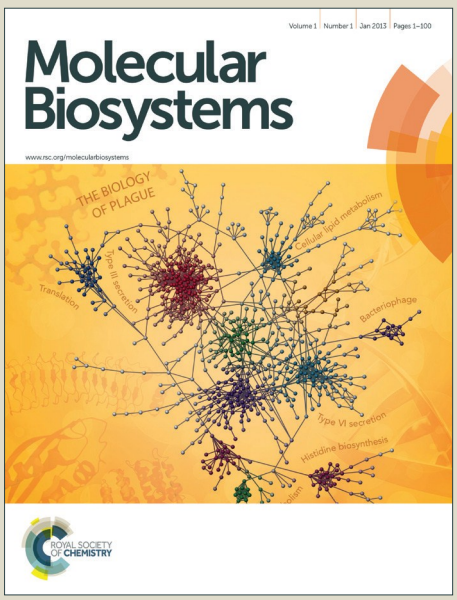
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Analyses of methyltransferases across the pathogenicity spectrum of different Mycobacterial species point to an extremophile connection

Sonam Grover[1#], Paras Gupta[1#], Parvinderdeep S. Kahlon[1,$], Sukriti Goyal[2], Abhinav Grover[2], Kuldeep Dalal[1], Sabeeha[1], Nasreen Z. Ehtesham[3], Seyed E. Hasnain[1,4*]

[1] Molecular Infection and Functional Biology Lab, Kusuma School of Biological Sciences, Indian Institute of Technology, Delhi, New Delhi-110016, India

[2] School of Biotechnology, Jawaharlal Nehru University, New Delhi-110067, India

[3] National Institute of Pathology, Safdarjang Hospital Campus, New Delhi-110029, India

[4] Dr Reddy's Institute of Life Sciences, University of Hyderabad Campus, Prof C R Rao Road, Hyderabad-500046, India

[$] Present Address: Institute of Cellular and Molecular Botany, University of Bonn, Bonn-153115, Germany

[#] These authors contributed equally to this work

* Correspondence should be addressed to:
Prof. Seyed E. Hasnain
Molecular Infection and Functional Biology Lab
Kusuma School of Biological Sciences,
Indian Institute of Technology (IIT) Delhi,
Hauz Khas New Delhi-110016, India
E-mail: seyedhasnain@gmail.com; seh@bioschool.iitd.ac.in

Abstract

Tuberculosis is a devastating disease, taking one human life every 20 seconds globally. We hypothesize that professional pathogens such as *M.tb* have acquired specific features that might assist in causing infection, persistence and transmissible pathology in its host. We have identified 121 methyltransferases (MTases) in the *M.tb* proteome which use a variety of substrates - DNA, RNA, protein, intermediates of mycolic acid biosynthesis and other fatty acids - which are involved in cellular maintenance within the host. Comparative analysis of the proteome of virulent H_{37}Rv and avirulent strain H_{37}Ra, identified 3 MTases which displayed significant variations in terms of N-terminal extension/deletion and point mutations, possibly impacting various physicochemical properties. The cross-proteomic comparison of MTases of *M.tb* H_{37}Rv with 15 different Mycobacterium species revealed the acquisition of novel MTases in MTB complex as a function of evolution. Phylogenetic analysis revealed that these newly acquired MTases showed common roots with certain extremophiles such as halophilic and acidophilic organisms. Our results establish an evolutionary relationship of *M.tb* with halotolerant organisms and also implicate MTases of *M.tb* in withstanding the host osmotic stress, thereby pointing to its likely role in pathogenesis, virulence and niche adaptation.

Keywords: Methyltransferase, Halotolerant, Mycobacterium, Extremophiles

Contact: seyedhasnain@gmail.com; seh@bioschool.iitd.ac.in

Introduction

Tuberculosis (TB) takes one human life every 20 seconds globally, the highest for any bacterial disease [1]. The causative agent for TB, *Mycobacterium tuberculosis (M.tb)*, has evolved the capability to persist in a dormant state for years under extreme conditions within its human host [2, 3].

In bacteria post-transcriptional modifications of rRNA serve a variety of purposes, from stabilizing ribosome structure to preserving its functional integrity. Nucleotide methylations in 16S rRNA occur during different stages of ribosome maturation and assembly, catalysed by specific MTases. Non-functioning of *tlyA* gene, which codes for a 2-*O*-methyltransferase and is involved in methylation of nucleotide C1409 of 16 S rRNA and C1920 of 23 S rRNA of *M.tb* leads to resistance to capreomycin and viomycin [4]. Inactivation of *GidB*, encoding 7-methylguanosine methyltransferase specific for 16S rRNA, confers streptomycin resistance in *M.tb* [5]. Rv2372c is proposed as a prospective 'high-confidence' drug target for *M.tb*, due to the presence of U1498 in the crucial helix 44 of the ribosome, the target of methylation by Rv2372c [6]. Significance of protein methylation is evident from the fact that recombinant HBHA (Heparin-binding Haemagglutinin); important vaccine candidate produced in *E.coli* is not immunogenic and methylation of HBHA is required for the full immunological properties of the protein [7, 8]. HBHA is an adhesin, which *via* its C-terminus interacts with sulphated glycoconjugates on epithelial cells and helps the mycobacterium to maintain cytoadherence. Methylation has been suggested to provide protection to HBHA against proteases present in brancheoalveolar lavage [9, 10].

The *M.tb* cell envelope plays a major role in both virulence and persistence and acts as a physiological barrier between the bacillus and its environment. Much of the subsequent structural diversity in mycolic acids is generated by a family of AdoMet dependent MTases, which use the unsaturated meromycolic acid as a substrate to generate *cis* and *trans* cyclopropanes and other mycolates [11]. Mycolic acids constitute the major lipid component of the envelope and form the external mycomembrane [12]. So far, 10 different MTases which are involved in the specific modifications of mycolic acid are characterised including MmaA1 to MmaA4, CmaA1, CmaA2, PcaA, and UmaA. The activity of the methyltransferase (Rv3263) MamA is required for normal survival of hypoxia, indicating that it is likely an important mediator of adaptation to this physiologically relevant stressor [13].

In the present study, we analysed the *M.tb* proteome and discovered the presence of diverse range of MTases. Comparative proteomic analyses revealed homologues of several *M.tb* MTases in non-pathogenic and opportunistic mycobacteria. Several unique MTases observed in this study were analysed for their ancestry through horizontal gene transfer. Our results highlight the acquisition of novel MTases in pathogenic mycobacterium specifically in MTB complex in course of evolution, probably as a strategy for the survival of these pathogens in extreme host environment and for niche adaptation.

Experimental

Identification of MTases in *M.tb* proteome

The presence of all possible characterized and uncharacterized MTases in *M.tb* H_{37}Rv proteome was searched by using the keywords: 'methyltransferase', 'methylase', 'S-adenosylmethionine-dependent methyltransferases' and 'AdoMet-MTase' in the gene annotation using custom Python scripts. The proteomes of the sixteen mycobacterial species used in the study were downloaded from NCBI ftp (Table 1).

Cross proteomic comparison of MTases across Mycobacteria

M.tb H_{37}Rv MTases identified in the previous step were compared with the rest of the fifteen species as listed in Table 1 for the identification of homologues using Blast-2.2.30+ [14]. For determining homology, Blastp algorithm was used. Two proteins were said to be homologous if they shared a minimum identity of 50% and the pairwise sequence alignment covered 70% of the query length and the corresponding e-value $\leq 10^{-4}$.

Antigenic Index

The Antigenic Index of proteins was predicted by the VaxiJen v2 a server [15] which uses default cut-off value of 0.4. VaxiJen uses an alignment independent algorithm, based on auto cross covariance based approach to calculate the antigenicity of a given amino acid sequence with a prediction accuracy of 70–89% [16].

Globularity and physicochemical analyses

GlobPlot (http://globplot.embl.de/) analysis was performed for exploring the globularity and disorder in the Mtase protein sequences. It uses a running total based approach for calculating the predisposition of an amino acid to exist in an ordered or a disordered state.

ProtParam tool of ExPASy (http://web.expasy.ogr/protparam/) was used to determine the GRAVY (Grand Average of Hydropathicity), Instability indices, Aliphatic indices and *in-vivo* half-life of proteins. GRAVY is calculated using the hydropathic scales defined by Kyte and Doolittle[17]. Instability index is calculated based on a correspondence between protein stability and the composition of its dipeptides [18]. Aliphatic indices correspond to the relative volume occupied by the side chains of the aliphatic amino acids and provide a measure of thermostability of globular proteins[19] *In-vivo* half-life predicts the time required by a protein to become half of its amount after synthesis in the cell.

