

Integrated Bioinformatics Analysis for Target Identification and Evaluation of Recombinant Protein as an Antigen for Intradermal Skin Test in Bovine Tuberculosis Diagnosis

Violetta Dias Pacce, Amanda Munari Guimarães, Frederico Schmitt Kremer, Gabriela Nascimento Ferreira, Jean Michel Dela Vedova-Costa, Aline Cristina dos Santos, Odir Antônio Dellagostin, Carlos Ricardo Soccol, and Vanete Thomaz-Soccol*

Cite This: *ACS Omega* 2025, 10, 9187–9196

Read Online

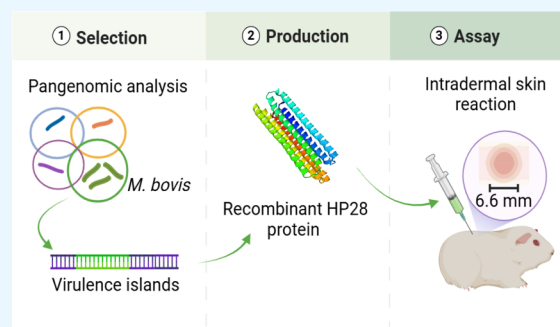
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Bovine tuberculosis (bTB) is a respiratory disease caused by *Mycobacterium bovis*, posing a significant threat to animal health and the livestock industry. Current control strategies for bTB rely on diagnostic tests and slaughter policies. However, the limitations of existing diagnostic methods, which depend on PPD antigens, necessitate the exploration of alternative antigens to enhance the accuracy and reliability of bTB diagnosis. This study aimed to identify, produce, and evaluate novel antigens for use in the intradermal skin test for bTB diagnosis. A pangenome analysis of four *Mycobacterium* species identified 12 unique genes specific to *M. bovis* SP38. Further integrated bioinformatic analysis revealed 224 genomic islands associated with virulence and pathogenesis. Among these, a highly antigenic protein, termed HP28, was selected for in vivo testing. The recombinant HP28 protein (rHP28) was expressed in *E. coli* and assessed for its ability to induce intradermal skin reactions in guinea pigs. The rHP28 protein elicited a skin reaction of 6.6 mm at 72 h post-injection, whereas negative controls showed no reaction. This study presents a pipeline for the selection of antigens using integrated bioinformatic analysis to identify diagnostic targets that can effectively distinguish between sensitized and non-sensitized animals, offering a promising approach for improving bTB diagnostics.



INTRODUCTION

Mycobacterium bovis is the causative agent of tuberculosis (TB) in animals, particularly in bovines. This bacterium is part of the *M. tuberculosis* Complex (MTC), characterized by slow evolution, strict clonality, and absence of gene transfer or recombination.¹ However, despite having >99% nucleotide identity, MTC species have variable host tropisms, phenotypes, and pathogenic degrees.² Clinical signs rarely manifest in animals, which often exhibit a healthy phenotype even if tested positive.³

Globally, bovine tuberculosis (bTB) causes an annual agricultural revenue loss of approximately 3 billion dollars^{4,5} owing to the loss of milk, meat, and carcass.⁶ The disease is also zoonotic, spreading to humans through close contact with infected cows or via the consumption of unpasteurized milk.⁷ A recent meta-analysis indicated that 702 (9.7%) of 7185 MTC isolates from human tuberculosis cases were caused by *M. bovis* subspecies.⁸ The disease burden is higher in marginalized populations and poor or rural communities living close to animals and lacking access to health care and/or safe food.^{9,10} For those reasons, the One Health concept has embraced the burden of bovine tuberculosis since the disease

can only be abolished from humans when it is completely eradicated from animals as well. Routine diagnostic tests are necessary for better disease control and epidemiological monitoring. Currently, the intradermal skin test (TST) using Purified Protein Derivative (PPD) as an antigen is widely applied and recognized as a reference technique by the World Organization for Animal Health.¹¹ However, bovine PPD can exhibit cross-reactivity due to the possible presence of proteins also found in nonpathogenic environmental mycobacteria. Additionally, the lack of antigen standardization can impact result consistency, and its production requires manipulation of a live virulent strain of *M. bovis*, raising biosafety concerns.^{12,13}

Although several antigens such as ESAT-6, CFP-10, Rv3615c, and MPB70 are being explored for bovine tuberculosis diagnostics, their performance can be limited due to

Received: October 14, 2024

Revised: January 13, 2025

Accepted: January 21, 2025

Published: February 24, 2025



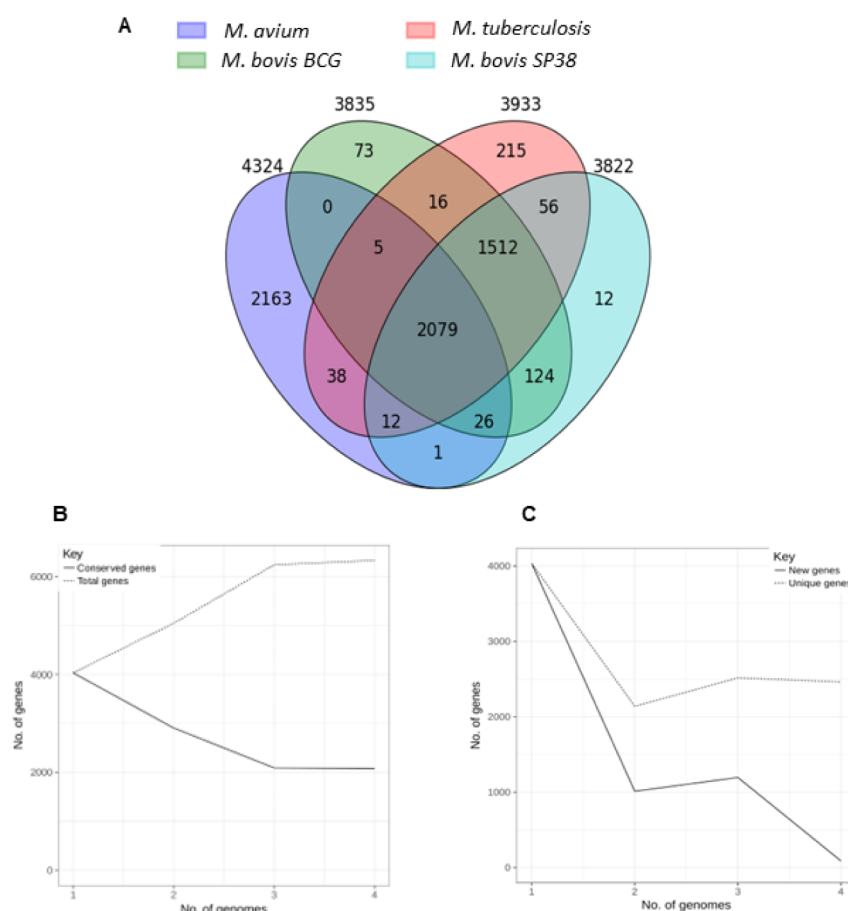


Figure 1. *Mycobacterium* pangenome. (A) Number of genes belonging to the core, the soft core, the shell or the cloud of the *Mycobacterium* strains; (B) continuous line (Conserved genes) shows how the number of conserved genes changes as more genomes are added to the analysis. In a pangenome, this line is expected to stabilize after a certain point, indicating that all essential or conserved genes have been identified. A horizontal continuous line suggests that these genes are highly conserved among the genomes. The dotted line (Total genes) illustrates how the total number of genes grows with the increasing number of genomes. Initially, the number of total genes rises rapidly because each new sequenced genome may contain unique genes not present in previous genomes. However, as more genomes are added, the new gene discovery rate decreases, and the line begins to stabilize. This indicates that most unique or variable genes have been identified. (C) Representation of the *Mycobacterium* gene content (extrapolated median-based line) according to how the pan-genome varies as genomes are added in random order to the analysis. The dotted line represents unique genes, whereas the continuous line represents new genes. The image of Figure 1A was generated through a custom Python script and graphics are derived from native output feature of Roary, which provides built-in visualizations of pangenome data.

variability in sensitivity and specificity depending on test methodology used.¹⁴ In this scenario, the search for bTB diagnosis of new antigens has been ongoing, either to introduce a new methodology or to improve existing ones. The goal is to use antigens that are only present in pathogenic bacteria, which can differentiate between infected and healthy animals and/or vaccinated individuals. For new antigens' search, pangenome analysis has proven effective in understanding the species' characteristics to develop vaccines or diagnostic strategies.¹⁵ Moreover, the accumulation of biological data and computational advances in recent decades has made pangenome analysis a fast and concise method for genotype-phenotype evaluation and understanding.¹⁶

From this perspective, this study aimed to identify exclusive genes from virulence clusters that could be further explored as diagnostic targets for bovine tuberculosis control.

