The *Mycobacterium tuberculosis* sRNA F6 regulates expression of *groEL/S*

Joanna Houghton[1], Angela Rodgers[2] Graham Rose[3], Kristine B. Arnvig[4]*

[1] Faculty of Infectious Tropical Diseases, London School of Hygiene and Tropical medicine, London WC1E 7HT, UK

[2] Mycobacterial Metabolism and Antibiotic Research and Host-Pathogen Interactions in Tuberculosis Laboratories, The Francis Crick Institute, London NW1 1AT, UK

[3] North Thames Genomic Laboratory Hub, Great Ormond Street Hospital for Children, London WC1N 3BH, UK

[4] Structural and Molecular Biology, University College London, London WC1E 6BT, UK

* To whom correspondence should be addressed. Tel: +44 2076792242 Email: k.arnvig@ucl.ac.uk

ABSTRACT

Almost 140 years after the identification of *Mycobacterium tuberculosis* as the etiological agent of tuberculosis, important aspects of its biology remain poorly described. Little is known about the role of post-transcriptional control of gene expression and RNA biology, including the role of most of the small RNAs (sRNAs) identified to date. We have carried out a detailed investigation of the *M. tuberculosis* sRNA, F6, and show it to be dependent on SigF for expression and significantly induced during *in vitro* starvation and in a mouse model of infection. However, we found no evidence of attenuation of a ΔF6 strain within the first 20 weeks of infection. A further exploration of F6 using *in vitro* models of infection suggests a role for F6 as a highly specific regulator of the heat shock repressor, HrcA. Our results point towards a role for F6 during periods of low metabolic activity similar to cold shock and associated with nutrient starvation such as that found in human granulomas in later stages of infection.

INTRODUCTION

In spite of claiming more than 3000 lives on a daily basis, the impact of *Mycobacterium tuberculosis* as a major human pathogen is currently being overshadowed by the ongoing COVID-19 pandemic. Notoriously difficult to work with and very different from both Gram-positive and Gram-negative model organisms, we still know preciously little about many aspects of its basic biology including RNA biology and post-transcriptional control of gene expression. However, the study of regulatory RNA in *M. tuberculosis* is slowly gaining momentum, aided by next-generation sequencing (NGS) applications, which have provided significant insights

into the abundance and dynamics of non-coding RNA over a range of growth conditions, the location of transcription start sites on a global scale and translated *versus* untranslated transcripts e.g. (1-5). Still, much remains to be uncovered about the post-transcriptional control exerted by regulatory RNA in *M. tuberculosis* in particular what role these molecules play in the pathogenesis, latency and persistence of *M. tuberculosis*. A multitude of *M. tuberculosis* small regulatory RNAs (sRNAs) have been identified in the last decade, but only few have been investigated and even less, potentially linked to latency and persistence e.g. (1,3,4,6-13).

F6 (ncRv10243) was one of the first *M. tuberculosis* sRNAs to be identified, not by RNA-seq but by cDNA cloning (6). F6 is conserved in a range of pathogenic mycobacteria and in *Mycobacterium smegmatis* with the 5' end showing the highest degree of conservation (6,7). The location of F6, i.e. between the convergent *fadA2* (encoding an acetyl-CoA transferase) and *fadE5* (encoding an acyl-CoA hydrogenase), is also highly conserved suggesting a regulatory role for F6 in lipid metabolism, which is critical for intracellular survival (14). Initial analysis, combining northern blotting with 5' and 3' RACE, revealed that F6 is expressed as a 102-nucleotide transcript, which is 3' processed to the dominant transcript of 58 nucleotides, comprising the more conserved part of the transcript (6). F6 is upregulated in stationary phase, during oxidative stress, and in low pH, while overexpression from a multicopy-number plasmid leads to a slow-growth phenotype on solid media (6). The F6 promoter contains a typical SigF promoter motif and exhibits the highest SigF occupancy according to ChIP-chip analysis, while overexpression of SigF leads to slow growth akin to overexpression of F6 (6,15). SigF is a non-essential sigma factor, conserved in most mycobacteria; its expression is induced by several stresses including anaerobiosis, nutrient starvation, oxidative stress and cold shock, while heat shock downregulates its expression, and deletion leads to attenuation in mice (16-18). (19).

In the current study, we have investigated *M. tuberculosis* F6 expression under different growth conditions and found that F6 is significantly upregulated in an *in vitro* model of persistence (nutrient starvation) and in mouse lungs. We have generated an F6 deletion strain (ΔF6) and using microarrays we have analysed differential gene expression which, in combination with *in silico* prediction of sRNA-mRNA interactions, has led to the identification of a potential regulatory target of F6. Finally, we have assessed the fitness of the ΔF6 strain *in vitro* and *in vivo*. Our results indicate that

although deletion of F6 does not lead to attenuation in macrophages or our specific mouse model, it may still be critical for certain aspects of infection. F6 is highly upregulated in mice, it does control *hrcA* expression and it does lead to impaired recovery from a Wayne model of hypoxia, suggesting a role in later stages of infection.

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli DH5α was used for plasmid construction and grown in LB Agar or Broth using kanamycin at 50 µg/ml and X-Gal where necessary at 200 µg/ml.

Mycobacterium tuberculosis was grown on 7H11 agar plus 10% OADC (Becton Dickinson). Liquid cultures were grown in standard Middlebrook 7H9 medium supplemented with 0.5% glycerol, 10% Middlebrook ADC (Becton Dickinson) and 0.05% Tween-80 at 37°C in a roller bottle (Nalgene) rolling at 2 rpm or 50 ml falcon tubes (Corning) in an SB3 Tube Rotator (Stewart) at 28 rpm. Kanamycin was added where required at 25 µg/ml and X-Gal where required at 50 µg/ml.

All plasmids and oligos used in this study are listed in Tables S1 and S2.

Wayne Model

Cultures were grown to exponential phase (OD_{600}=0.6-0.8) and subsequently diluted to OD_{600}=0.005 in 7H9 in triplicate in Wayne tubes, each containing a sterile stirring bar. Cultures were incubated at 37°C on a stirring platform and OD monitored over time until cultures reached NRP-2. Cultures were then either diluted into 7H9 OD_{600}=0.05 and OD monitored, or serially diluted onto 7H11 agar and CFU enumerated.