Phylogenetic analysis of distinct MTases

The protein sequence of the gene of interest was extracted from NCBI (http://www.ncbi.nlm.nih.gov/). A Blastp search [20] was performed against the non-redundant database excluding *M.tb* (Taxid id: 1773). The hits satisfying the threshold value of query coverage >70% and identity >35% were selected as homologues of the query gene for further analysis. FASTA sequences of all the homologues and the query gene were compiled and sequences were analysed for multiple sequence alignment using MUSCLE[21] and generation of phylogenetic tree using TREX server[22]. For phylogenetic analysis we used maximum likelihood algorithm and an analysis of 100 bootstraps was carried out to assess the reliability of tree so obtained. After careful examination of the tree, organism exhibiting same clade was taken into consideration. The members of the selected clade were predicted to be involved in horizontal gene transfer [23].

Results & Discussion

Classification of MTases in *M.tb*

Proteome of *M.tb* H_{37}Rv was explored for the presence of all characterised as well as uncharacterised (possible, probable as well as hypothetical) MTases as mentioned in material and methods section above and shown in Figure 1a and supplementary table 1. Based on our analyses we identified a total of 121 MTases in *M.tb* and these are depicted on a circular map with their Rv numbers (Figure 1b**)**. Interestingly, the MTases represent >3% of the proteome of *M.tb* and of these almost 70% are AdoMet dependent MTases. Of the 61 MTases which have been characterised so far, 17 are RNA MTases, 5 DNA MTases, 2 Protein MTases and 10 Mycolic acid MTases while the remaining are involved in different metabolic pathways. Functional categories of all the 121 MTases were retrieved from TubercuList database and were divided into eight categories based on functionality as represented in Figure 2. Highest

number of MTases (45) is implicated in intermediary metabolism and respiration category while 21 MTases are involved in lipid metabolism. These findings not only illustrate the variability of MTases in *M.tb* but also point to their likely functional importance.

Cross proteomic comparison of *M.tb* MTases

The MTases of *M.tb* were compared with various representative pathogenic, opportunistic and non-pathogenic species (total 15), based on homology search, at a sequence identity cut off of 50% and query coverage cut off of 70% [16, 24]. *M.bovis* was found to harbour the highest number of homologous MTases (118) while *M.leprae* had the lowest number of homologues (55) present in *M.tb* (Table 1, Figure 3). A reduction in the number of *M.tb* H_{37}Rv homologues was seen as we move from pathogenic to opportunistic to non-pathogenic bacteria. On an average strict pathogens have 100.5 MTases while opportunistic pathogens and non-pathogenic mycobacteria have 89.6 and 79.2 MTases, respectively.

Mycolic acid MTases- Forty percent of the dry weight of the *M.tb* cell is occupied by mycolic acids which contributes to its acid fastness and the comparative impermeability of the cell wall, specifically for drugs[25]. It is well known that in slow-growing pathogenic mycobacteria, various modifications of mycolic acids such as addition of cyclopropane rings, methyl branches, ketones, and methoxy groups takes place to generate a series of three major mycolic acids: alpha mycolates, methoxymycolates and ketomycolates [26, 27]. *M.tb* proteome harbours 10 different MTases involved in modification of mycolic acid moieties of the cell wall. All these MTases contain CMAS (Mycolic acid cyclopropane synthetase) domain. In a recent report, these enzymes were reported to be deleted by homologous recombination and a substantial attenuation of *M.tb* as well as feeble immune response of the host was observed[28]. We performed Blast analysis to look for the presence of these MTases in other Mycobacteria and observed that Rv0447 (UfaA) which has a cyclopropane-fatty-acyl-phospholipid synthase activity is restricted to only true pathogens and absent in all opportunistic and non-pathogenic mycobacterium species. This MTase was earlier shown to catalyze the biosynthesis of the tuberculostearic acid (10-methylstearic-acid, TSA), which constitutes a major lipid moiety of the mycobacterial cell wall and is a clinical marker of the disease [29]. UfaA is involved in the transfer of methyl group from SAM to the double bond of oleic acid in phosphatidylethanolamine or phosphatidylcholine to produce TSA [29]. These results point to the role of specific kinds of methylation modification in pathogenicity and modulation of immune response.

DNA MTases- DNA methylation is the only known mechanism of epigenetic regulation in prokaryotes. There are 5 different DNA MTases predicted in *M.tb* genome involved in maintaining gene expression pattern. *M.tb* has both cytosine and adenine methylation. Rv2756c which encodes type I restriction/modification system DNA methylase HsdM (Host specificity DNA Modification) activity is present only in pathogenic mycobacterium. Conversely, Rv3263 (MamA), an adenine MTase, was also restricted to pathogenic mycobacterium and exceptionally present in one of the non-pathogenic mycobacterium *Mycobacterium indicus pranii* (*MIP*) [30, 31]. In a recent report a strain specific adenine MTase was found to influence gene expression and its deletion reduces survival of *M.tb* in hypoxia thus suggesting that methylation mediates the regulatory pathways contributing to strain specific characteristics [13].

Antigenicity profiling of *M.tb* MTases

Antigenicity analysis of MTases, carried out using VaxiJen tool, revealed 86 probable antigens and 35 non-antigenic proteins. This showed that most of the MTase proteins (71%) of *M.tb* are antigenic in nature (Figure 4), reflecting the immunomodulatory nature of MTases. It was previously shown that Rv0470c, a SAM–dependent MTases is involved in proximal cyclopropanation of α mycolate and has been implicated in cording, persistence, and virulence[32]. The significance of methylation of mycolic acids in *M.tb* was also established by inhibition of Mycolic acid MTases that resulted in a viable but highly attenuated and hyper inflammatory response in mice model system[28]. Regions of high antigenicity within *M.tb* protein have been implicated with both humoral and T-cell responses not only *in vitro* but also in clinical samples[33-35].

Identification of distinct MTases in MTB Complex

At a sequence identity cut-off 35%, query coverage cut off 70% and e-value less than 10^{-4}, sequence of *M.tb* H_{37}Rv MTases protein was compared with proteomes of 15 different mycobacterium species, categorized as strict pathogens, opportunistic and non-pathogens [24] in order to find homologues. Results revealed that several MTases (Rv0326, Rv0329c, Rv1498c, Rv1509, Rv1515c, Rv1523, Rv1988, Rv2003c, Rv2492, Rv2954c, Rv2955c, Rv2990c, Rv3120, Rv3322 and Rv3729) were found to be exclusively present in *M.tb* complex and not present in any pathogenic, non-pathogenic or opportunistic pathogenic mycobacterium species (Supplementary table 2). Interestingly, majority of these MTases

(Rv1509, Rv1515c, Rv2492 Rv2990c, Rv2954c, Rv2955c, Rv2956 Rv2003c, Rv1498c Rv0329c) were predicted to be acquired through horizontal gene transfer [23]. These uniquely acquired MTases are highly polar in nature. Based on aliphatic indices cut off of >90, 7 of the distinct MTases showed high thermostability. All these proteins have an *in vivo* half life of >10hrs in *E.coli* (Supplementary Table 4). Remarkably, Rv1498c which is a predicted S-adenosylmethionine-dependent MTase contains ubiE/COQ5 MTase domain suggesting its role in quinone synthesis pathway. In an earlier study it was shown that Rv1498c expression was under the control of Mce3R, a TetR-type transcriptional repressor, whose regulon is involved in lipid metabolism [36]. Rv1506c was shown to be involved in the biosynthesis of acyltrehalose-containing glycolipids which play a critical role in the early intracellular fate of tubercle bacillus [37]. Rv3120 is present in RD5 region of *M.tb* genome [38] and has been reported as a potential candidate antigen for serodiagnosis of tuberculosis [39].

Among these proteins Rv2954c and Rv2955 are located in a 7kb genomic island called PGL locus, which is involved in the synthesis and modification of phenol glycolipids [40]. Rv2954c, Rv2955 and Rv2956, functionally characterized as O-methyltransferase, catalyze the O-methylation of the hydroxyl groups located at positions 3, 4 and 2, respectively of the terminal fucosyl residue of PGL-tb in a sequential process [41]. PGLs are produced exclusively by some slow growing mycobacterium (e.g. *M.tb, M.kansasii, M.marinum, M. ulcerans* and *M.leprae*) [42]. PGLs are complex lipids, believed to be noncovalently bound constituents of the outer leaflet of the unique mycobacterial outer membrane, and are known to be important effectors of virulence oxidative stress resistance and have a role in immunomodulation [42]. Rv2954c, Rv2955 and Rv2956 also figure in the list of virulence factor database of *M.tb*. Rv0329c lies in a region which is a *M.tb*-complex-specific genomic island. While both Rv2966c and Rv3720 are highly conserved in all the species of *M.tb*, Rv2966c is an RsmD-like [Ribosomal RNA small subunit methyltransferase D] MTase involved in 16sRNA methylation [43] whereas Rv3720, identified in the membrane fraction of *M.tb* H_{37}Rv, has a possible mycolic acid cyclopropane synthetase activity.

