RESEARCH ARTICLE

High Sequence Variability of the *ppE18* Gene of Clinical *Mycobacterium tuberculosis* Complex Strains Potentially Impacts Effectivity of Vaccine Candidate M72/AS01E

Susanne Homolka[1]*, Tanja Ubben[1], Stefan Niemann[1,2]

1 Molecular and Experimental Mycobacteriology, Research Center Borstel, Schleswig–Holstein, Germany,
2 German Centre for Infection Research (DZIF), Partner Site Borstel, Schleswig–Holstein, Germany

* shomolka@fz-borstel.de

Abstract

The development of an effective vaccine is urgently needed to fight tuberculosis (TB) which is still the leading cause of death from a single infectious agent worldwide. One of the promising vaccine candidates M72/AS01E consists of two proteins subunits PepA and PPE18 coded by Rv0125 and Rv1196. However, preliminary data indicate a high level of sequence variability among clinical *Mycobacterium tuberculosis* complex (MTBC) strains that might have an impact on the vaccine efficacy. To further investigate this finding, we determined *ppE18* sequence variability in a well-characterized reference collection of 71 MTBC strains from 23 phylogenetic lineages representing the global MTBC diversity. In total, 100 sequence variations consisting of 96 single nucleotide polymorphisms (SNPs), three insertions and one deletion were detected resulting in 141 variable positions distributed over the entire gene. The majority of SNPs detected were non-synonymous (n = 68 vs. n = 28 synonymous). Strains from animal adapted lineages, e.g., *M. bovis*, showed a significant higher diversity than the human pathogens such as *M. tuberculosis* Haarlem. SNP patterns specific for different lineages as well as for deeper branches in the phylogeny could be identified. The results of our study demonstrate a high variability of the *ppE18* gene even in the N-terminal domains that is normally highly conserved in *ppe* genes. As the N-terminal region interacts with TLR2 receptor inducing a protective anti-inflammatory immune response, genetic heterogeneity has a potential impact on the vaccine efficiency, however, this has to be investigated in future studies.

Introduction

Development of an effective vaccine to fight tuberlulosis (TB) which is still one of the leading infectious diseases worldwide is an ongoing endeavor. In 2014, an estimated 9.0 million people developed TB and 1.5 million died [1]. Infected people have a 10% lifetime risk to get active TB disease and every TB patient can infect up to 10–15 other people/year on average [2]. To break

this cycle, a new vaccine is an important component. So far, the only available prevention to develop severe forms of pediatric non-pulmonary TB is the live Bacille Calmette-Guérin (BCG) vaccine from 1921. However, it is virtually not effective against adult pulmonary TB, which is responsible for most of the TB disease burden worldwide[3]. During the last years, ongoing research work resulted in more than 16 promising TB vaccine candidates with different strategies like viral vectored or protein adjuvant booster vaccines as well as priming and therapeutic vaccines [4,5]. One of the in clinical phase II candidates is the M72/AS01E vaccine developed to boost BCG- or *M. tuberculosis* induced immune responses [4]. The M72 consist of the modified Mtb72F antigen and the AS01E adjuvant containing different immune stimulants [4,6]. Mtb72F comprises two proteins subunits coded by Rv1196 (*ppE18*, *mtb39a*) and Rv0125 (*pepA*, *mtb32a*) [7]. Although Mtb72F/AS01E showed already encouraging results in diverse animal models e.g. guinea pigs and cynomolgus monkeys [8,9] effectiveness of the vaccine candidates is potentially challenged by DNA polymorphisms in the *ppE18* gene [10,11].

TB is caused by members of the *M. tuberculosis* complex (MTBC) comprising twelve different species (*M. tuberculosis*, *M. africanum*, *M. canettii*, *M. bovis*, *M. caprae*, *M. microti*, *M. pinipedii*, *M. origys*, *M. mungi*, *M. suricattae*, the dassie bacillus and the chim-panzee bacillus) which can be further distinguished in at least two major phylogenetic groups (clade 1 and clade 2) and more than 20 phylogenetic lineages/genotypes [12,13]. In several studies, it was shown that the genetic variability of MTBC strains impacts the outcome of infection [14,15], the transmission probability [16,17], the success of spreading of resistant strains [18,19], virulence mechanism [20,21] and the performance of TB vaccines [10,22,23]. Especially with regard to the development of new vaccines, the investigation of the genetic diversity of target genes is indispensable for a maximum of vaccine efficacy. Hebert et al. investigated the genomic diversity of the *ppE18* gene among 225 clinical isolates from two different geographical regions [10]. Single nucleotide polymorphisms (SNPs) as well as insertions (INS) or deletions (DEL) occurred in more than 20% of all strains investigated. Sequence variations were correlated to the three principal genetic groups 1–3 defined by Sreevatsan et al. [24]. However, a thorough analysis of the sequence variability of the *ppE18* gene among strains representing the global diversity of the MTBC with several distinct phylogenetic groups and lineages/genotypes has not yet been performed.

