bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Arrayed CRISPRi and Quantitative Imaging Describe the Morphotypic Landscape of Essential Mycobacterial Genes

Timothy J. de Wet[1,2*], Kristin R. Winkler[1,2], Musa M. Mhlanga[2,3], Valerie Mizrahi[1,2,4], Digby F. Warner[1,2,4*]

[1]SAMRC/NHLS/UCT Molecular Mycobacteriology Research Unit, Department of Pathology, University of Cape Town, South Africa;

[2]Institute of Infectious Disease and Molecular Medicine, University of Cape Town, South Africa;

[3]Department of Integrative Biomedical Sciences, University of Cape Town, South Africa;

[4]Wellcome Centre for Infectious Diseases Research in Africa, University of Cape Town, South Africa.

*For correspondence: timdewet@gmail.com; digby.warner@uct.ac.za

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Mycobacterium tuberculosis possesses a large number of genes of unknown or merely predicted function, undermining fundamental understanding of pathogenicity and drug susceptibility. To address this challenge, we developed a high-throughput functional genomics approach combining inducible CRISPR-interference and image-based analyses of morphological features and sub-cellular molecular localizations in the related non-pathogen, *M. smegmatis*. Applying automated imaging and analysis to an arrayed library of 272 essential gene knockdown mutants, we derive robust, quantitative descriptions of bacillary morphologies consequent on gene silencing. Leveraging statistical-learning, we demonstrate that functionally related genes cluster by morphotypic similarity and that this information can be used to infer gene function. Exploiting this observation, we reveal a previously unknown restriction-modification system, and identify filamentation as a defining mycobacterial response to histidine starvation. Our results support the application of large-scale image-based analyses for mycobacterial functional genomics, simultaneously establishing the utility of this approach for drug mechanism-of-action studies.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

The acquisition of genomic data continues to exceed significantly the pace at which corresponding functional information can be obtained (Kemble, Nghe et al. 2019). This aphorism applies to *Mycobacterium tuberculosis*, etiological agent of tuberculosis (TB) and the leading cause of death globally from an infectious disease (World Health Organization 2019). Despite considerable activity (Satta, Lipman et al. 2018) in the two decades subsequent to the pioneering release of the first complete mycobacterial whole-genome sequence (Cole, Brosch et al. 1998), the *M. tuberculosis* genome still contains a large number of genes of unknown or hypothetical function (Mazandu and Mulder 2012, Satta, Lipman et al. 2018). Moreover, functional validation is lacking even for many annotated genes, undermining fundamental understanding of mycobacterial metabolic and cellular functions which impact pathogenicity and drug susceptibility. There is consequently a pressing need for tractable, high-throughput approaches that can inform mycobacterial gene function rapidly and at scale.

The most commonly applied functional genomics methods in bacteria – transposon-sequencing (Tn-Seq) (Opijnen and Camilli 2013) and, increasingly, CRISPR-interference (CRISPRi)-Seq (de Wet, Gobe et al. 2018, Wang, Guan et al. 2018, Lee, Ostrov et al. 2019) – combine pooled mutagenesis with next-generation sequencing, returning quantitative estimates of fitness (via relative abundance) of mutants in a particular growth condition. Common to these approaches is that they report on the capacity of the cells to produce biomass, thereby excluding other – potentially subtler – measures of the physiological consequences of target gene disruption. To overcome this limitation, some studies have utilized novel combinations of pooled approaches with cellular stains and sorting (Rego, Audette et al. 2017, Baranowski, Welsh et al. 2018), or with imaging and *in situ* sequencing (Camsund, Lawson et al. 2019) to provide alternative readouts of genetic function. However, these methods are restricted in their resolution and throughput.

As for many bacterial pathogens, mycobacterial functional genomics has focused primarily on the identification of essential genes, employing Tn-based investigations under a variety of conditions (Sassetti, Boyd et al. 2003, DeJesus, Gerrick et al. 2017), in different strain (Carey, Rock et al. 2018) and mutant (Kieser, Baranowski et al. 2015) backgrounds, and in models of infection (Sassetti and Rubin 2003, Rengarajan, Bloom et al. 2005). The development of genetic tools for conditional knockdown mutagenesis via targeted gene silencing (utilizing engineered hypomorphic strains carrying promoter replacement and/or targeted protein degradation mutations) has subsequently enabled gene-by-gene explorations of essential gene function (Kim, Wei et al. 2011, Kim, O'Brien et al. 2013, Schnappinger and Ehrt 2014). However, while these approaches have ensured increasingly refined (conditional) essentiality predictions (DeJesus, Gerrick et al. 2017), there is still an absence of whole-cell functional data for the majority of essential genes.

Here, we couple key recent advances in mycobacterial CRISPRi (Rock, Hopkins et al. 2017) with quantitative microscopic imaging (Ducret, Quardokus et al. 2016) in developing a platform enabling whole-cell characterization of essential gene knockdown in the non-pathogenic model mycobacterium, *M. smegmatis*. Leveraging available gene essentiality and CRISPRi guide efficacy data (de Wet, Gobe et al. 2018), we describe the construction of an arrayed collection of 272 validated CRISPRi mutants targeting essential *M. smegmatis* homologs of *M. tuberculosis* genes. Applying high-throughput quantitative imaging to the arrayed library, we implement a bespoke analytic pipeline to probe essential gene function

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

at scale. From the "phenoprints" generated utilizing this approach, we infer the functions of a number of putative genes and illustrate the capacity to analyze metabolic and macromolecular synthesis pathways using clustering analyses. Notably, these analyses reveal a previously unknown mycobacterial restriction-modification system, and identify filamentation as a defining mycobacterial response to histidine starvation but not to other amino acid auxotrophies. Finally, we demonstrate the potential utility of this approach in elucidating antimicrobial mechanism-of-action (MOA), supporting its potential incorporation as a complementary tool in current TB drug discovery pipelines. Consistent with our intention to generate a community-accessible resource, all data from the screen are made available via an interactive database (https://timdewet.shinyapps.io/MorphotypicLanscape/).

RESULTS

Construction of an arrayed CRISPRi library targeting essential mycobacterial genes.

We aimed to construct an arrayed library of inducible CRISPRi mutants (***Figure 1A***) targeting essential *M. smegmatis* genes (de Wet, Gobe et al. 2018). To this end, we leveraged an optimized mycobacterial CRISPRi system (Rock, Hopkins et al. 2017) and previously generated genome-wide CRISPRi-Seq data (de Wet, Gobe et al. 2018) from which we were able to identify the set of highest efficiency single-guide (sg)RNAs targeting 294 essential *M. smegmatis* genes with direct *M. tuberculosis* homologs (***Supplementary Table 1***).

CRISPRi produces polar effects on downstream genes in polycistronic operons (Rock, Hopkins et al. 2017). To ascertain operonic structure directly, we performed whole-genome RNA-Seq analyses on wild-type *M. smegmatis* mc^2155 during exponential growth in standard laboratory media. From visual inspection, the results indicated that for 176 of the 294 genes transcription was either monocistronic or the genes were located at the end of a polycistronic operon (***Supplementary Table 2***). For these genes, it was therefore expected that any knockdown phenotype should solely reflect the impact of silencing the targeted essential gene. The remaining 128 genes were either the first gene in a polycistronic operon or were located internally within an operon. Although the genomic context was likely to complicate the interpretation of any knockdown phenotype, we nevertheless considered these genes worth pursuing given our ability to evaluate the results in the context of the corresponding essentiality calls from Tn-Seq (Dragset, Ioerger et al. 2019) and in comparison to phenotypes observed for related (or "shared pathway") genes.

Following large-scale cloning, all 294 CRISPRi constructs were electroporated into an *M. smegmatis* ParB-mCherry reporter mutant (Santi and McKinney 2015). This strain was chosen as background to enable visualization of the *oriC* proximate region of the chromosome during imaging assays, thereby providing information about chromosome location dynamics and copy number as a function of essential gene knockdown. To validate the identity and essentiality of each *M. smegmatis* mutant strain, we performed Sanger sequencing of the inserted sgRNA and whole-cell spotting assays of ATc sensitivity. Cloning was repeated for mutants which failed the validation screen. This process yielded 272 validated CRISPRi mutants (***Figure 1B***), 93% of the initial target set. According to a review of the literature (***Supplementary Table 3***), approximately 70% of the strains in our library lacked whole-cell phenotypic characterization in either *M. tuberculosis* or *M. smegmatis*. Furthermore, almost 40% of the targeted genes and their protein products had no biochemical or structural information (***Supplementary Table 3***).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

