



In silico comparisons of lipid-related genes between *Mycobacterium tuberculosis* and BCG vaccine strains

Alice Sarno^{1,2}, Julia Bitencourt¹, Adriano Queiroz¹ and Sergio Arruda^{1,3} 

¹Fundação Oswaldo Cruz, Instituto Gonçalo Moniz, Laboratório Avançado em Saúde Pública, Salvador, BA, Brazil.

²Universidade Federal da Bahia, Salvador, BA, Brazil.

³Universidade do Estado da Bahia, Salvador, BA, Brazil.

Abstract

Despite highly variable efficacy, BCG (Bacillus Calmette-Guérin) is the only vaccine available to prevent the tuberculosis (TB). Genomic heterogeneity between attenuated BCG strains and virulent *Mycobacterium tuberculosis* might help to explain this vaccine's impaired capacity to induce long-term protection. Here, we investigate the lipid-related genes absent in attenuated BCG strains in order to correlate changes in both lipid metabolism and cell-wall lipid content to vaccine impairment. Whole genome sequences of *M. tuberculosis* H37Rv and the six most used BCG strains worldwide were aligned and the absent regions functionally categorized. Genomes of the BCG strains showed a total of 14 non-homologous lipid-related genes, including those belonging to *mce3* operon, as well as the gene *echaA1*, which encodes an enoyl-CoA hydratase, and the genes encoding phospholipases PlcA, PlcB and PlcC. Taken together, the depletion of these *M. tuberculosis* H37Rv genomic regions were associated with marked alterations in lipid-related genes of BCG strains. Such alterations may indicate a dormant-like state and can be determining factors to the vaccine's inability to induce long-term protection. These lipids can be further evaluated as an adjuvant to boost the current BCG-based vaccine.

Keywords: Genome comparison, BCG, lipid, cell-wall, tuberculosis.

Received: January 27, 2021; Accepted: July 22, 2021.

Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, is a leading cause of death worldwide: In 2019 alone, 1.2 million deaths and 10 million new cases were reported. About a quarter of the world population is estimated to harbor latent TB infection and are therefore at risk of developing active disease (WHO, 2020).

BCG (Bacillus Calmette-Guérin), a live-attenuated strain of *Mycobacterium bovis*, is currently the only vaccine available to prevent TB, typically administrated in endemic countries or in populations at high risk of infection (Cernuschi *et al.*, 2018). Currently, six strains account for more than 90% of the vaccines in use worldwide: early strains BCG Moreau, BCG Russian and BCG Tokyo, and late strains BCG Danish, BCG Glaxo and BCG Pasteur (WHO, 2012, 2017). Despite its capacity to protect against disease progression and disseminated forms of TB, the efficacy of BCG against pulmonary TB in adult populations varies from 0% to 80% (Mangtani *et al.*, 2014; Roy *et al.*, 2014).

One of the hypotheses to explain the variable protectiveness of BCG posits the genomic heterogeneity between vaccine and virulent *M. bovis* and *M. tuberculosis* strains (Behr 2002; Liu *et al.*, 2009; Angelidou *et al.*, 2020). Since it was first obtained and distributed, BCG has accumulated large sequence polymorphisms and has lost several virulence factor genes,

including deletion of the region RD1, which encodes antigenic proteins ESAT-6 and CFP-10 (Mahairas *et al.*, 1996; Lewis *et al.*, 2003). However, the expression of RD1 in recombinant BCG does not result in a complete restoration of protection against TB, which could indicate that other mechanisms may be involved in virulence (Pym *et al.*, 2003).

The genomic differences between BCG strains and virulent *M. bovis* and *M. tuberculosis*, as well as the remodeling of protein complexes, have been comprehensively explored through phylogenetic analysis (Brosch *et al.*, 2007; Zhang *et al.*, 2013; Abdallah *et al.*, 2015). However, the impact of genomic heterogeneity on virulence factors related to mycobacteria cell-wall lipid content and lipid metabolism has received less attention (Abdallah *et al.*, 2015). Discrepancies in lipid species in the cell walls of virulent and attenuated strains of mycobacteria might play a key role in host-pathogen interaction (Guenin-Macé *et al.*, 2009; Queiroz and Riley 2017; Mishra *et al.*, 2019). In BCG, genome polymorphisms and the absence of specific cell-wall lipid components have resulted in less-virulent strains that induce a restrained pro-inflammatory immune response and limit BCG-mediated T cell protection, with diminished immunological activity (Hayashi *et al.*, 2009; Tran *et al.*, 2016; Zhang *et al.*, 2016).

Here we compared the whole genome sequences of *M. tuberculosis* H37Rv and the six BCG strains more frequently used worldwide in an attempt to identify genomic differences related to lipid content and metabolism. By this approach, we established a comprehensive list of lipid-related genes absent in these BCG strains, in which the codified molecules may contribute to improve the BCG vaccines currently in circulation.