To address this question, we investigated the sequence diversity of the *ppE18* gene in a well characterized reference collection of 68 clinical MTBC isolates (Table 1) representing 23 phylogenetic lineages as well as the ATCC strains H37Rv (ATCC 27294), *M. africanum* (ATCC 25420), and *M. bovis* (ATCC 19210).

Materials and Methods

Strains were cultivated on Löwenstein/Jensen media at the National Reference Center for Mycobacteria in Borstel, Germany and DNA extraction was performed according to a standardized protocol [25]. All strains were analyzed by 24-loci-MIRU-VNTR (Mycobacterial Interspersed Repetitive Units-Variable Number of Tandem Repeats), spoligotyping (Spacer Oligonucleotide typing) as well as deletion analysis (region of difference; RD) to confirm phylogenetic classification (details at www.MIRU-VNTRplus.org). Molecular typing data were analyzed with the BioNumerics software (version 7.5, Applied Maths, Sint-Martens-Latem, Belgium) as instructed by the manufacturers.

For DNA sequence analysis of the *ppE18* gene (Rv1196) polymerase chain reaction (PCR) was performed for all strains of the reference collection. Primer Rv1196F (5`ccc gct gcc gat gag gtg t 3`) was located 223bp upstream of the *ppE18* gene, the reverse primer Rv1196R (5`gag acc gcc gat cac tgg t 3`) was located 47bp downstream of the stop codon of the gene. Direct

Table 1. Strains of the reference collection.

Sample Name	Species	Lineage/Genotype	Phylogenetic Group
1449/02	*M. africanum*	West African 1a	Clade 2
1473/02	*M. africanum*	West African 1a	Clade 2
5434/02	*M. africanum*	West African 1a	Clade 2
10473/01	*M. africanum*	West African 1b	Clade 2
10494/01	*M. africanum*	West African 1b	Clade 2
1443/02	*M. africanum*	West African 1b	Clade 2
10514/01	*M. africanum*	West African 2	Clade 2
10517/01	*M. africanum*	West African 2	Clade 2
5468/02	*M. africanum*	West African 2	Clade 2
9550/00	*M. africanum*	West African 2 ATCC 25420	Clade 2
4258/00	*M. bovis*	Bovis	Clade 2
751/01	*M. bovis*	Bovis	Clade 2
7540/01	*M. bovis*	Bovis	Clade 2
9564/00	*M. bovis*	Bovis ATCC 19210	Clade 2
3040/99	*M. canettii*	Canettii	Clade 2
3041/99	*M. canettii*	Canettii	Clade 2
3151/08	*M. canetti*	Canettii	Clade 2
1694/00	*M. caprae*	Caprae	Clade 2
8986/99	*M. caprae*	Caprae	Clade 2
9577/99	*M. caprae*	Caprae	Clade 2
416/01	*M. microti*	Llama	Clade 2
8753/00	*M. microti*	Llama	Clade 2
1479/99	*M. microti*	Vole	Clade 2
7011/02	*M. pinipedii*	Seal	Clade 2
7739/01	*M. pinipedii*	Seal	Clade 2
12594/02	*M. tuberculosis*	Beijing	Clade 1
1500/03	*M. tuberculosis*	Beijing	Clade 1
1934/03	*M. tuberculosis*	Beijing	Clade 1
3256/02	*M. tuberculosis*	Beijing	Clade 1
3329/02	*M. tuberculosis*	Beijing	Clade 1
4445/02	*M. tuberculosis*	Beijing	Clade 1
1417/02	*M. tuberculosis*	Cameroon	Clade 1
5390/02	*M. tuberculosis*	Cameroon	Clade 1
5400/02	*M. tuberculosis*	Cameroon	Clade 1
2637/02	*M. tuberculosis*	Delhi/CAS	Clade 1
7936/01	*M. tuberculosis*	Delhi/CAS	Clade 1
9915/01	*M. tuberculosis*	Delhi/CAS	Clade 1
1797/03	*M. tuberculosis*	EAI	Clade 2
4850/03	*M. tuberculosis*	EAI	Clade 2
947/01	*M. tuberculosis*	EAI	Clade 2
10469/01	*M. tuberculosis*	Ghana	Clade 1
10493/01	*M. tuberculosis*	Ghana	Clade 1
2570/02	*M. tuberculosis*	Ghana	Clade 1
9679/00	*M. tuberculosis*	H37Rv ATCC 27294	Clade 1
2336/02	*M. tuberculosis*	Haarlem	Clade 1
4130/02	*M. tuberculosis*	Haarlem	Clade 1
9532/03	*M. tuberculosis*	Haarlem	Clade 1