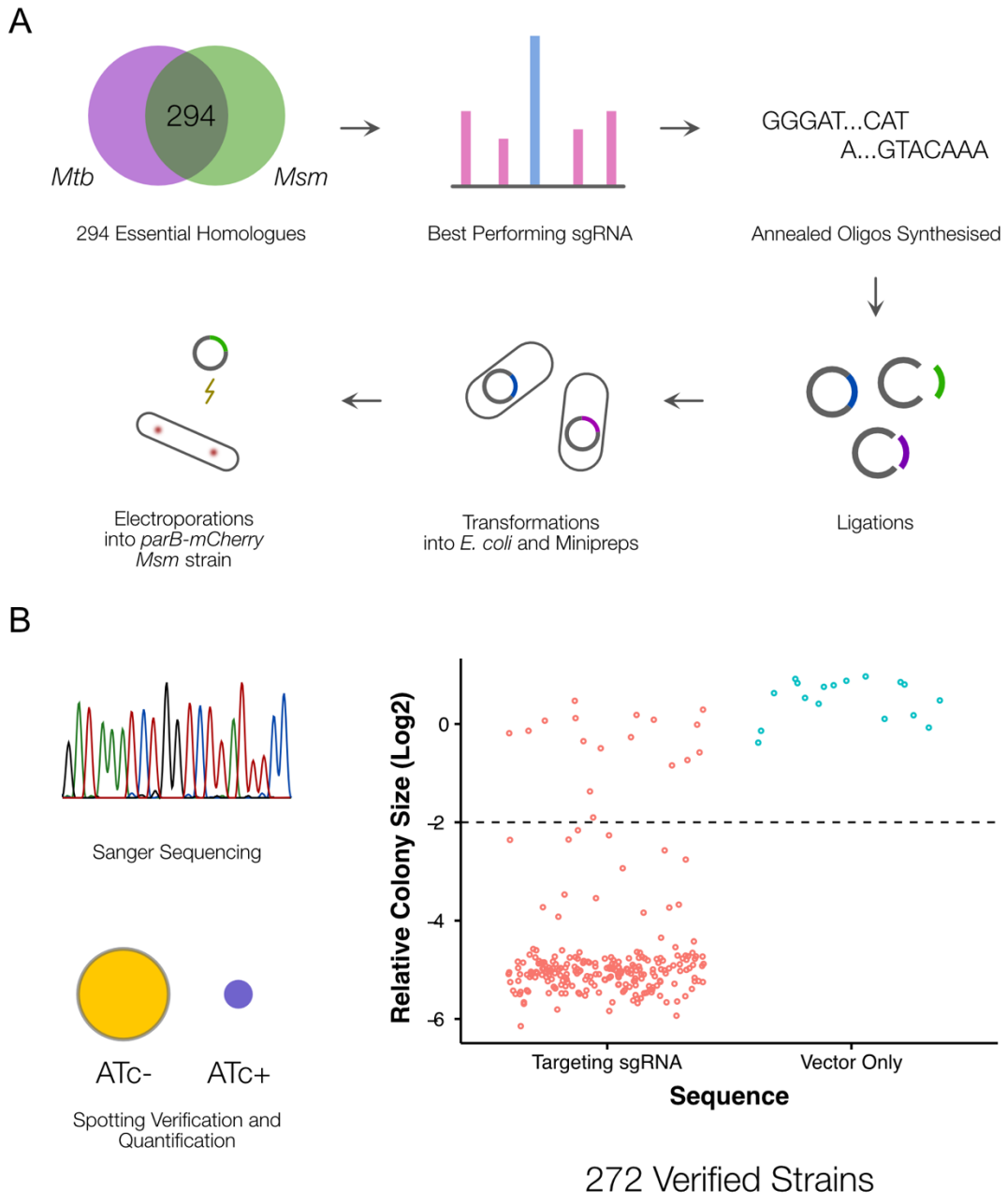


Figure 1. *A CRISPRi library targeting essential M. smegmatis homologs of M. tuberculosis genes.* ***(A)*** *An arrayed CRISPRi library was designed to target 294 essential M. smegmatis genes. For each gene, the highest efficiency sgRNA was identified from a previous pooled CRISPRi-Seq screen (de Wet, Gobe et al. 2018) and synthesized as an annealed oligonucleotide. Cloning was performed at scale, followed by electroporation into an M. smegmatis ParB-mCherry reporter strain (Santi and McKinney 2015).* ***(B)*** *A total of 272 sequence-verified transformants produced a greater than 2-fold decrease in colony size (Kritikos, Banzhaf et al. 2017) when spotted on 7H10 agar containing the inducer, anhydrotetracycline (ATc) compared to the same cells spotted onto solid medium without ATc, confirming ATc-dependent growth inhibition.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A quantitative imaging pipeline for mycobacteria

High-throughput, quantitative imaging has been productively utilized in a number of bacterial systems to rapidly characterize the impact of genetic alterations on cellular function (Peters, Colavin et al. 2016, Liu, Gallay et al. 2017, Campos, Govers et al. 2018) or to determine antimicrobial MOA (Nonejuie, Burkart et al. 2013). Until very recently (Smith, Pullen et al. 2020), and subsequent to the completion of this work, an equivalent approach had not been described for mycobacteria. Therefore, on initiating this study, we chose to investigate the use of imaging to phenotype the CRISPRi mutant library following ATc-dependent transcriptional silencing. This required the development of tools for the extraction of large-scale quantitative data describing mycobacterial cell morphology.

To determine the optimal duration of ATc exposure, we performed time-lapse microscopy of knockdown mutants in three well-characterized essential genes (***Supplementary Video 1 and 2***) – the cell-division mediator, *ftsZ* (Dziadek, Rutherford et al. 2003), the elongasome anchor, *wag31* (Kang, Nyayapathy et al. 2008), and the inosine monophosphate dehydrogenase, *guaB2* (Singh, Donini et al. 2013, Park, Pacitto et al. 2017). Time-lapse microscopy established that fully penetrant phenotypes were manifest at 18 hours post ATc exposure for all three genes. This represents approximately 6-7 doubling times of the wild-type *M. smegmatis* mc^2155 parental strain (Logsdon, Ho et al. 2017), or a dilution factor of pre-existing protein of approximately 64 to 128 (2^6 - 2^7)-fold. For both *wag31* and *ftsZ* mutants, lytic cell death was observable after the 18-hour time-point. While *guaB2* knockdown did not produce as distinctive a visual phenotype as the other two genes, the formation of minicells and a decline in growth rate was clearly evident from 18 hours. Although this test set comprised only three mutants, we were encouraged that, for these essential genes with distinct cellular functions, the 18-hour timepoint appeared adequate to ensure sufficient knockdown and dilution of pre-existing protein to yield a detectable phenotype while maintaining suitable cell numbers for imaging. We therefore proceeded to apply the same experimental conditions in phenotyping the full library, recognizing that the pragmatic utilization of a single, "terminal" endpoint might not be optimal for every mutant in the collection.

Adopting the 18-hour CRISPRi duration, we derived a protocol for medium-throughput imaging of batches of 24 strains (***Figure 2A***). Exponential-phase cultures were treated with ATc overnight for 18 hours prior to spotting onto 2% agarose pads for semi-automated imaging. Phase contrast and fluorescence images of ParB-mCherry localizations were captured for 264 mutant strains. Additionally, an empty vector control strain was treated and repeatedly imaged in the same manner. Moreover, to verify the reproducibility of the imaging workflow, replicate imaging was performed on separate days for 134 strains.

Image analysis tools such as MicrobeJ (Ducret, Quardokus et al. 2016) or Oufti (Paintdakhi, Parry et al. 2016) extract detailed quantitative descriptions of bacterial morphology and protein localization. However, automated image analysis is complicated by out-of-focus objects, or cells that are too close together to separate adequately – a particular concern for mycobacteria which have a proclivity to clump (Cheng, Porter et al. 2014). To overcome these complications, machine-learning-based classifiers have been utilized in *E. coli* for post-processing clean-up (Campos, Govers et al. 2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

