power. *Mb0025* overlaps with *Mb0024* and is the result of a frameshift mutation in the AF2122/97 genome. This mutation is also found in other assembled *M. bovis* genomes, and we could find no evidence for lack of conservation of this frameshift mutation in global collections (20). *Mb0024* and *Mb0025* represent orthologs of the 5' and 3' ends of *Rv0024*, respectively, which is annotated as a p60 homologue involved in cell-to-cell spread (21). The functionality of *Mb0024* and *Mb0025*, or the impact on the transposon insertion, is not known.

**Comparison with mutations known to cause attenuation in the MTBC.** Insertions in the RD1-encoded ESX-1 type VII secretion system secreting virulence factors and immunodominant antigens EsxA (CFP-10) and EsxB (ESAT-6) are expected to cause attenuation (22). The impacts of insertions in this region are summarized in Fig. 3 but are also available in Table S3 ("RD regions" tab) and Fig. S5. Insertions in genes encoding the structural components of the apparatus (*eccB1*, *eccCa1*, *eccCb1*, and *eccD1*) were significantly attenuated according to the criteria (adjusted  $P$  value cutoff of  $\leq 0.05$  and a  $\log_2$  fold change of  $-1.5$ ). Insertions in *eccA1*, which also codes for a structural component of the apparatus, were not attenuating despite good insertion saturation in this gene. This is supported by the work of others who have shown that deletion of *eccA1* in *Mycobacterium marinum* leads to only a partial secretion defect (23). There were no impacts seen due to insertions in



### ESX-1 locus (Mb3894-3913)

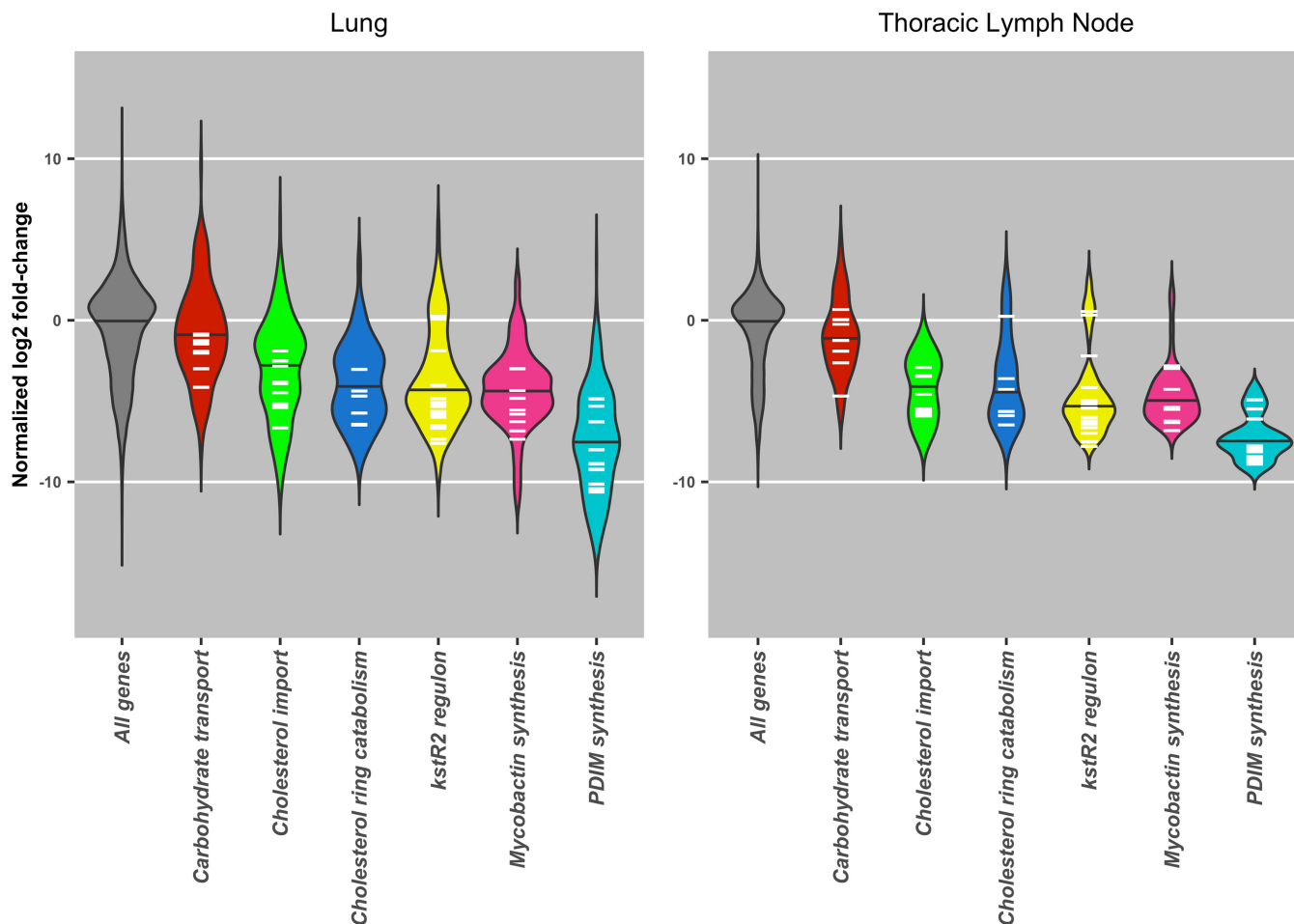


**FIG 3** Fold changes caused by transposon insertions in the ESX-1 secretion system in the lungs and lymph nodes of infected cattle. Asterisks indicate that genes were significantly attenuated in pooled output samples (adjusted  $P \leq 0.05$  in either lungs or node). A dollar sign indicates those genes that are attenuated in H37Rv in mice models. The genes are grouped according to function as indicated by the color scheme. The  $\log_2$  fold changes are indicated on a yellow-to-red scale and presented as a dot in the center of the gene.

accessory genes *espJ*, *espK*, and *espH*. The lack of attenuation seen in *espK* mutants is supported by other studies showing that this gene is dispensable for secretion through the apparatus and is not required for virulence of *M. bovis* in guinea pigs (24, 25). Insertions in *esxA* and *esxB* resulted in severe attenuation ( $\log_2$  fold change of  $-6$  to  $-7.5$ ) but did not reach significance cutoffs (adjusted  $P \leq 0.05$ ). This is likely to be due to the small number of TAs in these genes, which makes it challenging to measure mutant frequency, despite utilizing a pooled approach.

High levels of attenuation seen were in genes involved in the synthesis of the cell wall virulence lipids PDIMs (*ppsABCDE* and *mas* with  $\log_2$  fold changes of  $-7$  to  $7.5$  [Table S3, "Mycolipids" tab]). PDIM synthesis is well known to be required for the survival of *M. tuberculosis* and *M. bovis* in mice and guinea pigs (26, 27). MmpL7 is involved in PDIM transport, and there is evidence that it is phosphorylated by the serine-threonine kinase PknD (28). PknD-MmpL7 interactions are thought to be perturbed in *M. bovis* AF2122/97, as *pknD* is split into two coding sequences in the bovine pathogen by a frameshift mutation (29). The data presented here suggest that MmpL7 is required *in vivo* in cattle despite the frameshift mutation in *pknD*.