(Continued)

Table 1. (*Continued*)

Sample Name	Species	Lineage/Genotype	Phylogenetic Group
7968/03	*M. tuberculosis*	LAM	Clade 1
8885/03	*M. tuberculosis*	LAM	Clade 1
946/03	*M. tuberculosis*	LAM	Clade 1
10459/03	*M. tuberculosis*	New-1	Clade 1
12591/02	*M. tuberculosis*	New-1	Clade 1
8870/03	*M. tuberculosis*	New-1	Clade 1
2151/03	*M. tuberculosis*	S-type	Clade 1
2318/06	*M. tuberculosis*	S-type	Clade 1
6411/05	*M. tuberculosis*	S-type	Clade 1
11313/03	*M. tuberculosis*	Tur	Clade 1
10264/03	*M. tuberculosis*	Tur	Clade 1
10529/03	*M. tuberculosis*	Tur	Clade 1
2169/99	*M. tuberculosis*	Uganda I	Clade 1
2201/99	*M. tuberculosis*	Uganda I	Clade 1
2333/99	*M. tuberculosis*	Uganda I	Clade 1
2176/99	*M. tuberculosis*	Uganda II	Clade 1
2191/99	*M. tuberculosis*	Uganda II	Clade 1
2253/99	*M. tuberculosis*	Uganda II	Clade 1
1657/03	*M. tuberculosis*	Ural	Clade 1
2679/03	*M. tuberculosis*	Ural	Clade 1
8431/03	*M. tuberculosis*	Ural	Clade 1
4412/04	*M. tuberculosis*	X-type	Clade 1
8431/05	*M. tuberculosis*	X-type	Clade 1
9953/04	*M. tuberculosis*	X-type	Clade 1

ATCC: American Type Culture Collection; CAS: Central Asien; EAI: East African Indian; LAM: Latin American Mediterranean; Tur: Turkish
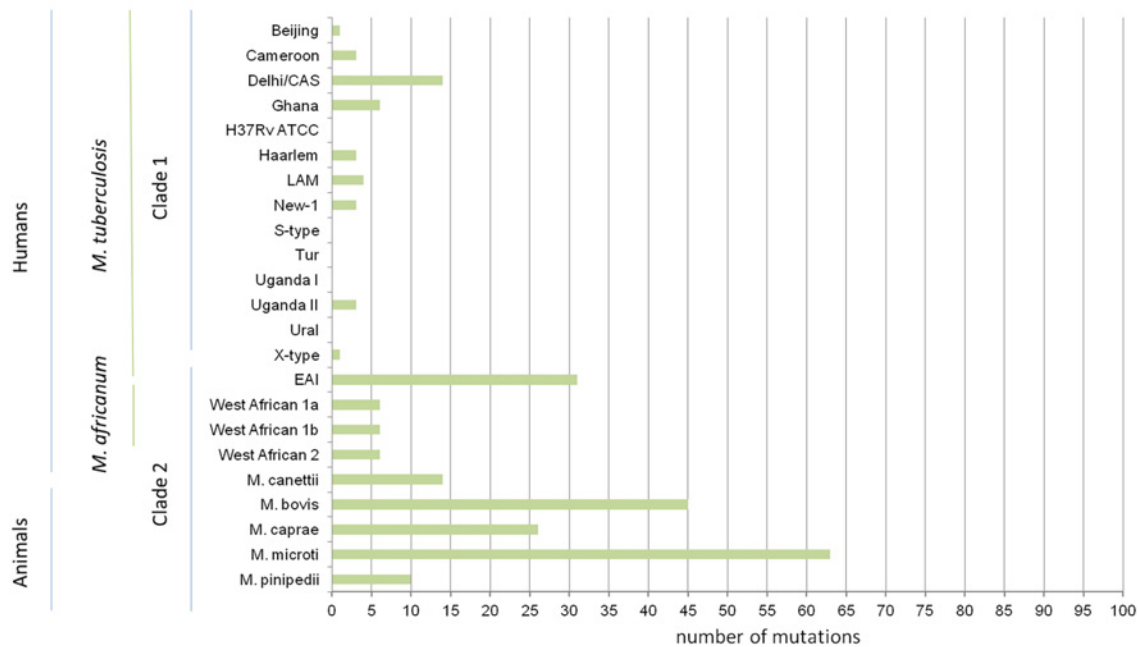
doi:10.1371/journal.pone.0152200.t001

sequencing of the 1446bp PCR fragments were carried out using a commercially available sequencing kit (BigDye terminator v1.1, Applied Biosystems, Foster City, USA) and the ABI 3500XL sequencer according to the manufactures instructions (Applied Biosystem).