Material and Methods

Whole genome sequence selection

The following whole genome sequences stored on GenBank were compared *in silico*: *M. tuberculosis* H37Rv (accession number NC_000962.3), early strains *M. bovis* BCG Moreau RDJ (accession number AM412059.2), *M. bovis* BCG Russian 368 (accession number CP009243.1) and *M. bovis* BCG Tokyo 172 (accession number AP010918.1), and late strains *M. bovis* BCG Danish 1331 (accession number CP039850.1), *M. bovis* BCG Glaxo (accession number NZ_CUWJ01000001.1) and *M. bovis* BCG Pasteur 1173P2 (accession number AM408590.1). The six BCG strains were selected for comparison with *M. tuberculosis* H37Rv, since these account for more than 90% of the vaccines in use worldwide.

Determination of homologous and non-homologous regions among sequences

Mauve software (Darling *et al.*, 2010) was used to align, identify and characterize homologous and non-homologous regions among the whole genomes. Regions were considered homologous if percent identity was > 60% and query cover was > 70%. After alignment, the gene annotations for homologous and non-homologous regions were obtained and exported as comma-separated values for further analysis. The number and percentage of homologous and non-homologous regions between each BCG strain and the *M. tuberculosis* H37Rv sequence were compared to measure the similarity among genomes. Finally, the gene annotations for non-homologous regions in each BCG strain were confirmed by BLASTN searches in the NCBI database (Morgulis *et al.*, 2008).

Functional category determination of non-homologous regions of early and late strains of BCG

The gene annotations in non-homologous regions confirmed by BLASTN searches were functionally categorized using the Mycobrowser database (Kapopoulou *et al.*, 2011). Early (BCG Moreau RDJ, BCG Russian and BCG Tokyo 172) and late (BCG Danish 1331, BCG Glaxo 1077 and BCG Pasteur 1173P2) BCG strains were compared to the *M. tuberculosis* H37Rv genome.

Results

Similarities among homologous and non-homologous regions

Sequence alignment was performed using Mauve software to investigate differences and similarities between the *M. tuberculosis* H37Rv and BCG strains genomes and to better visualize homology among the studied genomes. 4,034 genomic regions were identified in the *M. tuberculosis* H37Rv genome, 3,996 in BCG Danish, 3,993 in BCG Glaxo, 3,944 in BCG Moreau, 3,991 in BCG Pasteur, 4,297 in BCG Russian and 3,985 in BCG Tokyo (Figure 1).

The overlap between and the total number of homologous and non-homologous regions across *M. tuberculosis* H37Rv and all six BCG genome sequences, represented as Venn diagrams,

are illustrated in Figure 1. As expected, high homogeneity was observed between most BCG strains and *M. tuberculosis* H37Rv, with up to 94.6% of homologous regions identified in the BCG Moreau genome. BCG Russian was the strain with the greater number of non-homologous regions (16.7%), when compared to *M. tuberculosis* H37Rv. Together, the data shows a comparable genomic heterogeneity between each strain and *M. tuberculosis* H37Rv, as well as the overall similarity among attenuated vaccine strains (Figure S1).

Functional category identification of genes in non-homologous regions

The genes in non-homologous regions of all BCG strains identified in the alignment were confirmed by BLASTN searches and grouped according to functional category using the Mycobrowser database (Figure 2). The distribution of non-homologous regions – with no BLASTN similarity – in each functional category was similar among strains from the same phylogenetic groups: early strains (BCG Moreau, BCG Russian and BCG Tokyo) and late strains (BCG Danish, BCG Glaxo and BCG Pasteur).

Most of the identified non-homologous regions were associated with functional category “insertion sequences and phages”. In comparison to *M. tuberculosis* H37Rv, the number of non-homologous genes in this category was 28 among all BCG strains, representing between 31.11% and 37.33% of all non-homologous regions. In the early strains, the next most common identified category was “intermediary metabolism and respiration”, with 11 (between 14.1% and 14.67%) non-homologous regions, followed by “cell-wall and cell processes”, with 10 (between 12.82% and 13.33%) non-homologous regions. In the late strains, the opposite was identified: “cell-wall and cell processes” was the second most common identified category, with 14 (between 15.56% and 16.28%) non-homologous regions, followed by “intermediary metabolism and respiration”, with 11 (between 12.22% and 12.79%) non-homologous regions. Other categories also associated with non-homologous regions included “conserved hypotheticals” and “virulence, detoxification and adaptation”.

Highlighted in Table 1 are the 14 lipid-related genes absent in all six BCG vaccine strains most commonly used worldwide, when compared to *M. tuberculosis* H37Rv. The genes are associated with the functional categories “cell-wall and cell processes” (4), “virulence, detoxification and adaptation” (6), “lipid metabolism” (1) and “intermediary metabolism and respiration” (3). The complete list of absent genes, in all functional categories, is described in Table S1.

Ten of these 14 genes belonged to the *mce3* operon: four were in the “cell wall and cell processes” category (Rv1970 and Rv1972 to Rv1974) and six (Rv1965 to Rv1969 and Rv1971) in the “virulence, detoxification, adaptation” category. The gene encoding enoyl-CoA hydratase (Rv0222), which is part of the fatty acid degradation metabolism, was categorized as “lipid metabolism”. Finally, three genes encoding phospholipases PlcC, PlcB and PlcA (Rv2349c to Rv2351c, respectively), related to lipid metabolism, were included in the “intermediary metabolism and respiration” category.