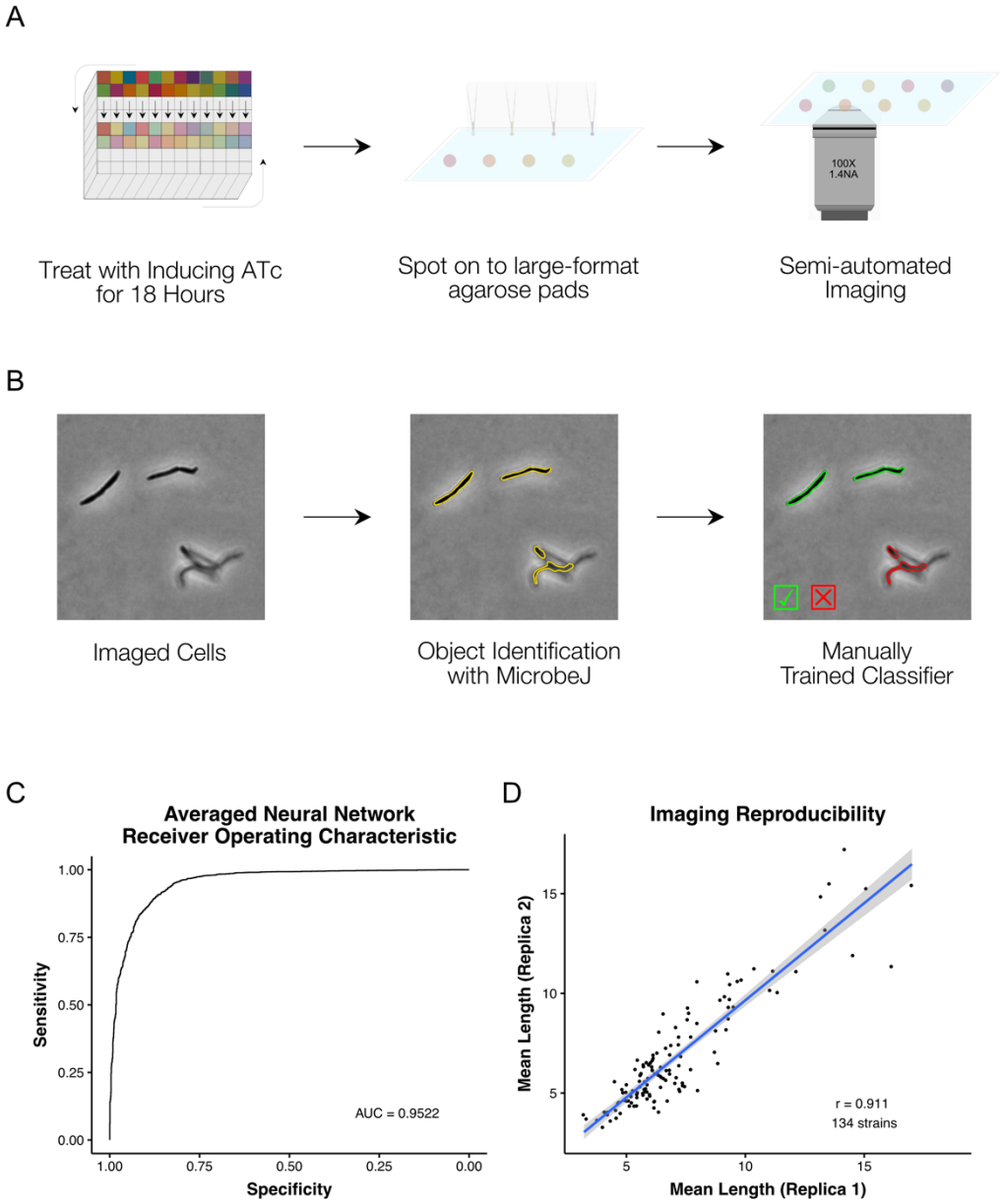


Figure 2*. A high-throughput, quantitative CRISPRi imaging pipeline for mycobacteria. (**A**) Cells were exposed to ATc inducer for 18 hours before spotting onto large-format agarose pads for semi-automated imaging. (**B**) Image processing in MicrobeJ (Ducret, Quardokus et al. 2016) was combined with a manually trained Averaged Neural Network classifier to extract quantitative descriptions of bacterial morphologies and ParB protein localizations for 163559 cells across 265 gene-specific CRISPRi mutants and 16 empty vector controls. (**C**) Classifier performance was measured by Receiver Operating Characteristic (ROC) Area Under the Curve (AUC), returning good performance metrics. (**D**) Mean cell lengths were compared for 134 strains imaged as biological replicates on two separate occasions, and showed high reproducibility (r = 0.911, Pearson's).*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To extract quantitative descriptions of mutant morphologies and ParB-mCherry localizations across our imaging dataset, we utilized the ImageJ package, MicrobeJ (Ducret, Quardokus et al. 2016), with a machine-learning based post-processing clean-up (***Figure 2B***). For this purpose, we trained an Averaged Neural Network with 22,936 manually classified objects, sampled from across our imaging dataset. Receiver Operator Characteristic (ROC) analysis of our classifier produced an Area Under the Curve (AUC) of 0.952 when applied to a reserved test set (***Figure 2C***). Our curated dataset contains morphological descriptions for 163559 cells across all 265 imaged mutant and empty vector control strains, with a mean of 568 cells per mutant (ranging from a minimum of 24 to a maximum of 3861 cells). Reassuringly, when comparing the mean cell lengths of replica-imaged knockdown mutants following ATc induction, the imaging and analytic pipeline showed high reproducibility (***Figure 2D***, r = 0.911, Pearson's).

An atlas of morphological changes consequent on essential gene silencing

There are relatively few published descriptions of the morphological impacts of essential gene silencing (***Supplementary Table 3***). We aimed to use our extracted imaging data to generate a comprehensive repository of morphological changes following CRISPRi-mediated essential gene knockdown in *M. smegmatis*. For each gene, a nearest-neighbor approach was applied to identify the representative morphotype which was in turn incorporated into an atlas of morphological changes consequent on essential gene silencing (***Figure 3***).

Visual inspection of the atlas revealed distinctive alterations in morphology in a large number of knockdown mutants. Moreover, it appeared that silencing certain cellular functions or pathways produced consistent morphological changes. In general, the responses could be broadly categorized into either elongation – occasionally with bulging and/or branching – or shortening, with a variety of alterations in cell width and roundness. For example, components of DNA metabolism (*dna*- prefix genes), the divisome (*fts*- genes), and protein translation (*rpl*- and *rps*- genes) produced cellular elongation. To our surprise, a filamentation response was also observed for components of the histidine biosynthesis pathway (*his*- genes), but not for other amino acid biosynthesis pathways. In contrast, components of cell wall homeostasis – including peptidoglycan (*ponA1*), arabinogalactan (*aftA*) and mycolic acid synthesis (*inhA* and *mmpL3*) – were associated with distinctive shortening of cells, often in combination with bulging.

To quantitate these observations, we adapted an approach developed for *E. coli* (Campos, Govers et al. 2018). Single-cell data were reduced to normalized, gene-level descriptors of morphology. For each mutant, the mean and coefficient of variation (CV) of each morphological feature were extracted and normalized to the distribution of empty vector control samples with a Z-score transformation (***Supplementary Figure 1A***). For a particular mutant and feature, the Z-score reflects how many standard deviations the mutant feature is away from the mean of the control strains. That is, it provides a measure of the extent to which transcriptional inhibition of the specific essential gene impacts a particular morphological characteristic. In total, 206 (78%) strains returned at least one Z-score>3, or <-3.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

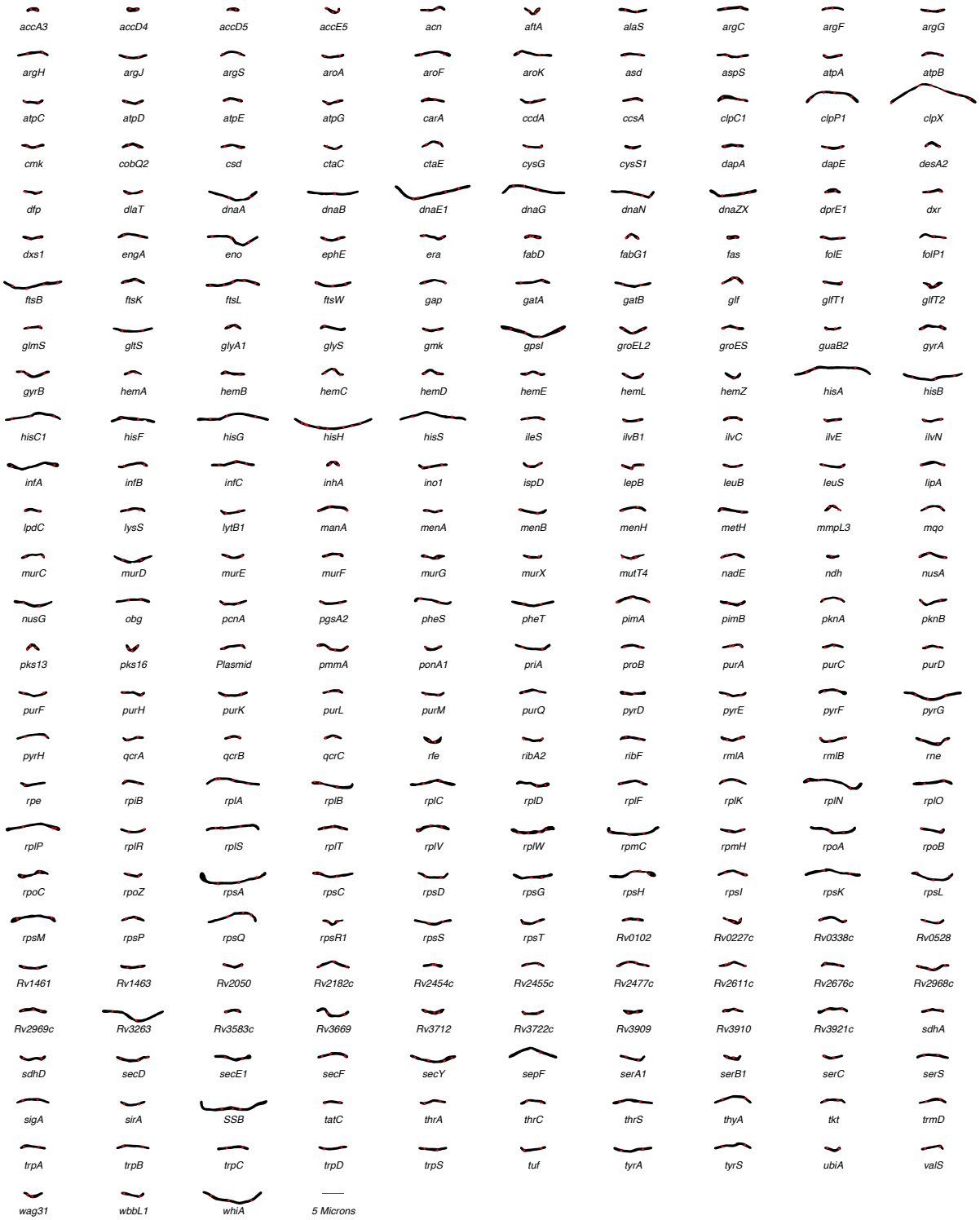


Figure 3. *An atlas of morphological changes consequent on essential gene silencing in M. smegmatis. Following imaging and extraction of quantitative data on cellular morphology, an approximate nearest neighbor approach was utilized to identify a single representative cell which was closest to the mean values of the measured morphological features for each specific (clonal) mutant population. Representative cell contours and ParB localization patterns were extracted from MicrobeJ and utilized to assemble the atlas of morphological changes.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Certain gross changes in morphology were more frequently observed than others (***Supplementary Figure 1B***). For example, many mutants exhibited marked increases in mean curvature, roundness and feret (the calliper diameter). Conversely, few mutants exhibited major alterations in the number of ParB foci (maxima) per cell, suggesting that the number of replication origins (*oriC*s) per cell – and, by implication, ploidy – was maintained even under essential gene knockdown.

To test for associations between morphological changes and functional classes of genes, we extracted Clusters of Orthologous Groups (COG) annotations for all 272 essential genes before testing for statistical enrichment with changes in morphology. For each morphological feature, we tested for COG enrichment of the mutants with Z-scores >3 or<-3. This analysis revealed that statistical enrichments were largely consistent with visual inspection of the cell atlas (***Supplementary Figure 2***). For example, the filamentation phenotype of *fts*-, *rpl*- and *dna*- genes was reflected in the enrichment of Cell Cycle Control (D), Translation (J) and DNA Replication and Repair (L) COGs with increases in length (means and CVs).