Analysis of sequence data (entire *ppE18* gene: 1176bp) and SNP detection was performed by using SeqScape v2.6 software (Applied Biosystems). Genome sequences of *M. tuberculosis* H37Rv (http://tuberculist.epfl.ch/) were used as a reference sequence.

Statistical analysis was performed using GraphPad Prism v.5 (GraphPad Software Inc., USA).

Results

To investigate the genomic diversity of the *ppE18* gene (Rv1196) coding for a protein subunit of the TB vaccine candidate M72/AS01E, direct sequencing of the entire gene (1176bp) was successfully performed in a reference collection of 71 MTBC strains which represent all major phylogenetic lineages described for the MTBC (three strains for most of the lineages, Table 1). This collection included 57 strains of human-adapted (*M. tuberculosis*, *M. africanum* and *M. canettii*) as well as 11 strains of animal adapted lineages (*M. bovis*, *M. microti*, *M. pinipedii*, *M. caprae*) and the ATCC strains H37Rv, *M. bovis* and *M. africanum*. All strains were previously classified into phylogenetic lineages based on spoligotyping, MIRU-VNTR and deletion typing [26]. Sequences obtained were compared to the H37Rv reference genome to identify possible variations e.g. SNPs, deletions or insertions.

Fig 1. Distribution of 100 sequence variation among different MTBC lineages. ATCC: American Type Culture Collection; CAS: Central Asien; EAI: East African Indian; LAM: Latin American Mediterranean; Tur: Turkish.

In total, we found 100 sequence gene variations consisting of 96 SNPs, three insertions (24bp, 6bp, 12bp) and one deletion (3bp). This translates to 141 variable positions in the 1176bp fragment of the *ppE18* gene resulting in an average mutation density of 1.69×10^{-3}/bp (S1 Table). According to previous findings, the majority of the 96 SNPs detected were non-synonymous (n = 68, 70.8%, S1 Table) leading to an amino acid replacement; 28 of them were synonymous (n = 28, 29.2%, S1 Table). Interestingly, transversions and transitions events were equally distributed (transversions n = 45, transitions n = 52, S1 Table), leading to this wide range of amino acid substitutions. However, 16 out of 68 non-synonymous SNPS detected are considered to be conservative mutation which might not impact the protein folding and function. Two of the insertions (6bp INS of Alanine [gcg] and Threonine [acg]; 12bp INS of Glutamine [cag], Leucine [ctg], Glycine [ggt] and Serine [tcg]) and the deletion (3bp DEL of Alanine [gcc]) could be detected at in-frame positions not interrupting the coding sequence (S1 Table and S2 Table). The 24bp INS could be defined as a frameshift mutation interrupting the coding sequence between position 341 and 342 after the start codon of the gene (S1 Table and S2 Table).

Overall, the level of sequence variability differs between strains of different phylogenetic lineages (Fig 1). Animal adapted strains showed a significant higher diversity than the human pathogens e.g. all *M. microti* strains investigated showed 60 lineage specific and three strain specific SNPs (Fig 1 and S1 Fig [p<0.05], S2 Table). In contrast to that, the closely related *M. pinipedii* strains just have ten SNPs in common. However, to compare the genetic diversity between *M. microti* and *M. pinnipedii* strains, further analysis of these specific lineages with more clinical isolates have to be performed. In general, all animal adapted strains share SNP patterns confirming the deep branching of clade 1 and clade II (S2 Table, highlighted in green).

When clade 1 strains are considered, overall sequence variability is lower. Most of the strains investigated do not carry more than 6 SNPs in comparison to the H37Rv reference genome. Strains of four lineages (*M. tuberculosis* S-type, Tur, Uganda I and Ural) showed no specific

mutation at all, confirming their close relationship (Fig 2). Shared SNP patterns between all human adapted strains are comparatively rare (S2 Table, highlighted in blue).

Interestingly, *M. africanum* and *M. canettii* strains, which both belong to the same phylogenetic group as the genetic divers animal adapted strains (clade 2), showed only 6 and 14 sequence variations comparable to the SNP occurrence in *M. tuberculosis* Dehli/CAS or Ghana strains which belong to the same phylogenetic group as the H37Rv reference strain. As *M. canettii* strains are considered as the precursor forms of modern *M. tuberculosis* strains, the low genetic diversity in this hypervariable gene is an unusual observation. Based on the evolutionary age of *M. canettii* strains, these strains normally showed a high amount of sequence variations in comparison to the H37Rv reference strain. In contrast to that, *M. tuberculosis* EAI strains showed 30 SNPs in total, whereas 26 of them were only acquired in one of three EAI strains investigated (S2 Table). Further analysis of the *ppE18* gene of more clinical isolates of this specific lineage is ongoing.

