**Comparative transcriptomics reveals PrrAB-mediated control of metabolic, respiration, energy-generating, and dormancy pathways in *Mycobacterium smegmatis***

**JASON D. MAARSINGH1, SHANSHAN YANG4, JIN G. PARK2, and SHELLEY E. HAYDEL1,3\***

*1School of Life Sciences, Arizona State University, Tempe AZ, USA*

*2The Biodesign Institute Virginia G. Piper Center for Personalized Diagnostics, Arizona State University, Tempe AZ, USA*

*3The Biodesign Institute Center for Immunotherapy, Vaccines and Virotherapy, Arizona State University, Tempe AZ, USA*

4*Bioinformatics Core, Knowledge Enterprise Development, Arizona State University, Tempe AZ, USA*

**Additional File 1: Supplemental figures**



**Figure S1.** Members of the mycobacterial *Abscessus-Chelonae* clade harbor unique PrrA amino acid “signatures”. Boxed residues correspond to amino acid residues only found in mycobacterial species belonging to the *Abscessus-Chelonae* clade (bottom row) compared to all other mycobacterial clades (top row) or *M. smegmatis* mc2155 and *M. tuberculosis* H37Rv (middle row). Numerical system below single-letter amino acid codes correspond to the residue position in *M. smegmatis* (top and middle rows) or *M. abscessus* (bottom row). Left box corresponds to PrrA residue S38 of *M. smegmatis* (top and middle rows) and N35 of *M. abscessus*. Right box corresponds to PrrA residue S49 of *M. smegmatis* (top and middle rows) and C46 of *M. abscessus*.



**Figure S2.** Members of the mycobacterial *Abscessus-Chelonae* clade harbor unique PrrB amino acid “signatures”. Boxed residues correspond to amino acid residues only found in mycobacterial species belonging to the *Abscessus-Chelonae* clade (bottom row) compared to all other mycobacterial clades (top row) or *M. smegmatis* mc2155 and *M. tuberculosis* H37Rv (middle row). Numerical system below single-letter amino acid codes correspond to the residue position in *M. smegmatis* (top and middle rows) or *M. abscessus* (bottom row). (**a**) PrrB residue T42 of *M. smegmatis* (top and middle rows) and E66 of *M. abscessus*. (**b**) PrrB residue G67 of *M. smegmatis* (top and middle rows) and V93 of *M. abscessus*. (**c**) PrrB residue V90 of *M. smegmatis* (top and middle rows) and 117 of *M. abscessus*. (**d**) PrrB residue M318 of *M. smegmatis* (top and middle rows) and D343 of *M. abscessus*. (**e**) PrrB residue A352 of *M. smegmatis* (top and middle rows) and K377 of *M. abscessus*. (**f**) PrrB residue R371 of *M. smegmatis* (top and middle rows) and V396 of *M. abscessus*.



**Figure S3:** *M. smegmatis*growth characteristics in M7H9 broth. Optical density (OD600) of mc2155 (open circles), FDL10 (red triangles), and FDL15 (green diamonds). The blue arrow shows the OD600 (~0.6) when cultures were collected for RNA isolation. Values represent the mean ±SEM of data collected from three independent cultures.



**Figure S4.** Multidimensional scaling (MDS) plot of triplicate *M. smegmatis* RNA-seq samples. Given the MDS-based distance separation of the mc2155\_2 sample (circled in the bottom-left corner of plot) from other mc2155 replicates, the mc2155\_2 sample was removed from differential expression analysis.



**Figure S5.** Principal component analysis (PCA) of *M. smegmatis* strains used for RNA-seq DEG analyses.



**Figure S6.** Global expression profile of DEGs (*q* <0.05). (**a**) Total DEGs (*q* < 0.05) induced (yellow) or repressed (blue) by PrrAB in mc2155 (WT) and FDL15 (Δ*prrAB* complementation) backgrounds (RNA-seq pair-wise comparisons to the Δ*prrAB* mutant). (**b**) Venn diagrams of DEGs (*q* < 0.05) demonstrating that 40 DEGs (*q* < 0.05) overlapped between RNA-seq pair-wise comparisons.



**Figure S7.** qRT-PCR verification of six randomly selected genes from the RNA-seq FDL10 vs. mc2155 comparison. All qRT-PCR measurements were performed from the same RNA samples used for RNA-seq analyses and each gene was tested in triplicate. Absolute fold-change ratios were calculated using the 2­-ΔΔCt method [1].



**Figure S8.** *M. smegmatis* extracellular ATP (supernatant) expressed as a percentage of whole culture normalized ATP (pM/CFU). Values represent the mean ±SEM of data collected from three independent cultures.



**Figure S9.** Multiple sequence alignment comparing the *M. smegmatis* and *M. tuberculosis* PrrA amino acid sequences. Secondary structures are represented by arrows (β-sheets) or bars (α-helices). Secondary structures colored in blue correspond to the N-terminal receiver domain while red corresponds to the C-terminal effector domain. The conserved phosho-receiving aspartate (D58) and DNA-binding recognition helix are labeled. Multiple sequence alignments were performed in JalView using default MUSCLE algorithms [2]. Secondary structure and DNA-binding recognition helix designations were adapted from Nowak et al. [3].

**REFERENCES**

1. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCt Method. Methods. 2001. 25(4):402-408.

2. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004. 5:113. doi:10.1186/1471-2105-5-113.

3. Nowak E, Panjikar S, Konarev P, Svergun DI, Tucker PA. The structural basis of signal transduction for the response regulator PrrA from *Mycobacterium tuberculosis*. J Biol Chem. 2006. 281(14):9659-9666.