Construction of an F6 deletion strain and complemented strain

Allelic replacement techniques were used to generate an *M. tuberculosis* knockout mutant as per published protocol (20). Briefly, approximately 1.5kb of the flanking region from either side of F6 (5' flank coordinates 292520-293605 and 3' flank 293709-294600) were cloned into the suicide vector pBackbone. The sRNA was removed by site directed mutagenesis to produce the targeting plasmid pJHP04. Electrocompetent H37Rv was transformed with pJHP04 and subsequent single cross overs (SCOs) and double cross overs (DCOs) selected.

To complement the ΔF6 mutant, electrocompetent cells were prepared and transformed with an integrating plasmid (pKP186) containing the cloned region 293428-293876 from H37Rv.

Preparation of starved cultures

M. tuberculosis H37Rv, ΔF6 and complemented strain were grown in Middlebrook 7H9 supplemented with 0.4% glycerol, 0.085% NaCl, 0.5% BSA and 0.05% Tyloxapol in roller bottle culture (2 rpm at 37°C). Exponentially growing bacteria were pelleted and washed 3 times with PBS supplemented with 0.025% Tyloxapol and finally resuspended in PBS with 0.025% Tyloxapol to an equivalent starting volume and incubated statically for 24 or 96 hours.

RNA isolation

RNA isolation from *in vitro* cultures was done as described previously (6). Briefly, cultures were harvested with rapid cooling by adding ice directly to the culture and subsequent centrifugation at 10,000 rpm for 10 minutes. RNA was isolated from the pellet using the FastRNA Pro Blue Kit from MP Biomedicals following manufacturer's instructions.

To isolate RNA from bacteria grown in mice, lung homogenates were spun at 13,000 rpm for 5 minutes to collect the bacteria. These were resuspended in 1 ml Trizol (Invitrogen) with 150 micron glass beads and the samples disrupted in a fast Prep (MPBio) at a setting of 6.0 for 40 secs. The RNA was extracted according to manufacturer's guidelines. RNA concentration was measured by nanodrop and RNA integrity measured by the 2100 Bioanalyzer using a Nano chip.

Quantitative RT-PCR

Total RNA was treated with Turbo DNase (Ambion) until DNA free. cDNA was synthesized using Superscript III (Invitrogen) and random hexamers. Primers were designed using the Applied Biosystems software Primer Express, and sequences are listed in Table S1. Each 20 µl qRT-PCR reaction, contained 16SYBRgreen

(Applied Biosystems), 900 nm each primer and 5 µl of template cDNA. Absolute quantitation was performed and all genes were normalised to 16S expression.

Microarrays

Whole genome *M. tuberculosis* microarray slides were purchased from Agilent Technologies through the Bacterial Microarray Group at St. George's (BmG@S), University of London. For cDNA synthesis 2 mg of wildtype H37Rv and ΔF6 knockout, isolated from 24 hour starved cultures was used. The cDNA was labelled individually with both Cy-3 and Cy-5 dyes (GE Healthcare) using Superscript III reverse transcriptase (Invitrogen). Dye swaps were performed, and the cDNA hybridized to an 8 Chamber Agilent slide at 65˚C for 16 hours before washing the slide with Oligo aCGH Wash Buffer 1 (Agilent) for 5 minutes at room temperature and Oligo aCGH Wash Buffer 2 (Agilent) for 1 minute at 37˚C. Slides were stabilized using Agilent Stabilisation and Drying Solution according to manufacturer's instructions.

Slides were scanned at 5 microns using an Agilent Technologies Microarray Scanner at BmG@S. Txt files created by the Agilent scanner were analysed using Genespring 12.0 filtering on flags and expression. T-test against zero was performed and p-value selected as p,0.05, correcting for multiple comparisons using Benjamini-Hochberg. Array design is available in BmG@Sbase (Accession No. A-BUGS-41; http://bugs.sgul.ac.uk/A-BUGS-41) and ArrayExpress (Accession No. A-BUGS-41). Microarray data have been deposited in ArrayExpress (Accession number E-MTAB-9327.

Mapping genome data and variant calling

WGS was performed as described elsewhere using the Illumina HiSeq platform (21). Sequencing reads for the ΔF6 strain were mapped against the *M. tuberculosis* H37Rv reference genome (AL123456) using BWA (22), and variants called with SAMtools (23). Variant filtering was performed by inclusion of only those variants with a minimum mapping quality of 10 and maximum read depth of 400. Finally, heterozygous calls or those found in repetitive or mobile elements (genes annotated as PE/PPE/insertions/phages) were removed.

Macrophage Infection

Bone marrow derived macrophages (BMDMs) were generated from 6-8 week old Balb/C mice in RPMI-1640 (Gibco) containing 10% foetal calf serum, 20 μM L-glutamine, 1 mM sodium pyruvate, 10 μM HEPES and 50 nM β-mercaptoethanol. The

cells were then grown and differentiated in complete RPMI-1640 supplemented with 20% L929 cell supernatant for 6 days at 37°C in 5% CO_2. The differentiated cells were seeded at a density of 2 x 105 cells/well in 1 ml complete RPMI-1640 supplemented with 5% L929 cell supernatant and incubated overnight prior to infection.

M. tuberculosis strains for infection were grown to an OD_{600} of 0.5-0.8 and inocula prepared by washing and resuspending the cultures in PBS, to produce a single cell suspension. This was used to infect BMDMs at a multiplicity of infection (MOI) of 0.1:1. After 4 hr, the cells were washed to remove all extracellular bacilli, and medium was replaced and incubation continued. Macrophages were lysed with water-0.05% Tween 80 to release intracellular bacteria after 4, 24, 72, 120 and 168 hr post-infection. Bacilli were serially diluted in PBS-Tween and plated on 7H11 with OADC. Plates were incubated for 3-4 weeks to obtain CFU counts.

Mice and ethics statement

Groups of 6–8 week old Balb/C mice were infected by low-dose aerosol exposure with *M. tuberculosis* H37Rv using a Glas-Col (Terre Haute, IN) aerosol generator calibrated to deliver approximately 100 bacteria into the lungs. Bacterial counts in the lungs (n = 5) at each time point of the study were determined by plating serial dilutions of individual lung homogenates on duplicate plates of Middlebrook 7H11 agar containing OADC enrichment. Colony-forming units were counted after 3–4 weeks incubation at 37˚C. Balb/C mice were bred and housed under specific pathogen free conditions at the Medical Research Council, National Institute for Medical Research (NIMR). All mouse studies and breeding were approved by the animal ethics committee at NIMR. Protocols for experiments were performed, under project license number 80/2236, in accordance with Home Office (United Kingdom) requirements and the Animal Scientific Procedures Act, 1986.