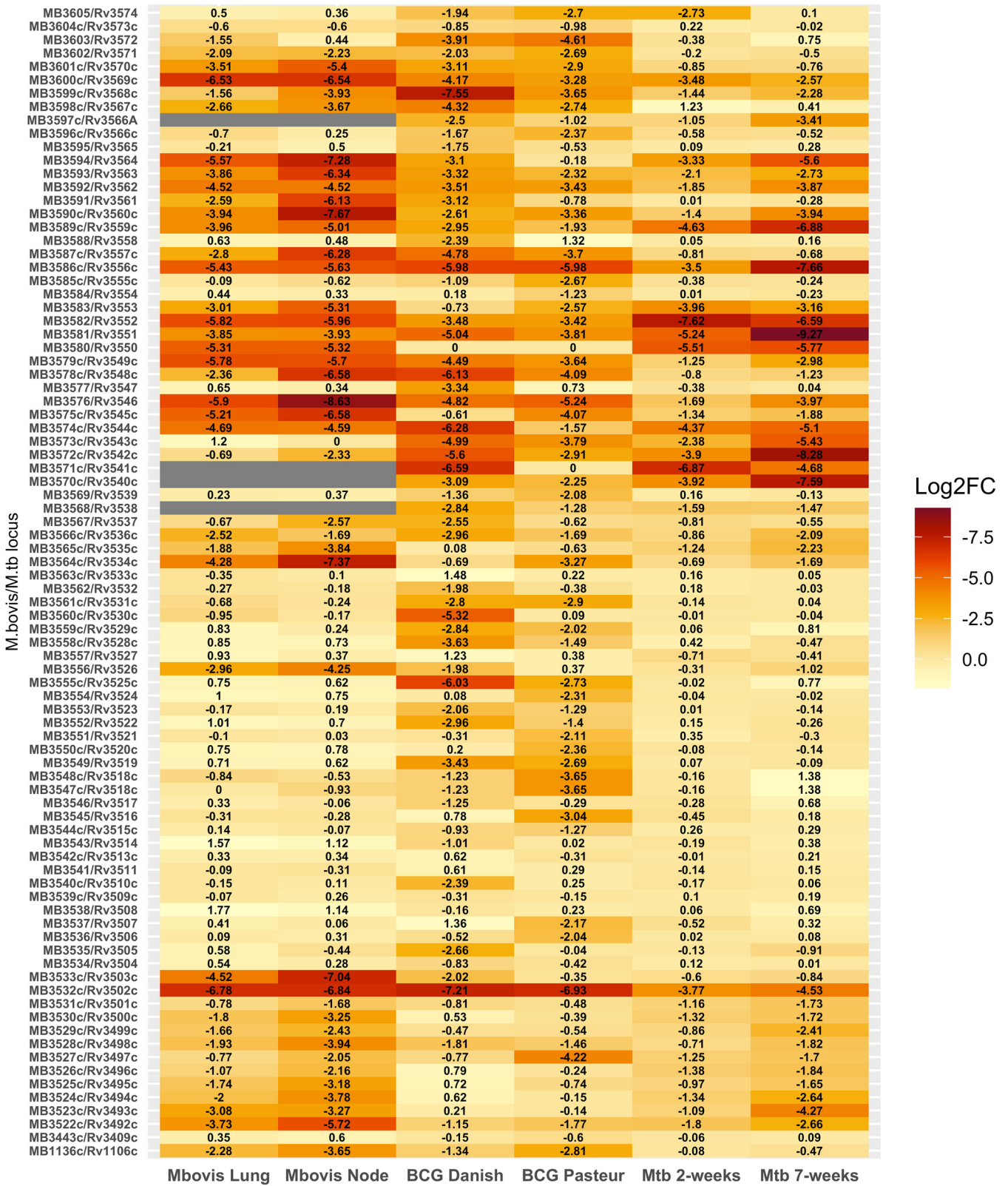
Iron restriction is thought to be a mechanism by which the host responds to mycobacterial infection, although different cellular compartments may be more restrictive than others (30). Insertion in many of the genes involved in mycobactin synthesis (*Mb2406-Mb2398c*, *mbtJ-mbtH*) were attenuating in cattle (Fig. 4; Table S3, "Mycobactin synthesis" tab). As mycobactin is required for the acquisition of iron, this confirms that, like other members of the MTBC, *M. bovis* needs to scavenge iron from the host for survival (10, 16).



**FIG 4** Violin plot of normalized  $\log_2$  fold changes in gene insertions recovered from bovine lung or thoracic lymph node tissue samples in selected gene groups. Black bars indicate overall median of normalized  $\log_2$  fold change among genes in grouping. White bars indicate mean  $\log_2$  fold change for each gene in the group across all samples in either lung or lymph node tissue.

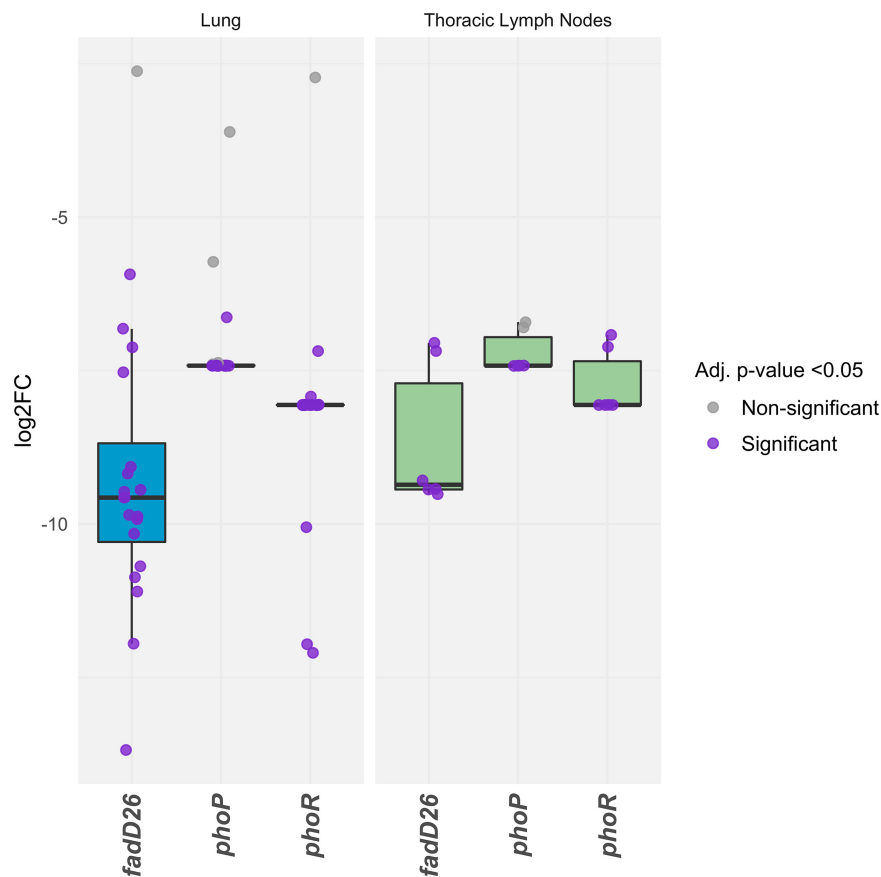
The role of cholesterol catabolism in *M. tuberculosis* is well documented, and it is required for both energy generation and manipulation of the immune response (31–33). Cholesterol uptake is mediated by the Mce4 transporter coded by the *mce4* operon *Rv3492c-Rv3501c* (*Mb3522c-MB3531c*) (34, 35). It has been suggested that an alternative cholesterol acquisition pathway operates in *M. bovis* BCG Danish, as, unlike insertions in genes in the downstream catabolic pathway, insertions in the *mce4* operon do not result in attenuation in this strain (16). In contrast, our study shows that cholesterol transport via the Mce4 transporter is required in *M. bovis* AF2122/97 (Fig. 4; Table S3, “*mce4* operons” tab). Interestingly, the significance cutoff (adjusted  $P \leq 0.05$ ) was only reached in the pooled lymph node samples, but it is difficult to say whether this indicates the requirement for Mce4 only occurs in the lymph nodes or if this is due to stochastic effects. The lymph nodes are the site of the engagement with the adaptive immune system and are the site of persistence for *M. tuberculosis* in nonhuman primates (36). The requirement for the Mce4 transporter corroborates work performed in *M. tuberculosis*, where Mce4 has been shown to be required for persistence in mice (8, 34). A comparison of the fitness impact on genes in the cholesterol catabolic pathway in *M. bovis* AF2122/97, *M. bovis* BCG Danish and Pasteur, and *M. tuberculosis* H37Rv is given in Fig. 5.