Discussion

TB still remains one of the leading infectious diseases worldwide. The World Health Organization (WHO) declared in 2001 to eliminate TB until 2050 [27]. However, in our current situation especially considering the emergence of multidrug and extremely drug resistant strains, this global aim cannot be achieved. Therefore it is all even more important and necessary do develop an efficient vaccine interrupting ongoing infection and transmission of the disease. However, development of an effective vaccine might be challenged by the genetic diversity within the MTBC that was shown to be higher than previously assumed [12,18,28,29]. Accordingly, the analysis of sequence variability's among MTBC genes coding for protein subunits of new vaccine candidates is urgently needed to estimate the potential to induce a protective
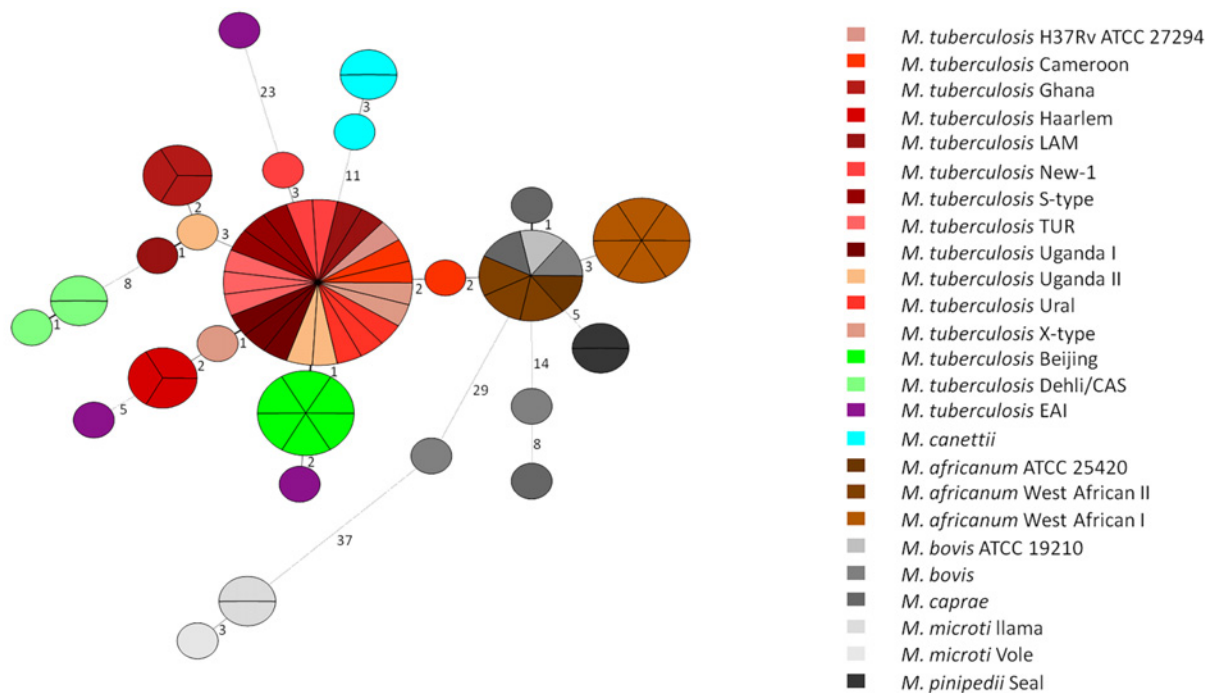
Fig 2. Minimum spanning tree based on 96 SNPs detected in the reference collection. Indicated numbers present SNP differences between the strains investigated. Nodes are highlighted in correspondence to the phylogenetic lineages. ATCC: American Type Culture Collection; CAS: Central Asien; EAI: East African Indian; LAM: Latin American Mediterranean; Tur: Turkish.

doi:10.1371/journal.pone.0152200.g002

immune response against a wide range of MTBC strains of different phylogenetic groups which occurrence has a strong geographical dependence. Here, we investigated the genetic diversity of the *ppE18* gene encoding for a subunit of the new vaccine candidate M72/AS01E. The analysis of a well-characterized reference collection of 71 MTBC strains representing 23 different phylogenetic lineages from different geographical settings provided a comprehensive insight into possible sequence variants which can be used to optimize the vaccine induced immunological response against a wide range of MTBC strains. With an average mutation density of 1.69×10^{-3}/bp, sequence variability was extraordinary high compared to an average variability observed in recent genome studies [30]. Although 16 out of 68 SNPs detected are conservative mutations which might not directly change the folding of the PPE18 protein, the high amount of sequence variations within the entire gene can lead to significant differences in the protein structure. However, to understand the impact of the mutation density on the protein function, molecular modeling approaches as well as functional analysis of this protein is urgently needed.