RESULTS

Expression of F6 is up-regulated by starvation and during infection

SigF is the suspected master regulator of F6 expression, and this was validated by northern blotting, which demonstrated a complete absence of F6 in a Δ*sigF* strain (Fig. S1). As SigF is highly expressed during nutrient starvation and infection, we speculated that upregulation of F6 under these conditions was highly likely (16). A

widely used *in vitro* model of *M. tuberculosis* latency and persistence is static incubation in PBS, effectively starving the bacteria for nutrients and to some extent oxygen, which leads to a significant up-regulation of SigF (16). We subjected cultures of *M. tuberculosis* H37Rv to starvation by washing and resuspending log-phase cultures in PBS and incubating them in standing cultures for 96 hours before isolating total RNA. F6 expression was measured by qRT-PCR (normalised to 16S rRNA) and compared to log-phase levels, and the results, shown in Fig. 1, demonstrate that F6 expression increased by almost two logs (>75-fold) in response to starvation, suggesting a role for this sRNA under limiting nutrient availability that may be seen during infection.

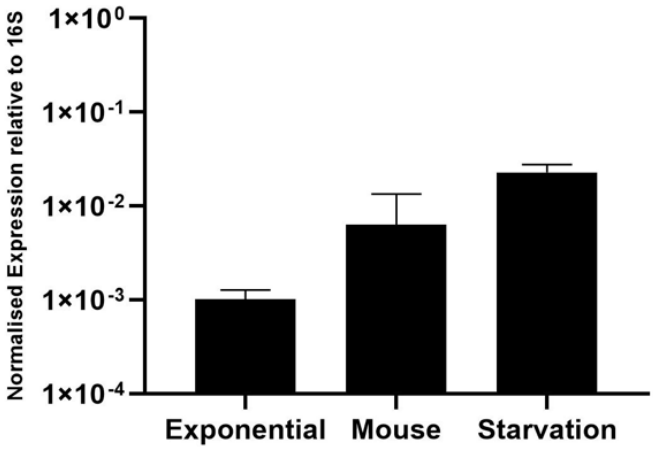

Figure 1: Expression of F6 in *M. tuberculosis*. F6 expression levels were measured using quantitative RT-PCT (qRT-PCR) normalised to 16S rRNA and compared to log-phase levels. Starvation refers to 96 hours in PBS. Mouse refers to *M. tuberculosis* RNA from the lungs of three-week infected mice. Data represents the mean and standard deviation of three biological replicates for each condition.

Next, we wanted to investigate F6 expression in a mouse model of infection. BALB/c mice were infected with ~100 CFU of H37Rv via aerosol route and left for three weeks before culling and isolation of *M. tuberculosis* total RNA from lung tissue. The RNA was analysed as described above demonstrating that F6 expression was up-regulated approximately 6-fold in mouse lung tissue compared to *in vitro* log-phase levels, but clearly not as dramatically as during *in vitro* starvation (Fig. 1). Together our results demonstrate that the expression of F6 is highly dynamic and that it may play a role during infection, more specifically in nutrient-poor environments.

Deletion of F6 in *M. tuberculosis* H37Rv

To identify a potential regulatory role of F6 with minimal impact on the two flanking *fad* genes, we generated an unmarked F6 deletion strain (ΔF6, see materials and methods for details). Candidates were screened by PCR amplification and sequencing of the PCR product, which confirmed the deletion of the sRNA from the *M. tuberculosis* genome (Fig. S2).

To further verify the deletion and to determine if there were any secondary mutations that might affect subsequent phenotypic analysis, ΔF6 strain was sequenced using high-throughput sequencing. Alignment of the genomic sequence of ΔF6 with H37Rv identified two single-nucleotide polymorphisms (SNPs) in addition to those generated by the deletion scar; one was C2864730T, resulting in an R102W substitution in Rv2541 while the other was C4178146T in the 5' UTR of Rv3729.

To ensure that these SNPs did not influence any potential phenotypes of the F6 deletion strain, we constructed a complemented strain, in which expression of F6 was driven by its native promoter from a single-copy plasmid integrated on the chromosome. F6 expression in wildtype, H37Rv, ΔF6 and the complemented strain was compared using qRT-PCR, and the results demonstrate that there was no statistical significance (unpaired t-test, $p<0.05$) between wildtype and the complemented strain, while there was no detectable expression in ΔF6 (Figure S3).

The ΔF6 strain was assessed for an *in vitro* growth phenotype by comparing its growth to that of wildtype and the complement strain in 7H9 roller bottle cultures over two weeks. Under these conditions, there was no significant difference between the three strains (Fig. 2).

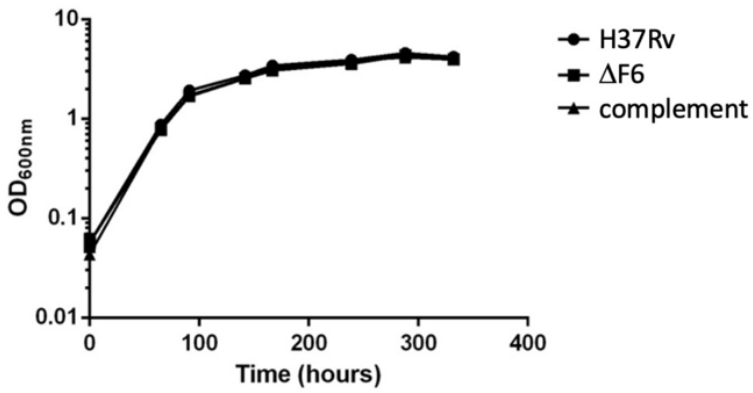

Figure 2: *In vitro* growth of ΔF6. Growth curves of wildtype *M. tuberculosis* H37Rv, ΔF6 and its complement in 7H9 in roller bottles. The results represent the average and standard deviation of three biological replicates.