Early stages of cholesterol catabolism involve the oxidation of cholesterol to cholestone, a reaction catalyzed by the  $3\beta$ -hydroxysteroid dehydrogenase (*hsd*) encoded by *Rv1106c/Mb1136c*. *Rv1106c* is not required for the survival of *M. tuberculosis* in immunocompetent mice or guinea pigs, and this is thought to be due to the availability of other



**FIG 5** Comparison of reported log<sub>2</sub> fold change in *M. bovis*, *M. bovis* BCG, and *M. tuberculosis* transposon insertion sequencing experiments for orthologous genes in the cholesterol catabolic pathway. Greatest attenuation (most negative log<sub>2</sub> fold change) is colored darkest red. Studies used for comparison include Mendum et al. (24) and Bellerose et al. (9). Gray bars represent genes for which there is no information, as they were either essential (ES) or had a growth disadvantage (GD) in the input library or had less than 5 insertions in any TA site in any sample (input and all output).





**FIG 6** Fold changes caused by transposon insertions in *phoP*, *phoR*, and *fadD26* in the lungs and lymph nodes of infected cattle. Samples with adjusted *P* values (Benjamini-Hochberg FDR corrected) of  $<0.05$  are indicated with purple points.

carbon sources, including glycolytic substrates, *in vivo* (8–10, 37). Insertions in *Mb1136c* in *M. bovis* AF2122/97 were attenuating (Table S3, “Cholesterol catabolism” tab), and this may be reflective of the inability of *M. bovis* AF2122/97 to utilize glycolytic substrates due to a disrupted pyruvate kinase (*pykA*) gene (38, 39). In a recent extended Tn screen utilizing diverse mouse genotypes, Tn insertions in *hsd* caused reduced fitness in a small panel of selected genotypes indicates there may be some host genetic component to the requirement for cholesterol oxidation by *hsd* (40). Given the potential for the use of host cholesterol metabolites, specifically cholestenone, as diagnostic biomarkers, this observation might have applications in the development of diagnostics (41).

**Genes that are differentially expressed between *Mycobacterium bovis* AF2122/97 and *Mycobacterium tuberculosis* H37Rv.** Several studies have identified key expression differences between *M. bovis* AF2122/97 and *M. tuberculosis* H37Rv (29, 42, 43). We examined the data set for insights into the role of differentially expressed genes and transcriptional regulators during infection. One important regulatory system in *M. tuberculosis* H37Rv is the two-component regulatory system PhoPR, and deletions in the *phoPR* genes alongside *fadD26* are attenuating mutations in the live vaccine MTBVAC (44–46). Our data show that insertions in both *phoPR* and *fadD26* caused attenuation (Fig. 6 and Table S3, “*phoPR* regulon” and “Mycolipids” tabs). This reinforces the role of this system in virulence despite the presence of a single nucleotide polymorphism (SNP) in the sensor kinase *phoR* that impacts signaling through the system in *M. bovis* AF2122/97 (44). However, care should be taken when using the data set to make inferences of the genetic requirements in field strains. For instance, *fadD26* contains nonsynonymous SNPs in global *M. bovis* collections (20). Signal potentiation via *phoR* is required for secretion of ESAT-6 through the ESX-1 secretory system, and *M. bovis* AF2122/97 is known to have compensatory mutations elsewhere in the genome, e.g.,

in the *espACD* operon that restores ESAT-6 secretion in the face of a deficient signaling system (44, 47). Our data also show that Tn insertions in *espA*, *espB*, and *espC* (required for ESAT-6 secretion) and in *mprA*, a transcriptional regulator of that operon (48), caused attenuation, emphasizing the relevance of ESAT-6 as a virulence factor.

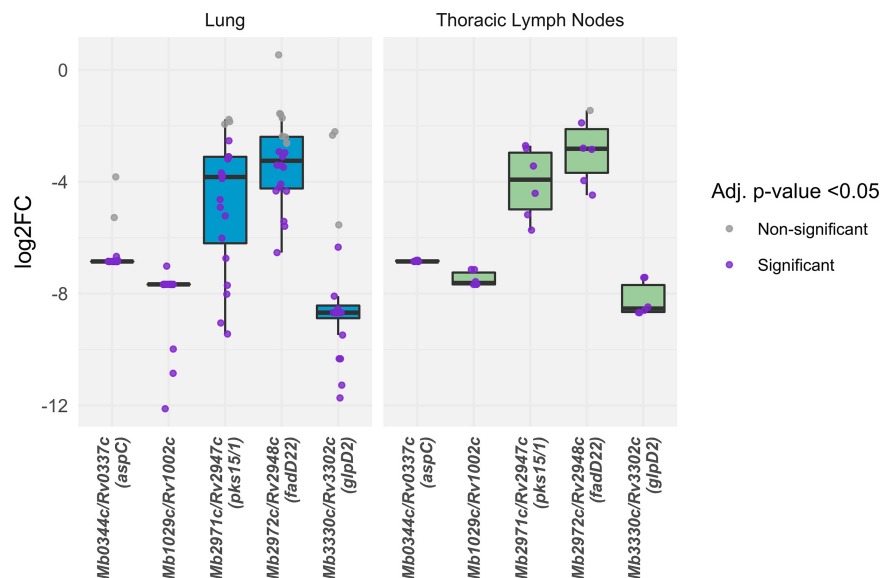
Studies comparing differences in expression during *in vitro* growth between *M. bovis* AF2122/97 and *M. tuberculosis* H37Rv show that genes involved in sulfolipid (SL-1) biosynthesis are expressed at lower levels in *M. bovis* AF2122/97 than *M. tuberculosis* H37Rv (29, 42). Interestingly, insertions in genes involved in SL-1 biosynthesis (*Mb3850* to *Mb3856*) were not attenuating *in vivo* (Table S3, "Mycolipids" tab), reinforcing the lack of importance of SL-1 for *M. bovis* AF2122/97 *in vivo*, at least at the stages of infection studied here.

One of the most highly attenuated insertions occurred in *Mb0222/Rv0216*. This gene has been shown to be highly (>10-fold) overexpressed in *M. bovis* AF2122/97 compared with *M. tuberculosis* H37Rv, but the physiological function of this gene is not currently known. The secreted antigens MPB70 and MPB83, encoded by *Mb2900* and *Mb2898*, are also overexpressed in *M. bovis* AF2122/97 and play a role in host-specific immune responses; however, insertions in these genes did not cause attenuation *in vivo* in our data set (49).

**Novel attenuated mutations.** We identified 80 genes that were required for survival of *M. bovis* AF2122/97 in cattle that had not been previously described as being essential *in vivo* through transposon mutagenesis screens of *M. tuberculosis* in standard laboratory mouse models (8–10) (see Table S3, "Significant gene" tab). Insertions in some of these genes have been shown to cause attenuation in standard mouse models in *M. tuberculosis* through the use of single mutants (50–53). While writing this publication, a large-scale Tn-seq study that utilized over 120 *M. tuberculosis* libraries and several diverse mouse genotypes was performed (the collaborative cross-mouse panel [54]). This study captured the genes required for survival under a greater variety of host microenvironments than those performed in the standard mouse models (40). In that study, a larger subset of "adaptive" virulence genes that are required in a small subset of mice were identified, including those genes that were required in immunodeficient mice. Interestingly, insertions in *hsd* were attenuating only in immunodeficient mice in this study. A direct comparison of our data set with the study by Smith et al. revealed that a further 15 genes were shown to be required in at least two mice strains, hence classified as "adaptive" virulence genes with specific host genetic components contributing to fitness. There remains a subset of 65 genes that are required for optimal fitness of *M. bovis* AF2122/97 during infection of cattle that have not been previously identified as required for survival of *M. tuberculosis* in any mice by using transposon mutagenesis screens.