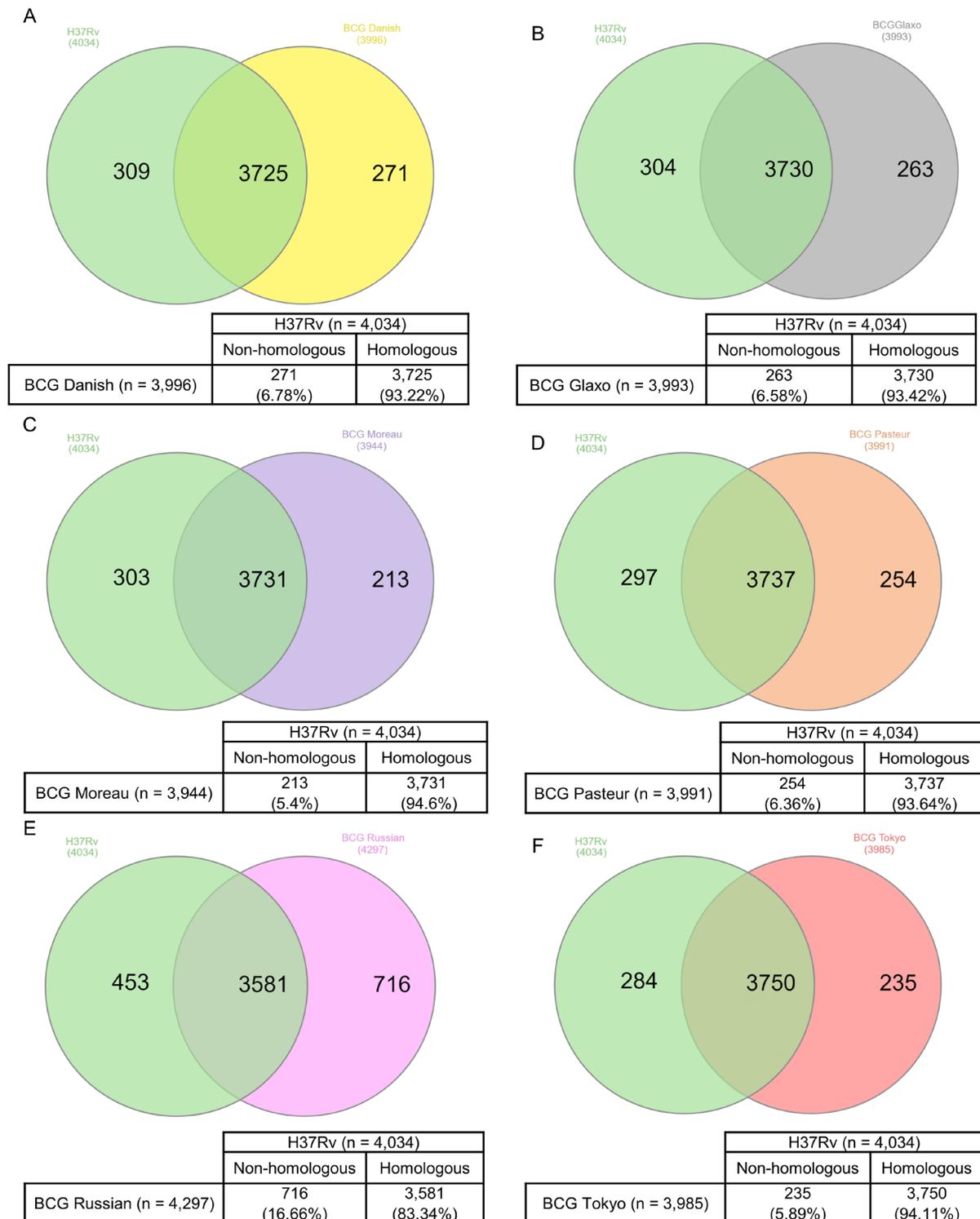
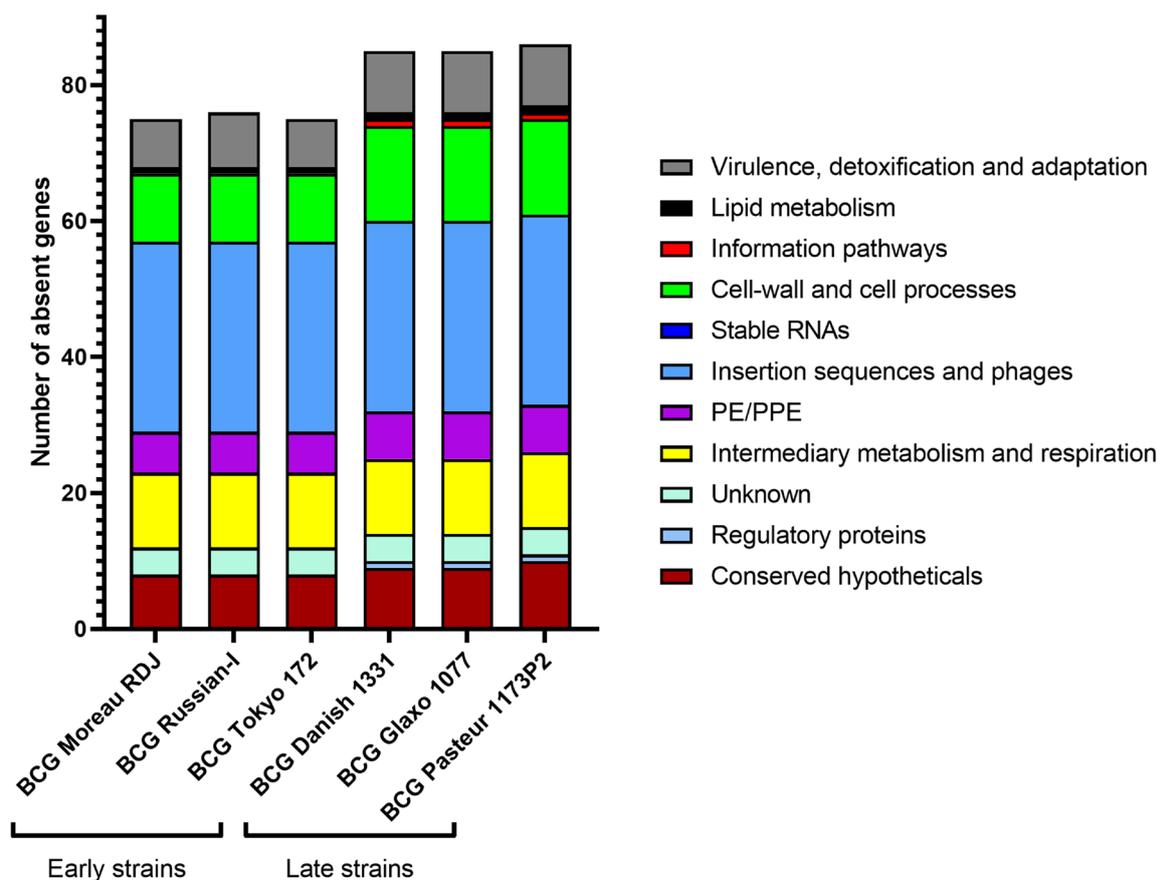