Recovery from Wayne Model hypoxia is impaired in ΔF6

In addition to the starvation model of latency, another widely used *in vitro*-model is the Wayne model, in which the amount of available oxygen is gradually limited in sealed cultures of *M. tuberculosis* (24). When oxygen concentrations decrease to the microaerobic level (1% oxygen saturation), the cells enter a state of non-replicating persistence (NRP-1) followed by NRP-2 when oxygen saturation reaches 0.06%. SigF is highly induced during anaerobiosis such as that observed the Wayne model, suggesting that its regulon, including F6, may play a role in low-oxygen conditions (18). We therefore decided to evaluate the fitness of the ΔF6 strain using the Wayne model. To measure potential differences in respiration rate, the depletion of oxygen for each strain was monitored by a methylene blue indicator tube set up for each strain at the start of incubation. After 14 days incubation, cultures were plated onto 7H11 agar for determination of colony forming units (CFU). In addition, NRP-2 cells were assessed for their ability to be resuscitated by dilution into fresh 7H9 media and monitoring of growth (OD_{600}). Enumeration was performed to ensure that an equal number of bacteria was inoculated.

According to the CFU counts, there was no significant difference in survival between wildtype H37Rv, ΔF6 and the complemented strain after incubation in the Wayne model indicating equal numbers of live bacteria per unit of OD (Fig. 3A). This also demonstrated that an equal number of bacteria were used in the inoculation of fresh media for regrowth.

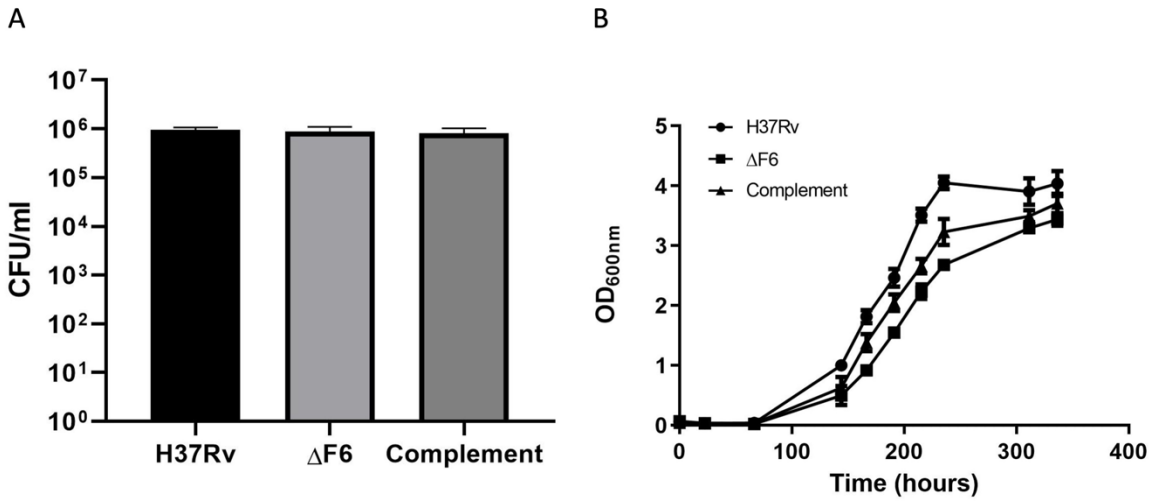

Figure 3: Recovery from NRP-2 is impaired in ΔF6. Wildtype *M. tuberculosis* H37Rv, ΔF6 and complement strains were grown in the Wayne model until NRP-2 was reached. Cultures were diluted and plated for CFU counts. (A) Number of viable bacteria after incubation in the Wayne model. All cultures were subsequently adjusted to OD_{600}=0.05 in fresh media and growth measured over 14 days (B). All data represents the averages and standard deviation of three biological replicates. Linear regression analysis indicates a significant difference in recovery of wildtype and ΔF6, as tested by $p<0.05$.

Conversely, the growth curve of the NRP-2 cells transferred to fresh media indicated that the ΔF6 strain was impaired for recovery/resuscitation following incubation in the Wayne model, while the complemented strain displayed an intermediate phenotype (Fig. 3B). These results suggest that F6 may play a role in resuscitation of non-replicating *M. tuberculosis*, associated with the reactivation of latent infection.

Deletion of F6 induces a heat shock response during starvation.

To establish a potential role for F6 during nutrient starvation, we performed microarray analysis on RNA isolated from PBS-starved cultures of wildtype H37Rv and ΔF6. To strike a balance between starvation and RNA quality, which showed significant degradation after 96 hours in PBS, we decided to reduce the PBS incubation to 24 hours. Experiments were performed in biological triplicates and technical (dye-swap) duplicates using Agilent gene expression microarrays (see materials and methods). To our surprise, only four genes showed ≥ 2-fold change (between 2.5 and 4 fold) and these were all upregulated in the mutant (Table 1). Strikingly, all of these are under the control of the transcriptional repressor HrcA (Rv2374c), although encoded in different parts of the genome (25). Two of these are Rv0440 (*groEL2*) and Rv3418c

(*groES*), i.e. heat-shock inducible chaperonins, the other two genes are conserved hypotheticals, Rv0990c and Rv0991c, which form an operon (25).

Gene number	Gene name	Fold change	P-value
Rv0440	*groEL2*	4.08	0.035
Rv3418c	*groES*	3.38	0.049
Rv0991c	-	2.81	0.032
Rv0990c	-	2.52	0.037

Table 1 Genes differentially expressed ≥ 2-fold in ΔF6 after 24 hrs starvation

To validate the microarray results and to ensure that the deletion strain could be complemented we performed qRT-PCR on RNA extracted from three biological replicates of each strain including the complemented KO strain.

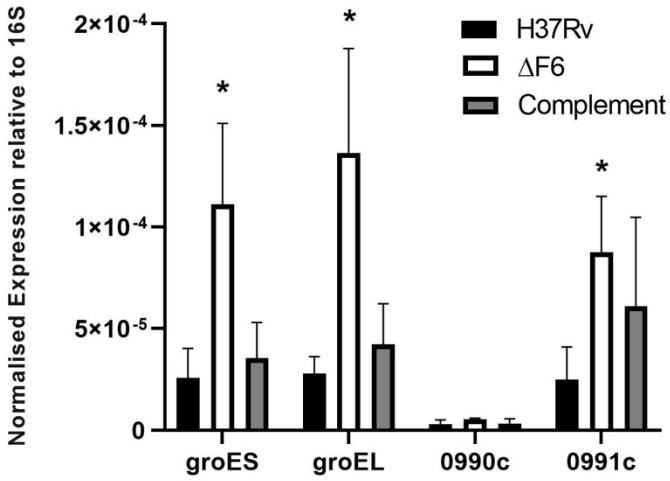

Figure 4: Validation of differential gene expression for ΔF6. qRT-PCR was performed on total RNA from *M. tuberculosis* H37Rv (wildtype), ΔF6 mutant and its complement. Each bar shows the expression level of the gene normalised to 16S. All data represents the mean and standard deviation of three biological replicates for each strain. * = p value <0.05 with significance tested with one-way ANOVA.