Genes required for phenolic glycolipid synthesis. Insertions in *Mb2971c/Rv2947c* (*pks15/1*) and in *Mb2972c/Rv2948c* (*fadD22*) were attenuating in *M. bovis* AF2122/97 (Fig. 7), but these genes are not required *in vivo* in *M. tuberculosis* H37Rv, including in the extended panel of mouse genotypes (8–10, 40). Both *pks15/1* and *fadD22* are involved in the early stages of synthesis of phenolic glycolipids (PGLs) and are involved in virulence (55). The requirement for these genes in *M. bovis* AF2122/97 but not in *M. tuberculosis* H37Rv is consistent with the observation that Tn-seq studies in *M. tuberculosis* are often carried out using lineage 4 strains (H37Rv and CDC1551) that harbor a frameshift mutation in the *pks15/1* gene, which renders them unable to synthesize PGLs. This removes the requirement for these genes *in vivo* in lineage 4 strains of *M. tuberculosis*. *pks15/1* has been previously reported to be required for survival of an *M. bovis* isolated in New Zealand in a guinea pig model of infection (56).

Genes that are involved in post-translational modifications such as glycosylation were attenuating in *M. bovis* AF2122/97 but not required *in vivo* in *M. tuberculosis* H37Rv. *Rv1002c* is thought to add mannose groups to secreted proteins, and overexpression of this protein in *M. smegmatis* was recently shown to enhance survival *in vivo* and inhibit proinflammatory cytokine production (57). The substrates of the protein mannosyltransferase are thought to be several secreted lipoproteins, including LpqW,



**FIG 7** Fold changes caused by transposon insertions in *pks15/1*, *fadD22*, *Rv1002c*, *aspC*, and *glpD2* in the lungs and lymph nodes of infected cattle. Samples with adjusted *P* values (Benjamini-Hochberg FDR corrected) of <0.05 are indicated with purple points.

which is involved in the insertion of the virulence lipid LAM at the mycobacterial cell surface (57, 58).

Finally, insertions in *aspC* and *glpD2* were attenuating in *M. bovis* AF2122/97 but not required *in vivo* in *M. tuberculosis* H37Rv. An examination of the *in vitro* essentiality literature showed that both of these genes are essential *in vitro* in *M. tuberculosis* H37Rv when grown on standard media but not in *M. bovis* AF2122/97 (11, 12, 19, 59). Information regarding *aspC* and *glpD2* from Tn-seq approaches is likely to be lacking in *M. tuberculosis* H37Rv because Tn mutants in these genes will not be represented in the input pool. The absence of insertion mutants in these genes in the most recent large-scale *M. tuberculosis* H37Rv Tn-seq study supports this (40). *aspC* (Mb0344c/Rv0337c) is an aspartate aminotransferase involved in the utilization of amino acids (aspartate) as a nitrogen source (60). This provides evidence that *M. bovis* utilizes aspartate *in vivo*, as is observed in *M. tuberculosis* (61). *glpD2* (Mb3330c/Rv3302c) is a membrane-bound glycerol-phosphate dehydrogenase. In *Escherichia coli*, *glpD2* is an essential enzyme, functioning at the central junction of respiration, glycolysis, and phospholipid biosynthesis, and catalyzes the oxidation of dihydroxyacetone phosphate (DHAP) from glycerol-3-phosphate, resulting in the donation of electrons to the electron transport chain (62). Its essentiality *in vitro* in *M. tuberculosis* H37Rv might be explained by the usage of glycerol during *in vitro* growth in this species. The contribution of the membrane-bound *glpD2* to the donation of electrons in the electron transport chain has been suggested but not yet explored in the MTBC (63). Given the interest in the electron transport chain as a chemotherapeutic target in *M. tuberculosis*, the data presented here suggest that inhibition of *glpD2* might be a fruitful approach in the development of new drugs for the treatment of TB in humans (64). The role of this gene in *M. bovis* AF2122/97 *in vivo* is perhaps surprising given the disruptions in glycerol phosphate uptake and pathways that phosphorylate glycerol in *M. bovis* AF2122/97 (65). However, *M. tuberculosis* is thought to engage in catabolism of membrane-derived glycerophospholipids, which may be an alternative potential source of glycerol-3-phosphate in members of the complex (66).

In this study, we have identified the genes required for survival in *M. bovis* AF2122/97 in cattle. The data set correlates well with preexisting knowledge. However, in addition to known requirements, we have uncovered novel virulence factors that had not previously been described in members of the complex. In this way, we both corroborate and expand our current knowledge of tuberculosis.

## MATERIALS AND METHODS

**Bacterial strains and culture methods.** *M. bovis* strain AF2122/97 was maintained on modified Middlebrook 7H11 (BD Difco) medium (65). Liquid cultures of *M. bovis* AF2122/97 were grown in Middlebrook 7H9 media (BD Difco) containing 75 mM sodium pyruvate, 0.05% (vol/vol), Tween 80, and 10% Middlebrook albumin-dextrose-catalase (ADC) (BBL BD Biosciences). Kanamycin at 25  $\mu\text{g}/\text{mL}$  was used for selection where appropriate.

**Generation of input transposon mutant library and preparation of the inoculum.** Transposon libraries in *M. bovis* AF2122/97 were generated as previously described using the MycomarT7 phagemid system as per Majumdar et al. with modifications (18). Approximately 66,000 kanamycin-resistant transductants were scraped and homogenized in 7H9 medium and stored frozen at  $-80^{\circ}\text{C}$  in 1-mL aliquots. CFU counting was performed on the homogenized culture to inform inoculum dosage.

**Cattle infection.** Experiments were carried out according to the UK Animal (Scientific Procedures) Act 1986 under project license PPL70/7737. Ethical permission was obtained from the APHA Animal Welfare Ethical Review Body (AWERB) (UK Home Office PCD number 70/6905). All animal infections were carried out within the APHA large animal biocontainment level 3 facility. Twenty-four Holstein-Friesian crosses of 6 months of age were sourced from an officially TB-free herd. An infectious dose of  $7 \times 10^4$  CFU was targeted for the “input” library, allowing each mutant to be represented in the library  $\sim 2.5$ -fold. Retrospective counting of the inoculum revealed the actual inoculum for infection contained  $4 \times 10^4$  CFU. The inoculum was delivered endobronchially in 2 mL of 7H9 medium.

**Infection monitoring with the IGRA.** Blood was collected by jugular venipuncture from animals on the day of the infectious challenge and 2 weeks after infection. Heparinized whole blood (250  $\mu\text{L}$ ) was incubated with purified protein derivative (PPD) from *M. avium* (PPD-A) or PPD from *M. bovis* (PPD-B) (Prionics), respectively, at 25 IU and 30 IU final. Pokeweed mitogen was used as the positive control at 10  $\mu\text{g}/\text{mL}$  and a medium-only negative control. After 24 h of incubation in 5% (vol/vol)  $\text{CO}_2$ , 95% humidity,  $37^{\circ}\text{C}$  atmosphere, blood was centrifuged ( $400 \times g$  for 5 min); 120  $\mu\text{L}$  of supernatant was removed and stored at  $-80^{\circ}\text{C}$  for subsequent IFN- $\gamma$  quantification using the Bovigam kit (Prionics) in accordance with the manufacturer's instructions.