Figure 1 – Homologous and non-homologous regions between *M. tuberculosis* H37Rv, early BCG strains and late BCG strains genome sequences. Venn diagrams showing overlap between and the total number of homologous and non-homologous regions across (A) *M. tuberculosis* H37Rv and BCG Danish, (B) *M. tuberculosis* H37Rv and BCG Glaxo, (C) *M. tuberculosis* H37Rv and BCG Moreau, (D) *M. tuberculosis* H37Rv and BCG Pasteur, (E) *M. tuberculosis* H37Rv and BCG Russian and (F) *M. tuberculosis* H37Rv and BCG Tokyo.

Differences between *M. tuberculosis* H37Rv and BCG strains, previously established in the literature, were also verified. The absence of the five genes encoding the ESAT-6 secretion system-1 (ESX-1) in all six strains: Rv3874 (*esxB*), Rv3875 (*esxA*), Rv3876 (*espI*), Rv3877 (*eccDI*) and Rv3878 (*espJ*), comprised the “cell wall and cell processes”

category (Table S1) (Hsu *et al.*, 2003; Lewis *et al.*, 2003). In addition, the mutation in the Rv2930 (*fadD26*) and Rv2931 (*ppsA*) loci, which impairs the biosynthesis of phthiocerol dimycocerosates (PDIMs) and phenolic glycolipids (PGLs) in BCG Moreau (Chen *et al.*, 2007; Leung *et al.*, 2008), was verified (data not shown).



Functional Category	Early strains			Late strains		
	BCG Moreau	BCG Russian	BCG Tokyo	BCG Danish	BCG Glaxo	BCG Pasteur
	n (%)					
Virulence, detoxification and adaptation	7 (9.33)	8 (10.53)	7 (9.33)	9 (10.59)	9 (10.59)	9 (10.47)
Lipid metabolism	1 (1.33)	1 (1.32)	1 (1.33)	1 (1.18)	1 (1.18)	1 (1.16)
Information pathways	0	0	0	1 (1.18)	1 (1.18)	1 (1.16)
Cell-wall and cell processes	10 (13.33)	10 (13.16)	10 (13.33)	14 (16.47)	14 (16.47)	14 (16.28)
Stable RNAs	0	0	0	0	0	0
Insertion sequences and phages	28 (37.33)	28 (36.84)	28 (37.33)	28 (32.94)	28 (32.94)	28 (32.56)
PE/PPE	6 (8)	6 (7.89)	6 (8)	7 (8.24)	7 (8.24)	7 (8.14)
Intermediary metabolism and respiration	11 (14.67)	11 (14.47)	11 (14.67)	11 (12.94)	11 (12.94)	11 (12.79)
Unknown	4 (5.33)	4 (5.26)	4 (5.33)	4 (4.71)	4 (4.71)	4 (4.65)
Regulatory proteins	0	0	0	1 (1.18)	1 (1.18)	1 (1.16)
Conserved hypotheticals	8 (10.67)	8 (10.53)	8 (10.67)	9 (10.59)	9 (10.59)	10 (11.63)
TOTAL	75 (100)	76 (100)	75 (100)	85 (100)	90 (100)	86 (100)

Figure 2 – Functional categories of non-homologous genes in early and late BCG strains compared to *M. tuberculosis* H37Rv. n (%): number and percentage of non-homologous genes in each functional category relative to the total number of non-homologous genes per genome.