The results confirmed that *groES, groEL2,* and *Rv0991c* were significantly upregulated in ΔF6 during starvation, and that the phenotype could be largely

complemented by providing a copy of F6 in *trans* (Fig. 4). The changes for *Rv0990c* (co-transcribed with *Rv0991c*) were not significant according to qRT-PCR.

In silico **prediction of F6-*hrcA* RNA interaction**

The HrcA-regulon comprises very few genes in *M. tuberculosis*, i.e. the four in table 1 as well as *groEL1*, which is co-transcribed with *groES* (25). Deletion of *hrcA* leads to upregulation of those specific genes, and the most likely explanation for the observed coordinated upregulation in the ΔF6 strain, is a down-regulation of HrcA expression leading to de-repression of its regulon. In other words, F6 activates HrcA expression, directly or indirectly. *hrcA* was not among the significantly differentially expressed genes in the microarray analysis, but a direct interaction between mRNA and sRNA does not necessarily affect mRNA levels. Importantly, a lack of significant change in mRNA levels does not exclude regulation by a given sRNA.

The 5' untranslated region (UTR) of the *hrcA* mRNA is relatively short (34 nucleotides), but nevertheless has the potential to form a stable stem-loop that sequesters the start codon and the region immediately upstream, excluding the potential Shine-Dalgarno (Fig. 5A). As our results were largely consistent with a role for HrcA in the coordinated upregulation of the two operons, we decided to investigate a potential interaction between F6 and the *hrcA* mRNA *in silico*, using the targetRNA2 webserver (26). We changed the default of 80 nucleotides upstream to 34 nucleotides in line with the mapped transcription start site for *hrcA* (2) and used *hrcA* as single target with a hybridisation seed of six nucleotides. The result shown in Fig. 5B suggests that F6 can interact with the first 12 nucleotides of the *hrcA* mRNA, i.e. the exact region that sequesters part of the translation initiation region. This interaction was confirmed by RNAhybrid (Fig. 5) (27). Further studies are required to validate this prediction, but is seems plausible that F6 is involved in activating *hrcA* translation by opening an otherwise autoinhibitory structure similar to what has been described for the interaction between the *E. coli rpoS* mRNA and several sRNAs (28).

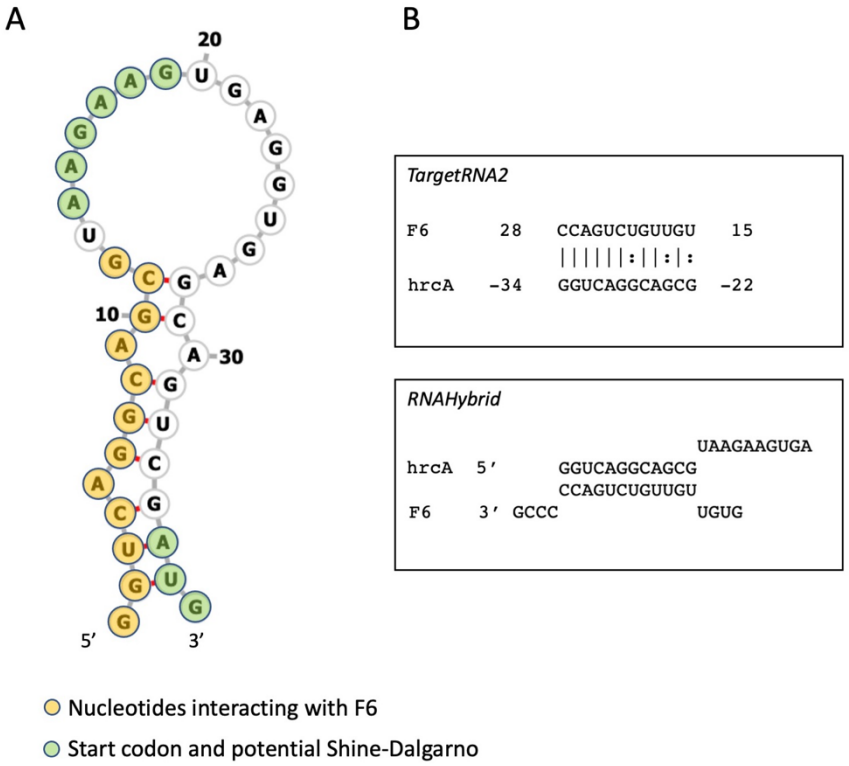

Figure 5: Predicted interactions between F6 and the *hrcA* 5' UTR. A: Predicted auto-inhibitory structure of the *hrcA* 5' UTR including the AUG start codon, which is partially sequestered by the upper end of the UTR. Nucleotides predicted to interact with F6 are indicated in yellow. A putative Shine-Dalgarno sequence and start codon are indicated in green. B: Predicted basepairing between F6 and *hrcA* 5'UTR, according to TargetRNA2 (top) and RNAhybrid (bottom).

Fitness of the ΔF6 strain in macrophages and mouse models of infection

M. tuberculosis is an intracellular pathogen that is subjected to a variety of stresses within the host, including but not limited to oxidative stress and nutrient deprivation. F6 is induced under oxidative stress, starvation and infection, and expression is under control of SigF, which is critical for intracellular survival in mice and guinea pigs (17,29). Moreover, the ΔF6 strain displayed impaired recovery from the Wayne model.

Wildtype *M. tuberculosis* H37Rv, ΔF6 and the complemented strain, were used to infect naïve and activated murine BMDMs. Macrophages were pre-activated with 10 ng/ml IFN-γ and incubated overnight and infected alongside naive macrophages at an MOI of 0.1:1 (bacteria:macrophages). The infection was allowed to continue for 7 days with five time points. At each time point the macrophages were lysed and plated for CFUs. For both naïve and activated macrophages we observed no significant differences between the three strains, suggesting that under these conditions, F6 does not result in attenuation in the mouse macrophage model (Fig. 6).