The morphotypic landscape of essential gene silencing in mycobacteria

In a complementary approach, we visualized the mean population Z-scores for key morphological features by COG (***Figure 4***). Again, consistent with the cell atlas and enrichment results, specific COGs (***Supplementary Table 3***) produced distinct morphological signatures (or "phenoprints"). Together, these results highlighted the utility of multi-parameter morphological descriptors for characterizing mutants and, moreover, signaled the potential to link genes according to whole-cell phenotype independent of known or predicted (annotated) function.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

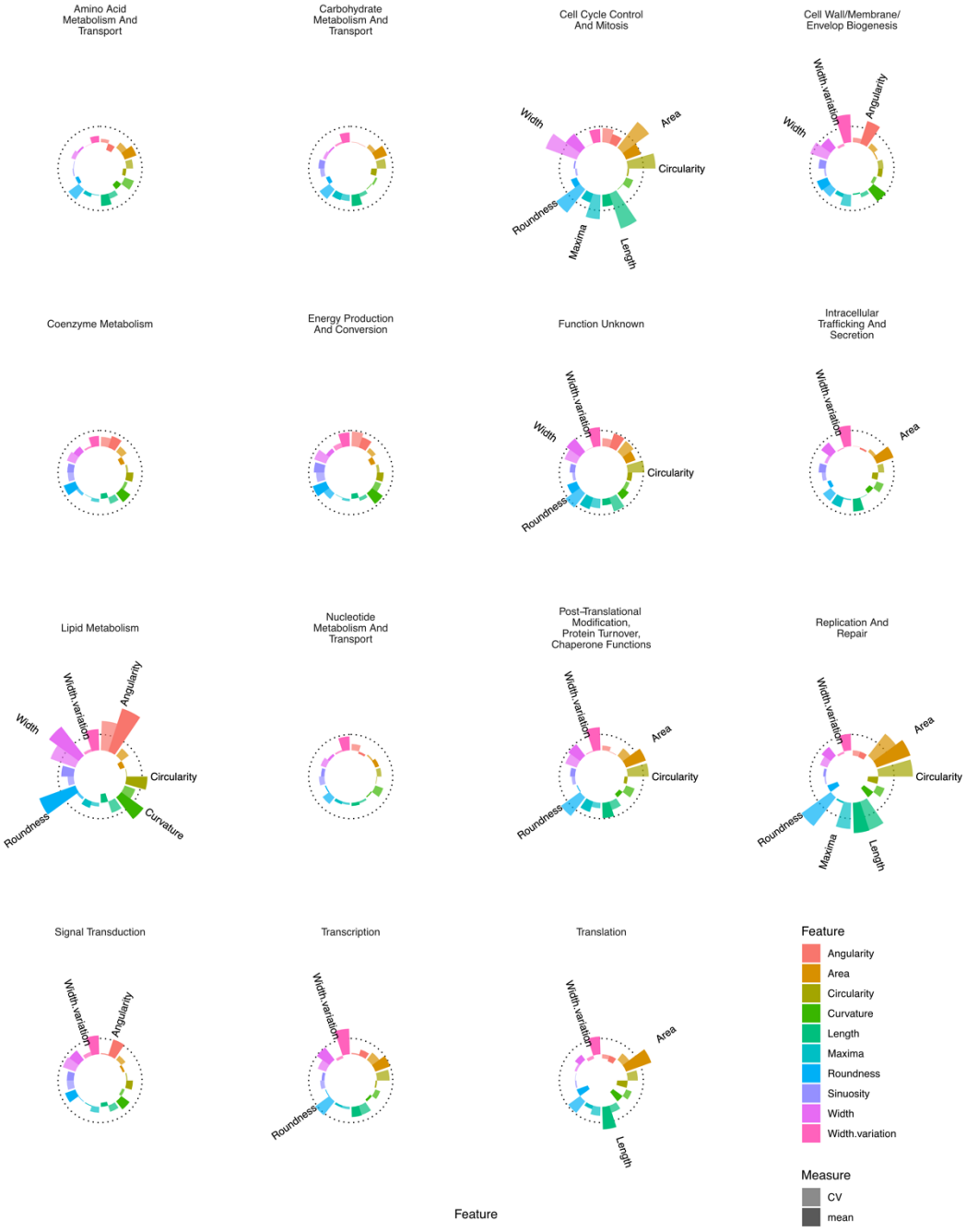


Figure 4. *Essential gene silencing produces distinctive phenoprints of particular genetic function. Mean Z-scores were calculated and visualized for all genes assigned to a particular COG, highlighting distinctive changes in morphology. The dotted line indicates a Z-score of 3. Features with mean Z-scores >3 are labelled.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Based on prior cytological profiling studies in eukaryotic (Caicedo, Cooper et al. 2017) and prokaryotic (Nonejuie, Burkart et al. 2013, Huang 2015, Campos, Govers et al. 2018) systems, we applied an unsupervised learning approach that combined dimensionality reduction with clustering of our multidimensional dataset. For this, we utilized Uniform Manifold Approximation and Projection (UMAP) (Lel, McInnes et al. 2018) and optimized hyper-parameters to produce consistent visual clusters of our control strains in combination with density-based clustering using hdscan (Lel, McInnes et al. 2017) (***Figure 5***).

This process produced a two-dimensional UMAP space consisting of three reproducible clusters (***Figure 5A***). The scale of different morphological features varied across the 2-dimensional space with visible patterns, such as regions of decreased length, or increased width (***Figure 5D***). To explore the three clusters, we tested for COG enrichment and visualized mean feature Z-scores for each cluster (***Figures 5B and C***). Cluster 1, which was enriched for lipid metabolism (I) and cell wall metabolism (M), was associated with increases in width and curvature and decreases in area and aspect ratio. Conversely, Cluster 3, which was associated with translation (J) and DNA replication and repair (L), was characterized by increases in area and length. Cluster 2, which contained the control, non-targeting strains, was enriched for energy production and conversion (C) and nucleotide metabolism (F), highlighting the comparatively small impact that silencing of these genes has on cellular morphology.

To analyze further the spatial arrangements within the three large clusters, and across UMAP space, we adapted a spatial analysis of functional enrichment (SAFE) approach (Baryshnikova 2016, Campos, Govers et al. 2018) to identify sub-clusters of either COG or KEGG enrichment. This yielded a number of distinct sub-clusters that were not originally apparent (***Figures 5E and F***). For example, Cluster 1 contained regions enriched for either Lipid Metabolism (I) or the Cell Wall (M), while Cluster 3 could be subdivided into a region enriched for DNA metabolism (L), ribosomal genes (K) and genes involved in aminoacyl-tRNA biosynthesis (by KEGG enrichment). Using this approach, even Cluster 2 could be subdivided into functionally enriched regions despite the apparent similarity of these strains (on visual inspection) to wild-type morphologies.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

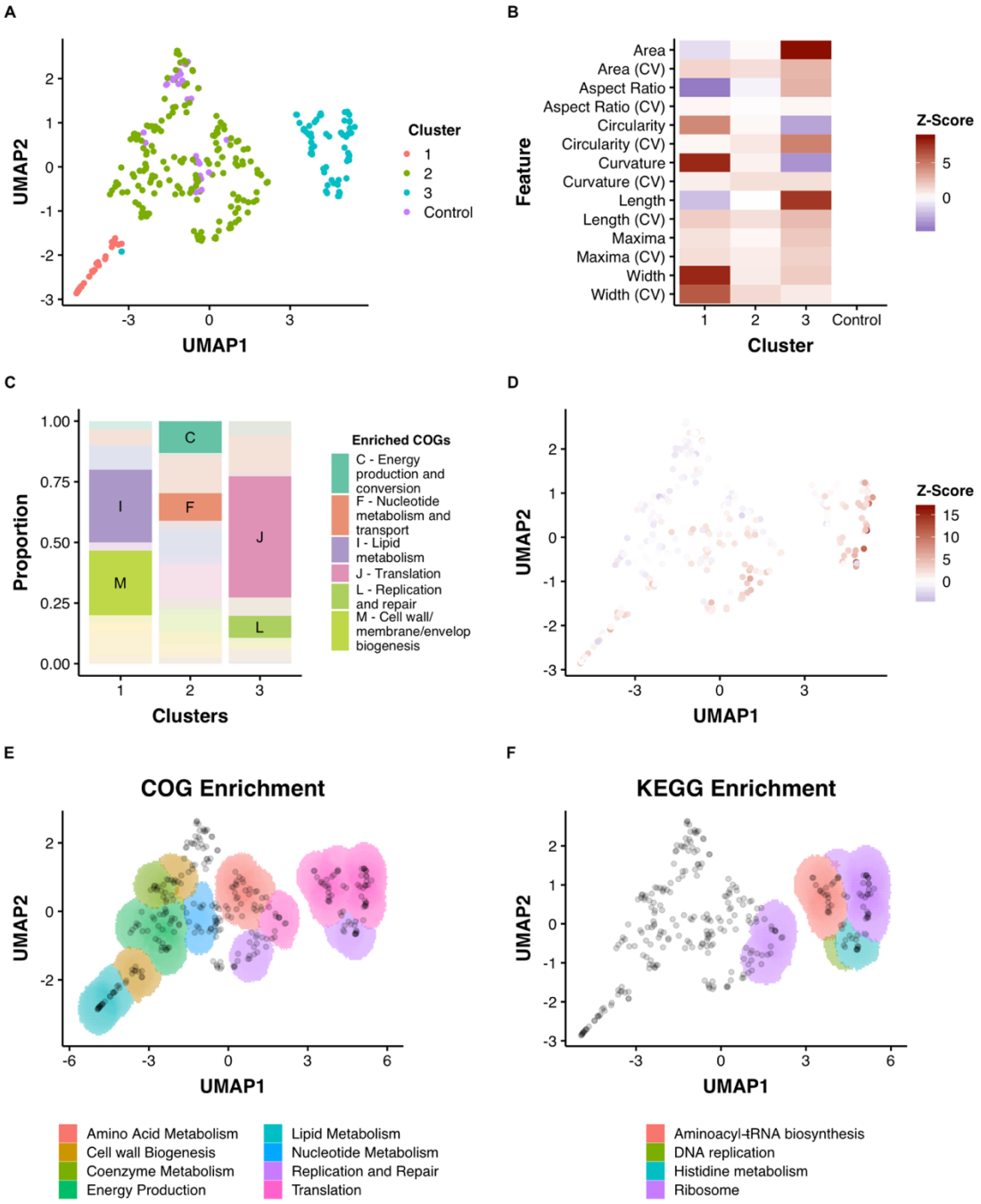


Figure 5. *Dimensionality reduction reveals clusters of morphological change that associate with function. (**A**) A combination of UMAP dimensionality reduction (Lel, McInnes et al. 2018) and hdbscan clustering (Lel, McInnes et al. 2017) was applied to processed morphological data. Parameters were optimized to ensure a uniform clustering of control strains. A consistent seed was used for visualization. (**B**) The mean Z-Scores of each cluster highlights the dominant features determining clustering. (**C**) Clusters are enriched for particular COGs. (**D**) Overlaid Z-score data for mean aspect ratio, length, maxima and width demonstrate heterogeneity within clusters. (**E**) and (**F**) An adapted SAFE (Baryshnikova 2016) approach demonstrates that functionally enriched sub-clusters are present in UMAP space.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