**Collection of tissues and gross pathology scores.** Six weeks after the initial infection, animals were subjected to postmortem examination. Initially, the experiment was designed with two time points, an early time point (6 weeks) and a later time point of 8 weeks. However, due to the unexpected high levels of pathology seen at the earlier time points, all animals were culled at 6 weeks. Gross pathology and evidence of TB-like granuloma lesions were scored using a modified methodology to that previously described in reference 67. Tissue from head and neck lymph nodes (from the right and left submandibular lymph nodes, the right and left medial retropharyngeal lymph nodes), thoracic lymph nodes (the right and left bronchial lymph nodes, the cranial tracheobronchial lymph nodes, and the cranial and caudal mediastinal lymph nodes), and from lung lesions was collected into sterile containers and frozen at  $-80^{\circ}\text{C}$  until further processing. Frozen tissues were thawed and homogenized in phosphate-buffered saline (PBS) using a Seward Stomacher paddle blender.

**Recovery of the output transposon mutant library from tissues.** Tissue macerates collected from study animals were thawed at room temperature, diluted in PBS, and plated on modified 7H11 agar to determine bacterial loads. Colony counts were performed after 3 to 4 weeks of growth. For recovery of the library from tissue macerates,  $\sim 10^5$  to  $10^6$  CFU were plated from lung lesions and thoracic lymph node lesions onto modified 7H11 media containing 25  $\mu\text{g}/\text{mL}$  kanamycin. The colonies were plated over several 140-mm petri dishes to minimize competition between mutants. The colonies were harvested after 4 to 6 weeks of growth and genomic DNA extracted.

**Genomic DNA extraction.** Genomic DNA from the input and recovered libraries was isolated by an extended bead beating procedure with detergent-based lysis, phenol-chloroform DNA extraction, and precipitation as previously described (19). DNA quality was assessed by nanospectrometry (DeNovix) and gel electrophoresis and quantified by Qubit analysis using the broad-range assay kit (Thermo Scientific).

**Library preparation for transposon-directed insertion sequencing.** DNA (2  $\mu\text{g}$ ) was resuspended in 50  $\mu\text{L}$  distilled water and sheared to approximately 550-bp fragments using an S220 focused ultrasonicator (Covaris) according to the manufacturer's protocol. Fragmented DNA was repaired using NEBNext blunt-end repair kit (New England Biolabs) and purified using Monarch PCR cleanup kit (NEB). Blunted DNA was A-tailed using NEBNext dA-tailing kit (NEB) and column purified. Custom transposon sequencing adaptors (Table S4) were generated by heating an equimolar mix of Com\_AdaptorPt1 primer and Com\_AdaptorPt2 (P7+index) primers to  $95^{\circ}\text{C}$  for 5 min, followed by cooling by  $1^{\circ}\text{C}$  every 40 s to a final temperature of  $4^{\circ}\text{C}$  in a thermocycler. Adaptors were ligated to A-tailed library fragments using NEBNext quick ligase kit. Transposon-containing fragments were enriched by PCR using the ComP7 primer (10  $\mu\text{M}$ ) and an equimolar mix of primers P5-IR2a-d primer (10  $\mu\text{M}$ ) in a reaction with 50 ng of adaptor-ligated template and Phusion DNA polymerase (NEB) in a thermocycler with the following program:  $98^{\circ}\text{C}$  for 3 min; 4 cycles of  $98^{\circ}\text{C}$  for 20 s,  $70^{\circ}\text{C}$  for 20 s, and  $72^{\circ}\text{C}$  for 1 min; 20 cycles of  $98^{\circ}\text{C}$  for 20 s,  $67^{\circ}\text{C}$  for 20 s, and  $72^{\circ}\text{C}$  for 1 min; and  $72^{\circ}\text{C}$  for 3 min. Transposon-enriched libraries were subsequently purified with AMPure XP beads (Beckman), pooled, and further purified using AMPure XP beads.

**Data analysis.** Indexed libraries were combined, spiked with 20% PhiX, and sequenced on the Illumina HiSeq 3000 platform, using v2 chemistry, generating single-end reads of 250 bp. Raw FASTQ sequencing files were analyzed for quality and preprocessed using the TRANSIT TPP tool (68) set to default Sasseti protocol in order to remove transposon tags and adapter sequences and to map reads using BWA-MEM to TA sites to the *M. bovis* AF2122/97 genome (GenBank accession no. [NC\\_002945](#)). The TRANSIT tseq\_stats tool was run on each sample to assess insertion density, skew, kurtosis, and potential amplification bias.

The *M. bovis* AF2122/97 genome was scanned for the nonpermissive Himar1 transposon insertion motif (SGNTANC, where S is either G or C, and N is any base) as previously described (11). We identified 6,605 sites as nonpermissive (approximately 9% of total TA sites) and excluded them from resampling analysis. A custom annotation, prot-table for TRANSIT, was created from the *M. bovis* AF2122/97 annotation file (GenBank accession number [LT708304](#), version [LT708304.1](#)). TRANSIT HMM was run on the input library using the default normalization (TTR) with locally estimated scatterplot smoothing (LOESS) correction for genomic position bias. Each TA site was assigned an essentiality state, and genes were assigned an essentiality call based on the assigned state of the TA sites within annotated gene boundaries.

Resampling between the input library and each of the output sample libraries was performed independently using the TRANSIT resampling algorithm and the complete prot-table. TTR normalization was used for 23 of the samples, and betageom normalization was used for the three samples with skew of greater than 50. The initial resampling output files were evaluated to identify genes with very few, or no, reads at any TA site within the gene boundaries in both the input library and output sample libraries. Genes with no read counts greater than 4 at any TA site, in any sample, and with a sum of all reads at any TA site across the 26 samples less than 55 were flagged. Essential and unchanged genes were removed from the prot-table prior to further evaluation. Resampling was further limited to protein-coding genes. Resampling was rerun for each sample using the edited prot-table and an edited TRANSIT resampling script to return the left-tail *P* value, as the data were expected to reflect attenuation. Resampling was performed in a separate analysis ("pooled") with all of the sample insertion files as replicates using the edited TRANSIT resampling script and prot-table with betageom normalization. All *P* values were corrected for multiple testing with false-discovery rate (FDR) adjustment.

All analysis and plots were performed using R and R packages tidyverse and circlize (69–71). Orthologous TB genes were obtained from supplementary data files published by Malone et al. (29). All scripts, prot-tables, and insertion files are available at [https://github.com/jenjane118/Mbovis\\_in-vivo\\_Tnseq](https://github.com/jenjane118/Mbovis_in-vivo_Tnseq) and <https://doi.org/10.5281/zenodo.6576716>.

**Data availability.** Sequencing files (FASTQ) were deposited into BioProject under accession no. [PRJNA816175](#).

## ACKNOWLEDGMENTS

We acknowledge the help and support of APHA colleagues from the BAC4 workgroup and the Pathology Unit and, in particular, acknowledge the care and support provided to animals under experimentation by members of APHA's Animals Sciences Unit. We also thank Andrew Waller for useful discussions regarding *in vivo* Tn-seq experimental design. We thank Dany Beste for useful discussions surrounding the role of *glpD2* in mycobacterial metabolism.

This work was funded by the BBSRC grant reference BB/N004590/1 (awarded to S.L.K. [PI], D.W. (Co-I), B.W.W. (Co-I), and SE3297 to B.V.-R. as part of the joint BBSRC-DEFRA EradTB Consortium). A.J.G., I.J.P., and S.W. were supported by the funding. V.F. was in receipt of an RVC PhD studentship. A.J.G. currently holds a Sêr Cymru II Lectureship funded by the European Research Development Fund and Welsh Government. B.V.-R. is a Ser Cymru II Professor of Immunology at Aberystwyth University. J.S. is supported by a Bloomsbury Colleges PhD Studentship (LIDO program).