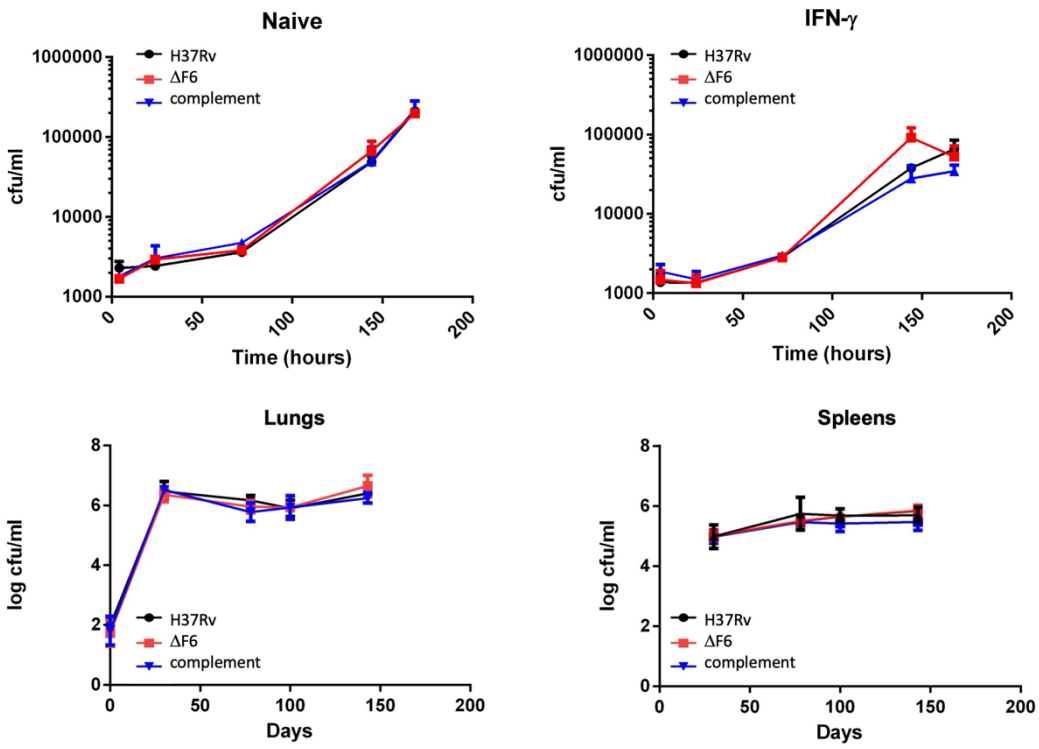

Figure 6: Survival of the ΔF6 in models of infection. Top panels show survival of Wildtype *M. tuberculosis* H37Rv, ΔF6, and complementing strains within naive and IFN-γ activated BMDM over a time course of infection. Data is the mean and standard deviation of triplicate infections. Significance was tested with One-way ANOVA $p<0.05$. Bottom panels show survival of the three strains within the lungs and spleens of BALB/c mice. Data represents the averages and standard deviations from 5 mice per time point. Significance was tested with Two-way ANOVA.

A macrophage model of infection does not entirely replicate the conditions encountered by the bacteria in more complex animal models of infection. In order to assess the role of F6 in pathogenesis in a more representative infection model, wildtype *M. tuberculosis* H37Rv, ΔF6, and the complemented strain were used in a mouse model of infection. BALB/c mice were infected with approximately 100 CFU via the aerosol route and the infection followed for 143 days. At time points 0, 30, 78, 100 and 143 days, lungs and spleens were harvested for bacterial enumeration. No significant difference was observed in CFUs between the wildtype *M. tuberculosis* H37Rv and ΔF6 at any of the time points (Fig. 6).

DISCUSSION

M. tuberculosis F6 is a SigF-controlled sRNA, which is well-expressed in exponential phase, but up-regulated by a variety of stresses (6).

In this study we have shown that F6 is also highly up-regulated after three weeks of murine infection and to an even higher extent in the starvation model of 96 hours starvation in PBS, suggesting that this sRNA plays a role in the early stages of pathogenesis but possibly a more significant one for survival in nutrient deficient environments such as those encountered in human granulomas in later stages of infection. *M. tuberculosis* has the ability to resuscitate after long periods of low metabolic activity/dormancy (24,30), and using the Wayne model of dormancy, we found that recovery from NRP-2 was impaired for ∆F6 *versus* wildtype/complement, although initial survival appeared to be unchanged, according to CFU counts.

Our study indicates that starvation for 24 hours leads to an upregulation of three *loci* (encoding *groES, groEL2,* and *Rv0991c*) in the ΔF6 strain compared to the wildtype and complemented strains. Together, these *loci* make up the HrcA regulon suggesting that this heat shock repressor is likely coordinating and responsible for the observed differences (25). We observed no significant change in the levels of *hrcA* mRNA, but we did identify a potential autoinhibitory structure within the 5' UTR of the transcript, which could be sequestered by interaction with F6 to expose the entire translation initiation region. SigF has previously been shown to be downregulated during heat shock (18), offering the possibility that F6 represents a critical regulatory link between heat/cold shock and the expression of essential chaperonins by directly targeting the *hrcA* mRNA and thereby GroEL/S (Fig. 7).

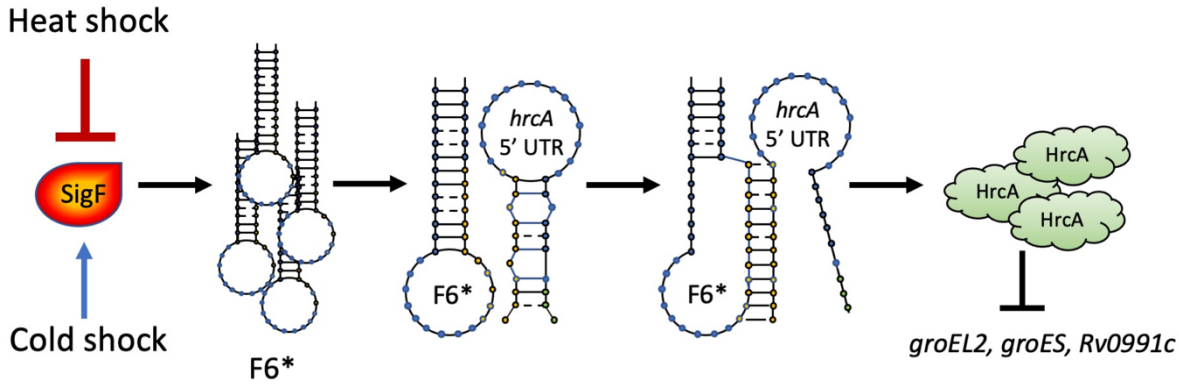

Figure 7: Model for F6-mediated heat shock control. The schematic shows how heat and cold shock regulate SigF expression. SigF in turn controls the expression of F6, shown here in its processed form, F6* (residues 9-50 rather than1-55), which can potentially interact with a large portion of the *hrcA* 5' UTR, thereby relieving an autoinhibitory structure to ensure translation of HrcA. Broken lines in RNA indicate G:U basepairs.