Table 1 – *M. tuberculosis* H37Rv lipid-related genes corresponding to non-homologous regions in all six BCG strains.

Functional category	H37Rv gene locus	Gene product
Cell wall and cell processes	Rv1970	mce3E
	Rv1972-Rv1974	Mce associated membrane proteins
Virulence, detoxification, adaptation	Rv1965-Rv1969	YrbE3B, Mce3A, Mce3B, Mce3C and Mce3D
	Rv1971	Mce3F
Lipid Metabolism	Rv0222	Enoyl-CoA hydratase (EchA1)
Intermediary metabolism and respiration	Rv2349c-Rv2351c	Phospholipases C (PlcC, PlcB and PlcA)

Discussion

The present study identified and categorized a comprehensive list of absent lipid-related genes shared by the most used BCG strains worldwide compared to the *M. tuberculosis* H37Rv genome. The *M. tuberculosis* H37Rv genome, and not *M. bovis*, was used as reference genomic sequence to assure comparison between vaccine strains and the best curated sequence of the causative agent of TB. The cell wall lipid content of mycobacteria plays an important role in the pathogen-host interaction and inflammation (Forrellad *et al.*, 2013; Dulberger *et al.*, 2020; Petrilli *et al.*, 2020). Thus, these genes can be further explored as common targets for virulence and efficacy improvement of the BCG vaccine currently in circulation.

Since the sequencing of the *M. tuberculosis* H37Rv genome in 1998, the *in silico* comparisons of genus, species and isolates have resulted in abundant data of mycobacterial sequences (Cole *et al.*, 1998, 2001; Fraser *et al.*, 2000; Gordon *et al.*, 2001). Furthermore, comparative genomic analyses have supported many molecular based hypotheses regarding the impact of protein virulence factors on the protection induced by BCG strains (Behr *et al.*, 1999; Lewis *et al.*, 2003; Sherman *et al.*, 2004; Zhang *et al.*, 2013). However, the role of mycobacterial cell-wall content and lipid metabolism on virulence has received less attention, with analyses often limited in number of lipid antigens and BCG strains included, as well as comparisons with *M. bovis* sequences (Rhoades *et al.*, 2005; Layre *et al.*, 2014; Abdallah *et al.*, 2015; Tran *et al.*, 2016; Gonzalo-Asensio *et al.*, 2017; Jia *et al.*, 2017; Wright *et al.*, 2017).

Genes at loci Rv1965 to Rv1974, which encode proteins from the *mce3* operon, were found to be absent in all BCG strains, accounting for 10 of the 14 non-homologous genes identified. The *mce3* operon is an important virulence factor, since *M. tuberculosis* strains disrupted on this operon displayed longer survival and lower colony-forming units (CFU) in mice and guinea pig models (Gioffr e *et al.*, 2005; Obreg on-Henao *et al.*, 2011). Proteins Mce3A (Rv1966), Mce3D (Rv1969) and Mce3E (Rv1970) also induced antibody response serum samples from TB patients (Ahmad *et al.*, 2004). Similar to other *mce* operons, the products of *mce3* has been shown to affect the internalization process of mycobacteria (El-Shazly *et al.*, 2007) and are possibly involved in cholesterol and fatty acids transport across the cell wall (Pandey and Sasseti 2008; Mohn *et al.*, 2008; Perkowski *et al.*, 2016). Interestingly, this intake of fatty acids seems to be greatly reduced in the BCG strains that do not produce PDIM (Nazarova *et al.*, 2017), namely BCG Moreau, Tokyo and Glaxo.

A protein involved in the degradation of fatty acid was also found to be absent in all BCG strains compared to *M. tuberculosis* H37Rv. The gene *echaA1* (Rv0222) encodes an enoyl-CoA hydratase (EchA1) involved in energy production via β -oxidation, essential for mycobacterial survival and adaptation in environments with distinct fatty acids as the only carbon sources (Mu oz-El as and McKinney 2006; Srivastava *et al.*, 2015). Despite the gene redundancy involved in five pathways of β -oxidation, EchA1 is secreted to the host cytosol and impairs the production of pro-inflammatory

cytokines, by inhibiting TRAF6 (tumor necrosis factor (TNF)-receptor-associated factor 6) activation (Wang *et al.*, 2020).

With regard to synthesis of glycerolipids, phospholipase C (PlcC) (Rv2349c), PlcB (Rv2350c) and PlcA (Rv2351c) were absent in all BCG strains. These enzymes facilitate the hydrolysis of phosphatidylcholine, phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) to produce diacylglycerol (DAG) (Srinivas *et al.*, 2008) and their absence has been associated with reduced CFU in mice (Raynaud *et al.*, 2002). Furthermore, pre-existing DAG is used for production of triacylglycerol (TAG), which is essential for the survival of *M. tuberculosis* in the host (Garton *et al.*, 2008).