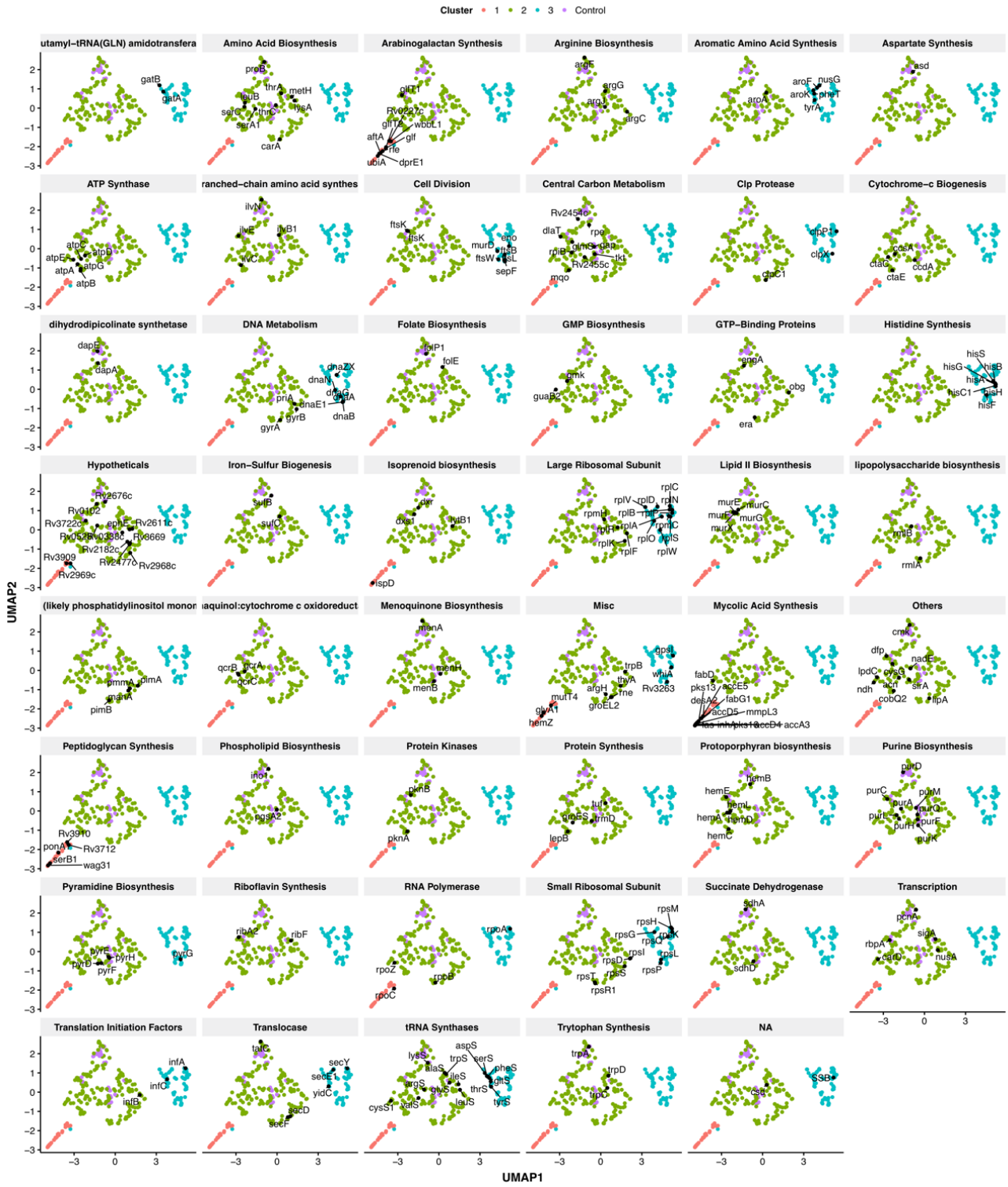


Figure 6*. Functionally related genes are frequently found associated in UMAP space. Genes were manually classified following a review of the literature (**Supplementary Table 3**) and visualized in UMAP space.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As UMAP maintains the essential topological structure of data, we wondered if biological connections would be reflected in the distances between genes in UMAP space. Given the acknowledged propensity of CRISPRi to produce polar effects (Rock, Hopkins et al. 2017), we predicted that these would be reflected in the Euclidean distance between operonic genes. Consistent with our hypothesis, genes in verified operons (Martini, Zhou et al. 2019) were located more proximally than random pairs of genes (***Supplementary Figure 3***). Similarly, predicted protein-protein interaction pairs (Cong, Anishchenko et al. 2019) were found to be non-randomly proximate (***Supplementary Figure 3***).

To demonstrate further the utility of UMAP space for understanding the link between gene function and morphology, we performed more granular analysis of manually annotated functional groupings. Again, consistent clustering of many genes with similar biological functions was observed (***Figure 6***), a striking finding given that the close associations in UMAP space were driven by similarities in cytological characteristics (phenoprints) independent of annotated gene function.

The observation that subtle morphological changes appeared to associate with particular genetic functions implied the potential to utilize the pipeline to elucidate unexpected biological relationships and as an additional tool in MOA studies. Four examples are presented below illustrating the application of this approach in validating (i) putative genetic function, (ii) pathway-specific phenotypes, (iii) macromolecular biosynthetic phenotypes, and (iv) chemical-genetic approaches to antimycobacterial MOA identification.

Elucidating gene function: a mycobacterial restriction-modification system

MSMEG_3213 is annotated as a putative DNA methylase but lacks biological validation. The *M. tuberculosis* homolog, *Rv3263*, has been implicated in bacillary fitness during hypoxia (Shell, Prestwich et al. 2013) but, in contrast to its *M. smegmatis* homolog, is not included in the essential gene set in the pathogen (DeJesus, Gerrick et al. 2017). In our imaging pipeline, *MSMEG_3213* produced a marked filamentation phenotype, clustering with components of DNA replication and repair in UMAP space (***Figures 7A and B***). Utilizing the ParB-mCherry marker, we noted a significant disturbance in normal cell cycle progression, consistent with the interpretation that knockdown of *MSMEG_3213* disrupted DNA replication (***Figure 7C***). Moreover, Nanopore-based RNA-Seq confirmed the specificity of *MSMEG_3213* knockdown, eliminating potential off-target effects on downstream or upstream genes (***Figure 7D***). Notably, the whole-genome transcriptional profile on *MSMEG_3213* knockdown showed marked overlap with the well-characterized DNA damage (SOS) regulon (Davis, Dullaghan et al. 2002, Boshoff, Reed et al. 2003) induced by treatment with genotoxins such as the DNA crosslinking agent, mitomcyin C (MMC, (Müller, Imkamp et al. 2018) (***Figure 7E***). Based on this intriguing observation, we hypothesized that, since depletion of the methylase appeared to trigger a DNA damage response, *MSMEG_3213* might be part of a restriction-modification (R-M) system (Loenen, Dryden et al. 2014). Following an extensive bioinformatic search, this conclusion gained further credence by the observation that REBASE (Roberts, Vincze et al. 2014) lists *MSMEG_3213* as putative DNA methylase of a type II R-M system, with *MSMEG_3214* the associated restriction endonuclease (***Figure 7F***).

We reasoned that, if *MSMEG_3213-MSMEG_3214* does constitute a cryptic R-M system, knockdown of *MSMEG_3214* should render *MSMEG_3213* non-essential. To test this, a dual-targeting CRISPRi construct was generated to enable simultaneous silencing of both genes. Knockdown of *MSMEG_3214* alone produced no growth phenotype whereas dual silencing

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of *MSMEG_3213* and *MSMEG_3214* alleviated the lethal impact of *MSMEG_3213* depletion (***Figure 7G***). Taken together, these observations support the identification of *MSMEG_3213-MSMEG_3214* as a previously undescribed R-M system in *M. smegmatis*.

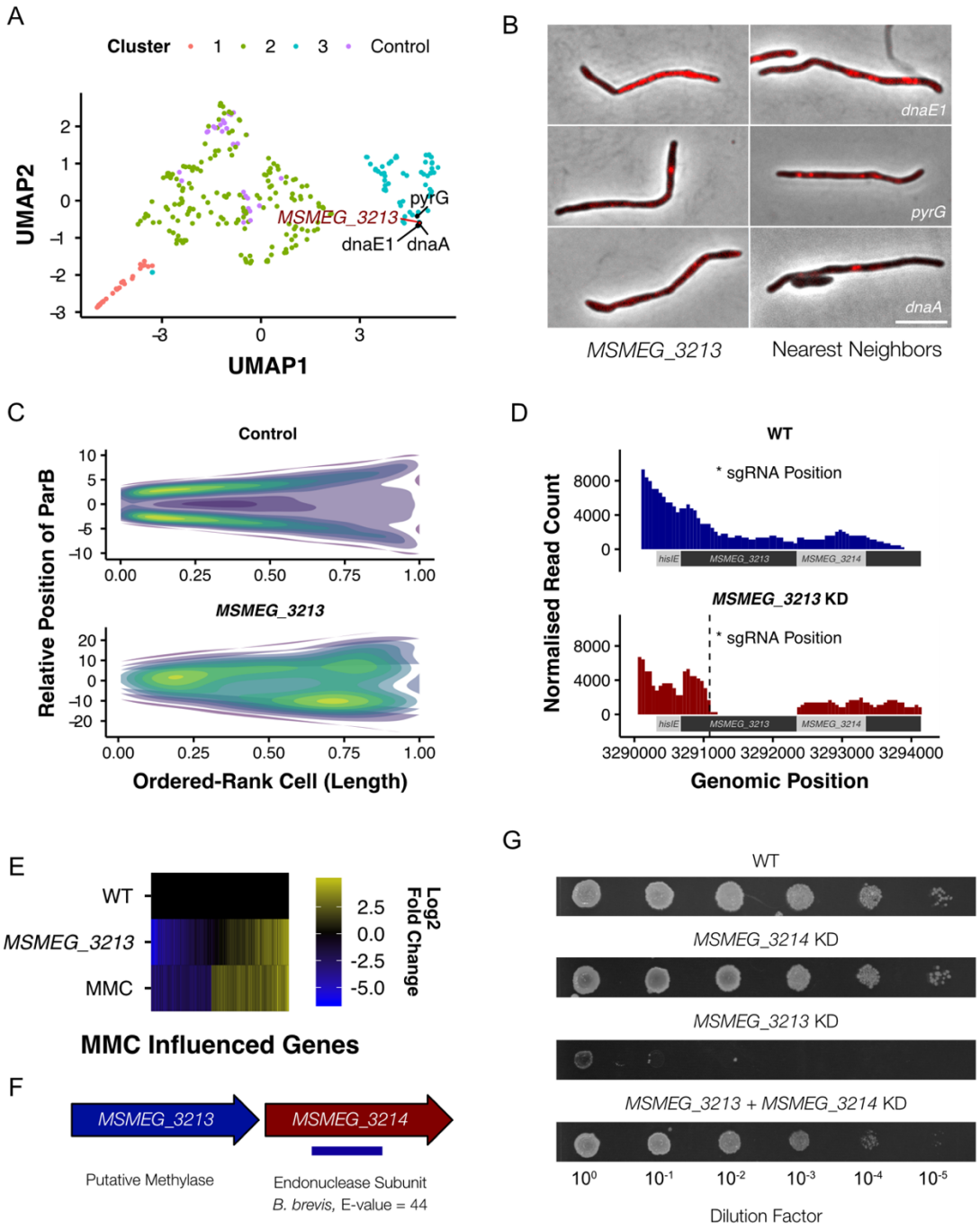


Figure 7*. M. smegmatis encodes a previously undescribed restriction-modification system. (**A**) MSMEG_3213, a putative DNA methylase, is associated with genes involved in DNA replication and repair in UMAP space. (**B**) Knockdown of MSMEG_3213 leads to cellular filamentation. The*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