The slow growth phenotype associated with F6 overexpression could therefore be the result of a detrimental repression of these essential chaperonins, suggesting in turn that overexpression of F6 may result in a heat shock sensitive phenotype in addition to the slow growth (6). The lack of a significant downregulation of *hrcA* mRNA levels in the ΔF6 strain could be due to a stabilising effect of the closed *hrcA* stem-loop on the transcript, similar to what we have reported for the DrrS sRNA (10). Having an sRNA link between stimulus and expression of effector proteins, possibly from and already existing, stable transcript, could provide the system with a faster response time. Moreover, RNA is a ubiquitous temperature-sensing molecule and several heat shock responses employ RNA in their control mechanisms, including the eukaryotic heat shock transcription factor 1 (HSF1), *E. coli rpoH* and *Deinococcus radiodurans* sRNA DnrH, which activates the expression of Hsp20 (31-33).

Whether the observed changes in gene expression are the result of a direct interaction between F6 and the *hrcA* mRNA could be validated using translational *hrcA*-reporter fusions and by determining *hrcA* mRNA stability. Assessing fitness of ΔF6 during cold shock (or overexpression of F6 combined with heat shock) may elaborate further on the role of F6. It is possible that longer periods of starvation will provide more significant differences between wildtype and mutant, although the rapid decline in RNA integrity during starvation has to be taken into consideration. An alternative approach could be to use short-term induced overexpression of F6, combined with RNA-seq.

Infection of BMDMs for 7 days with H37Rv, ΔF6 and complement revealed no attenuation of the deletion strain. Similarly, we observed no difference in the CFUs recovered from the lungs and spleens of BALB/c mice after 143 days of infection. As F6 expression is not only directed by SigF, but also dominant in terms of promoter occupancy, one might expect some parallels between infection with ΔF6 and Δ*sigF*.

Indeed, a lack of attenuation has been observed in a study of a Δ*sigF* strain in human monocyte derived macrophages, while late stages of infection and time to death experiments result in a strong phenotype of a CDC1551 Δ*sigF* strain (17,34). Given that starvation is a potent inducer of SigF/F6 expression, it seems possible that a more long-term infection and/or the ability to form granulomas, are required to observe strong phenotypes associated with the deletion of F6. Moreover, the mouse model of

infection does not replicate all aspects of an *M. tuberculosis* infection in humans. This is mainly apparent in the lack of ability of the mouse to form granulomas. Granulomas provide a unique niche for *M. tuberculosis* in which oxygen and nutrients are limited, the latter being the strongest inducer of F6 expression in our study, and a different animal model or e.g. the more susceptible C3HeB/FeJ mouse strain which develop necrotic lung granulomas, may provide more answers. Our results point towards a role for F6 during periods of low metabolic activity similar to cold shock and associated with nutrient starvation such as that found in human granulomas in later stages of infection, and it remains a possibility that F6 plays a role in surviving in this nutrient poor environment as well as during resuscitation from dormancy.

ACCESSION NUMBERS

Array design is available in BmG@Sbase (Accession No. A-BUGS-41; http://bugs.sgul.ac.uk/A-BUGS-41) and ArrayExpress (Accession No. A-BUGS-41). Microarray data have been deposited in ArrayExpress (Accession number E-MTAB-9327.

ACKNOWLEDGEMENT

We thank Ruben Hartkoorn and Stewart Cole for the H37Rv Δ*sigF* mutant, the Biological Service staff at NIMR, Mill Hill for help with maintaining mice, the BµG@S (the Bacterial Microarray Group at St. George's, University of London) for supplying *M. tuberculosis* microarrays and Finn Werner for critical reading of the manuscript

FUNDING

This work was supported by the British Medical Research Council (MRC) [programme number U117581288; grant numbers MR/L018519/1; MR/S009647/1]. Funding for open access charge: University College London.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CONFLICT OF INTEREST

None to declare

REFERENCES

1. Arnvig, K.B., Comas, I., Thomson, N.R., Houghton, J., Boshoff, H.I., Croucher, N.J., Rose, G., Perkins, T.T., Parkhill, J., Dougan, G. *et al.* (2011) Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog*, **7**, e1002342.
2. Cortes, T., Schubert, O.T., Rose, G., Arnvig, K.B., Comas, I., Aebersold, R. and Young, D.B. (2013) Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in Mycobacterium tuberculosis. *Cell Rep*, **5**, 1121-1131.
3. Houghton, J., Cortes, T., Schubert, O., Rose, G., Rodgers, A., De Ste Croix, M., Aebersold, R., Young, D.B. and Arnvig, K.B. (2013) A small RNA encoded in the Rv2660c locus of Mycobacterium tuberculosis is induced during starvation and infection. *PLoS One*, **8**, e80047.
4. Miotto, P., Forti, F., Ambrosi, A., Pellin, D., Veiga, D.F., Balazsi, G., Gennaro, M.L., Di Serio, C., Ghisotti, D. and Cirillo, D.M. (2012) Genome-wide discovery of small RNAs in Mycobacterium tuberculosis. *PLoS One*, **7**, e51950.
5. Shell, S.S., Wang, J., Lapierre, P., Mir, M., Chase, M.R., Pyle, M.M., Gawande, R., Ahmad, R., Sarracino, D.A., Ioerger, T.R. *et al.* (2015) Leaderless Transcripts and Small Proteins Are Common Features of the Mycobacterial Translational Landscape. *PLoS Genet*, **11**, e1005641.
6. Arnvig, K.B. and Young, D.B. (2009) Identification of small RNAs in Mycobacterium tuberculosis. *Mol Microbiol*, **73**, 397-408.
7. DiChiara, J.M., Contreras-Martinez, L.M., Livny, J., Smith, D., McDonough, K.A. and Belfort, M. (2010) Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res*, **38**, 4067-4078.
8. Gerrick, E.R., Barbier, T., Chase, M.R., Xu, R., Francois, J., Lin, V.H., Szucs, M.J., Rock, J.M., Ahmad, R., Tjaden, B. *et al.* (2018) Small RNA profiling in Mycobacterium tuberculosis identifies MrsI as necessary for an anticipatory iron sparing response. *Proc Natl Acad Sci U S A*, **115**, 6464-6469.
9. Mai, J., Rao, C., Watt, J., Sun, X., Lin, C., Zhang, L. and Liu, J. (2019) Mycobacterium tuberculosis 6C sRNA binds multiple mRNA targets via C-rich loops independent of RNA chaperones. *Nucleic Acids Res*, **47**, 4292-4307.
10. Moores, A., Riesco, A.B., Schwenk, S. and Arnvig, K.B. (2017) Expression, maturation and turnover of DrrS, an unusually stable, DosR regulated small RNA in Mycobacterium tuberculosis. *PLoS One*, **12**, e0174079.
11. Pelly, S., Bishai, W.R. and Lamichhane, G. (2012) A screen for non-coding RNA in Mycobacterium tuberculosis reveals a cAMP-responsive RNA that is expressed during infection. *Gene*, **500**, 85-92.
12. Salina, E.G., Grigorov, A., Skvortsova, Y., Majorov, K., Bychenko, O., Ostrik, A., Logunova, N., Ignatov, D., Kaprelyants, A., Apt, A. *et al.* (2019) MTS1338, A Small Mycobacterium tuberculosis RNA, Regulates Transcriptional Shifts Consistent With Bacterial Adaptation for Entering Into Dormancy and Survival Within Host Macrophages. *Front Cell Infect Microbiol*, **9**, 405.
13. Solans, L., Gonzalo-Asensio, J., Sala, C., Benjak, A., Uplekar, S., Rougemont, J., Guilhot, C., Malaga, W., Martin, C. and Cole, S.T. (2014) The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in Mycobacterium tuberculosis. *PLoS Pathog*, **10**, e1004183.