RESULTS

Pangenome Analysis. Roary analysis of *Mycobacterium* strains pangenome identified 2079 core, 1555 soft core, 235 shell, and 2463 cloud genes, respectively, out of 6332 total

genes (Figure 1A). The large number of cloud genes implies that a large heterogeneity exists among the four strains considered, highlighting the “open” nature of *Mycobacterium* pan-genome (Figure 1B,C).

Genomic Island Analysis. For the genome *M. avium*, 32 genomic islands associated with virulence and pathogenesis were identified. As for *M. tuberculosis* H37Rv and *M. tuberculosis* variant *bovis*, 240 and 224 genomic islands related to virulence and pathogenesis were identified, respectively. Lastly, 159 islands were identified for the *M. tuberculosis* variant *bovis* BCG str. Pasteur. The respective islands for each organism can be found in [Supplementary Material 1](#).

Unique Protein Analysis and Antigenicity Evaluation. The Multiple Blast alignment for identities with less than 60% generated a single protein for *M. avium* (MAP01). For *M. bovis* BCG str. Pasteur, two proteins were found (MBCGP02 and MBCGP03). Alignment for *M. tuberculosis* H37Rv and *M. bovis* SP38 resulted in two proteins for both organisms, named MHR04, MHR05, MSP06, and MSP07, respectively. These proteins were classified as antigenic (≥ 0.5) or non-antigenic (< 0.5), following a 0.5 threshold (Table 1).

Table 1. Antigenicity Values for Unique Proteins Identified^a

Protein	Organism	Antigenicity value
MAP01	<i>M. avium</i>	0.0657
MBCGP02	<i>M. bovis</i> BCG str. Pasteur	0.0747
MBCGP03	<i>M. bovis</i> BCG str. Pasteur	0.4330
MHR04	<i>M. tuberculosis</i> H37Rv	0.1982
MHR05	<i>M. tuberculosis</i> H37Rv	0.1278
MSP06	<i>M. bovis</i> SP38	0.3290
MSP07	<i>M. bovis</i> SP38	0.9272

^aAntigenicity value is generated by ANTIGENPro software.

It is observed that protein MSP07 showed the highest antigenicity value, being classified as antigenic. This protein was termed by the authors “rHP28” protein, meaning Hypothetical Protein 28, with 28 representing the vector where it was further cloned into (pET28a+).

Protein Structure. Examination of the pLDDT plots derived from the prediction of *Mycobacterium* strains suggests the existence of high-confidence regions. All regions identified in the plots can be seen in the final model as a complex formation of alpha-helices (Figure 2). These helices compose

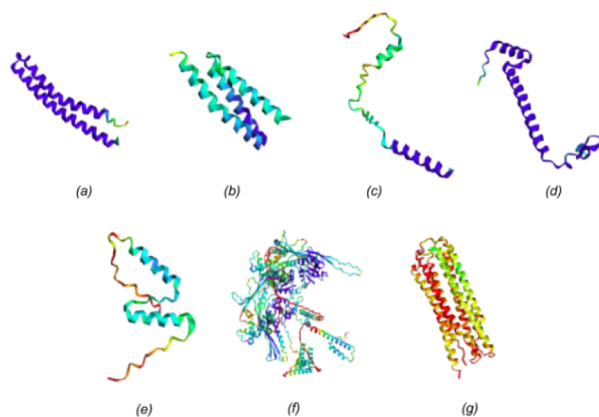


Figure 2. ColabFold predictions analysis for *M. avium*, *M. tuberculosis* H37Rv, *M. bovis*, and *M. bovis* BCG str. Pasteur. The red region corresponds to pLDDT very low confidence values (<50). The yellow region corresponds to pLDDT low confidence values (60). The green region corresponds to pLDDT ok confidence values (70). The light blue region corresponds to pLDDT confidence values (80). The dark blue region corresponds to pLDDT very high confidence values (>90). (a) *M. avium* MAP01 protein. (b) *M. bovis* BCG MBCGP02 protein. (c) *M. bovis* BCG MBCGP03 protein. (d) *M. tuberculosis* H37Rv MHR04 protein. (e) *M. tuberculosis* H37Rv MHR05 protein. (f) *M. bovis* SP38 MSP06 protein. (g) *M. bovis* SP38 MSP07 protein.

the transmembrane part of the protein. The comparison between the ColabFold model with the corresponding structure of the online AlphaFold database produces a superposition with a TM-score of 0.74 and an RMSD of 2.57 Å. This deviation is caused purely by the misalignment of the low pLDDT regions. However, all the high-confidence regions are precisely aligned.

Cloning, Transformation, Expression, and Purification of Recombinant Protein. The HP28 protein gene sequence was successfully cloned in the pET-28a vector (Figure 3).

Transformation in *E. coli* was confirmed by the formation of colonies in plates with kanamycin and by colony PCR using

specific primers for the T7 promoter, which resulted in size amplification of 879 bp (Figure 4).

Comparative analysis of the pre- and post-induction extract showed the presence of a band with the expected size of 30 kDa, indicating efficient production of the protein of interest. To determine in which fraction the protein was expressed, soluble and insoluble fractions collected during the expression process were analyzed by SDS-PAGE (Figure 5A). After protein purification and dialysis, the presence of the protein of interest was observed on the gel (Figure 5B). Protein quantity was measured using micro-BCA, and the production yield was 34 mg/L.

Intradermal Skin Test. The delayed-type hypersensitivity response was evaluated in vivo using guinea pigs. None of the animals exhibited cutaneous reactions or any other signs associated with the saline solution injections.

In accordance with Brazilian Technical Regulations, animals are considered positive when they exhibit a reaction greater than 4 mm after 72 h post-injection.¹⁷ In this study, induration was observed at all three measurement intervals, with PPD-B showing the highest reading, approximately 21 mm (Figure 6). The response to the rHP28 protein varied with concentration: at 40 μg/100 μL, the average diameters were 6.6 mm at 24 h and 4.4 mm at 48 h, with the reaction disappearing by 72 h. At 60 μg, rHP28 produced diameters of 7.2 mm at 24 h and 3.8 mm at 48 h. Injection of 80 μg resulted in average diameters of 8.6 mm and 7.2 mm, which decreased to 6 mm across the three measurement intervals. The control group (NS) showed an initial reaction at 24 h for all antigens, which decreased or disappeared in subsequent readings.

A two-way ANOVA multiple comparisons test was conducted to assess the statistical difference between sensitized and non-sensitized groups. The results indicate that the rHP28 protein at 80 μg effectively distinguishes sensitized animals from non-sensitized ones ($p < 0.006$), with a 6 mm difference in reaction at the 72 h reading. Additionally, in the non-sensitized group at 72 h, the rHP28 protein showed no noticeable reaction, unlike PPD-B, which still produced a detectable, though weak, reaction.