Together, these 14 non-homologous genes may signal a lipid-dependent dormant-like state in all six BCG strains. The absence of *mce3* and *echa1* indicates an overall decline of cholesterol and fatty acid intake in BCG, that could result in lower carbon sources for lipid and energy production. In addition, the absence of *plcC*, *plcB* and *plcA* seems to be associated with lower levels of lipids upstream of DAG and higher levels of TAG. This condition has been previously described in BCG Pasteur (Layre *et al.*, 2011, 2014) and related to long-term dormancy in *M. tuberculosis* (Daniel *et al.*, 2004; Galagan *et al.*, 2013). Therefore, while the lipids increased in level in *M. tuberculosis* H37Rv (such as PE, PG and trehalose-containing lipids) induce a more pro-inflammatory immune response, the accumulation of TAG could be favoring a dormant state in BCG strains.

The identification and study of genes related to cell-wall lipid content and lipid metabolism in BCG strains can contribute to elucidating the impact of attenuation on vaccine virulence and protection efficacy. We suggest that the *M. tuberculosis* lipid-related genes and its products that are absent in BCG should be explored as adjuvants alongside new vaccine candidates due to their capacity to enhance immune response.

Acknowledgments

This work was supported by Bahia State Research Support Foundation, Bahia, Brazil (grant numbers BOL 0172/2019 and BOL0264/2018).

Conflict of Interest

The authors declare no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Author Contributions

AS, JB and AQ were involved in the conceptualization and formal analysis of the study. AS was responsible for data curation and writing of the original draft. AQ and SA supervised the findings of the work. All authors were involved in reviewing and editing the final manuscript. All authors read and approved the final version.

References

- Abdallah AM, Hill-Cawthorne GA, Otto TD, Coll F, Guerra-Assun o JA, Gao G, Naeem R, Ansari H, Malas TB, Adroub SA *et al.* (2015) Genomic expression catalogue of a global collection of BCG vaccine strains show evidence for highly diverged metabolic and cell-wall adaptations. *Sci Rep* 5:15443.