Overlap of MTases with Virulence factors and Essential genes

In an attempt to understand the likely functions of different MTases, a search of virulence factor database (VFDB), using BLAST, revealed the presence of 10 MTases of *M.tb* and these are: Rv0470c, Rv3720 and Rv503c (all three are cyclopropane mycolic acid synthase), Rv0642c (hydroxymycolate synthase MmaA4), Rv2952 (phthiotriol/phenolphthiotriol

dimycocerosates methyltransferase), Rv2954c, Rv2955, Rv2956 (all the three are involved in PGL biosynthesis), and Rv2959c (rhamnosyl O-methyltransferase). These results reveal that majority of MTases associated with virulence are involved in mycolic acid biosynthetic processes.

Nine of the MTases also lie in the category of essential genes based on bibliometric approach [44]. Correspondence of MTases with virulence factors and Essential genes is depicted in figure 5. Essential genes represent the minimal set of genes that an organism needs for maintaining critical metabolic processes and are identified based on a combination of high density transposon mediated mutagenesis and high throughput sequencing methods [45]. These essential genes are: Rv0208c (tRNA (guanine-N(7))-methyltransferase), Rv0224c, Rv0511 (uroporphyrin-iii c-methyltransferase), Rv0558 (ubiquinone/menaquinone biosynthesis methyltransferase), Rv1133 (5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase), Rv2165c (S-adenosyl-methyltransferase MraW), Rv2211c (glycine cleavage system aminomethyltransferase T), Rv2225 (3-methyl-2-oxobutanoate hydroxymethyltransferase), Rv2906c (tRNA (guanine-N(1)-)-methyltransferase), Rv3579c (possible trna/rrna methyltransferase). All these MTases were confirmed as essential genes by experimental validation through high density mutagenesis [46].

Comparison of MTases of *M.tb* H_{37}Rv with *M.tb* H_{37}Ra

M.tb H_{37}Ra is an avirulent strain of *M.tb* attenuated after serial passage of H_{37}Rv strain [47] and has been earlier shown by us to differ from each other in terms of genomic insertion, deletion, frame shift, phosphorylation sites etc. [48] Extending this earlier approach, we attempted to analyse protein sequence differences between the MTases of *M.tb* H_{37}Rv and H_{37}Ra. A total of three MTases were found to have an extension or deletion in the sequence length as well as point mutations at other locations (Table 3). Rv1703c codes for probable catechol-O-methyltransferase while its homologue in H_{37}Ra, MRA_1712, has an N-terminal deletion and V54M point mutation consequently resulting in an 18% increase in hydrophilicity of the protein. Rv0380c harbours possible RNA methyltransferase activity and its homologue MRA_0388 contains N-terminal extension and M38V point mutation thereby causing a 44% increase in hydrophilicity of the protein and also causing loss of a globular domain. Rv0503c is a cyclopropane mycolic acid synthase 2 and is involved in mycolic acid biosynthesis. The H_{37}Ra homologue of this protein has acquired an N-terminal extension causing increased instability index. Although these MTases are not well characterised functionally but they are

anticipated to play an important role in virulence of *M.tb* H_{37}Rv and could be a possible cause of attenuation of *M.tb* H_{37}Ra.

Comparison of MTases of *M.tb versus M.leprae*

M.leprae, a pathogenic mycobacterium that causes leprosy, has undergone genome reduction and psudogenisation [49]. A comparison of the MTases of *M.tb* with *M.leprae* revealed that of the 121 MTases present in *M.tb* only 55 proteins were found to have homologues while 66 genes did not show any homology. This was the highest number of nonhomologues compared to any other mycobacterium species. We identified several proteins namely Rv2847c, Rv3699, Rv3342, Rv3701c, Rv3204, Rv2847c, Rv3037c, Rv2751, Rv2689c, Rv2751, Rv1407, Rv1317c, Rv0881, Rv0764c, Rv0567, Rv0560c, Rv0294 and Rv3720 which are interestingly absent in *M.leprae* but are present in all other species of mycobacterium. Out of these, Rv3204 is a culture filtrate protein which induces high levels of IFNγ and IL-12 production [50] and contains DNA repair protein O6-alkylguanine-DNA MTase activity. O6-methylguanine DNA MTase helps in converting DNA lesion O6-methylguanine to guanine and thus helps in preventing replication and transcription errors. Rv2847c has a multifunctional uroporphyrin-III C-methyltransferase/precorrin-2 oxidase/ferrochelatase activity and belongs to the TP-methylase superfamily which catalyzes the transformation of uroporphyrinogen III into precorrin-2. The MTase Rv3037c is known to be involved in the biosynthesis of methylglucose lipopolysaccharides and is considered to be essential for *M.tb* growth [51]. Rv2689c contains a TRAM domain which has SAM-dependent tRNA (uracil-5-)-MTase activity while Rv0764c was earlier characterized as a lanosterol 14-alpha demethylase belonging to the cytochrome P450 superfamily [52]. Rv0560c possesses possible benzoquinone MTase activity and has been found be upregulated under rifampicin-induced transcriptome response in rifampicin-resistant *M.tb* [53] and is also over expressed under anaerobic conditions [54]. Protein Rv0560c contains three conserved motifs characteristic of SAM-dependent MTases, while Rv1407 has SUN domain which is involved in tRNA and rRNA cytosine-C5-methylation as shown in Conserved Domain database.

Analysis of horizontal gene transfer for distinct MTases of *M.tb*

A phylogenetic tree aims to model evolutionary substitutions over the time period and represents evolutionary relationship between sequences. In order to delineate the acquisition of distinct MTases of *M.tb*, we performed multiple sequence alignment of protein sequences sharing more than 35% identity and more than 70% query coverage. The phylogenetic tree

and sequences sharing same clade were taken into consideration for comparative analysis (Supplementary Table 3). The distance values show the number of substitutions as a proportion of the length of the alignment (excluding gaps). Rv1498c was found to be unique in *M.tb* and *M.canettii*. On the other hand it is more conserved in *Bacillus* species such as *Bacillus cereus* and *Cyanothece* sp. PCC 7424 (a unicellular, diazotrophic, oxygenic photosynthesizing cyanobacteria) [55]. It also shares common ancestry with *Clostridiales* bacterium which is a part of human microbiome, thus suggesting the exchange of gene during the process of adaptation within the human host. Within the genus mycobacterium, Rv2954c has its homologue only in *M. canettii* but also shares common ancestry with *Mycrocystis aeruginosa* which is involved in the formation of harmful algal blooms and is involved in production of neurotoxins [56]. Rv2954c was also found to be homologous with *Oscillatoriales* Cynobacterium, a marine cynobacterium.

Phylogenetic analysis of Rv0521 reveals its evolutionally correlation with certain marine bacteria like *Ponticoccus sp. UMTAT08, Labrenzia alexandrii* and *Thalasobaculam salexigen*. Rv0521 is also related to *Roseivivax halodurans* [57], a bacterium isolated from saline lakes and *Sediminimonas qiahouensis* isolated from salt mines [58], reflecting the possible role of this protein in adapting to high salt conditions. Rv1509 which is a unique protein present only in members of the MTB complex (*M.caprae*, *M canettii* 100% identity, and *M. rhodesiae* 67% identity), shares homology with *Salinarimonas rosea* (37% identity) which is a halotoralent bacterial strain isolated from a salt mine in China [59] and with *Caenispirillum salinarum*, a bacterial strain isolated from solar saltern of southern east part of India [60]. It also shares homology with *Neorhizobium galegae*, a recently discovered host specific N_2 fixing bacteria [61].