DISCUSSION

The diagnosis of bovine tuberculosis currently faces challenges due to the lack of specific antigens for pathogenic bacteria, with the existing PPD antigen presenting limitations in both specificity and production. To address this issue, there is an ongoing search for novel targets that could replace PPD in the TST methodology.^{14,18} In this study, we present a comprehensive analysis of selected *Mycobacterium* strains, including their pangenomes and genomic islands. The focus was on identifying unique genes in each strain, particularly those from pathogenic *M. bovis*, to produce and test new antigens. Our bioinformatic analysis led to the identification of two unique genes, MSP06 and MSP07, from *M. bovis* SP38. We successfully produced one of these proteins and tested it as a TST antigen in guinea pigs. This approach has the potential to improve disease management in the livestock production industry.

Four *Mycobacterium* genomes of interest were chosen for this study. Among these, *M. bovis* SP38, the first strain isolated in Brazil, was chosen due to its relevance in developing new diagnostic methods to control bovine tuberculosis (bTB) in the country. This strain is particularly significant given the high prevalence of bTB in Brazil, coupled with widespread

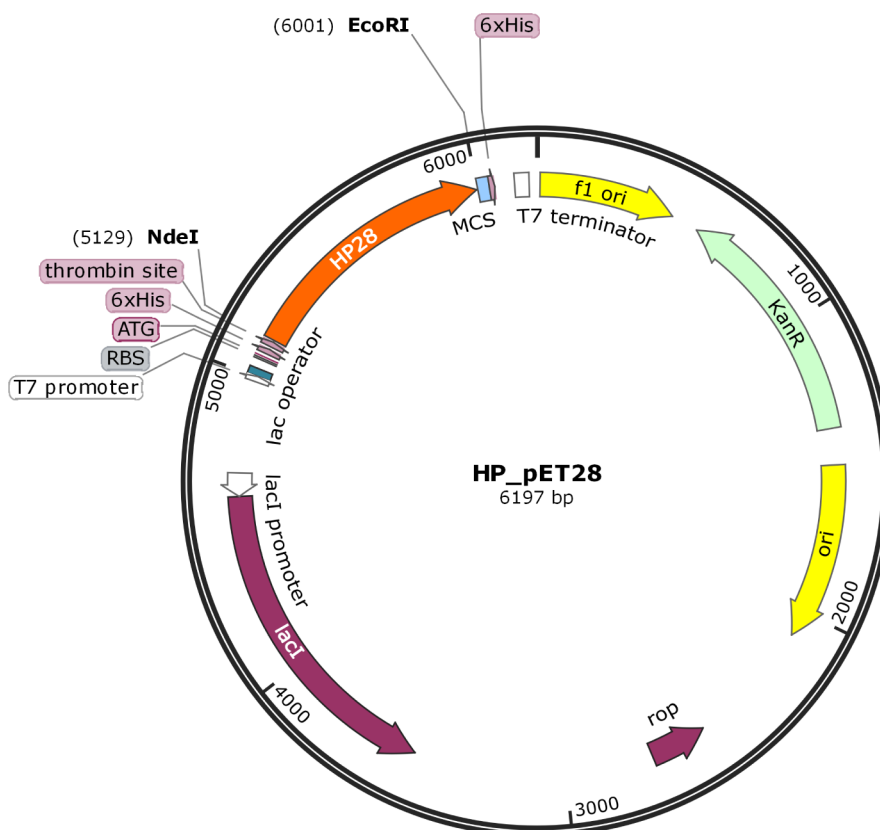


Figure 3. Construction map of the gene that origins the HP28 protein in plasmid pET28a, visualized in SnapGene 1.1.3 software.

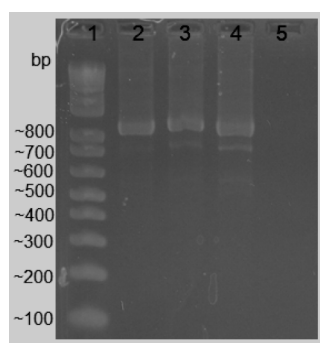


Figure 4. Electrophoresis of PCR reaction using primer T7 in agarose gel 2%. 1—Invitrogen 100 bp DNA Ladder. 2–4—Amplification of transformed colonies. 5—Negative control (colony not transformed).

underreporting.¹⁹ *M. tuberculosis* H37Rv was selected because it is the reference strain in most *in silico* and *in vitro* studies; and *M. bovis* BCG strain Pasteur was chosen to serve as a reference from a vaccine strain. Finally, the *M. avium* strain was integrated for comparison.

The dissection of the *Mycobacterium* pangenome into four different gene categories (“core”, “softcore”, “shell”, and “cloud”) can facilitate strategies for vaccine and the choice of the diagnostic target. Our results revealed variable numbers of cloud genes depending on the strain evaluated, and 12 were identified for *M. bovis* SP38, giving us different possibilities of unique genes to explore. The use of pangenome analysis has already been reported for several organisms, including *Enterococcus faecium*,²⁰ Epstein-Barr virus,²¹ *Leptospira*,²² *Mycobacteroides abscessus*,²³ *Salmonella enterica*,²⁴ and *Bordetella pertussis*.²⁵ Moreover, our results also indicated that the

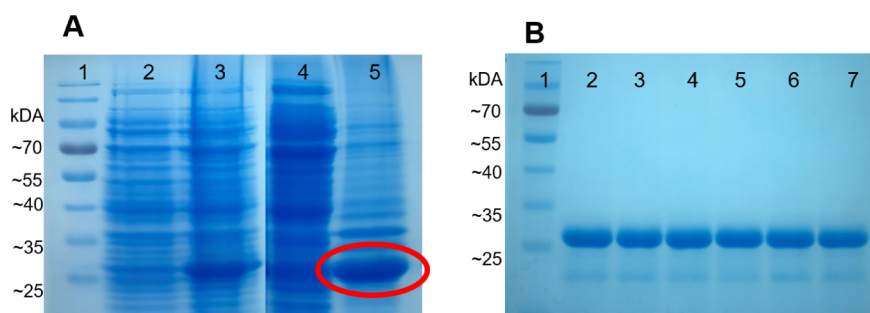


Figure 5. (A) SDS-PAGE 12.5% gel of protein expression. 1—PageRuler Prestained Protein Ladder molecular marker, 10 to 180 kDa; 2—Un-induced culture; 3—Culture induced with 1 mM IPTG; 4—Soluble Fraction; 5—Insoluble Fraction. (B) SDS-PAGE of protein purification. 1—PageRuler Prestained Protein Ladder molecular marker, 10 to 180 kDa; 2–7—Fractions of rHP28 protein eluted during purification.

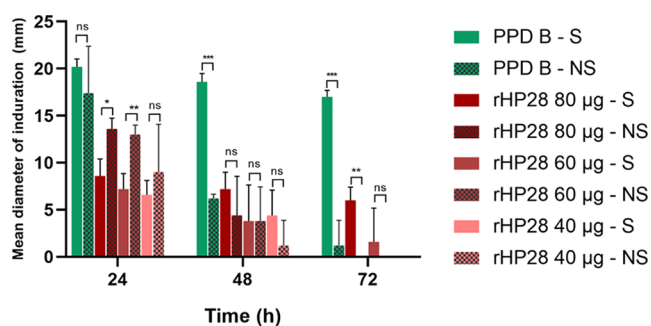


Figure 6. Analysis of intradermal skin test for PPD-B and rHP28 antigens in sensitized (S) and non-sensitized (NS) animals. Error bars demonstrate the standard deviations (SD). Statistical significance was determined using ANOVA *, **, and *** indicate significant differences between groups at $p < 0.05$, $p < 0.005$, and $p < 0.0001$, respectively; ns indicates non-significance.