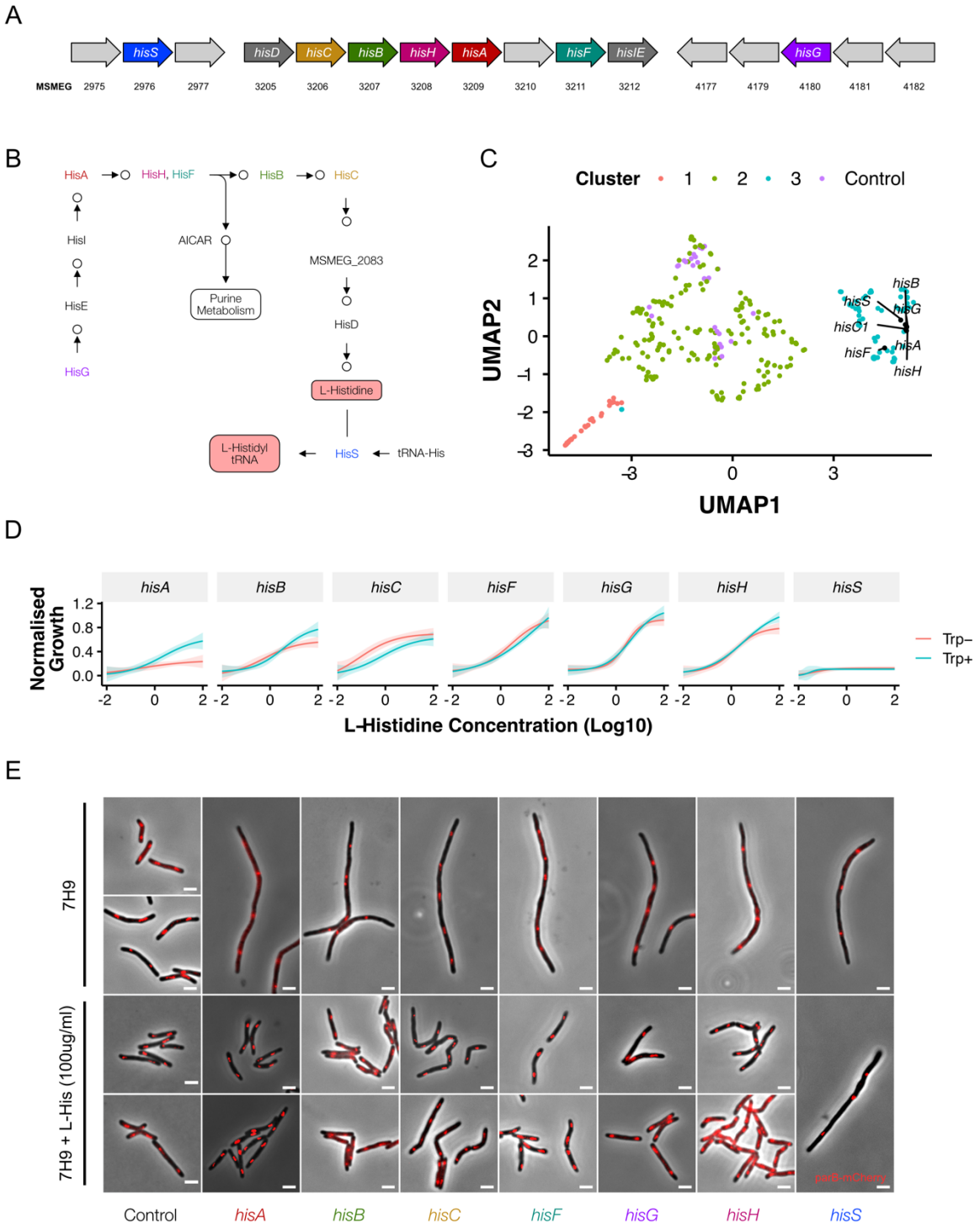
*morphological nearest-neighbors are dnaE1, pyrG and dnaA. (**C**) Consensus heatmaps of ParB-mCherry localization demonstrate that oriC positioning is disrupted in the MSMEG_3213 knockdown mutant. (**D**) Nanopore-based RNA-Seq confirmed that knockdown of MSMEG_3213 was specific to the targeted sgRNA position. (**E**) MSMEG_3213 produces a transcriptional response comparable to treatment with the DNA damaging agent, mitomycin C (MMC). (**F**) MSMEG_3213 is located upstream of MSMEG_3214, a gene with weak homology, according to HHpred (Hildebrand, Remmert et al. 2009), to an endonuclease subunit (REBASE : 3098 BbvCI). (**G**) Lethality of MSMEG_3213 knockdown is suppressed by simultaneous CRISPRi-mediated knockdown of MSMEG_3214.*

Identifying pathway-specific phenotypes: histidine auxotroph filamentation

A feature of the library is that many metabolic pathways are represented by multiple mutants, allowing for pathway-level analyses of metabolic function. *M. smegmatis* possesses a full complement of genes for the biosynthesis of the essential amino acid, L-histidine (***Figure 8A and B***). In the absence of histidine supplementation, the majority of the *his*- prefix genes are predicted to be essential, but most have not been validated individually (Lunardi, Nunes et al. 2013). Surprisingly, CRISPRi-mediated knockdown of histidine biosynthesis genes produced a filamentation phenotype, with tight clustering in UMAP space (***Figure 8C***). The possibility existed that the operonic arrangement of some genes in the pathway (*hisB*, *hisC1*, *hisH*, *hisA*, *hisF, hisI; hisE, hisG*) had confounded interpretation of the CRISPRi phenoprints a result of polar effects. However, our transcriptional data and those of others (Martini, Zhou et al. 2019) appeared to eliminate this possibility, indicating that *hisD, hisC, hisB, hisH and hisF* have nested promoters which were likely to be impervious to CRISPRi-mediated silencing of upstream genes. Moreover, CRISPRi knockdown of *hisG* and *hisS*, which are distally located in the *M. smegmatis* genome, produced the same phenotype, suggesting that filamentation might be a consistent, and previously unreported, consequence of disruptions to histidine biosynthesis.

To test the specificity of this phenotype, we supplemented ATc-containing media with L-histidine. The reversal of the growth inhibitory phenotype verified all *his* knockdowns as histidine auxotrophs, with the notable exceptions of *hisA* and *hisS*. Previous reports indicate that *hisA* likely encodes a bifunctional HisA/TrpF enzyme (Due, Kuper et al. 2011). Consistent with this annotation, growth of the *hisA* CRISPRi mutant was rescued in medium supplemented with tryptophan (***Figure 8D***). Intriguingly, for all his knockdowns other than *hisS* (encoding the histidine tRNA synthetase), histidine supplementation reversed the filamentation phenotype (***Figure 8E***). This was also the case for *hisA*, even when the lack of tryptophan impaired growth. Morphologically, the filaments observed under histidine starvation differ from those which result from depletion of components of the divisome: that is, the *his* mutants display minimal branching. Moreover, staining of peptidoglycan with the D-alanine analogue, NADA (Botella, Yang et al. 2017), failed to resolve any clear septa (***Supplementary Figure 5***). While the mechanism underlying the filamentation response to histidine deficiency remains unclear, these results support the utility of the CRISPRi imaging pipeline to discover, and then validate, pathway-specific phenotypes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



392 ***Figure 8***. *Disrupting components of histidine biosynthesis produces a filamentation response.*
393 *(**A**) Histidine biosynthetic genes are located throughout the chromosome. (**B**) Histidine is*
394 *synthesized through a multi-step pathway (KEGG). (**C**) Cluster analysis demonstrated a*
395 *consistent filamentation phenotype associated with genes involved in histidine synthesis. (**D**)*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Growth could be rescued with histidine supplementation. As HisA has an additional role in tryptophan synthesis, full rescue required simultaneous tryptophan supplementation (50ug/ml). (**E**) The addition of histidine to ATc-containing growth media rescued the filamentation phenotype in all mutants other than hisS.*

Validating macromolecular pathway phenotypes: morphotypic consequences of disrupting mycolic acid biosynthesis

Mycobacteria possess a multilayered cell envelope, the outer mycolic acid layer of which is synthesized via a well-characterized pathway containing a number of established and experimental drug targets (Jankute, Cox et al. 2015) (***Figure 9A***). CRISPRi-mediated inhibition of the components of this pathway produced consistent changes in cellular morphology (***Figure 9B and C***), with the majority of genes implicated in mycolic acid synthesis clustering tightly together in 2-dimensional UMAP space (***Figure 9B***). Notably, the fact that clustering was observed despite the location of these genes throughout the chromosome appeared to eliminate polar effects as a potential confounder.