S.L.K., D.W., B.W.W., B.V.-R., and S.B. undertook funding acquisition and designed the study. A.J.G., V.F., J.M., S.W., I.J.P., and M.C. carried out the experimental work. Data analysis was done by I.N. and J.S. A.J.G., J.S., and S.L.K. wrote the first draft of the manuscript. All authors contributed to the manuscript revision and read and approved the submitted version.

## REFERENCES

- Müller B, Dürr S, Alonso S, Hattendorf J, Laise CJM, Parsons SDC, van Helden PD, Zinsstag J. 2013. Zoonotic *Mycobacterium bovis*-induced tuberculosis in humans. *Emerg Infect Dis* 19:899–908. <https://doi.org/10.3201/eid1906.120543>.
- Mablesen HE, Okello A, Picozzi K, Welburn SC. 2014. Neglected zoonotic diseases—the long and winding road to advocacy. *PLoS Negl Trop Dis* 8:e2800. <https://doi.org/10.1371/journal.pntd.0002800>.
- Bayissa B, Sirak A, Worku A, Zewude A, Zeleke Y, Chanyalew M, Gumi B, Berg S, Conlan A, Hewinson RG, Wood JLN, Vordermeier HM, Ameni G. 2021. Evaluation of the efficacy of BCG in protecting against contact challenge with bovine tuberculosis in Holstein-Friesian and Zebu crossbred calves in Ethiopia. *Front Vet Sci* 8:702402. <https://doi.org/10.3389/fvets.2021.702402>.
- Srinivasan S, Conlan AJK, Easterling LA, Herrera C, Dandapat P, Veerasami M, Ameni G, Jindal N, Raj GD, Wood J, Juleff N, Bakker D, Vordermeier M, Kapur V. 2021. A meta-analysis of the effect of *Bacillus Calmette-Guérin* vaccination against bovine tuberculosis: is perfect the enemy of good? *Front Vet Sci* 8:637580. <https://doi.org/10.3389/fvets.2021.637580>.
- Vordermeier HM, Whelan A, Cockle PJ, Farrant L, Palmer N, Hewinson RG. 2001. Use of synthetic peptides derived from the antigens ESAT-6 and CFP-10 for differential diagnosis of bovine tuberculosis in cattle. *Clin Diagn Immunol* 8:571–578. <https://doi.org/10.1128/CDLI.8.3.571-578.2001>.
- Whelan AO, Clifford D, Upadhyay B, Breadon EL, McNair J, Hewinson GR, Vordermeier MH. 2010. Development of a skin test for bovine tuberculosis for differentiating infected from vaccinated animals. *J Clin Microbiol* 48:3176–3181. <https://doi.org/10.1128/JCM.00420-10>.
- Cain AK, Barquist L, Goodman AL, Paulsen IT, Parkhill J, van Opijnen T. 2020. A decade of advances in transposon-insertion sequencing. *Nat Rev Genet* 21:526–540. <https://doi.org/10.1038/s41576-020-0244-x>.
- Bellerose MM, Proulx MK, Smith CM, Baker RE, Ioegeer TR, Sasseti CM. 2020. Distinct bacterial pathways influence the efficacy of antibiotics