Mycobacterium pangenome is open and the number of accessory genes will continue to increase, which aligns with the conclusions of other studies.^{26,27}

An important goal in pathogen genomics is to identify and characterize genomic islands (GIs) in pathogenic microorganisms since certain virulence factors are strongly associated with GIs.²⁸ Virulence factors are associated with horizontal gene transitions between species. As for *Mycobacterium*, ancient genomic islands from *M. tuberculosis* may be rooted to the pool of mobile genetic vectors distributed among *Mycobacteria*, as observed by another genomic island study.²⁹ In this study, genomic islands (GIs) were identified for each genome, with the highest numbers found in *M. tuberculosis* and *M. bovis* and the lowest in *M. avium*. For diagnostic and vaccine development, it is crucial that the identified proteins are associated with pathogenicity and/or virulence, as they are more likely to elicit an immunogenic response in the host.³⁰

Our objective of finding unique proteins for each strain was successful when combining the pangenome results with the GIs, considering alignments of less than 60% as the previously indicated threshold.³¹ The analysis generated one protein for *M. avium* and two proteins for *M. tuberculosis* H37Rv, *M. bovis*, and *M. bovis* BCG. Antigenicity analysis was also performed since the goal of this study was to test a protein for bTB diagnosis. From the two proteins identified for *M. bovis*, MSP07 (named HP28) showed an antigenicity score of 0.9272. Therefore, it was chosen for further cloning and expression. All sequences were submitted to ColabFold for protein structure visualization. The resulting models revealed that all proteins were composed of alpha-helices, the most common structural motif in transmembrane proteins.³² These proteins are considered suitable targets for therapeutics,³³ alongside secreted and cell surface membrane proteins.^{34,35} An example of a widely studied alpha-helix protein, known for its virulence and potential as a vaccine or diagnostic candidate, is the ESAT-6 protein.³⁶ A portion of the HP28 protein sequence has been identified in the literature as originating from the ESAT-6 family (UNIPROT accession number: P64094) and is excreted by the bacterium. The protein was selected for recombinant expression. Protein production was successfully obtained with a yield of 34 mg/mL. For *in vivo* testing, we estimated the initial protein concentration for *M. tuberculosis* according to studies previously conducted by our group,^{37,38} thus evaluating concentrations of recombinant protein ranging from 0.0004 to 0.04 µg per 100 µL. The injection point with

0.04 µg proved to be the formulation with the most efficient response, and for this reason, it was selected as our starting point. The ideal protein concentration in the standardization of an intradermal test is variable, and concentrations from 10 to 400 µg have already been reported.^{39–41} Here, we started at 40 µg, followed by 60 and 80 µg. Their response in the group sensitized with *M. bovis* (S) varied from 8.6 to 1.6 mm. Using 80 µg, rHP28 showed a desirable reaction response, within the threshold that differentiate positive from negative tests.¹⁷ We could also observe significant differences in response magnitude between PPD-B and recombinant proteins, which has also been reported in previous studies. Consequently, lower cutoff points should be considered when evaluating tests based on recombinant proteins or peptides.^{42,43}

The most significant reaction in terms of diameter occurred when using 80 µg, reaching 6.6 mm. Three reading times were evaluated here. At the 24-hour period, non-sensitized animals exhibited a significant higher response than sensitized animals. This may be attributed to an initial innate immune response or nonspecific skin irritation in non-sensitized animals. In contrast, sensitized animals, previously exposed to the antigen, likely exhibited a more regulated immune response, with reduced reactivity due to tolerance or diminished inflammation resulting from prior exposure.⁴⁴

Injections of 40 and 60 µg resulted in reactions of 7.2 mm and 6.6 mm, respectively, which decreased over time to 1.6 mm and 0 mm at the final reading. This reduction in reaction size between 24 and 48 h is commonly reported.⁴⁵ During the 72 h period, only the 80-µg injection maintained a reaction size close to the initial 24 h measurement. As mentioned, initial reaction at 24 h was observed for all antigens in both groups, but it completely disappeared in the control group (NS), suggesting that a 24 h reading may not be recommended. Across all responses, the 72 h reading proved to be more accurate, as the PPD-B response was lower (2 mm) in non-sensitized animals. The protein tested here for the first time showed an even lower response, with no perceptible reaction in control animals, clearly differentiating between sensitized and non-sensitized groups. The 72 h induration reaction reading is also recommended by the bTB diagnostic manual⁷ and by Brazilian PNCEBT regulations.¹⁷ Additionally, readings taken after 72 h have been described in studies as ideal, as recombinant proteins such as ESAT-6 and CFP10 reach their peak reaction during this period.^{40,41}

The ability of the rHP28 protein to elicit an induration response and distinguish between sensitized and non-sensitized animals suggests its potential use in bovine tuberculosis diagnosis. To our knowledge, this is the first study to identify and evaluate MSP07 (HP28) as a diagnostic antigen for bovine tuberculosis. This novel finding provides an important contribution to the field, as it demonstrates significant potential to improve diagnostic accuracy. Therefore, we propose its inclusion in diagnostic cocktails alongside ESAT-6, CFP10, and/or EsxI, for example, as antigens combined have been reported to enhance the host's immune response, improve induration reactions, and consequently increase disease detection accuracy.^{37,42}

CONCLUSION

The bioinformatic tools used in this study provided valuable insights into the genetic composition and potential virulence factors of these *Mycobacterium* species. Additionally, these tools demonstrated significant potential in identifying

immunogenic targets for the development and improvement of diagnostic tests, applicable not only to bovine tuberculosis but also to a variety of other infectious diseases. Moreover, the rHP28 protein was successfully cloned and expressed in *E. coli*. When tested intradermally in guinea pigs, the protein effectively differentiated between sensitized and non-sensitized animals, indicating its potential as an antigen for tuberculin skin tests (TST).

MATERIALS AND METHODS

The project planning phase involved identifying novel targets for bovine respiratory diseases using two distinct approaches: (1) identifying unique genes through pangenome analysis and (2) detecting genomic islands associated with virulence. This was followed by the production and testing of the recombinant protein HP28. The workflow can be visualized in Figure 7. Detailed methodology is further described:

Pangenome Analysis. The GFF3 genome files of four *Mycobacterium* strains (*M. bovis* SP38, *M. bovis* BCG, *M. bovis*, and *M. tuberculosis* H37Rv) were retrieved from the National Center for Biotechnology Information (NCBI) database (Table 1). To investigate homologous protein clusters across these genomes, Roary v3.13.0 was used to cluster orthologous genes (OGs) based on a minimum percentage identity of 90% with BLASTp.⁴⁶ The analysis categorized genes into three types: core genes present in all strains, genes present in at least two strains, and strain-specific genes. A binary matrix representing the presence or absence of each gene across all strains was used to estimate the sizes of the pan-genome and core-genome using PanGP v1.0.1 with the totally random sampling algorithm.⁴⁶ Gene distribution analysis was conducted using a Python script, which utilized the Roary gene_presence_absence.csv output files to perform genomic association analysis.

Genomic Island Analysis. The software IslandViewer 4⁵⁰ was used to predict GIs in the four strains of *Mycobacterium* selected (Table 2). GIs < 10 kb were considered genomic islands unless otherwise stated. Some GIs predicted by SIGI-HMM⁵¹ and IslandPath-DIMOB,⁵² two components of IslandViewer 4, were partially superimposed, increasing the number of predictions. After the identification of each predicted GI by IslandViewer 4, the regions were selected and visualized using Artemis software to generate .GBK and .FASTA files for each GI.⁵³ These files contained the amino acid sequence, the name, and the protein ID of each product of each GI, and were used to create multi-fasta files for each microorganism.

Single Protein Analysis and Antigenicity Evaluation. To analyze the individual protein present in each strain, a database of SwissProt bacterial proteomes was downloaded from the UniProt database using a bash script. The output from the genomic island analysis was aligned with the bacterial proteome database through multiple alignments using the BLAST tool in Python. E-value parameters of 1e-50 and BLASTp for alignment were applied. Alignments with an identity of less than 60% were selected to identify unique proteins in each microorganism.