Rv1515c is also unique to members of MTB complex and *M.triplex*. On the other hand it also displays homolgy with *Ectothiorhodospira haloalkaliphila*, a photosynthetic purple sulfur bacteria from steppe soda lakes [62]. Rv3729 is present exclusively in members of MTB complex and two of the pathogenic mycobacterium *M.gastri* and *M.kansasii*, revealing its distinctiveness and possible role in adaptation to extremely adverse conditions present within the host. Rv2955c is also exclusively present in 5 members of MTB complex and is absent in other species of the genus mycobacterium. It shows common lineage with *Microcystys sp.T1-4* which is a fresh water cyanobacteria. Rv2990c exists in 3 members of *M.tb* complex and two of the pathogenic species *M.arupense* and *M.elephantis*. Interestingly, it lies in the

lineage of *Halomonas sp KM-1,* a halophilic bacteria isolated from Osaka, Japan and is reported as a potential bioplastic PHB (Polyhydroxybutyrate) producer. It is also closely related (58% identity) to *Thioalkali vibrio sp. ALJ6***,** a haloalkaliphilic chemolithoautotrophic sulfur-oxidizing bacterium.

Rv3120 is specifically found in three species of MTB complex. Interestingly, it has common roots with *Actinomadura madurae*, a potent agent of actinomycotic mycetomas and recently been reported as an opportunistic pathogen to cause pneumonia and bacteremia in immunosupressed AIDS patients. It also shows common ancestry with *Actinomadura rifamycini* and *Actinomadura oligospora* which are involved in the production of Rifamycin and a new polyether antibiotic, respectively. These findings suggest the probable lateral gene transfer of Rv3120 from genus Actinomadura to support *M.tb* in evolving as a hide and seek player with the human immune system, perhaps as an immune quorum sensor [63]. Rv1506c is present in three species of MTB complex and two of the pathogenic species *M. orygis* and *M. marinum.* It was found to have common parentage with *Pandoraea pnomenusa,* a soil bacterium was recently reported to be the cause of a recent sepsis related mortality after lung transplantation [64]. It also shows descendence with *Neorhizobium galegae*, which is involved in host specific nitrogen fixation [65].

Rv2492 shows phylogenetic correlation with 48% identity with *Parcubacteria bacterium* which has been identified in anoxic environments and harbours a reduced genome.[66] Rv0329c is present in *M.canettii,* a member of MTB complex and in one of the opportunistic pathogen *M.kansasii.* It also shares ancestry with *Jiangella alkaliphila,* a salt tolerant soil bacteria isolated from a cave in Korea and *Kribbella catacumbae,* a bacteria isolated from Roman catacombs. Rv329c also lies in the common lineage of salt tolerant soil bacteria *Nocardioides halotolerans*; and *Ruegeria pomeroyi* from coastal Gorgia sea water. Phylogenetic analysis of this gene suggests that gene acquisition have taken place from relatively distinct bacterial species favouring the survival of *M.tb* in extreme environment of lung macrophages. Rv2003 is exclusively present in two species of MTB complex, *M.bovis* and *M.canattii*. Interestingly, it has common ancestry with certain extremophiles such as *Acidothermus cellulolyticus* and *Aciduliprofundum boonei* found in oceanic hydrothermal vents and known to produce antibiotic against common pathogenic bacteria.

Regulation of expression at gene and protein level is essential for adaptation to stressful circumstances. Numerous features of Mycobacteria, such as ability to propagate inside phagolysosomes of host macrophages and their broad resistance to a range of antibiotics, have been attributed to the fact that these pathogens have in-built complex regulation machinery at the level of DNA, RNA and protein. These post transcriptional and post-translational modifications enhance the chemical and structural interactions and function of these biological macromolecules beyond the capabilities of their basic structure. In current years, a great deal of attention has been given to methyltransferases. *S*-Adenosylmethionine (AdoMet)-dependent methylation represents a major class of biological process and includes methyl group transfer from AdoMet yielding *S*-adenosylhomocysteine and a methylated molecular target. Many of these reactions are fundamental to an extent that deficiency of the gene product can rigorously, if not wholly, abrogate the normal working of an organism [67-69]. DNA methylation has foremost effect by activation or repression of gene expression in a locus specific mode. *M.tb* reportedly lacks methylated Dam and Dcm recognition sequences pointing to the absence of these MTases, however it contains considerable levels of 6-methyladenine and 5-methylcytosine suggesting the presence of other DNA MTases [70]. Presence of methyl cytosine in virulent strain of *M.tb* H_{37}Rv and absence in avirulent strain *M.tb* H_{37}Ra is suggestive of the role of DNA methylation in virulence [71].

MTases are important cellular proteins involved in modifications of a vast variety of substrates. In this study we have mined the *M.tb* proteome and identified 121 MTases in *M.tb* proteome covering 3% of whole proteome. Interestingly the number of MTases decreases as we move from pathogenic to opportunistic to non-pathogenic mycobacteria. SAM-dependent MTase genes are a multigene family and 70% of MTases of *M.tb* are SAM dependent, however, only a few have been characterized at the functional level. A pointer to the role of methyltransferases in niche adaptation was the restricted presence of DNA MTase Rv3263 (MamA) in pathogenic mycobacterium and *MIP*. MamA methylates the adenine base in the genomic DNA and regulates the expression of a number of proteins involved in hypoxic stress response including Rv0142, CorA and WhiB [13]. Another DNA Mtase HsdM which harbours Type I restriction-modification system, is also present only in pathogenic mycobacterium, reflecting the role of differential gene expression in pathogens as a tool to cope up with the host defence mechanism.

Mycobacterium avium subsp. paratuberculosis K10 (MAP) shares 91 MTases with *M.tb* . A study on the evolution of MAP, has identified 275 genes acquired through lateral gene transfer. The analysis showed that 53 of the 275 genes were acquired after the divergence of MAP from *M. avium subsp. avium*, whereas the remaining 222 genes were possibly acquired by a common ancestor of MAP and *M.avium subsp. avium* after its divergence from the ancestor of MTB Complex [72]. It differs from other members of the *M.avium* complex in having 14-18 copies of IS900 inserted into conserved loci in its genome. This region also includes 6-O-methylguanine methyltransferase at locus 9[73].

As evident from phylogenetic analysis, many of the MTases acquired by MTB complex have their common ancestry with halophilic/marine/salt tolerant and acidophilic organisms. Bacteria are repeatedly exposed to sharply changing solute concentrations in their surroundings. In order to survive under osmotic stress, bacteria have developed numerous approaches, for example adjusting their intracellular osmolality or enhancing their cell wall stability thus, enabling them to grow in a wide array of solute concentrations. High salinity or osmotic stress poses an extreme surroundings which relatively, a small number of organisms have been able to acclimatize. In a recent report, Mycobacteria could grow even at 0.800 water-activity indicating that they are, with the sole exception of halophiles, more xerotolerant than other bacteria (or any Archaea) [74]. Water activity is a measure of the amount of water within a substrate an organism can use to support its metabolic functions. In case of halophiles, salt solutes chemically bind to water and thus bring down the water activity[75]. In addition, when mycobacterium cells were grown over a range of salt concentration, they were able to resist the osmotic stress and very less change was observed in the fatty-acid composition, signifying a high level of toughness despite the stress load [74]. This has a pathophysiological significance in terms of adaptation of *M.tb* to changes in environmental osmolarity because of its transitions between air-borne droplet nuclei, mucosal epithelia, alveolar macrophages, necrotic cells, and caseous granulomas [76]. Osmotic alterations modify turgor pressure, which can damage protein folding and its metabolic function [77]. Bacteria usually counter such fluctuations by the compensatory addition or exclusion of compatible solutes which restore osmotic balance to the cell [78]. In a recent study, it was discovered that osmotic stress stimulates a signalling pathway in *M.tb* regulated by the receptor STPK PknD [79].

Conclusion

Results presented above provide substantial insight into the role of MTases in the biology and evolution of mycobacterial pathogens. Further our findings delineate the role of MTases, acquired by horizontal gene transfer, in physiological adaptations relevant to pathogenesis. It will be interesting to further examine the role of these distinct MTases in adaptation of the *M.tb* in *in-vivo* systems. To our knowledge it is the first report establishing an evolutionary relationship of *M.tb* with halotolerant organisms and also the role of MTases of *M.tb* in withstanding the host osmotic stress.

Acknowledgements

This work was supported by a Center of Excellence Grant (BT/PR12817/COE/34/23/2015) from Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India. SEH is a JC Bose National Fellow, Department of Science and Technology, Government of India and Robert Koch Fellow, Robert Koch Institute, Berlin, Germany. SG is a recipient of the Indian Institute of Technology, Delhi Post Doctoral Fellowship. PG and S are recipients of Junior Research Fellowship from Department of Biotechnology (DBT); KD is a Junior Research Fellow of UGC; PSK was a Summer Intern supported by the Indian Academy of Sciences.