14. Lovewell, R.R., Sassetti, C.M. and VanderVen, B.C. (2016) Chewing the fat: lipid metabolism and homeostasis during M. tuberculosis infection. *Curr Opin Microbiol*, **29**, 30-36.
15. Hartkoorn, R.C., Sala, C., Uplekar, S., Busso, P., Rougemont, J. and Cole, S.T. (2012) Genome-wide definition of the SigF regulon in Mycobacterium tuberculosis. *J Bacteriol*, **194**, 2001-2009.
16. Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A. and Duncan, K. (2002) Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. *Mol Microbiol*, **43**, 717-731.
17. Geiman, D.E., Kaushal, D., Ko, C., Tyagi, S., Manabe, Y.C., Schroeder, B.G., Fleischmann, R.D., Morrison, N.E., Converse, P.J., Chen, P. *et al.* (2004) Attenuation of late-stage disease in mice infected by the Mycobacterium tuberculosis mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis. *Infect Immun*, **72**, 1733-1745.
18. Michele, T.M., Ko, C. and Bishai, W.R. (1999) Exposure to antibiotics induces expression of the Mycobacterium tuberculosis sigF gene: implications for chemotherapy against mycobacterial persistors. *Antimicrob Agents Chemother*, **43**, 218-225.
19. DeMaio, J., Zhang, Y., Ko, C., Young, D.B. and Bishai, W.R. (1996) A stationary-phase stress-response sigma factor from Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A*, **93**, 2790-2794.
20. Hinds, J., Mahenthiralingam, E., Kempsell, K.E., Duncan, K., Stokes, R.W., Parish, T. and Stoker, N.G. (1999) Enhanced gene replacement in mycobacteria. *Microbiology*, **145 (Pt 3)**, 519-527.
21. Rose, G., Cortes, T., Comas, I., Coscolla, M., Gagneux, S. and Young, D.B. (2013) Mapping of genotype-phenotype diversity among clinical isolates of mycobacterium tuberculosis by sequence-based transcriptional profiling. *Genome Biol Evol*, **5**, 1849-1862.
22. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760.
23. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
24. Wayne, L.G. and Hayes, L.G. (1996) An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infect Immun*, **64**, 2062-2069.
25. Stewart, G.R., Wernisch, L., Stabler, R., Mangan, J.A., Hinds, J., Laing, K.G., Young, D.B. and Butcher, P.D. (2002) Dissection of the heat-shock response in Mycobacterium tuberculosis using mutants and microarrays. *Microbiology*, **148**, 3129-3138.
26. Kery, M.B., Feldman, M., Livny, J. and Tjaden, B. (2014) TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res*, **42**, W124-129.
27. Rehmsmeier, M., Steffen, P., Hochsmann, M. and Giegerich, R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*, **10**, 1507-1517.
28. Battesti, A., Majdalani, N. and Gottesman, S. (2011) The RpoS-mediated general stress response in Escherichia coli. *Annu Rev Microbiol*, **65**, 189-213.
29. Karls, R.K., Guarner, J., McMurray, D.N., Birkness, K.A. and Quinn, F.D. (2006) Examination of Mycobacterium tuberculosis sigma factor mutants using low-

dose aerosol infection of guinea pigs suggests a role for SigC in pathogenesis. *Microbiology*, **152**, 1591-1600.
30. Rosser, A., Stover, C., Pareek, M. and Mukamolova, G.V. (2017) Resuscitation-promoting factors are important determinants of the pathophysiology in Mycobacterium tuberculosis infection. *Crit Rev Microbiol*, **43**, 621-630.
31. Shamovsky, I., Ivannikov, M., Kandel, E.S., Gershon, D. and Nudler, E. (2006) RNA-mediated response to heat shock in mammalian cells. *Nature*, **440**, 556-560.
32. Xue, D., Chen, Y., Li, J., Han, J., Liu, Y., Jiang, S., Zhou, Z., Zhang, W., Chen, M., Lin, M. *et al.* (2019) Targeting Hsp20 Using the Novel Small Non-coding RNA DnrH Regulates Heat Tolerance in Deinococcus radiodurans. *Front Microbiol*, **10**, 2354.
33. Yura, T. (2019) Regulation of the heat shock response in Escherichia coli: history and perspectives. *Genes Genet Syst*, **94**, 103-108.
34. Chen, P., Ruiz, R.E., Li, Q., Silver, R.F. and Bishai, W.R. (2000) Construction and characterization of a Mycobacterium tuberculosis mutant lacking the alternate sigma factor gene, sigF. *Infect Immun*, **68**, 5575-5580.