Specifically, for *M. bovis* pathogenicity, the pangenome and virulence island results were compared using a Python script to identify any unique genes within virulence genomic islands. An additional BLAST alignment was performed to ensure that the proteins were not similar to any other *Mycobacterium* species. The protein sequences were then submitted to antigenicity

evaluation using ANTIGENPro (Scratch Protein), where a threshold of 0.5 was considered indicative of an antigenic protein, following the standard criterion established by the software.⁵⁴

Protein Structures. The open-source software ColabFold⁵⁵ was used for the computational analysis of protein structures (versions 1.3.0 and 1.5.2). The notebook “ColabFold: AlphaFold2 using MMseqs2” was used through the Google Collaboratory platform to run AlphaFold with the following parameters:

```
use_templates = false, use_amber = true, msa_mode = "MMseqs2 (UniRef + Environmental)", model_type = "AlphaFold2-ptm", num_models = 5, model_order = [1, 2, 3, 4, 5], num_recycles = 6, rank_by = "pLDDT", max_msa = null, pair_mode = "unpaired + paired".
```

The wild-type sequences of the seven proteins of interest identified in step 3 (MAP01, MBP02, MBP03, MHP04, MHP05, MSP06, and MSP07) were submitted to ColabFold for structure prediction using the specified parameters. The predictions were made without the use of templates. The AlphaFold algorithm generated five structural models for each input, ranking them based on the overall predicted Local Distance Difference Test (pLDDT), AlphaFold's per-residue confidence metric. Additionally, the Predicted Aligned Error (PAE) of each model was visualized in a matrix, indicating confidence in the relative positioning of residue pairs. All structural models underwent relaxation within the Amber99sb force field⁵⁶ as part of the AlphaFold framework. To enhance the accuracy of the results, the number of recycles was increased to six, doubling the default three cycles.

Bacterial Strains. The *M. bovis* strain was obtained from the Reference Bacteria Collection for Sanitary Surveillance at the Oswaldo Cruz Foundation (FIOCRUZ), Brazil. The bacteria were cultured in glass tubes containing Lowenstein-Jensen solid medium for 3–4 weeks at 37 °C.

Cloning, Transformation, Expression and Purification of Recombinant Protein. The full sequence of HP28 was cloned and inserted into the pET28a+ expression plasmid. The plasmid was then used to transform *E. coli* BL21(DE3) via heat shock, using 50 μ L of competent bacteria and 2 μ L of plasmid under the following conditions: 20 min at 4 °C, 45 seconds at 42 °C, and 5 min at 4 °C. Transformation was confirmed through colony PCR using specific primers targeting the T7 promoter.

To express the protein, a pre-inoculum of 20 mL Terrific Broth (TB) medium with 30 μ g/mL kanamycin was incubated overnight at 150 rpm and 37 °C. For the inoculum, 10 mL of the previous culture was transferred to 1 L of fresh TB medium, and the optical density (O.D.) was monitored. When the O.D. reached 0.5, the culture was induced with 1 mM IPTG and incubated overnight with agitation at 16 °C. The cells were harvested by centrifugation at 5000 g for 20 min, and the resulting pellets were sonicated in lysis buffer (50 mM Tris pH 8 and 50 μ g/mL lysozyme). Following a second centrifugation under the same conditions, the pellet was subjected to further sonication in urea buffer (50 mM Tris, 300 mM NaCl, 10 mM imidazole, and 8 M urea). After another centrifugation cycle, the supernatant (insoluble fraction) was collected for purification.

Protein purification was performed using HisTrap HP 5 mL Nickel affinity chromatography column (Cytiva). The column was equilibrated with binding buffer (8 M urea, 10 mM imidazole, 300 mM NaCl, and 20 mM NaH₂PO₄, pH 7.2).

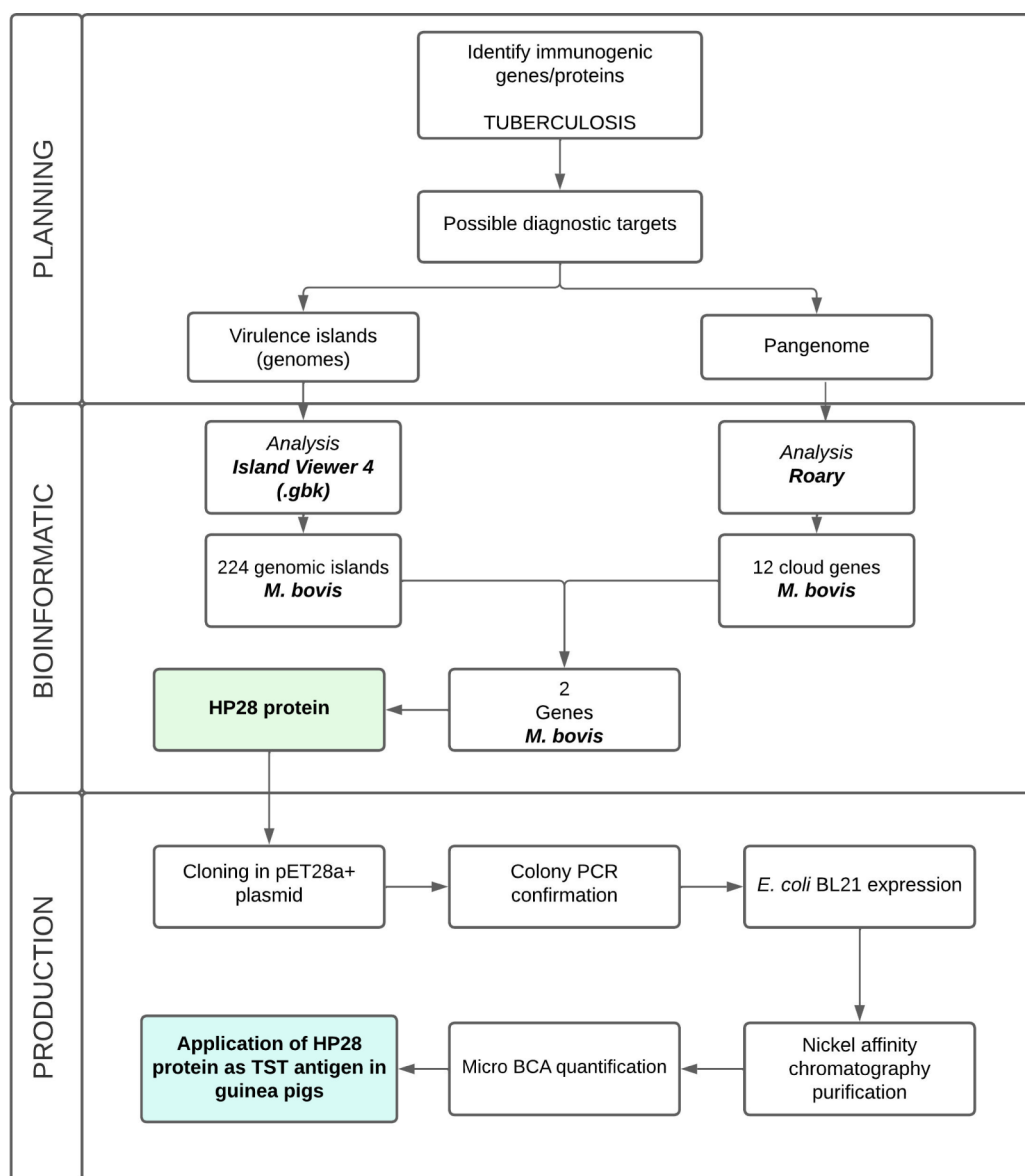


Figure 7. Overall workflow on the study, which was divided into three steps. The first phase consisted of developing the work plan; the second phase consisted of carrying out the bioinformatic study; the third and last phase consisted of testing the protein selected as the antigen to be used by intradermal reaction.