PE-PPE proteins (Pro-Glu [PE] and Pro-Pro-Glu [PPE]), comprising 10% of the coding capacity of the MTBC genome [31], and are involved in several pathobiological processes [32–35]. It was shown that in contrast to the N-terminal end of the PE-PPE protein superfamily the C-terminal remains highly variable [35] which contributes to e.g. antigenic variation and helps the pathogen to escape from antimicrobial defense mechanism of the host [33]. However, in our study, sequence analysis of the *ppE18* gene revealed a quite high diversity among the entire gene whereas most of the mutations leading to an amino acid exchange. In comparison to other MTBC gene classes like housekeeping genes or genes encoding for surface molecules, the ratio of non- synonymous SNPs (nSNPs) to synonymous SNPs (sSNPs) is extremely high (nSNPs:sSNPs housekeeping 1.59; nSNPs:sSNPs surface 1.60; nSNPs:sSNPs *ppE18* 2.43) [29]. Most of the mutations with potential functional impact were detected in animal adapted strains as well as in *M. tuberculosis* EAI strains which might explain discrepancies to previous findings described by Mortier et al. postulating a high level of sequence conservation among MTBC strains in this particular gene [36]. However, Mortier and colleagues excluded all animal adapted strains as well as the *M. canettii* lineage. In contrast to human restricted strains, a higher genetic diversity within the animal adapted lineages might reflect a strong and enduring selection to escape the immune response induced by a bright range of hosts cells of different animal reservoirs [37]. Another possible explanation might be an increased rate of homologous recombination events enforced by the adaptation to different host environments in comparison to classical *M. tuberculosis* strains [33]. Whether the genetic diversity and immunity of host cells shape the MTBC pathogen evolution should be investigated in future studies. Furthermore, the evolutionary impact of the host-pathogen-interaction on vaccine efficacy has to be elaborated. Based on our sequence analysis, an epitope prediction study which was perfomed by McNamara et al. in 2010 [11] has to be repeated to understand the impact of the genetic diversity of the host cells as well as of different MTBC lineages on the vaccine efficacy. McNamara and colleagues predicted that the MTB72f vaccine is less effective for several DRB1 genotypes occuring especially in high burden TB countries due to limited vaccine epitope binding or to binding primarily by unconserved PPE18 epitopes [11].

Nair et al. showed that PPE18 inhibits proinflammatory cytokine production (e.g. Il12p-40) by binding to TLR2 receptors and by inducing Il-10 secretion in macrophages[38,39]. The TLR2 interacting domain is localized within the N-terminal PPE domain whereas amino acid 1–180 are essential [39]. In our study, we detected 36 mutations in this specific area, most of them leading to an amino acid exchange (S1 Table). Especially animal adapted strains show a genetic variability in this N-terminal domain which might explain the comparable low immunogenicity of the *M. tuberculosis* antigen in BCG vaccinated and *M bovis* infected cattles and

argue for specific immune adaptation due to a wider range of host species [40]. However, if the genetic diversity observed also plays a significant role on the modulation of a protective anti-inflammatory immune response in humans has to be further investigated.

Supporting Information

S1 Fig. Box-Whisker-Plot indicate the mutation distribution among MTBC strains. Mann-Whitney test showed significant differences between human and animal adapted MTBC lineages (p<0.05).
(TIF)

S1 Table. Mutations detected in the reference collection.
(XLSX)

S2 Table. Mutations detected in each strain of the reference collection. Shared SNP patterns are highlighted. Green: among animal adapted strains; Blue: among human pathogens; Pink: among human and animal adapted strains
(XLSX)

Acknowledgments

Author Contributions

Conceived and designed the experiments: SH. Performed the experiments: TU SH. Analyzed the data: SH SN. Contributed reagents/materials/analysis tools: SN. Wrote the paper: SH SN.

References

1. World Health Organisation. Global tuberculosis report 2014. 2014. Available: http://www.who.int/tb/publications/global_report/en/

2. Ottenhoff THM, Kaufmann SHE. Vaccines against tuberculosis: where are we and where do we need to go? PLoS Pathog. 2012; 8: e1002607. doi: 10.1371/journal.ppat.1002607 PMID: 22589713

3. Marais BJ, Raviglione MC, Donald PR, Harries AD, Kritski AL, Graham SM, et al. Scale-up of services and research priorities for diagnosis, management, and control of tuberculosis: a call to action. Lancet Lond Engl. 2010; 375: 2179–2191. doi: 10.1016/S0140-6736(10)60554-5

4. Montagnani C, Chiappini E, Galli L, de Martino M. Vaccine against tuberculosis: what's new? BMC Infect Dis. 2014; 14: S2. doi: 10.1186/1471-2334-14-S1-S2

5. Frick M. The tuberculosis vaccine pipeline. Back to basic science. 2014. Available: http://www.treatmentactiongroup.org/sites/g/files/g450272/f/201407/2014%20Pipeline%20Report%20Full.pdf