- against *Mycobacterium tuberculosis*. *mSystems* 5:e00396-20. <https://doi.org/10.1128/mSystems.00396-20>.
9. Zhang YJ, Reddy MC, Ioerger TR, Rothchild AC, Dartois V, Schuster BM, Trauner A, Wallis D, Galaviz S, Huttenhower C, Sacchetti JC, Behar SM, Rubin EJ. 2013. Tryptophan biosynthesis protects mycobacteria from CD4 T-cell-mediated killing. *Cell* 155:1296–1308. <https://doi.org/10.1016/j.cell.2013.10.045>.
  10. Sassetti CM, Rubin EJ. 2003. Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* 100:12989–12994. <https://doi.org/10.1073/pnas.2134250100>.
  11. Dejesus MA, Gerrick ER, Xu W, Park SW, Long JE, Boutte CC, Rubin EJ, Schnappinger D, Ehrst S, Fortune SM, Sassetti CM, Ioerger TR. 2017. Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *mBio* 8:e02133-16. <https://doi.org/10.1128/mBio.02133-16>.
  12. Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM. 2011. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* 7:e1002251. <https://doi.org/10.1371/journal.ppat.1002251>.
  13. Patil S, Palande A, Lodhiya T, Pandit A, Mukherjee R. 2021. Redefining genetic essentiality in *Mycobacterium tuberculosis*. *Gene* 765:145091. <https://doi.org/10.1016/j.gene.2020.145091>.
  14. Van Rhijn I, Moody DB. 2015. CD1 and mycobacterial lipids activate human T cells. *Immunol Rev* 264:138–153. <https://doi.org/10.1111/imr.12253>.
  15. Dutta NK, Mehra S, Didier PJ, Roy CJ, Doyle LA, Alvarez X, Ratterree M, Be NA, Lamichhane G, Jain SK, Lacey MR, Lackner AA, Kaushal D. 2010. Genetic requirements for the survival of tubercle bacilli in primates. *J Infect Dis* 201:1743–1752. <https://doi.org/10.1086/652497>.
  16. Mendum TA, Chandran A, Williams K, Vordermeier HM, Villarreal-Ramos B, Wu H, Singh A, Smith AA, Butler RE, Prasad A, Bharti N, Banerjee R, Kasibhata SM, Bhatt A, Stewart GR, McFadden J. 2019. Transposon libraries identify novel *Mycobacterium bovis* BCG genes involved in the dynamic interactions required for BCG to persist during in vivo passage in cattle. *BMC Genomics* 20:431. <https://doi.org/10.1186/s12864-019-5791-1>.
  17. Smith AA, Villarreal-Ramos B, Mendum TA, Williams KJ, Jones GJ, Wu H, McFadden J, Vordermeier HM, Stewart GR. 2020. Genetic screening for the protective antigenic targets of BCG vaccination. *Tuberculosis (Edinb)* 124:101979. <https://doi.org/10.1016/j.tube.2020.101979>.
  18. Majumdar G, Mbau R, Singh V, Warner DF, Dragset MS, Mukherjee R. 2017. Genome-wide transposon mutagenesis in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. *Methods Mol Biol* 1498:321–335. [https://doi.org/10.1007/978-1-4939-6472-7\\_21](https://doi.org/10.1007/978-1-4939-6472-7_21).
  19. Gibson AJ, Passmore IJ, Faulkner V, Xia D, Nobeli I, Stiens J, Willcocks S, Clark TG, Sobkowiak B, Werling D, Villarreal-Ramos B, Wren BW, Kendall SL. 2021. Probing differences in gene essentiality between the human and animal adapted lineages of the *Mycobacterium tuberculosis* complex using TnSeq. *Front Vet Sci* 8:760717. <https://doi.org/10.3389/fvets.2021.760717>.
  20. Zwyyer M, Çavusoglu C, Ghielmetti G, Pacciarini ML, Scaltriti E, Van Soolingen D, Dötsch A, Reinhard M, Gagneux S, Brites D. 2021. A new nomenclature for the livestock-associated *Mycobacterium tuberculosis* complex based on phylogenomics. *Open Res Europe* 1:100. <https://doi.org/10.12688/openreseurope.14029.1>.
  21. Pilgrim S, Kolb-Mäurer A, Gentschev I, Goebel W, Kuhn M. 2003. Deletion of the gene encoding p60 in *Listeria monocytogenes* leads to abnormal cell division and loss of actin-based motility. *Infect Immun* 71:3473–3484. <https://doi.org/10.1128/IAI.71.6.3473-3484.2003>.
  22. Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, Sherman DR. 2003. Deletion of RD1 from *Mycobacterium tuberculosis* mimics *Bacille Calmette-Guérin* attenuation. *J Infect Dis* 187:117–123. <https://doi.org/10.1086/345862>.
  23. Phan TH, van Leeuwen LM, Kuijij C, Ummels R, van Stempvoort G, Rubio-Canalejas A, Piersma SR, Jiménez CR, van der Sar AM, Houben ENG, Bitter W. 2018. EspH is a hypervirulence factor for *Mycobacterium marinum* and essential for the secretion of the ESX-1 substrates EspE and EspF. *PLoS Pathog* 14:e1007247. <https://doi.org/10.1371/journal.ppat.1007247>.
  24. Inwald J, Jahans K, Hewinson RG, Gordon SV. 2003. Inactivation of the *Mycobacterium bovis* homologue of the polymorphic RD1 gene Rv3879c (Mb3909c) does not affect virulence. *Tuberculosis (Edinb)* 83:387–393. <https://doi.org/10.1016/j.tube.2003.08.018>.
  25. Bold TD, Davis DC, Penberthy KK, Cox LM, Ernst JD, de Jong BC. 2012. Impaired fitness of *Mycobacterium africanum* despite secretion of ESAT-6. *J Infect Dis* 205:984–990. <https://doi.org/10.1093/infdis/jir883>.
  26. Hotter GS, Wards BJ, Mouat P, Besra GS, Gomes J, Singh M, Bassett S, Kawakami P, Wheeler PR, De Lisle GW, Collins DM. 2005. Transposon mutagenesis of Mb0100 at the ppe1-nrp locus in *Mycobacterium bovis* disrupts phthiocerol dimycoserate (PDIM) and glycosylphenol-PDIM biosynthesis, producing an avirulent strain with vaccine properties at least equal to those of *M. bovis* BCG. *J Bacteriol* 187:2267–2277. <https://doi.org/10.1128/JB.187.7.2267-2277.2005>.
  27. Cox JS, Chen B, McNeil M, Jacobs WR. 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* 402:79–83. <https://doi.org/10.1038/47042>.
  28. Pérez J, García R, Bach H, de Waard JH, Jacobs WR, Av-Gay Y, Bubis J, Takiff HE. 2006. *Mycobacterium tuberculosis* transporter MmpL7 is a potential substrate for kinase PknD. *Biochem Biophys Res Commun* 348:6–12. <https://doi.org/10.1016/j.bbrc.2006.06.164>.
  29. Malone KM, Rue-Albrecht K, Magee DA, Conlon K, Schubert OT, Nalpas NC, Browne JA, Smyth A, Gormley E, Abersold R, Machugh DE, Gordon SV. 2018. Comparative ‘omics analyses differentiate *Mycobacterium tuberculosis* and *Mycobacterium bovis* and reveal distinct macrophage responses to infection with the human and bovine tubercle bacilli. *Microb Genom* 4:e000163. <https://doi.org/10.1099/mgen.0.000163>.
  30. Pisu D, Huang L, Grenier JK, Russell DG. 2020. Dual RNA-Seq of Mtb-infected macrophages in vivo reveals ontologically distinct host-pathogen interactions. *Cell Rep* 30:335–350.e4. <https://doi.org/10.1016/j.celrep.2019.12.033>.
  31. Nesbitt NM, Yang X, Fontán P, Kolesnikova I, Smith I, Sampson NS, Dubnau E. 2010. A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infect Immun* 78:275–282. <https://doi.org/10.1128/IAI.00893-09>.
  32. Van Der Geize R, Yam K, Heuser T, Wilbrink MH, Hara H, Anderton MC, Sim E, Dijkhuizen L, Davies JE, Mohn WW, Eltis LD. 2007. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc Natl Acad Sci U S A* 104:1947–1952. <https://doi.org/10.1073/pnas.0605728104>.
  33. Bandyopadhyay U, Chadha A, Gupta P, Tiwari B, Bhattacharyya K, Popli S, Raman R, Brahmachari V, Singh Y, Malhotra P, Natarajan K. 2017. Suppression of Toll-like receptor 2-mediated proinflammatory responses by *Mycobacterium tuberculosis* protein Rv3529c. *J Leukoc Biol* 102:1249–1259. <https://doi.org/10.1189/jlb.4A0217-042R>.
  34. Pandey AK, Sassetti CM. 2008. Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A* 105:4376–4380. <https://doi.org/10.1073/pnas.0711159105>.
  35. Mohn WW, Van Der Geize R, Stewart GR, Okamoto S, Liu J, Dijkhuizen L, Eltis LD. 2008. The actinobacterial mce4 locus encodes a steroid transporter. *J Biol Chem* 283:35368–35374. <https://doi.org/10.1074/jbc.M805496200>.
  36. Ganchua SKC, Cadena AM, Maiello P, Gideon HP, Myers AJ, Junecko BF, Klein EC, Lin PL, Mattila JT, Flynn JAL. 2018. Lymph nodes are sites of prolonged bacterial persistence during *Mycobacterium tuberculosis* infection in macaques. *PLoS Pathog* 14:e1007337. <https://doi.org/10.1371/journal.ppat.1007337>.
  37. Yang X, Gao J, Smith I, Dubnau E, Sampson NS. 2011. Cholesterol is not an essential source of nutrition for *Mycobacterium tuberculosis* during infection. *J Bacteriol* 193:1473–1476. <https://doi.org/10.1128/JB.01210-10>.
  38. Keating LA, Wheeler PR, Mansoor H, Inwald JK, Dale J, Hewinson RG, Gordon SV. 2005. The pyruvate requirement of some members of the *Mycobacterium tuberculosis* complex is due to an inactive pyruvate kinase: implications for in vivo growth. *Mol Microbiol* 56:163–174. <https://doi.org/10.1111/j.1365-2958.2005.04524.x>.
  39. Lofthouse EK, Wheeler PR, Beste DJV, Khatri BL, Wu H, Mendum TA, Kierzek AM, McFadden J. 2013. Systems-based approaches to probing metabolic variation within the *Mycobacterium tuberculosis* complex. *PLoS One* 8:e75913. <https://doi.org/10.1371/journal.pone.0075913>.
  40. Smith CM, Baker RE, Proulx MK, Mishra BB, Long JE, Park SW, Lee H-N, Kiritsy MC, Bellerose MM, Olive AJ, Murphy K, Papavinasandaram K, Boehm FJ, Reames CJ, Meade RK, Hampton BK, Linnertz CL, Shaw GD, Hock P, Bell TA, Ehrst S, Schnappinger D, Pardo-Manuel de Villena F, Ferris MT, Ioerger TR, Sassetti CM. 2022. Host-pathogen genetic interactions underlie *Mycobacterium tuberculosis* susceptibility in genetically diverse mice. *Elife* 11:e74419. <https://doi.org/10.7554/eLife.74419>.
  41. Chandra P, Coullon H, Agarwal M, Goss CW, Philips JA. 2022. Macrophage global metabolomics identifies cholestenone as host/pathogen cometabolite present in human *Mycobacterium tuberculosis* infection. *J Clin Invest* 132:e152509. <https://doi.org/10.1172/JCI152509>.
  42. Golby P, Hatch KA, Bacon J, Cooney R, Riley P, Allnutt J, Hinds J, Nunez J, Marsh PD, Hewinson RG, Gordon SV. 2007. Comparative transcriptomics reveals key gene expression differences between the human and bovine