Abbreviations: MTases- Methyltransferases, *M.tb*- *Mycobacterium tuberculosis*, TB-Tuberculosis

Competing Interest: None declared

Authors Contribution:

SG, PG and SEH designed the study. SG, PG, PSK, SG2, KD and SA analyzed and interpreted the data. SG, PG, AG, NZE and SEH drafted the manuscript. All authors read and approved the final manuscript.

Legends to Figures

Figure 1: a) Classification of MTases found in *M.tb* H_{37}Rv. All the MTases were classified based on the different substrates to which they transfer a methyl group from the methyl donor. b) Genomic map showing the coordinates of the 121 MTases found in *M.tb* H_{37}Rv.

The map was created using the CGView tool. Methyltransferases constitute 3% of *M.tb* proteome.

Figure 2: Functional categorization of *M.tb* H_{37}Rv MTases. The functional class of each MTase was taken from the tuberculist database (http://tuberculist.epfl.ch/).

Figure 3: Comparison of *M.tb* H_{37}Rv MTases with other Mycobacteria. The homology between *M.tb* MTases and the homologues in the rest of the 15 species was checked by Blastp with the following cutoffs: 50% identity, 70% query coverage and 10^{-4} e-value.

Figure 4: Antigenicity profile of MTases found in *M.tb* H_{37}Rv. The antigenicity index was predicted using the VaxiJen v2 tool (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and the threshold was kept at the default value of 0.4.

Figure 5: Comparison of *M.tb* H_{37}Rv MTases with essential genes and virulence factors found in *M.tb* H_{37}Rv. 10 MTases also act as a virulence factor and 9 MTases have also been designated as essential genes but however none of the MTases belonged to the above mentioned two classes simultaneously.

Table 1: List of species used in the analysis

Sr. No.	Species	Abbreviations	Classification based on virulence	Proteome size
1	*Mycobacterium tuberculosis $H_{37}Rv$*	*M.tb*	True pathogen	4034
2	*Mycobacterium bovis subsp. bovis AF2122/97*	mbo	True pathogen	5149
3	*Mycobacterium leprae TN*	mlep	True pathogen	1605
4	*Mycobacterium ulcerans Agy99*	mulc	True pathogen	3918
5	*Mycobacterium marinum M*	mmar	True pathogen	5423
6	*Mycobacterium canettii CIPT 140010059*	mcan	True pathogen	3861
7	*Mycobacterium africanum GM041182*	maf	True pathogen	3970
8	*Mycobacterium avium 104*	mav104	Opportunistic pathogen	5120
9	*Mycobacterium intracellulare MOTT-64*	mint	Opportunistic pathogen	4160
10	*Mycobacterium avium subsp. paratuberculosis K-10*	mavk10	Opportunistic pathogen	4326
11	*Mycobacterium indicus pranii MTCC 9506*	mip	Non-pathogen	5254
12	*Mycobacterium smegmatis str. MC2 155*	msmeg	Non-pathogen	6689
13	*Mycobacterium gilvum PYR-GCK*	mgil	Non-pathogen	5241
14	*Mycobacterium vanbaalenii PYR-1*	mvac	Non-pathogen	5979
15	*Mycobacterium vaccae*	mvan	Non-pathogen	5812
16	*Mycobacterium tuberculosis $H_{37}Ra$*	*M.tb $H_{37}Ra$*	Non-pathogen	4036

Table 2: Methyltransferases homologous to *Mycobacterium tuberculosis* $H_{37}Rv$

Sr. No.	Species	No. of Homologous Proteins	No. of Non-homologous Proteins
1	*Mycobacterium tuberculosis $H_{37}Rv$*	121	0
2	*Mycobacterium bovis subsp. bovis AF2122/97*	118	3
3	*Mycobacterium leprae TN*	55	66
4	*Mycobacterium ulcerans Agy99*	88	33
5	*Mycobacterium marinum M*	100	21
6	*Mycobacterium canettii CIPT 140010059*	121	0
7	*Mycobacterium africanum GM041182*	121	0
8	*Mycobacterium avium 104*	89	32
9	*Mycobacterium intracellulare MOTT-64*	89	32
10	*Mycobacterium avium subsp. paratuberculosis K-10*	91	30
11	*Mycobacterium indicus pranii MTCC 9506*	91	30
12	*Mycobacterium smegmatis str. MC2 155*	79	42
13	*Mycobacterium gilvum PYR-GCK*	74	47
14	*Mycobacterium vanbaalenii PYR-1*	73	48
15	*Mycobacterium vaccae*	79	42

Table 3: Comparison of Methyltransferases of *M.tb* $H_{37}Rv$ with $H_{37}Ra$

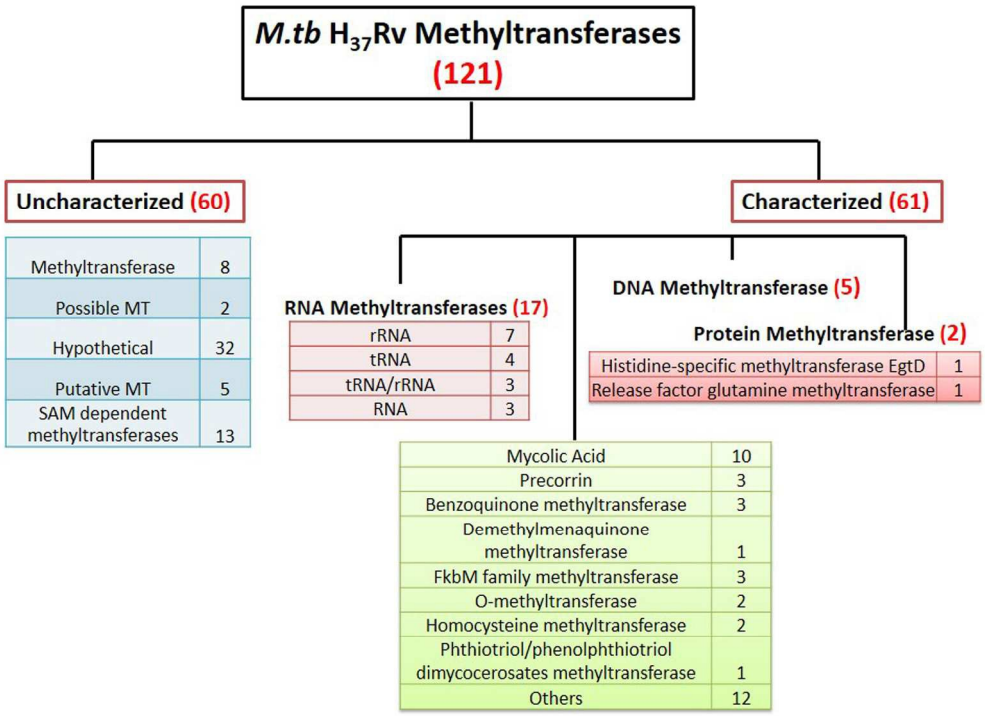
H37Rv	Sequence length	H37Ra	Sequence length	Kind of mutation	Function	Conserved Domain
Rv0380c	183	MRA_0388	220	N-terminal Extension & M38V point mutation	Possible RNA methyltransferase	SPoU_methylase superfamily
Rv1703c	249	MRA_1712	196	N-terminal Deletion & V54M point mutation	Putative methyl transferase	Adomet Mtases superfamily
Rv0503c	302	NA	322	N-terminal extension	Mycolic acid methyltransferase	Adomet Mtases superfamily