6. Leroux-Roels I, Forgus S, De Boever F, Clement F, Demoitié M- A, Mettens P, et al. Improved CD4+ T cell responses to Mycobacterium tuberculosis in PPD-negative adults by M72/AS01 as compared to the M72/AS02 and Mtb72F/AS02 tuberculosis candidate vaccine formulations: a randomized trial. Vaccine. 2013; 31: 2196–2206. doi: 10.1016/j.vaccine.2012.05.035 PMID: 22643213

7. Von Eschen K, Morrison R, Braun M, Ofori-Anyinam O, De Kock E, Pavithran P, et al. The candidate tuberculosis vaccine Mtb72F/AS02A: Tolerability and immunogenicity in humans. Hum Vaccin. 2009; 5: 475–482. PMID: 19587528

8. Brandt L, Skeiky YAW, Alderson MR, Lobet Y, Dalemans W, Turner OC, et al. The protective effect of the Mycobacterium bovis BCG vaccine is increased by coadministration with the Mycobacterium tuberculosis 72-kilodalton fusion polyprotein Mtb72F in M. tuberculosis-infected guinea pigs. Infect Immun. 2004; 72: 6622–6632. doi: 10.1128/IAI.72.11.6622–6632.2004 PMID: 15501795

9. Reed SG, Coler RN, Dalemans W, Dalemans W, Tan EV, DeLa Cruz EC, et al. Defined tuberculosis vaccine, Mtb72F/AS02A, evidence of protection in cynomolgus monkeys. Proc Natl Acad Sci U S A. 2009; 106: 2301–2306. doi: 10.1073/pnas.0712077106 PMID: 19188599

10. Hebert AM, Talarico S, Yang D, Durmaz R, Marrs CF, Zhang L, et al. DNA polymorphisms in the pepA and PPE18 genes among clinical strains of Mycobacterium tuberculosis: implications for vaccine efficacy. Infect Immun. 2007; 75: 5798–5805. doi: 10.1128/IAI.00335-07 PMID: 17893137

11. McNamara LA, He Y, Yang Z. Using epitope predictions to evaluate efficacy and population coverage of the Mtb72f vaccine for tuberculosis. BMC Immunol. 2010; 11: 18. doi: 10.1186/1471-2172-11-18 PMID: 20353587

12. Wirth T, Hildebrand F, Allix-Béguec C, Wölbeling F, Kubica T, Kremer K, et al. Origin, spread and demography of the Mycobacterium tuberculosis complex. PLoS Pathog. 2008; 4: e1000160. doi: 10.1371/journal.ppat.1000160 PMID: 18802459

13. Coscolla M, Gagneux S. Consequences of genomic diversity in Mycobacterium tuberculosis. Semin Immunol. 2014; 26: 431–444. doi: 10.1016/j.smim.2014.09.012 PMID: 25453224

14. Stavrum R, PrayGod G, Range N, Faurholt-Jepsen D, Jeremiah K, Faurholt-Jepsen M, et al. Increased level of acute phase reactants in patients infected with modern Mycobacterium tuberculosis genotypes in Mwanza, Tanzania. BMC Infect Dis. 2014; 14: 309. doi: 10.1186/1471-2334-14-309 PMID: 24903071

15. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M, Adetifa IM, et al. Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. J Infect Dis. 2008; 198: 1037–1043. doi: 10.1086/591504 PMID: 18702608

16. Gehre F, Antonio M, Otu JK, Sallah N, Secka O, Faal T, et al. Immunogenic Mycobacterium africanum strains associated with ongoing transmission in The Gambia. Emerg Infect Dis. 2013; 19: 1598–1604. doi: 10.3201/eid1910.121023 PMID: 24050158

17. Warner DF, Koch A, Mizrahi V. Diversity and disease pathogenesis in Mycobacterium tuberculosis. Trends Microbiol. 2015; 23: 14–21. doi: 10.1016/j.tim.2014.10.005 PMID: 25468790

18. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. Nat Genet. 2015; 47: 242–249. doi: 10.1038/ng.3195 PMID: 25599400

19. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. Science. 2006; 312: 1944–1946. doi: 10.1126/science.1124410 PMID: 16809538

20. Reiling N, Homolka S, Walter K, Brandenburg J, Niwinski L, Ernst M, et al. Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. mBio. 2013; 4. doi: 10.1128/mBio.00250-13

21. Ribeiro SCM, Gomes LL, Amaral EP, Andrade MRM, Almeida FM, Rezende AL, et al. Mycobacterium tuberculosis strains of the modern sublineage of the Beijing family are more likely to display increased virulence than strains of the ancient sublineage. J Clin Microbiol. 2014; 52: 2615–2624. doi: 10.1128/JCM.00498-14 PMID: 24829250