By superimposing our imaging results and quantitative analyses of morphology on the mycolic acid biosynthetic pathway, we noted highly analogous morphotypes irrespective of the gene targeted for silencing (***Figure 9C and D***). It was also evident that knockdown of components involved in early steps produced similar results to inhibition of the final step, involving the flippase, MmpL3 (Xu, DeJesus et al. 2017, Su, Klenotic et al. 2019). Given the consistency observed following genetic disruption of mycolic acid synthesis, we wondered whether chemical inhibition would produce a comparable morphological change. To test this possibility, we treated cells with isoniazid (INH), a frontline anti-TB drug which inhibits the essential enoyl-ACP reductase, InhA (Timmins and Deretic 2006, Vilchèze and Jacobs 2007). Exposure to 1X Minimum Inhibitory Concentration (MIC_{90}) INH triggered increases in cell width and morphological alterations which were closely similar to those produced by genetic silencing of *inhA* (***Figure 9C and D*****).** This was an important observation since it suggested the possibility that, by analogy with bacterial cytological profiling (Nonejuie, Burkart et al. 2013, Nonejuie, Trial et al. 2016), drug-induced morphological changes could be utilized to inform antimycobacterial MOA more broadly; that is, by exploiting the CRISPRi imaging database of correlated genotype-phenotype connections with chemical (drug-induced) phenotype.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

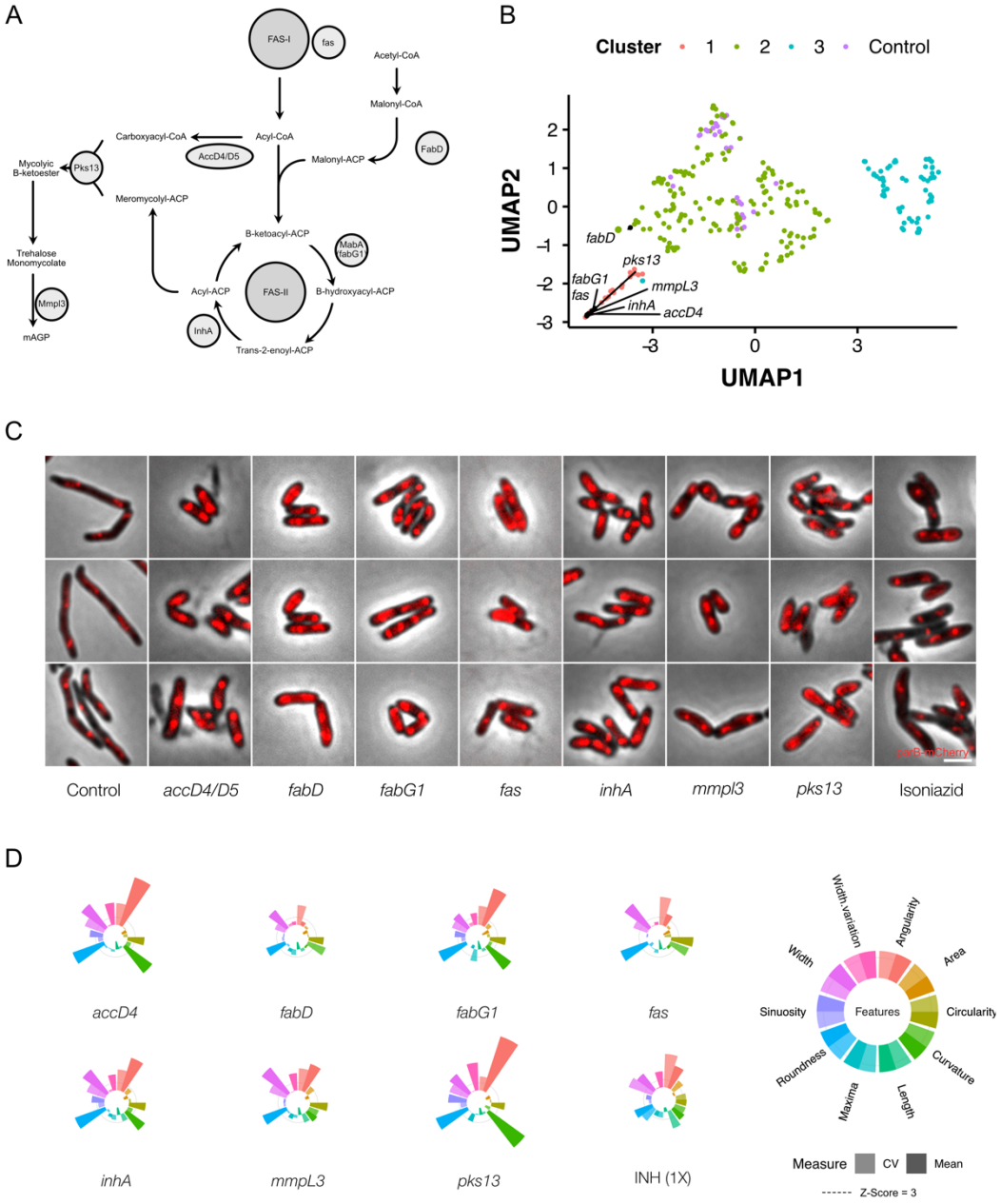


Figure 9. *Genetic or chemical inhibition of mycolic acid biosynthesis produces a distinctive change in bacillary morphology. (**A**) The multi-step pathway of mycolic acid biosynthesis. (**B**) Components of mycolic acid synthesis, visualized in UMAP space, cluster closely and are morphologically similar (**C**) with overlapping phenoprints (**D**). Exposure to the frontline anti-TB drug, isoniazid (INH), at 1X MIC produces a similar morphological change and an analogous phenoprint to silencing of inhA, the validated target of INH (Vilchèze and Jacobs 2007).*

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CRISPRi chemical-genetics for antimycobacterial MOA identification

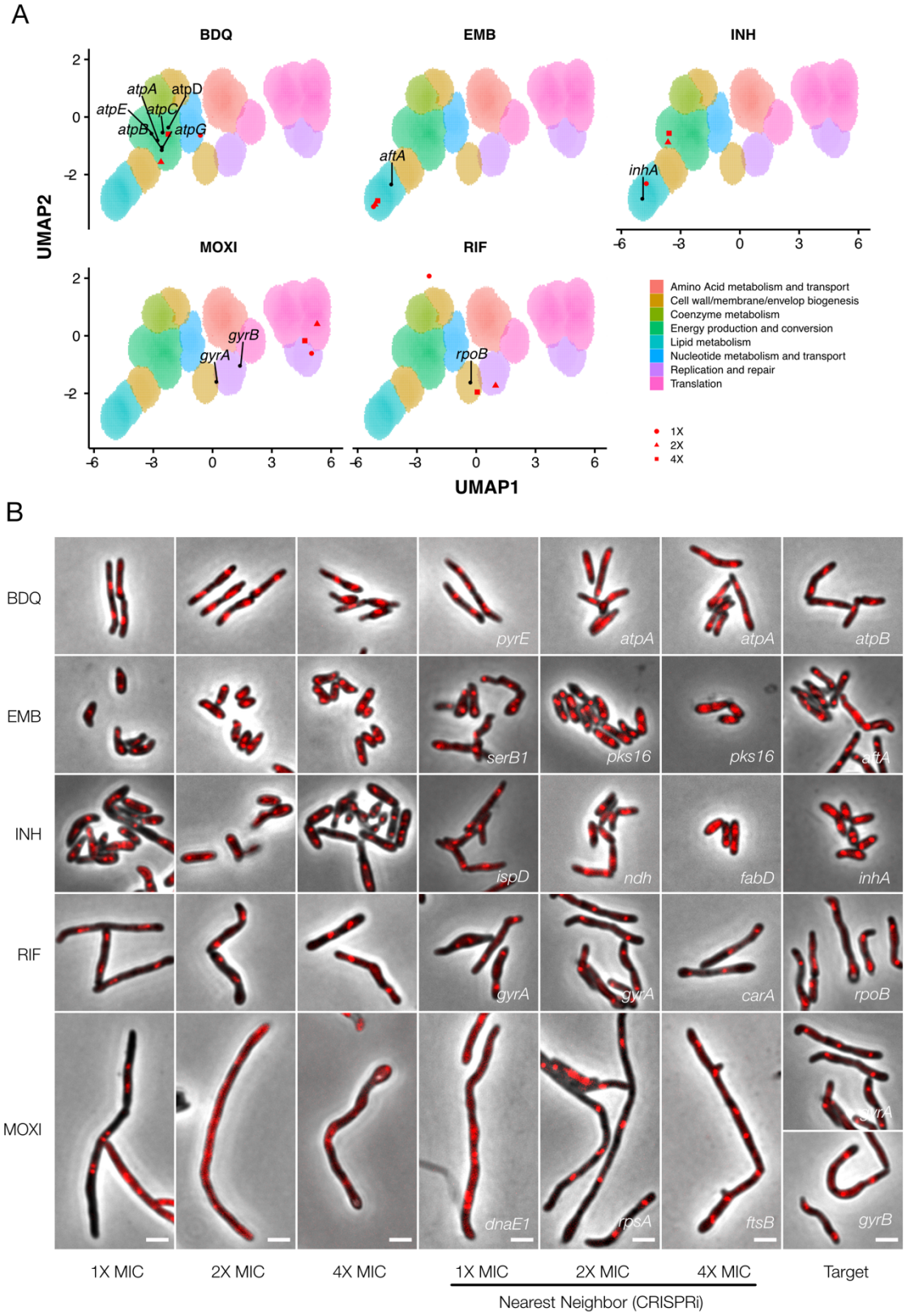
To determine whether the CRISPRi dataset could be used to determine antimycobacterial MOA, we investigated a panel of compounds from different classes that inhibit various essential mycobacterial processes (***Figure 10A***). For proof-of-concept assays, the initial selection was limited to 5 clinically used anti-TB drugs: the first-line agents, INH and ethambutol (EMB), which disrupt cell wall biosynthesis (Abrahams and Besra 2018); RIF, which binds the DNA-dependent RNA polymerase subunit, RpoB (Koch, Mizrahi et al. 2014); the second-line fluoroquinolone, moxifloxacin (MOXI), which inhibits DNA gyrase (Kumar, Madhumathi et al. 2014); and bedaquiline (BDQ), an inhibitor of mycobacterial ATP biosynthesis recently approved for the treatment of MDR-TB (Sarathy, Gruber et al. 2019). Thereafter, the panel was expanded to include an additional 12 known and experimental antimycobacterial compounds (***Supplementary Figure 6***).