Table 2. *Mycobacterium* Genomes Used in this Study and their Respective References

Name	Genome accession number	Reference
<i>Mycobacterium bovis</i> SP38	NZ_CP015773	Guimarães et al., 2015 ⁴⁷
<i>Mycobacterium bovis</i> BCG Pasteur	NC_008769	Brosch et al., 2007 ⁴⁸
<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i>	NZ_CP033910	Unpublished
<i>Mycobacterium tuberculosis</i> H37Rv	AL123456	Cole et al., 1998 ⁴⁹

After sample loading, 50 mL of washing buffer (8 M urea, 20 mM imidazole, 300 mM NaCl, and 20 mM NaH₂PO₄, pH 7.2) was applied to remove nonspecific proteins. For elution, 10 to 15 mL of elution buffer (8 M urea, 500 mM imidazole, 300 mM NaCl, and 20 mM NaH₂PO₄, pH 7.2) were collected in 2 mL fractions for electrophoresis analysis. After purification, the most concentrated protein aliquots were dialyzed to remove urea and imidazole. Buffer exchanges with 50 mM Tris/0.05% Triton X-100, pH 8.5, were performed every hour to gradually remove urea. Finally, the samples were evaluated by SDS-

PAGE. Protein quantification was carried out using the Micro BCA Protein Assay Kit (Thermo Scientific), and *E. coli* LPS was removed using Pierce High-Capacity Endotoxin Removal Spin Columns (Thermo Scientific).

Intradermal Skin Test. For the production of *M. bovis* inocula, 100 mg of wet mass from mycobacteria cultured on Löwenstein-Jensen solid medium was homogenized with 25 mL of sterilized mineral oil. The mixture was inactivated by autoclaving in flowing steam for 1 h. The final concentration of the prepared inoculum was 4 mg/mL.

The use of ten guinea pigs weighing between 250 and 350 g was approved by the Ethics Committee on Animal Use (CEUA) of the Instituto de Tecnologia do Paraná (TECPAR) under protocol 009/2022. Five female animals were sensitized by intramuscular injection of 0.5 mL (2 mg/mL) of inactivated *M. bovis* and were classified as the sensitized group (S). The remaining five animals served as the control group (NS – not sensitized to *M. bovis*). Both groups were observed for 30 days.

After the 30-day period, the guinea pigs were shaved at the intercostal region, and each antigen was injected intradermally in a volume of 100 μ L. In addition to recombinant protein antigens, standard bovine PPD tuberculin (Laboratorios Microsules) was injected as a positive control, and pure diluent was used as a negative control. The reactions to the injections were assessed 24, 48, and 72 hours post-injection, based on the diameter of the palpable induration area. In Brazil, intradermal reactions are classified according to criteria defined in the Technical Regulation of the Ministry of Agriculture.¹⁷ Animals considered positive exhibit a reaction greater than 4 mm 72 h post-injection.

Statistical analysis of the experiment was conducted using GraphPad Prism 8.0 software.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.4c09374>.

Figure S1. Genomic islands of virulence and pathogenicity predicted by IslandViewer 4. The islands marked with a dark purple dot are genes for virulence factors. Light purple dots are homologous to virulence factors; dark pink dots are resistance genes; light pink dots are the homologous resistance genes, while the ones marked in yellow are pathogen-associated genes. (A) *M. avium*. (B) *M. tuberculosis* H37Rv. (C) *M. bovis* SP38. (D) *M. bovis* BCG str. Pasteur (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vanete Thomaz-Soccol – Laboratório de Biologia Molecular, Programa de Pós Graduação em Engenharia de Bioprocessos e Biotecnologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil; orcid.org/0000-0001-9301-541X; Email: vanetesoccol@gmail.com

Authors

Violetta Dias Pacce – Laboratório de Biologia Molecular, Programa de Pós Graduação em Engenharia de Bioprocessos e Biotecnologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil; orcid.org/0000-0003-1576-5949

Amanda Munari Guimarães – Programa de Pós Graduação em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul 96160-000, Brazil

Frederico Schmitt Kremer – Programa de Pós Graduação em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul 96160-000, Brazil; orcid.org/0000-0001-7284-6054

Gabriela Nascimento Ferreira – Laboratório de Biologia Molecular, Programa de Pós Graduação em Engenharia de

Bioprocessos e Biotecnologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil

Jean Michel Dela Vedova-Costa – Laboratório de Biologia Molecular, Programa de Pós Graduação em Engenharia de Bioprocessos e Biotecnologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil; orcid.org/0000-0002-5656-7239

Aline Cristina dos Santos – Laboratório Provas Biológicas, Instituto de Tecnologia do Paraná, Curitiba, Paraná 80035-060, Brazil

Odir Antônio Dellagostin – Programa de Pós Graduação em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul 96160-000, Brazil

Carlos Ricardo Soccol – Laboratório de Biologia Molecular, Programa de Pós Graduação em Engenharia de Bioprocessos e Biotecnologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil; orcid.org/0000-0001-7630-6864

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acsomega.4c09374>

Author Contributions

V.D.P., O.A.D., and V.T.-S. gave the original idea. A.M.G. performed bioinformatic analysis data. V.D.P. and G.N.F. performed the expression and purification of recombinant protein. V.D.P. and J.M.D.V.-C. provided support in assays in animal experimentation. A.C.S. performed in vivo assays in animal experimentation and provided laboratory space for animals. V.D.P. analyzed experimental data. V.D.P. and A.M.G. wrote the draft of paper. F.S.K. and C.R.S. supervised and edited the paper. V.T.-S. and C.R.S. provided funding, project administration, supervision and for all resources used in this work. Authors have read and agreed to the published version of the manuscript.

Funding

The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil (ROR identifier: 00x0ma614).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support from the Coordination for the Improvement of Higher Education Personnel (CAPES), under grant number 88881.505279/2020-01.

■ ABBREVIATIONS

bTB, bovine tuberculosis; GI, genomic islands; kDa, kilodalton; Mtb, *Mycobacterium tuberculosis*; MTC, *M. tuberculosis* complex; NCBI, National Center for Biotechnology Information; OG, orthologous genes; PPD, purified protein derivative; TST, tuberculin skin test

■ REFERENCES

- Reis, A. C.; Cunha, M. V. Genome-Wide Estimation of Recombination, Mutation and Positive Selection Enlightens Diversification Drivers of *Mycobacterium bovis*. *Sci. Rep.* **2021**, *11* (1), 18789.
- Brosch, R.; Gordon, S. V.; Marmiesse, M.; Brodin, P.; Buchrieser, C.; Eiglmeier, K.; Garnier, T.; Gutierrez, C.; Hewinson,