22. Cohen T, Colijn C, Murray M. Modeling the effects of strain diversity and mechanisms of strain competition on the potential performance of new tuberculosis vaccines. Proc Natl Acad Sci U S A. 2008; 105: 16302–16307. doi: 10.1073/pnas.0808746105 PMID: 18849476

23. Davila J, Zhang L, Marrs CF, Durmaz R, Yang Z. Assessment of the genetic diversity of Mycobacterium tuberculosis esxA, esxH, and fbpB genes among clinical isolates and its implication for the future immunization by new tuberculosis subunit vaccines Ag85B-ESAT-6 and Ag85B-TB10.4. J Biomed Biotechnol. 2010; 2010: 208371. doi: 10.1155/2010/208371 PMID: 20617139

24. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. Proc Natl Acad Sci U S A. 1997; 94: 9869–9874. PMID: 9275218

25. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol. 1993; 31: 406–409. PMID: 8381814

26. Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. MIRU-VNTRplus: a web tool for polyphasic genotyping of Mycobacterium tuberculosis complex bacteria. Nucleic Acids Res. 2010; 38: W326–331. doi: 10.1093/nar/gkq351 PMID: 20457747

27. World Health Organisation. Stop TB Partnership. The Global Plan to Stop TB 2011–2015. Available: http://www.stoptb.org/assets/documents/global/plan/tb_globalplantostoptb2011-2015.pdf

28. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat Genet. 2013; 45: 1176–1182. doi: 10.1038/ng.2744 PMID: 23995134

29. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, et al. High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography. PLoS Biol. 2008; 6: e311. doi: 10.1371/journal.pbio.0060311 PMID: 19090620

30. Sherman DR, Gagneux S. Estimating the mutation rate of Mycobacterium tuberculosis during infection. Nat Genet. 2011; 43: 400–401. doi: 10.1038/ng.815 PMID: 21522177

31. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998; 393: 537–544. doi: 10.1038/31159 PMID: 9634230

32. Ahmed A, Das A, Mukhopadhyay S. Immunoregulatory functions and expression patterns of PE/PPE family members: Roles in pathogenicity and impact on anti-tuberculosis vaccine and drug design. IUBMB Life. 2015; 67: 414–427. doi: 10.1002/iub.1387 PMID: 26104967

33. Karboul A, Mazza A, Gey van Pittius NC, Ho JL, Brousseau R, Mardassi H. Frequent homologous recombination events in Mycobacterium tuberculosis PE/PPE multigene families: potential role in antigenic variability. J Bacteriol. 2008; 190: 7838–7846. doi: 10.1128/JB.00827-08 PMID: 18820012

34. Bhat KH, Ahmed A, Kumar S, Sharma P, Mukhopadhyay S. Role of PPE18 protein in intracellular survival and pathogenicity of Mycobacterium tuberculosis in mice. PloS One. 2012; 7: e52601. doi: 10.1371/journal.pone.0052601 PMID: 23300718

35. Sampson SL. Mycobacterial PE/PPE proteins at the host-pathogen interface. Clin Dev Immunol. 2011; 2011: 497203. doi: 10.1155/2011/497203 PMID: 21318182

36. Mortier M-C, Jongert E, Mettens P, Ruelle J-L. Sequence conservation analysis and in silico human leukocyte antigen-peptide binding predictions for the Mtb72F and M72 tuberculosis candidate vaccine antigens. BMC Immunol. 2015; 16: 63. doi: 10.1186/s12865-015-0119-7 PMID: 26493839

37. Cobey S. Pathogen evolution and the immunological niche. Ann N Y Acad Sci. 2014; 1320: 1–15. doi: 10.1111/nyas.12493 PMID: 25040161

38. Nair S, Pandey AD, Mukhopadhyay S. The PPE18 protein of Mycobacterium tuberculosis inhibits NF-κB/rel-mediated proinflammatory cytokine production by upregulating and phosphorylating suppressor of cytokine signaling 3 protein. J Immunol Baltim Md 1950. 2011; 186: 5413–5424. doi: 10.4049/jimmunol.1000773

39. Nair S, Ramaswamy PA, Ghosh S, Joshi DC, Pathak N, Siddiqui I, et al. The PPE18 of Mycobacterium tuberculosis interacts with TLR2 and activates IL-10 induction in macrophage. J Immunol Baltim Md 1950. 2009; 183: 6269–6281. doi: 10.4049/jimmunol.0901367

40. Mustafa AS, Skeiky YA, Al-Attiyah R, Alderson MR, Hewinson RG, Vordermeier HM. Immunogenicity of Mycobacterium tuberculosis Antigens in Mycobacterium bovis BCG-Vaccinated and M. bovis-Infected Cattle. Infect Immun. 2006; 74: 4566–4572. doi: 10.1128/IAI.01660- PMID: 16861643