- Ahmad S, El-Shazly S, Mustafa AS and Al-Attayah R (2004) Mammalian cell-entry proteins encoded by the *mce3* operon of *Mycobacterium tuberculosis* are expressed during natural infection in humans. *Scand J Immunol* 60:382–391.
- Angelidou A, Conti M-G, Diray-Arce J, Benn CS, Shann F, Netea MG, Liu M, Potluri LP, Sanchez-Schmitz G, Husson R *et al.* (2020) Licensed Bacille Calmette-Guérin (BCG) formulations differ markedly in bacterial viability, RNA content and innate immune activation. *Vaccine* 38:2229–2240.
- Behr MA (2002) BCG--different strains, different vaccines? *Lancet Infect Dis* 2:86–92.
- Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S and Small PM (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284:1520–1523.
- Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C *et al.* (2007) Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* 104:5596–5601.
- Cernuschi T, Malvolti S, Nickels E and Friede M (2018) Bacillus Calmette-Guérin (BCG) vaccine: A global assessment of demand and supply balance. *Vaccine* 36:498–506.
- Chen JM, Islam ST, Ren H and Liu J (2007) Differential productions of lipid virulence factors among BCG vaccine strains and implications on BCG safety. *Vaccine* 25:8114–8122.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537–544.
- Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D *et al.* (2001) Massive gene decay in the leprosy bacillus. *Nature* 409:1007–1011.
- Daniel J, Deb C, Dubey VS, Sirakova TD, Abomoelak B, Morbidoni HR and Kolattukudy PE (2004) Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in *Mycobacterium tuberculosis* as it goes into a dormancy-like state in culture. *J Bacteriol* 186:5017–5030.
- Darling AE, Mau B and Perna NT (2010) ProgressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e1114.
- Dulberger CL, Rubin EJ and Boutte CC (2020) The mycobacterial cell envelope – a moving target. *Nature* 18:47–59.
- El-Shazly S, Ahmad S, Mustafa AS, Al-Attayah R and Krajei D (2007) Internalization by HeLa cells of latex beads coated with mammalian cell entry (Mce) proteins encoded by the *mce3* operon of *Mycobacterium tuberculosis*. *J Med Microbiol* 56:1145–1151.
- Forrellad MA, Klepp LI, Gioffré A, Sabio y García J, Morbidoni HR, de la Paz Santangelo M, Cataldi AA and Bigi F (2013) Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence* 4:3–66.
- Fraser CM, Eisen J, Fleischmann RD, Ketchum KA and Peterson S (2000) Comparative genomics and understanding of microbial biology. *Emerg Infect Dis* 6:505–512.
- Galagan JE, Minch K, Peterson M, Lyubetskaya A, Azizi E, Sweet L, Gomes A, Rustad T, Dolganov G, Glotova I *et al.* (2013) The *Mycobacterium tuberculosis* regulatory network and hypoxia. *Nature* 499:178–183.
- Garton NJ, Waddell SJ, Sherratt AL, Lee SM, Smith RJ, Senner C, Hinds J, Rajakumar K, Adegbola RA, Besra GS *et al.* (2008) Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med* 5:0634–0645.
- Gioffré A, Infante E, Aguilar D, De La Paz Santangelo M, Klepp L, Amadio A, Meikle V, Etchehoury I, Romano MI, Cataldi A *et al.* (2005) Mutation in *mce* operons attenuates *Mycobacterium tuberculosis* virulence. *Microbes Infect* 7:325–334.
- Gonzalo-Asensio J, Marinova D, Martin C and Aguilo N (2017) MTBVAC: Attenuating the human pathogen of tuberculosis (TB) toward a promising vaccine against the TB epidemic. *Front Immunol* 8:1803.
- Gordon SV, Eiglmeier K, Garnier T, Brosch R, Parkhill J, Barrell B, Cole ST and Hewinson RG (2001) Genomics of *Mycobacterium bovis*. *Tuberculosis (Edinb)* 81:157–163.
- Guenin-Macé L, Siméone R and Demangel C (2009) Lipids of Pathogenic Mycobacteria: Contributions to virulence and host immune suppression. *Transbound Emerg Dis* 56:255–268.
- Hayashi D, Takii T, Fujiwara N, Fujita Y, Yano I, Yamamoto S, Kondo M, Yasuda E, Inagaki E, Kanai K *et al.* (2009) Comparable studies of immunostimulating activities *in vitro* among *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) substrains. *FEMS Immunol Med Microbiol* 56:116–128.
- Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M *et al.* (2003) The primary mechanism of attenuation of bacillus Calmette-Guérin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci U S A* 100:12420–12425.
- Jia X, Yang L, Dong M, Chen S, Lv L, Cao D, Fu J, Yang T, Zhang J, Zhang X *et al.* (2017) The bioinformatics analysis of comparative genomics of *Mycobacterium tuberculosis* Complex (MTBC) provides insight into dissimilarities between intraspecific groups differing in host association, virulence, and epitope diversity. *Front Cell Infect Microbiol* 7:88.
- Kapopoulou A, Lew JM and Cole ST (2011) The MycoBrowser portal: A comprehensive and manually annotated resource for mycobacterial genomes. *Tuberculosis* 91:8–13.
- Layre E, Lee HJ, Young DC, Martinot AJ, Buter J, Minnaard AJ, Annand JW, Fortune SM, Snider BB, Matsunaga I *et al.* (2014) Molecular profiling of *Mycobacterium tuberculosis* identifies tuberculosis nucleoside products of the virulence-associated enzyme Rv3378c. *Proc Natl Acad Sci U S A* 111:2978–2983.
- Layre E, Sweet L, Hong S, Madigan CA, Desjardins D, Young DC, Cheng T-YY, Annand JW, Kim K, Shamputa IC *et al.* (2011) A comparative lipidomics platform for *Mycobacterium tuberculosis* provides chemotaxonomic analysis for biomarker discovery. *Chem Biol* 18:1537–1549.
- Leung AS, Tran V, Wu Z, Yu X, Alexander DC, Gao GF, Zhu B and Liu J (2008) Novel genome polymorphisms in BCG vaccine strains and impact on efficacy. *BMC Genomics* 9:413.
- Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA and Sherman DR (2003) Deletion of RD1 from *Mycobacterium tuberculosis* mimics Bacille Calmette-Guérin attenuation. *J Infect Dis* 187:117–123.
- Liu J, Tran V, Leung AS, Alexander DC and Zhu B (2009) BCG vaccines: Their mechanisms of attenuation and impact on safety and protective efficacy. *Hum Vaccin* 5:70–78.
- Mahairas GG, Sabo PJ, Hickey MJ, Singh DC and Stover CK (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J Bacteriol* 178:1274–1282.
- Mangtani P, Abubakar I, Ariti C, Beynon R, Pimpin L, Fine PEM, Rodrigues LC, Smith PG, Lipman M, Whiting PF *et al.* (2014) Protection by BCG Vaccine against tuberculosis: A systematic review of randomized controlled trials. *Clin Infect Dis* 58:470–480.
- Mishra M, Adhyapak P, Dadhich R and Kapoor S (2019) Dynamic remodeling of the host cell membrane by virulent mycobacterial sulfolipid-1. *Sci Rep* 9:12844.
- Mohn WW, Van Der Geize R, Stewart GR, Okamoto S, Liu J, Dijkhuizen L and Eltis LD (2008) The actinobacterial *mce4* locus encodes a steroid transporter. *J Biol Chem* 283:35368–35374.

- Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R and Schäffer AA (2008) Database indexing for production MegaBLAST searches. *Bioinformatics* 24:1757-1764
- Muñoz-Elias EJ and McKinney JD (2006) Carbon metabolism of intracellular bacteria. *Cell Microbiol* 8:10–22.
- Nazarova EV, Montague CR, La T, Wilburn KM, Sukumar N, Lee W, Caldwell S, Russell DG and VanderVen BC (2017) Rv3723/LucA coordinates fatty acid and cholesterol uptake in *Mycobacterium tuberculosis*. *eLife* 6:e26969
- Obregón-Henao A, Shanley C, Bianco MV, Cataldi AA, Basaraba RJ, Orme IM and Bigi F (2011) Vaccination of guinea pigs using mce operon mutants of *Mycobacterium tuberculosis*. *Vaccine* 29:4302–4307.
- Pandey AK and Sasseti CM (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc Natl Acad Sci U S A* 105:4376–4380
- Perkowski EF, Miller BK, McCann JR, Sullivan JT, Malik S, Allen IC, Godfrey V, Hayden JD and Braunstein M (2016) An orphaned Mce-associated membrane protein of *Mycobacterium tuberculosis* is a virulence factor that stabilizes Mce transporters. *Mol Microbiol* 100:90–107
- Petrilli JD, Müller I, Araújo LE, Cardoso TM, Carvalho LP, Barros BC, Teixeira M, Arruda S, Riley LW and Queiroz A (2020) Differential host pro-inflammatory response to mycobacterial cell wall lipids regulated by the Mce1 Operon. *Front Immunol* 11:1848.
- Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C and Cole ST (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat Med* 9:533–539.
- Queiroz A and Riley LW (2017) Bacterial immunostat: *Mycobacterium tuberculosis* lipids and their role in the host immune response. *Rev Soc Bras Med Trop* 50:9–18.
- Raynaud C, Guillhot C, Rauzier J, Bordat Y, Pelicic V, Manganelli R, Smith I, Gicquel B and Jackson M (2002) Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* 45:203–217.
- Rhoades ER, Geisel RE, Butcher BA, McDonough S and Russell DG (2005) Cell wall lipids from *Mycobacterium bovis* BCG are inflammatory when inoculated within a gel matrix: Characterization of a new model of the granulomatous response to mycobacterial components. *Tuberculosis* 85:159–176.
- Roy A, Eisenhut M, Harris RJ, Rodrigues LC, Sridhar S, Habermann S, Snell L, Mangtani P, Adetifa I, Lalvani A *et al.* (2014) Effect of BCG vaccination against *Mycobacterium tuberculosis* infection in children: systematic review and meta-analysis. *BMJ* 349:g4643
- Sherman DR, Guinn KM, Hickey MJ, Mathur SK, Zakel KL and Smith S (2004) *Mycobacterium tuberculosis* H37Rv:ΔRD1 is more virulent than *M. bovis* Bacille Calmette-Guérin in long-term murine infection. *J Infect Dis* 190:123–126.
- Srinivas M, Rajakumari S, Narayana Y, Joshi B, Katoch VM, Rajasekharan R and Balaji KN (2008) Functional characterization of the phospholipase C activity of Rv3487c and its localization on the cell wall of *Mycobacterium tuberculosis*. *J Biosci* 33:221–230.
- Srivastava S, Chaudhary S, Thukral L, Shi C, Gupta RD, Gupta R, Priyadarshan K, Vats A, Haque AS, Sankaranarayanan R *et al.* (2015) Unsaturated lipid assimilation by mycobacteria requires auxiliary cis-trans Enoyl CoA isomerase. *Chem Biol* 22:1577–1587.
- Tran V, Ahn SK, Ng M, Li M and Liu J (2016) Loss of lipid virulence factors reduces the efficiency of the BCG vaccine. *Sci Rep* 6:29076.
- Wang L, Wu J, Li J, Yang H, Tang T, Liang H, Zuo M, Wang J, Liu H, Liu F *et al.* (2020) Host-mediated ubiquitination of a mycobacterial protein suppresses immunity. *Nature* 577:682–688.
- WHO (2012) Information Sheet - observed rate of vaccine reactions Bacille Calmette-Guérin (BCG) Vaccine. World Health Organization, Geneva.
- WHO (2017) Report on BCG vaccine use for protection against mycobacterial infections including tuberculosis, leprosy, and other nontuberculous mycobacteria (NTM) infections. World Health Organization, Geneva, 77 p.
- WHO (2020) Global Tuberculosis Report. World Health Organization, Geneva, 208 p.
- Wright CC, Hsu FF, Arnett E, Dunaj JL, Davidson PM, Pacheco SA, Harriff MJ, Lewinsohn DM, Schlesinger LS and Purdy GE (2017) The *Mycobacterium tuberculosis* MmpL11 cell wall lipid transporter is important for biofilm formation, intracellular growth, and nonreplicating persistence. *Infect Immun* 85:e00131-17.
- Zhang L, Ru H, Chen F, Jin C, Sun R, Fan X, Guo M, Mai J, Xu W, Lin Q *et al.* (2016) Variable virulence and efficiency of BCG vaccine strains in mice and correlation with genome polymorphism. *Mol Ther* 24:398–405.
- Zhang W, Zhang Y, Zheng H, Pan Y, Liu H, Du P, Wan L, Liu J, Zhu B, Zhao G *et al.* (2013) Genome sequencing and analysis of BCG vaccine strains. *PLoS One* 8:e71243

Internet Resources

- Mauve software, <http://darlinglab.org/mauve/mauve.html> (accessed 12 August 2020)
- Nucleotide Basic Local Alignment Search Tool, <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed 16 November 2020)
- Mycobrowser database, <https://mycobrowser.epfl.ch> (accessed 16 November 2020)

Supplementary material

The following online material is available for this article:
 Table S1 – Complete list of H37Rv lipid-related genes corresponding to non-homologous regions in BCG-Moreau, -Danish, -Glaxo, -Pasteur, -Russian or -Tokyo.
 Figure S1 – Homologous and non-homologous regions between *M. tuberculosis* H37Rv, early BCG strains and late BCG strains genome sequences.

Associate Editor: Ana Tereza R. Vasconcelos

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.