- G.; Kremer, K.; Parsons, L. M.; Pym, A. S.; Samper, S.; Van Soolingen, D.; Cole, S. T. A New Evolutionary Scenario for the *Mycobacterium tuberculosis* Complex. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (6), 3684–3689.
- (3) Skuce, R. A.; Allen, A. R.; McDowell, S. W. J. Herd-Level Risk Factors for Bovine Tuberculosis: A Literature Review. *Vet. Med. Int.* **2012**, *2012*, 1–10.
- (4) Rue-Albrecht, K.; Magee, D. A.; Killick, K. E.; Nalpas, N. C.; Gordon, S. V.; MacHugh, D. E. Comparative Functional Genomics and the Bovine Macrophage Response to Strains of the *Mycobacterium* Genus. *Front. Immunol.* **2014**, *5*, 536.
- (5) Schiller, I.; Oesch, B.; Vordermeier, H. M.; Palmer, M. V.; Harris, B. N.; Orloski, K. A.; Buddle, B. M.; Thacker, T. C.; Lyashchenko, K. P.; Waters, W. R. Bovine Tuberculosis: A Review of Current and Emerging Diagnostic Techniques in View of Their Relevance for Disease Control and Eradication. *Transboundary Emerging Dis.* **2010**, *57*, 205–220.
- (6) Thoen, C. O.; Steele, J. H.; Gilsdorf, M. J. *Mycobacterium bovis* Infection in Animals and Humans, Blackwell Publishing, 2006, Vol. 5.
- (7) WOA. Bovine Tuberculosis. In *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*; WOA, 2018.
- (8) Hawult, T.; Kassahun, A.; Adane, M.; James, L. N. W.; Ziv, S.; Stefan, B.; Abraham, A. Global Prevalence of *Mycobacterium bovis* Infections among Human Tuberculosis Cases: Systematic Review and Meta-Analysis. *Zoonoses Public Health* **2021**, *68* (7), 704–718.
- (9) Inlamea, O. F.; Soares, P.; Ikuta, C. Y.; Heinemann, M. B.; Achá, S. J.; Machado, A.; Neto, J. S. F.; Correia-Neves, M.; Rito, T.; Yang, R. Evolutionary Analysis of *Mycobacterium bovis* Genotypes across Africa Suggests Co-Evolution with Livestock and Humans. *PLoS Neglected Trop. Dis.* **2020**, *14*, No. e0008081.
- (10) Thoen, C.; LoBue, P.; De Kantor, I. The Importance of *Mycobacterium bovis* as a Zoonosis. *Vet. Microbiol.* **2006**, *112* (2–4), 339–345.
- (11) World Organisation for Animal Health (WOAH). *Bovine tuberculosis*. 2024. <https://www.woah.org/en/disease/bovine-tuberculosis/>.
- (12) Echeverría, G.; Zumárraga, M. J.; Proaño-Pérez, F.; Blasco, F. B.; de Waard, J. H. Assessing the Impact of Various Tuberculin PPD Brands on Bovine Tuberculosis Diagnosis. *Sci. Rep.* **2024**, *14* (1), 5155.
- (13) Humblet, M.; Boschiroli, M.; Saegerman, C. Classification of Worldwide Bovine Tuberculosis Risk Factors in Cattle: A Stratified Approach. *Vet. Res.* **2009**, *40* (5), 50.
- (14) Pacce, V. D.; Oliveira, N. R. D.; Costa, J. M. D.; Dellagostin, O. A.; Soccol, V. T. Peptides and current methods on bovine tuberculosis diagnosis. *Biotechnol. Res. Innov.* **2022**, *6* (1), No. e2020002.
- (15) Kim, Y.; Gu, C.; Kim, H. U.; Lee, S. Y. Current Status of Pan-Genome Analysis for Pathogenic Bacteria. *Curr. Opin. Biotechnol.* **2020**, *63*, 54–62.
- (16) The Computational Pan-Genomics Consortium. The Computational Pan-Genomics Consortium. Computational pan-genomics: Status, promises and challenges. *Brief. Bioinform.* **2018**, *19* (1), 118–135.
- (17) MAPA. *Instrução normativa SDA No10, de 3 de março de 2017*, 2017. <https://www.gov.br/agricultura/pt-br/assuntos/saude-animal-e-vegetal/saude-animal/programas-de-saude-animal/pncebt/principais-normas-pncebt/in-10-de-3-de-marco-de-2017-aprova-o-regulamento-tecnico-do-pncebt.pdf/view>.
- (18) Srinivasan, S.; Jones, G.; Veerasami, M.; Steinbach, S.; Holder, T.; Zewude, A.; Fromsa, A.; Ameni, G.; Easterling, L.; Bakker, D.; et al. A defined antigen skin test for the diagnosis of bovine tuberculosis. *Sci. Adv.* **2019**, *5* (7), No. eaax4899.
- (19) Zimpel, C. K.; Brandão, P. E.; de Souza, A. F.; de Souza, R. F.; Ikuta, C. Y.; Ferreira, N. J. S.; Camargo, N. C. S.; Heinemann, M. B.; Guimarães, A. M. S. Complete Genome Sequencing of *Mycobacterium bovis* SP38 and Comparative Genomics of *Mycobacterium bovis* and *M. tuberculosis* Strains. *Front. Microbiol.* **2017**, *8*, 2389.
- (20) Alotaibi, G.; Khan, K.; Al Mouslem, A. K.; Ahmad Khan, S.; Naseer Abbas, M.; Abbas, M.; Ali Shah, S.; Jalal, K. Pan Genome Based Reverse Vaccinology Approach to Explore *Enterococcus faecium* (VRE) Strains for Identification of Novel Multi-Epitopes Vaccine Candidate. *Immunobiology* **2022**, *227* (3), 152221.
- (21) Priyamvada, P.; Ramaiah, S. Pan-Genome and Reverse Vaccinology Approaches to Design Multi-Epitope Vaccine against Epstein-Barr Virus Associated with Colorectal Cancer. *Immunol. Res.* **2023**, *71* (6), 887–908.
- (22) Zeng, L.; Wang, D.; Hu, N.; Zhu, Q.; Chen, K.; Dong, K.; Zhang, Y.; Yao, Y.; Guo, X.; Chang, Y.-F.; Zhu, Y. A Novel Pan-Genome Reverse Vaccinology Approach Employing a Negative-Selection Strategy for Screening Surface-Exposed Antigens against Leptospirosis. *Front. Microbiol.* **2017**, *8*, 198575.
- (23) Dar, H. A.; Ismail, S.; Waheed, Y.; Ahmad, S.; Jamil, Z.; Aziz, H.; Hetta, H. F.; Muhammad, K. Designing a Multi-Epitope Vaccine against *Mycobacteroides abscessus* by Pangenome-Reverse Vaccinology. *Sci. Rep.* **2021**, *11* (1), 11197.
- (24) Ho, C.; Wu, A. K. L.; Tse, C. W. S.; Yuen, K.-Y.; Lau, S. K. P.; Woo, P. C. Y. Automated Pangenomic Analysis in Target Selection for PCR Detection and Identification of Bacteria by Use of ssGeneFinder Webserver and Its Application to *Salmonella enterica* Serovar Typhi. *J. Clin. Microbiol.* **2012**, *50* (6), 1905–1911.
- (25) Juscamayta-López, E.; Valdivia, F.; Soto, M. P.; Nureña, B.; Horna, H. A Pangenome Approach-Based Loop-Mediated Isothermal Amplification Assay for the Specific and Early Detection of *Bordetella pertussis*. *Sci. Rep.* **2023**, *13* (1), 4356.
- (26) Hurtado-Páez, U.; Álvarez Zuluaga, N.; Arango Isaza, R. E.; Contreras-Moreira, B.; Rouzaud, F.; Robledo, J. Pan-Genome Association Study of *Mycobacterium tuberculosis* Lineage-4 Revealed Specific Genes Related to the High and Low Prevalence of the Disease in Patients from the North-Eastern Area of Medellín, Colombia. *Front. Microbiol.* **2023**, *13*, 1076797.
- (27) Negrete-Paz, A. M.; Vázquez-Marrufo, G.; Gutiérrez-Moraga, A.; Vázquez-Garcidueñas, M. S. Pangenome Reconstruction of *Mycobacterium tuberculosis* as a Guide to Reveal Genomic Features Associated with Strain Clinical Phenotype. *Microorganisms* **2023**, *11* (6), 1495.
- (28) Ho Sui, S. J.; Fedynak, A.; Hsiao, W. W. L.; Langille, M. G. I.; Brinkman, F. S. L. The Association of Virulence Factors with Genomic Islands. *PLoS One* **2009**, *4* (12), No. e8094.
- (29) Reva, O.; Korotetskiy, I.; Ilin, A. Role of the Horizontal Gene Exchange in Evolution of Pathogenic Mycobacteria. *BMC Evol. Biol.* **2015**, *15* (Suppl1), S2.
- (30) Kroesen, V. M.; Madacki, J.; Frigui, W.; Sayes, F.; Brosch, R. Mycobacterial Virulence: Impact on Immunogenicity and Vaccine Research. *F1000Research* **2019**, *8*, 2025.
- (31) Mira, A.; Martín-Cuadrado, A. B.; D'auria, G.; Rodríguez-Valera, F. The Bacterial Pangenome: A New Paradigm in Microbiology. *Int. Microbiol.* **2010**, *13*, 45–57.
- (32) Branden, C. I.; Tooze, J. *Introduction to Protein Structure*; Garland Science, 2012.
- (33) Josic, D.; Clifton, J. G.; Kovac, S.; Hixson, D. C. Membrane Proteins as Diagnostic Biomarkers and Targets for New Therapies. *Curr. Opin. Mol. Ther.* **2008**, *10* (2), 116–123.
- (34) Bonin-Debs, A. L.; Boche, I.; Gille, H.; Brinkmann, U. Development of Secreted Proteins as Biotherapeutic Agents. *Expert Opin. Biol. Ther.* **2004**, *4* (4), 551–558.
- (35) Maffei, B.; Francetic, O.; Subtil, A. Tracking Proteins Secreted by Bacteria: What's in the Toolbox? *Front. Cell. Infect. Microbiol.* **2017**, *7*, 269489.
- (36) Peng, X.; Sun, J. Mechanism of ESAT-6 Membrane Interaction and Its Roles in Pathogenesis of *Mycobacterium tuberculosis*. *Toxicon* **2016**, *116*, 29–34.
- (37) Malaghini, M.; Thomaz-Soccol, V.; Probst, C. M.; Krieger, M. A.; Preti, H.; Kritski, A.; Soccol, C. R. Recombinant Antigen Production for Assays of Intradermoreaction for Diagnosis and Surveillance of Tuberculosis. *J. Biotechnol.* **2011**, *156* (1), 56–58.
- (38) Araújo, L. D. C. T.; Rodriguez-Fernández, D. E.; Karp, S. G.; Marquez, E. D. S.; Santos, A. C. D.; Hospinal-Santiani, M.; Boschero, R. A.; Ramos, E. L. P.; Thomaz-Soccol, V. Formulation and Validation