- pathogens of the *Mycobacterium tuberculosis* complex. *Microbiology (Reading)* 153:3323–3336. <https://doi.org/10.1099/mic.0.2007/009894-0>.
43. Rehren G, Walters S, Fontan P, Smith I, Zárrega AM. 2007. Differential gene expression between *Mycobacterium bovis* and *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 87:347–359. <https://doi.org/10.1016/j.tube.2007.02.004>.
  44. Gonzalo-Asensio J, Malaga W, Pawlik A, Astarie-Dequeker C, Passemar C, Moreau F, Laval F, Daffé 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. <https://doi.org/10.1073/pnas.1406693111>.
  45. 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. <https://doi.org/10.3389/fimmu.2017.01803>.
  46. White AD, Sibley L, Sarfas C, Morrison A, Gullick J, Clark S, Gleeson F, McIntyre A, Arlehamn CL, Sette A, Salguero FJ, Rayner E, Rodriguez E, Puentes E, Laddy D, Williams A, Dennis M, Martin C, Sharpe S. 2021. MTBVAC vaccination protects rhesus macaques against aerosol challenge with *M. tuberculosis* and induces immune signatures analogous to those observed in clinical studies. *NPJ Vaccines* 6:4. <https://doi.org/10.1038/s41541-020-00262-8>.
  47. 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. <https://doi.org/10.1371/journal.ppat.0040033>.
  48. 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. <https://doi.org/10.1128/JB.01067-12>.
  49. Queval CJ, Fearnas A, Botella L, Smyth A, Schnettger L, Mitermite M, Wooff E, Villarreal-Ramos B, Garcia-Jimenez W, Heunis T, Trost M, Werling D, Salguero FJ, Gordon SV, Gutierrez MG. 2021. Macrophage-specific responses to human- and animal-adapted tubercle bacilli reveal pathogen and host factors driving multinucleated cell formation. *PLoS Pathog* 17:e1009410. <https://doi.org/10.1371/journal.ppat.1009410>.
  50. Cowley S, Ko M, Pick N, Chow R, Downing KJ, Gordhan BG, Betts JC, Mizrahi V, Smith DA, Stokes RW, Av-Gay Y. 2004. The *Mycobacterium tuberculosis* protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. *Mol Microbiol* 52:1691–1702. <https://doi.org/10.1111/j.1365-2958.2004.04085.x>.
  51. Liu CF, Tonini L, Malaga W, Beau M, Stella A, Bouyssié D, Jackson MC, Nigou J, Puzo G, Guilhot C, Bulet-Schiltz O, Rivière M. 2013. Bacterial protein-O-mannosylating enzyme is crucial for virulence of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 110:6560–6565. <https://doi.org/10.1073/pnas.1219704110>.
  52. Maksymiuk C, Ioerger T, Balakrishnan A, Bryk R, Rhee K, Sacchetti J, Nathan C. 2015. Comparison of transposon and deletion mutants in *Mycobacterium tuberculosis*: the case of rv1248c, encoding 2-hydroxy-3-oxoadipate synthase. *Tuberculosis (Edinb)* 95:689–694. <https://doi.org/10.1016/j.tube.2015.08.009>.
  53. Venugopal A, Bryk R, Shi S, Rhee K, Rath P, Schnappinger D, Ehrh S, Nathan C. 2011. Virulence of *Mycobacterium tuberculosis* depends on lipamide dehydrogenase, a member of three multienzyme complexes. *Cell Host Microbe* 9:21–31. <https://doi.org/10.1016/j.chom.2010.12.004>.
  54. Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, et al. 2004. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet* 36:1133–1137. <https://doi.org/10.1038/ng1104-1133>.
  55. Constant P, Perez E, Malaga W, Lanéelle MA, Saurel O, Daffé 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. <https://doi.org/10.1074/jbc.M206538200>.
  56. Collins DM, Skou B, White S, Bassett S, Collins L, For R, Hurr K, Hotter G, De Lisle GW. 2005. Generation of attenuated *Mycobacterium bovis* strains by signature-tagged mutagenesis for discovery of novel vaccine candidates. *Infect Immun* 73:2379–2386. <https://doi.org/10.1128/IAI.73.4.2379-2386.2005>.
  57. Yang S, Sui S, Qin Y, Chen H, Sha S, Liu X, Deng G, Ma Y. 2022. Protein O-mannosyltransferase Rv1002c contributes to low cell permeability, biofilm formation in vitro, and mycobacterial survival in mice. *APMIS* 130:181–192. <https://doi.org/10.1111/apm.13204>.
  58. Crellin PK, Kovacevic S, Martin KL, Brammananth R, Morita YS, Billman-Jacobe H, McConville MJ, Coppel RL. 2008. Mutations in pimE restore lipoarabinomannan synthesis and growth in a *Mycobacterium smegmatis* lppW mutant. *J Bacteriol* 190:3690–3699. <https://doi.org/10.1128/JB.00200-08>.
  59. Butler RE, Smith AA, Mendum TA, Chandran A, Wu H, Lefrançois L, Chambers M, Soldati T, Stewart GR. 2020. *Mycobacterium bovis* uses the ESX-1 type VII secretion system to escape predation by the soil-dwelling amoeba *Dictyostelium discoideum*. *ISME J* 14:919–930. <https://doi.org/10.1038/s41396-019-0572-z>.
  60. Agapova A, Serafini A, Petridis M, Hunt DM, Garza-Garcia A, Sohaskey CD, de Carvalho LPS. 2019. Flexible nitrogen utilisation by the metabolic generalist pathogen *Mycobacterium tuberculosis*. *Elife* 8:e41129. <https://doi.org/10.7554/eLife.41129>.
  61. Gozuy A, Larrouy-Maumus G, Wu T-D, Peixoto A, Levillain F, Lugo-Villarino G, Guerin-Kern J-L, Gerquin-Kern J-L, de Carvalho LPS, Poquet Y, Neyrolles O. 2013. *Mycobacterium tuberculosis* nitrogen assimilation and host colonization require aspartate. *Nat Chem Biol* 9:674–676. <https://doi.org/10.1038/nchembio.1355>.
  62. Yeh JI, Chinte U, Du S. 2008. Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism. *Proc Natl Acad Sci U S A* 105:3280–3285. <https://doi.org/10.1073/pnas.0712331105>.
  63. Cook GM, Hards K, Dunn E, Heikal A, Nakatani Y, Greening C, Crick DC, Fontes FL, Pethe H, Hasenoehrl E, Berney M. 2017. Oxidative phosphorylation as a target space for tuberculosis: success, caution, and future directions. *Microbiol Spectr* 5. <https://doi.org/10.1128/microbiolspec.TBTB2-0014-2016>.
  64. Bald D, Villellas C, Lu P, Koul A. 2017. Targeting energy metabolism in *Mycobacterium tuberculosis*, a new paradigm in antimycobacterial drug discovery. *mBio* 8:e00272-17. <https://doi.org/10.1128/mBio.00272-17>.
  65. Garnier T, Eiglmeier K, Camus J-C, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler PR, Parkhill J, Barrell BG, Cole ST, Gordon SV, Hewinson RG. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* 100:7877–7882. <https://doi.org/10.1073/pnas.1130426100>.
  66. Larrouy-Maumus G, Biswas T, Hunt DM, Kelly G, Tsoodikov OV, de Carvalho LPS. 2013. Discovery of a glycerol 3-phosphate phosphatase reveals glycerophospholipid polar head recycling in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 110:11320–11325. <https://doi.org/10.1073/pnas.1221597110>.
  67. Vordermeier HM, Chambers MA, Cockle PJ, Whelan AO, Simmons J, Hewinson RG. 2002. Correlation of ESAT-6-specific gamma interferon production with pathology in cattle following *Mycobacterium bovis* BCG vaccination against experimental bovine tuberculosis. *Infect Immun* 70:3026–3032. <https://doi.org/10.1128/IAI.70.6.3026-3032.2002>.
  68. DeJesus MA, Ambadipudi C, Baker R, Sasseti C, Ioerger TR. 2015. TRANSIT—a software tool for Himar1 TnSeq analysis. *PLoS Comput Biol* 11:e1004401. <https://doi.org/10.1371/journal.pcbi.1004401>.
  69. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize implements and enhances circular visualization in R. *Bioinformatics* 30:2811–2812. <https://doi.org/10.1093/bioinformatics/btu393>.
  70. R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
  71. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the Tidyverse. *J Open Source Softw* 4:1686. <https://doi.org/10.21105/joss.01686>.