References

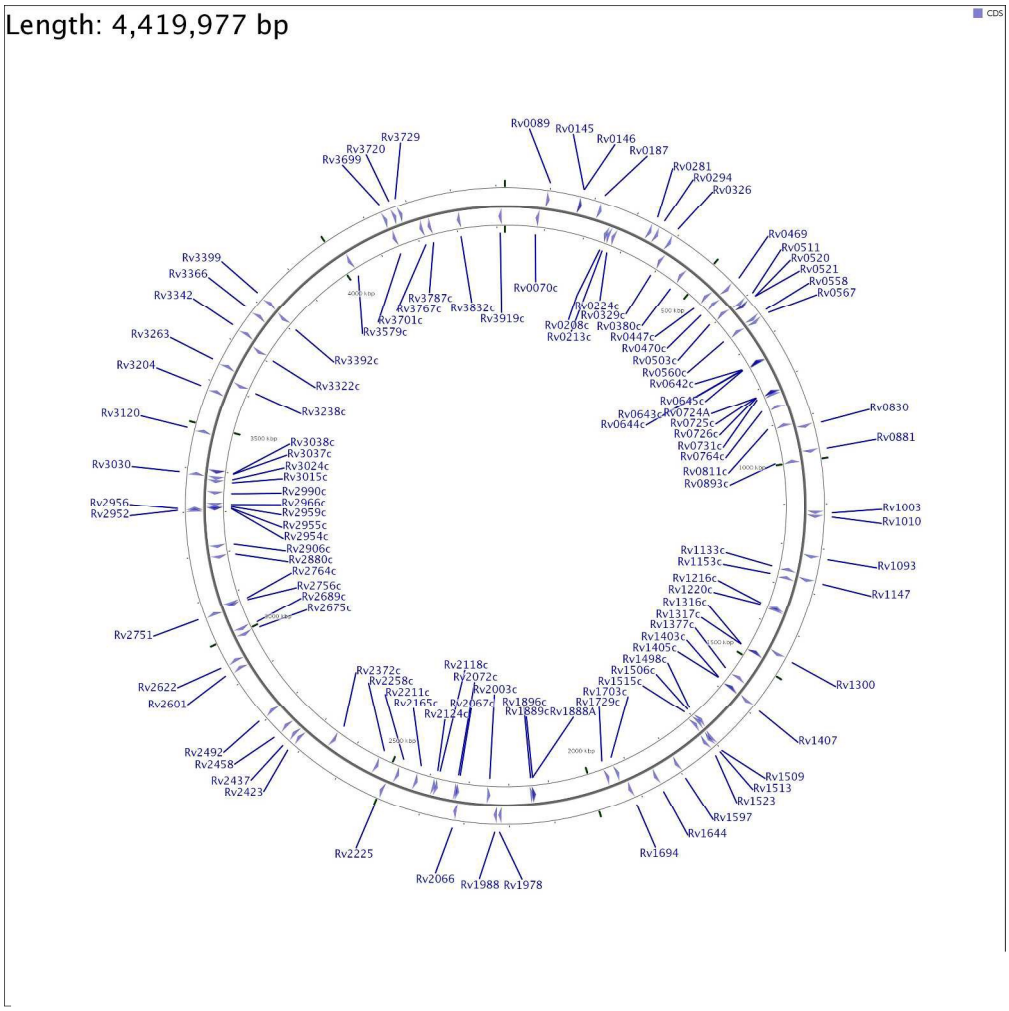
1. A. Zumla, A. George, V. Sharma, R. H. N. Herbert, B. Masham Ilton, A. Oxley and M. Oliver, *Lancet Glob Health*, 2015, **3**, E10-E12.
2. M. Gengenbacher and S. H. E. Kaufmann, *Fems Microbiol Rev*, 2012, **36**, 514-532.
3. A. Kumar, A. Lewin, P. S. Rani, I. A. Qureshi, S. Devi, M. Majid, E. Kamal, S. Marek, S. E. Hasnain and N. Ahmed, *Cytokine*, 2013, **64**, 258-264.
4. S. K. Johansen, C. E. Maus, B. B. Plikaytis and S. Douthwaite, *Mol Cell*, 2006, **23**, 173-182.
5. S. Okamoto, A. Tamaru, C. Nakajima, K. Nishimura, Y. Tanaka, S. Tokuyama, Y. Suzuki and K. Ochi, *Mol Microbiol*, 2007, **63**, 1096-1106.
6. A. Kumar, S. Kumar and B. Taneja, *Acta Crystallogr D*, 2014, **70**, 821-832.
7. G. Delogu, A. Bua, C. Pusceddu, M. Parra, G. Fadda, M. J. Brennan and S. Zanetti, *Fems Microbiol Lett*, 2004, **239**, 33-39.
8. S. M. Zhao, Y. Zhao, F. F. Mao, C. Q. Zhang, B. Bai, H. Zhang, C. H. Shi and Z. K. Xu, *Plos One*, 2012, **7**.
9. K. Pethe, S. Alonso, F. Biet, G. Delogu, M. J. Brennan, C. Locht and F. D. Menozzi, *Nature*, 2001, **412**, 190-194.
10. K. Pethe, P. Bifani, H. Drobecq, C. Sergheraert, A. S. Debrie, C. Locht and F. D. Menozzi, *P Natl Acad Sci USA*, 2002, **99**, 10759-10764.
11. C. E. Barry, R. E. Lee, K. Mdluli, A. E. Sampson, B. G. Schroeder, R. A. Slayden and Y. Yuan, *Prog Lipid Res*, 1998, **37**, 143-179.
12. B. Zuber, M. Chami, C. Houssin, J. Dubochet, G. Griffiths and M. Daffe, *J Bacteriol*, 2008, **190**, 5672-5680.
13. S. S. Shell, E. G. Prestwich, S. H. Baek, R. R. Shah, C. M. Sassetti, P. C. Dedon and S. M. Fortune, *Plos Pathog*, 2013, **9**.
14. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden, *BMC Bioinformatics*, 2009, **10**, 421.
15. I. A. Doytchinova and D. R. Flower, *BMC Bioinformatics*, 2007, **8**, 4.
16. Y. Singh, S. Kohli, D. T. Sowpati, S. A. Rahman, A. K. Tyagi and S. E. Hasnain, *Int J Med Microbiol*, 2014, **304**, 742-748.
17. J. Kyte and R. F. Doolittle, *J Mol Biol*, 1982, **157**, 105-132.
18. K. Guruprasad, B. V. B. Reddy and M. W. Pandit, *Protein Eng*, 1990, **4**, 155-161.
19. A. Ikai, *J Biochem-Tokyo*, 1980, **88**, 1895-1898.
20. S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *J Mol Biol*, 1990, **215**, 403-410.
21. R. C. Edgar, *Nucleic Acids Res*, 2004, **32**, 1792-1797.
22. B. Alix, D. A. Boubacar and M. Vladimir, *Nucleic Acids Res*, 2012, **40**, W573-W579.
23. F. Veyrier, D. Pletzer, C. Turenne and M. A. Behr, *Bmc Evol Biol*, 2009, **9**.
24. S. A. Rahman, Y. Singh, S. Kohli, J. Ahmad, N. Z. Ehtesham, A. K. Tyagi and S. E. Hasnain, *Mbio*, 2014, **5**.
25. D. Barkan, Z. Liu, J. C. Sacchettini and M. S. Glickman, *Chem Biol*, 2009, **16**, 499-509.
26. M. Watanabe, Y. Aoyagi, M. Ridell and D. E. Minnikin, *Microbiol-Sgm*, 2001, **147**, 1825-1837.
27. M. Arish, A. Husein, M. Kashif, P. Sandhu, S. E. Hasnain, Y. Akhter and A. Rub, *Biochimie*, 2015, **113C**, 111-124.
28. D. Barkan, D. Hedhli, H. G. Yan, K. Huygen and M. S. Glickman, *Infect Immun*, 2012, **80**, 1958-1968.
29. L. S. Meena and P. E. Kolattukudy, *Biotechnol Appl Bioc*, 2013, **60**, 412-416.
30. V. Saini, S. Raghuvanshi, J. P. Khurana, N. Ahmed, S. E. Hasnain, A. K. Tyagi and A. K. Tyagi, *Nucleic Acids Res*, 2012, **40**, 10832-10850.
31. V. Saini, S. Raghuvanshi, G. P. Talwar, N. Ahmed, J. P. Khurana, S. E. Hasnain, A. K. Tyagi and A. K. Tyagi, *Plos One*, 2009, **4**.
32. M. S. Glickman, J. S. Cox and W. R. Jacobs, *Mol Cell*, 2000, **5**, 717-727.