- of Recombinant Antigens CFP10 and ESAT6 for Tuberculosis Diagnosis. *Braz. Arch. Biol. Technol.* **2021**, *64* (spe), No. e21210127.
- (39) Meneses, I. I. F. D. S.; Melo, E. S. D. P.; Ramos, C. A. D. N.; Osório, A. L. A. R.; Jorge, K. D. S. G.; Vidal, C. E. S.; Silva, M. R.; Pellegrin, A. O.; Araújo, F. R. D. Avaliação de Proteínas Recombinantes Para Detecção de Anticorpos Contra *Mycobacterium bovis* por ELISA. *Bol. Tec. -Cent. Pesqui. Desenvolvimento* **2015**, *20*, 26.
- (40) Pollock, J. M.; McNair, J.; Bassett, H.; Cassidy, J. P.; Costello, E.; Aggerbeck, H.; Rosenkrands, I.; Andersen, P. Specific Delayed-Type Hypersensitivity Responses to ESAT-6 Identify Tuberculosis-Infected Cattle. *J. Clin. Microbiol.* **2003**, *41* (5), 1856–1860.
- (41) Whelan, A. O.; Clifford, D.; Upadhyay, B.; Breadon, E. L.; McNair, J.; Hewinson, G. R.; Vordermeier, M. H. Development of a Skin Test for Bovine Tuberculosis for Differentiating Infected from Vaccinated Animals. *J. Clin. Microbiol.* **2010**, *48* (9), 3176–3181.
- (42) Casal, C.; Bezos, J.; Díez-Guerrier, A.; Álvarez, J.; Romero, B.; de Juan, L.; Rodríguez-Campos, S.; Vordermeier, M.; Whelan, A.; Hewinson, R. G.; Mateos, A.; Domínguez, L.; Aranaz, A. Evaluation of Two Cocktails Containing ESAT-6, CFP-10 and Rv-3615c in the Intradermal Test and the Interferon- γ Assay for Diagnosis of Bovine Tuberculosis. *Prev. Vet. Med.* **2012**, *105* (1–2), 149–154.
- (43) Jones, G. J.; Hewinson, R. G.; Martin Vordermeier, H. Screening of Predicted Secreted Antigens from *Mycobacterium bovis* Identifies Potential Novel Differential Diagnostic Reagents. *Clin. Vaccine Immunol.* **2010**, *17* (9), 1344–1348.
- (44) de Martino, M.; Lodi, L.; Galli, L.; Chiappini, E. Immune Response to *Mycobacterium tuberculosis*: A Narrative Review. *Front. Pediatr.* **2019**, *7*, 350.
- (45) Fernández-Veiga, L.; Fuertes, M.; Geijo, M. V.; Val, B. P. D.; Vidal, E.; Michelet, L.; Boschiroli, M. L.; Gómez-Buendía, A.; Bezos, J.; Jones, G. J.; et al. Differences in Skin Test Reactions to Official and Defined Antigens in Guinea Pigs Exposed to Non-Tuberculous and Tuberculous Bacteria. *Sci. Rep.* **2023**, *13*, 2936.
- (46) Page, A. J.; Cummins, C. A.; Hunt, M.; Wong, V. K.; Reuter, S.; Holden, M. T. G.; Fookes, M.; Falush, D.; Keane, J. A.; Parkhill, J. Roary: Rapid Large-Scale Prokaryote Pan Genome Analysis. *Bioinformatics* **2015**, *31* (22), 3691–3693.
- (47) Guimarães, A.; Zimpel, C.; Ikuta, C.; Do Nascimento, N.; Dos Santos, A.; Messick, J.; Heinemann, M.; Ferreira Neto, J.; Brandão, P. Draft Genome Sequence of *Mycobacterium bovis* Strain SP38, a Pathogenic Bacterium Isolated from a Bovine in Brazil. *Genome Announc.* **2015**, *3* (3), 10–1128.
- (48) Brosch, R.; Gordon, S. V.; Garnier, T.; Eiglmeier, K.; Frigui, W.; Valenti, P.; Dos Santos, S.; Duthoy, S.; Lacroix, C.; Garcia-Pelayo, C.; Inwald, J. K.; Golby, P.; Garcia, J. N.; Hewinson, R. G.; Behr, M. A.; Quail, M. A.; Churcher, C.; Barrell, B. G.; Parkhill, J.; Cole, S. T. Genome Plasticity of BCG and Impact on Vaccine Efficacy. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (13), 5596–5601.
- (49) Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G. Deciphering the Biology of *Mycobacterium tuberculosis* from the Complete Genome Sequence. *Nature* **1998**, *393* (6685), 537–544.
- (50) Bertelli, C.; Laird, M. R.; Williams, K. P.; Lau, B. Y.; Hoad, G.; Winsor, G. L.; Brinkman, F. S. IslandViewer 4: Expanded Prediction of Genomic Islands for Larger-Scale Datasets. *Nucleic Acids Res.* **2017**, *45* (W1), W30–W35.
- (51) Waack, S.; Keller, O.; Asper, R.; Brodag, T.; Damm, C.; Fricke, W. F.; Surovcik, K.; Meinicke, P.; Merkl, R. Score-Based Prediction of Genomic Islands in Prokaryotic Genomes Using Hidden Markov Models. *BMC Bioinf.* **2006**, *7* (1), 142.
- (52) Hsiao, W.; Wan, I.; Jones, S. J.; Brinkman, F. S. L. IslandPath: Aiding Detection of Genomic Islands in Prokaryotes. *Bioinformatics* **2003**, *19* (3), 418–420.
- (53) Rutherford, K.; Parkhill, J.; Crook, J.; Horsnell, T.; Rice, P.; Rajandream, M.; Barrell, B. Artemis: Sequence Visualization and Annotation. *Bioinformatics* **2000**, *16* (10), 944–945.
- (54) Magnan, C. N.; Zeller, M.; Kayala, M. A.; Vigil, A.; Randall, A.; Felgner, P. L.; Baldi, P. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* **2010**, *26* (23), 2936–2943.
- (55) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19* (6), 679–682.
- (56) Showalter, S.; Brüschweiler, R. Validation of Molecular Dynamics Simulations of Biomolecules Using NMR Spin Relaxation as Benchmarks: Application to the AMBER99SB Force Field. *J. Chem. Theory Comput.* **2007**, *3* (3), 961.