33. S. Tundup, N. Pathak, M. Ramanadham, S. Mukhopadhyay, K. J. R. Murthy, N. Z. Ehtesham and S. E. Hasnain, *Plos One*, 2008, **3**.
34. S. Banerjee, A. Nandyala, R. Podili, V. M. Katoch, K. J. R. Murthy and S. E. Hasnain, *P Natl Acad Sci USA*, 2004, **101**, 12652-12657.
35. P. Chakhaiyar, Y. Nagalakshmi, B. Aruna, K. J. R. Murthy, V. M. Katoch and S. E. Hasnain, *J Infect Dis*, 2004, **190**, 1237-1244.
36. M. D. Santangelo, L. Klepp, J. Nunez-Garcia, F. C. Blanco, M. Soria, M. D. Garcia-Pelayo, M. V. Bianco, A. A. Cataldi, P. Golby, M. Jackson, S. V. Gordon and F. Bigi, *Microbiol-Sgm*, 2009, **155**, 2245-2255.
37. P. Brodin, Y. Poquet, F. Levillain, I. Peguillet, G. Larrouy-Maumus, M. Gilleron, F. Ewann, T. Christophe, D. Fenistein, J. C. Jang, M. S. Jang, S. J. Park, J. Rauzier, J. P. Carralot, R. Shrimpton, A. Genovesio, J. A. Gonzalo-Asensio, G. Puzo, C. Martin, R. Brosch, G. R. Stewart, B. Gicquel and O. Neyrolles, *Plos Pathog*, 2010, **6**.
38. R. C. Huard, L. C. D. Lazzarini, W. R. Butler, D. van Soolingen and J. L. Ho, *J Clin Microbiol*, 2003, **41**, 1637-1650.
39. M. M. Zhang, J. W. Zhao, Z. Q. Sun, J. Liu, X. K. Guo, W. D. Liu and S. L. Zhang, *Clin Dev Immunol*, 2012, DOI: 10.1155/2012/738043.
40. J. Becq, M. C. Gutierrez, V. Rosas-Magallanes, J. Rauzier, B. Gicquel, O. Neyrolles and P. Deschavanne, *Mol Biol Evol*, 2007, **24**, 1861-1871.
41. R. Simeone, G. Huet, P. Constant, W. Malaga, A. Lemassu, F. Laval, M. Daffe, C. Guilhot and C. Chalut, *Plos One*, 2013, **8**.
42. K. C. Onwueme, C. J. Vos, J. Zurita, J. A. Ferreras and L. E. N. Quadri, *Prog Lipid Res*, 2005, **44**, 259-302.
43. A. Kumar, K. Saigal, K. Malhotra, K. M. Sinha and B. Taneja, *J Biol Chem*, 2011, **286**, 19652-19661.
44. G. Y. Xu, B. Liu, F. Wang, C. G. Wei, Y. Zhang, J. Y. Sheng, G. Q. Wang and F. Li, *Bmc Infect Dis*, 2013, **13**.
45. H. Luo, Y. Lin, F. Gao, C. T. Zhang and R. Zhang, *Nucleic Acids Res*, 2014, **42**, D574-D580.
46. C. M. Sassetti, D. H. Boyd and E. J. Rubin, *Mol Microbiol*, 2003, **48**, 77-84.
47. W. Steenken, Jr. and L. U. Gardner, *American review of tuberculosis*, 1946, **54**, 62-66.
48. S. Kohli, Y. Singh, K. Sharma, A. Mittal, N. Z. Ehtesham and S. E. Hasnain, *Nucleic Acids Res*, 2012, **40**, 7113-7122.
49. N. Ahmed, U. Dobrindt, J. Hacker and S. E. Hasnain, *Nat Rev Microbiol*, 2008, **6**, 387-394.
50. J. H. Lim, H. J. Kim, K. S. Lee, E. K. Jo, C. H. Song, S. B. Jung, S. Y. Kim, J. S. Lee, T. H. Paik and J. K. Park, *Fems Microbiol Lett*, 2004, **232**, 51-59.
51. V. Mendes, A. Maranha, S. Alarico and N. Empadinhas, *Natural Product Reports*, 2012, **29**, 834-844.
52. A. Bellamine, A. T. Mangla, W. D. Nes and M. R. Waterman, *P Natl Acad Sci USA*, 1999, **96**, 8937-8942.
53. G. J. de Knegt, O. Bruning, M. T. ten Kate, M. de Jong, A. van Belkum, H. P. Endtz, T. M. Breit, I. A. J. M. Bakker-Woudenberg and J. E. M. de Steenwinkel, *Tuberculosis*, 2013, **93**, 96-101.
54. J. Starck, G. Kallenius, B. I. Marklund, D. I. Andersson and T. Akerlund, *Microbiol-Sgm*, 2004, **150**, 3821-3829.
55. T. J. Mueller, B. M. Berla, H. B. Pakrasi and C. D. Maranas, *Bmc Syst Biol*, 2013, **7**.
56. R. Alexova, T. C. Dang, M. Fujii, M. J. Raftery, T. D. Waite, B. C. Ferrari and B. A. Neilan, *Environ Microbiol*, 2015, DOI: 10.1111/1462-2920.12958.
57. T. Suzuki, Y. Muroga, M. Takahama and Y. Nishimura, *Int J Syst Bacteriol*, 1999, **49**, 629-634.
58. Y. X. Wang, Z. G. Wang, J. H. Liu, Y. G. Chen, X. X. Zhang, M. L. Wen, L. H. Xu, Q. Peng and X. L. Cui, *Int J Syst Evol Micr*, 2009, **59**, 1561-1567.
59. J. H. Liu, Y. X. Wang, X. X. Zhang, Z. G. Wang, Y. G. Chen, M. L. Wen, L. H. Xu, Q. Peng and K. L. Cui, *Int J Syst Evol Micr*, 2010, **60**, 55-60.

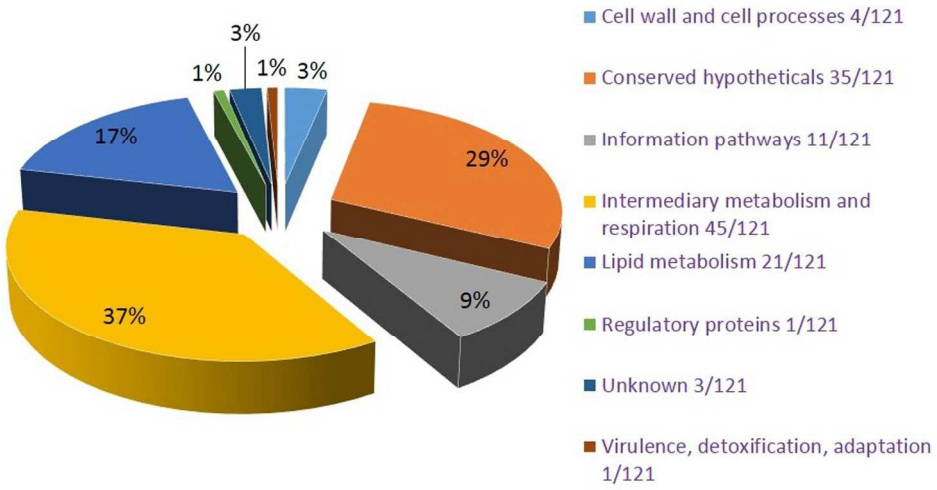
60. C. Ritika, K. Suresh and P. A. Kumar, *Int J Syst Evol Micr*, 2012, **62**, 1698-1702.
61. S. A. Mousavi, J. Osterman, N. Wahlberg, X. Nesme, C. Lavire, L. Vial, L. Paulin, P. de Lajudiee and K. Lindstrom, *Syst Appl Microbiol*, 2014, **37**, 208-215.
62. V. M. Gorlenko, I. A. Bryantseva, S. Rabold, T. P. Tourova, D. Rubtsova, E. Smirnova, V. Thiel and J. F. Imhoff, *Int J Syst Evol Microbiol*, 2009, **59**, 658-664.
63. S. Tundup, K. Mohareer and S. E. Hasnain, *Febs Open Bio*, 2014, **4**, 822-828.
64. M. E. Stryjewski, J. J. LiPuma, R. H. Messier, L. B. Reller and B. D. Alexander, *J Clin Microbiol*, 2003, **41**, 2255-2257.
65. J. Osterman, S. A. Mousavi, P. Koskinen, L. Paulin and K. Lindstrom, *Bmc Genomics*, 2015, **16**.
66. W. C. Nelson and J. C. Stegen, *Front Microbiol*, 2015, **6**.
67. X. D. Mao, B. Schwer and S. Shuman, *Mol Cell Biol*, 1996, **16**, 475-480.
68. D. Lafontaine, J. Delcour, A. L. Glasser, J. Desgres and J. Vandenhaute, *J Mol Biol*, 1994, **241**, 492-497.
69. J. Anderson, L. Phan and A. G. Hinnebusch, *P Natl Acad Sci USA*, 2000, **97**, 5173-5178.
70. K. C. Hemavathy and V. Nagaraja, *Fems Immunol Med Mic*, 1995, **11**, 291-296.
71. R. Srivastava, K. P. Gopinathan and T. Ramakrishnan, *J Bacteriol*, 1981, **148**, 716-719.
72. P. R. Marri, J. P. Bannantine, M. L. Paustian and G. B. Golding, *Can J Microbiol*, 2006, **52**, 560-569.
73. T. J. Bull, J. Hermon-Taylor, I. Pavlik, F. El-Zaatari and M. Tizard, *Microbiol-Uk*, 2000, **146**, 3285-3285.
74. R. Santos, C. C. de Carvalho, A. Stevenson, I. R. Grant and J. E. Hallsworth, *Environ Microbiol Rep*, 2015, DOI: 10.1111/1758-2229.12306.
75. W. D. Grant, *Philos T Roy Soc B*, 2004, **359**, 1249-1266.
76. C. T. Price, A. Bukka, M. Cynamon and J. E. Graham, *J Bacteriol*, 2008, **190**, 3955-3961.
77. R. D. Sleator and C. Hill, *Fems Microbiol Rev*, 2002, **26**, 49-71.
78. L. N. Csonka, *Microbiol Rev*, 1989, **53**, 121-147.
79. S. K. Hatzios, C. E. Baer, T. R. Rustad, M. S. Siegrist, J. M. Pang, C. Ortega, T. Alber, C. Grundner, D. R. Sherman and C. R. Bertozzi, *Proc Natl Acad Sci U S A*, 2013, **110**, E5069-5077.



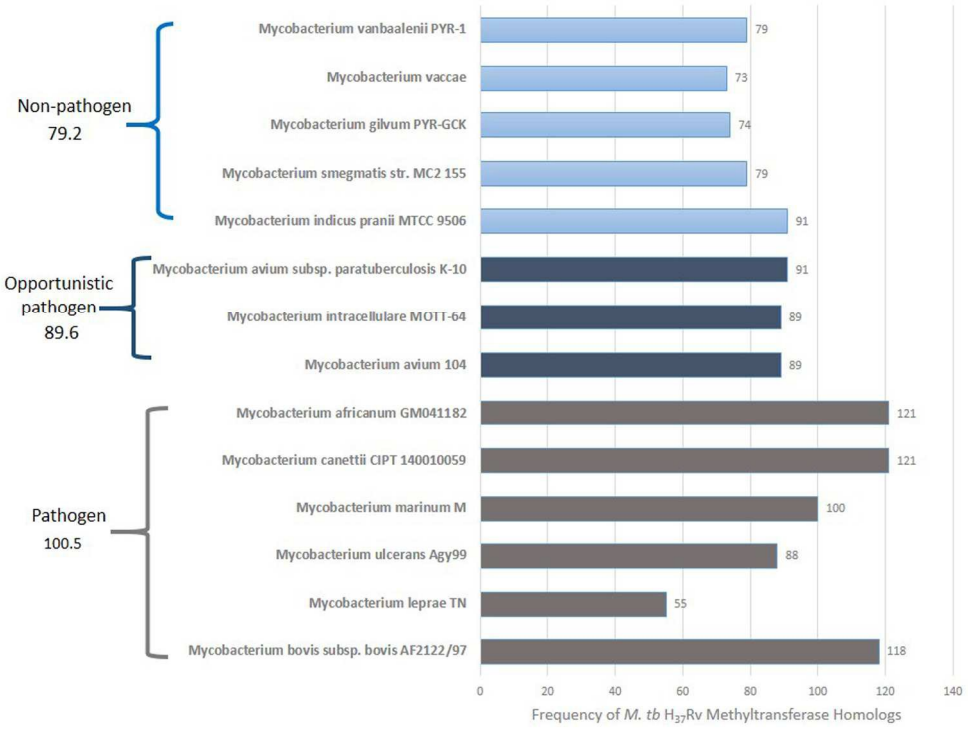
190x142mm (600 x 600 DPI)



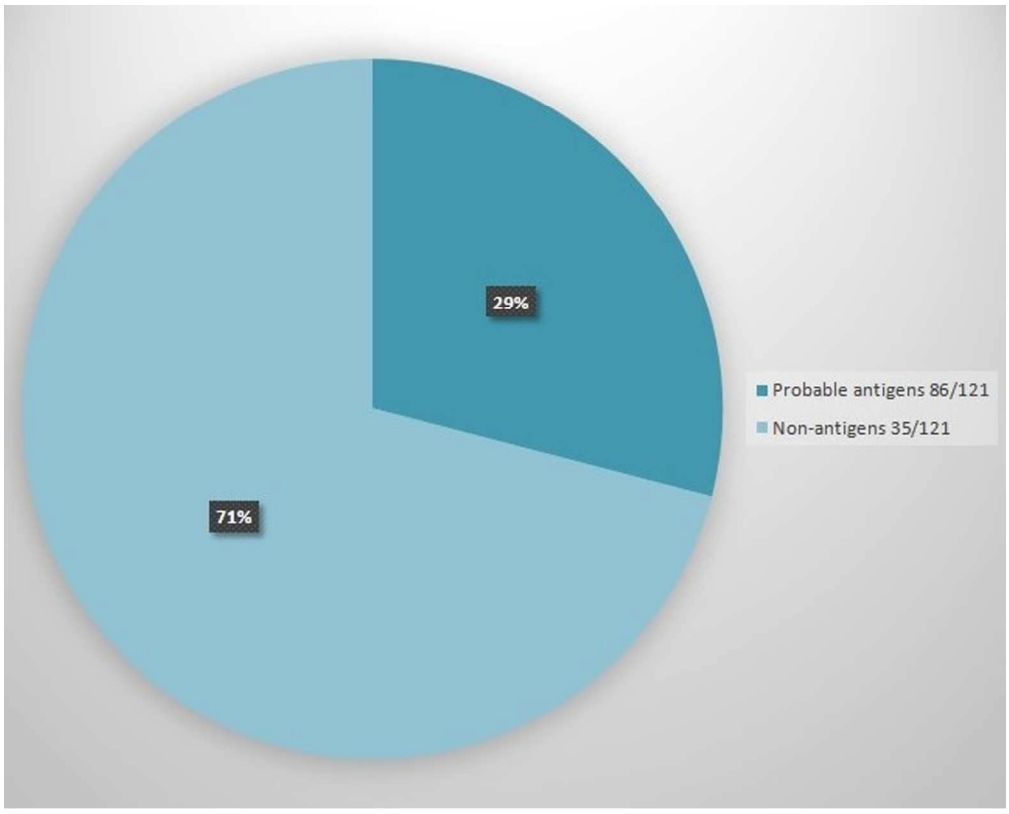
529x529mm (300 x 300 DPI)



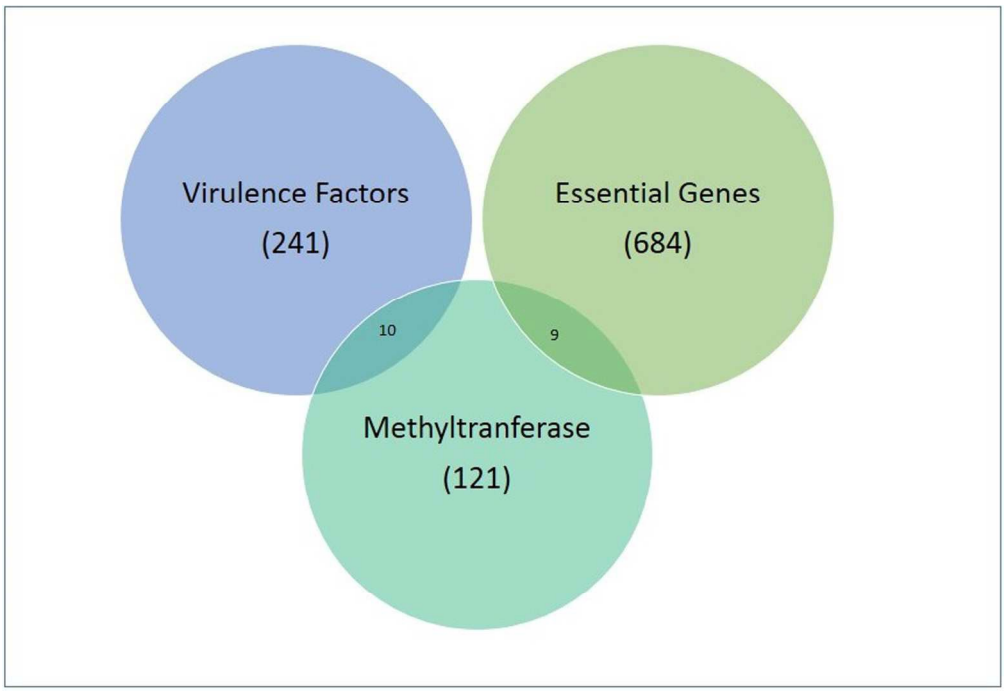
99x58mm (600 x 600 DPI)



190x142mm (600 x 600 DPI)



90x72mm (600 x 600 DPI)



103x71mm (600 x 600 DPI)