Morphotypes observed following drug treatment were often positioned in regions of UMAP space consistent with their known MOAs. For example, cells exposed to BDQ at 2X and 4X MIC clustered tightly with components of the mycobacterial ATP synthase and mapped to a region enriched for genes involved in energy metabolism (***Figure 10A***); the nearest neighbor for both 2X and 4X BDQ-treated cells was *atpA* (***Figure 10B***). The cell-wall targeting compounds similarly produced closely correlating profiles: EMB-treated cells clustered tightly together, irrespective of applied concentration, and were positioned in lipid metabolism-enriched UMAP space. The nearest neighbor at 2X and 4X MIC was *pks16*, a gene involved in mycolic acid synthesis; moreover, consistent with inhibition of arabinogalactan synthesis, *aftA*, *ubiA*, *dprE1*, *glfT2* were proximal to the EMB-exposed cells at all three applied drug concentrations. INH-treated cells were situated in lipid metabolism-enriched space at 1X MIC but, at higher concentrations, fell in a region associated with energy production and conversion. Notably, however, the nearest neighbor at 4X MIC was *fabD*, a component of mycolic acid synthesis. In contrast, compounds targeting peptidoglycan, including vancomycin and D-cycloserine, did not yield definitive profiles (***Supplementary Figure 6***). Manual inspection of the data indicated this was probably due to the induction of lysis, highlighting a limitation inherent in applying a single, 18-hour endpoint for analysis of all drug treatments.

Compounds targeting DNA replication generally produced filamentation, and clustered accordingly. Notably, MOXI-treated cells did not co-localize with *gyrA* and *gyrB* knockdowns (the DNA gyrase subunits, and targets of MOXI)in UMAP space and were instead associated with DNA metabolism, an observation common to all fluoroquinolones. In contrast, novobiocin (NVB), a gyrase B inhibitor (Chatterji, Unniraman et al. 2001), did not trigger filamentation, instead positioning closely with *gyrA* and *gyrB* genetic knockdowns. Like the fluoroquinolones, the experimental agents nargenicin (putative DnaE1 inhibitor; (Painter, Adam et al. 2015)) and griselimycin (which binds to the β-clamp, DnaN; (Kling, Lukat et al. 2015)) were associated with DNA metabolism pathways, in this case mapping much closer to their known or predicted targets.

While compounds disrupting DNA replication and cell envelope biogenesis were readily associated with a target class, this was not universally true for drugs with other MOAs. At 4X MIC, RIF was closely associated with *rpoB*. However, at the two lower concentrations, RIF-treated cells were positioned at dramatically different locations in UMAP space.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



478

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 10. *Phenoprinting can inform antimycobacterial MOA. Cells were exposed to varying supra-MIC concentrations (1X, 2X, 4X MIC) of the selected antimycobacterial compounds for 18 hours, imaged, and analyzed using the pipeline developed for CRISPRi imaging. The resulting profiles were visualized in CRISPRi-generated UMAP space (**A**), and the morphological nearest-neighbor identified (**B**). Known targets were simultaneously visualized for comparative purposes. BDQ, bedaquiline; EMB, ethambutol; INH, isoniazid; RIF, rifampicin; MOXI, moxifloxacin.*

It was noticeable, too, that compounds targeting mycobacterial protein synthesis – for example, linezolid (Leach, Brickner et al. 2011), streptomycin or kanamycin (Riska, Jacobs et al. 2000) – similarly failed to produce responses analogous to their cognate gene knockdowns. Like RNA polymerase, the protein translation machinery operates as a large, multi-protein complex, perhaps exposing a deficiency in this approach when applied to the macromolecular machines responsible for transcription and translation. Overall, however, the results supported the utility of CRISPRi-enabled cytological profiling to inform compound MOA, especially for agents targeting DNA replication, cell wall biosynthesis, and energy metabolism.

DISCUSSION

We have described the construction and validation of an arrayed library of CRISPRi mutants targeting essential *M. smegmatis* homologs of *M. tuberculosis* genes. By coupling the library with a high-throughput imaging pipeline, we confirmed the utility of quantitative methodologies to describe morphological changes and, moreover, established the capacity to harness morphotypic information to construct a database of cytological phenotypes, or mycobacterial phenoprints. Then, in demonstrating the utility of the resulting phenoprint atlas, we described examples illustrating the assignment of function to a putative methylase of previously unknown function, the identification of distinct morphotypes associated with biosynthetic (histidine) and macromolecular (mycolic acid) pathway disruptions, and the potential to exploit CRISPRi morphotypes as a chemical-genetic tool to illuminate antimycobacterial MOA determination.

Advances in imaging technologies (Camsund, Lawson et al. 2019) and sample preparation (Shi, Colavin et al. 2017), together with the availability of software tools such as MicrobeJ (Ducret, Quardokus et al. 2016), have enabled the application of imaging in bacterial systems biology (Ducret, Quardokus et al. 2016). Our pipeline was optimized for medium-throughput imaging of mycobacteria; however, increasing the throughput of imaging, for example with the Strain Library Imaging Protocol (Shi, Colavin et al. 2017), would be advantageous for larger collections of mutants. An immediate priority is to expand the library to include all 423 predicted essential *M. smegmatis* genes from our pooled screening data (de Wet, Gobe et al. 2018). There appears to be significant value, too, in establishing an equivalent library in *M. tuberculosis*, a challenging undertaking but one which would benefit from the lessons learnt from this study. In this context, we note with interest key recent advances in mycobacterial cytological profiling (Smith, Pullen et al. 2020) which suggest the potential to combine these complementary approaches in applying CRISPRi-imaging to *M. tuberculosis*.

In developing our analytical pipeline, we relied heavily on the pioneering work of Jacobs-Wagner and colleagues (Campos, Govers et al. 2018). In particular, we adapted a

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

machine-learning based cleanup that leverages the power of MicrobeJ but adds a *Mycobacterium*-specific post-processing step. This classifier should have downstream utility for future imaging studies in *M. smegmatis* but will require retraining if utilized for *M. tuberculosis*. For analysis, we applied an unsupervised learning approach that combined UMAP dimensionality reduction with density-based clustering of our multidimensional dataset. UMAP is a generally applicable algorithm, but has enjoyed especially rapid uptake for the visualization of single-cell sequencing results (Becht, McInnes et al. 2018). In our pipeline, we also used UMAP for data visualization and as a preprocessing step prior to the application of density-based clustering approaches, analogous to the use of t-SNE by Jacobs-Wagner and colleagues (Campos, Govers et al. 2018). While non-linear dimensionality reduction techniques must be used with caution, our extensive validation supports the utility of the approach. Additionally, our dataset should be amenable to further analysis using alternative analytic approaches, including deep-learning (Chamier, Laine et al. 2019).

Using data obtained from our CRISPRi screen, we generated a quantitative atlas of morphological changes associated with essential gene silencing in *M. smegmatis*. For ~70% of the strains in our library, no prior morphological data were available at the outset of the study. Moreover, approximately 40% of the genes and their protein products had not been characterized at all, either functionally (biochemically or microbiologically) or structurally. It was notable, therefore, that almost 78% of the essential genes assayed here resulted in at least one dramatically aberrant morphological feature on CRISPRi knockdown. This number exceeds the ~60% of *B. subtilis* mutants that produce clear terminal phenotypes (Peters, Colavin et al. 2016) and may reflect increased sensitivity of our analytic approach and/or the differential compositions of the respective libraries. However, given that only 20% of the non-essential *E. coli* Keio Collection displays significant morphological changes (Campos, Govers et al. 2018), it is evident that silencing of essential genes is more likely to disrupt core processes affecting maintenance of general bacillary morphology.

A number of CRISPRi mutants produced morphological alterations that were unexpected. For example, while filamentation is a well-described phenomenon consequent on interference in cell-division (Dziadek, Rutherford et al. 2003, Campos, Govers et al. 2018, Wu, Zhang et al. 2018), DNA replication (Greendyke, Rajagopalan et al. 2002, Justice, Hunstad et al. 2008) and depletion of ClpP protease components (Li, Zeng et al. 2010), it is not typically associated with genes involved in protein synthesis or export. Nevertheless, we recorded filamentation responses following silencing of ribosomal subunits, the Sec translocase, and histidine biosynthesis. These phenotypes contrast with *B. subtilis* where filamentation is not observed on silencing of the corresponding genes (Peters, Colavin et al. 2016). In *Salmonella typhimurium* (Murray and Hartman 1972) and *E. coli* (Frandsen and D'Ari 1993), mutants overexpressing *hisH* and *hisF* are known to produce filaments, possibly through a scarcity of PBP3 substrates required for cell division (Cano, Mouslim et al. 1998). To our knowledge, though, filamentation of histidine auxotrophs has not been reported in any bacterial species, including mycobacteria. The mechanism underlying the filamentation phenotype is unclear, though potentially attributable to interference in cell division through an unknown mechanism. Moreover, the fact that this phenotype manifests on knockdown of the L-histidine-tRNA ligase, *hisS*, suggests the involvement of a regulatory mechanism dependent on the presence of his-tRNAs (Raina and Ibba 2014). However, further work is required to explain this observation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

While interfering in cell envelope biogenesis might be expected to cause aberrations in morphology, the increases in mycobacterial cell width observed on silencing essential steps in mycolic acid synthesis have limited precedent in the literature. One example is the association of temperature-sensitive mutants of *inhA* with analogous phenotypes (Vilchèze, Morbidoni et al. 2000). The phenotypic consistency across the pathway was nevertheless surprising; again, however, the precise mechanism remains elusive. It is possible that the cell wall synthetic machinery acts together in a spatiotemporally related complex that requires the presence of all proteins for function. It might be instructive, for example, that fluorescently-tagged versions of InhA and MabA co-localize (Vilchèze, Morbidoni et al. 2000) and may be present in the same specialized membrane domain (Hayashi, Richardson et al. 2018). Knockdown of individual components of the pathway might therefore impact multiple protein-protein interactions, collapsing numerous components of cell wall synthesis to produce the same phenotypic outcome. Alternatively, it is possible that biosynthetic precursors accumulate, producing consistent alterations in cell structure. A further possibility is that mycolic acids might play a more important structural role in the maintenance of mycobacterial cell width – an intriguing prospect considering the absence of the prokaryotic actin homolog, MreB, in mycobacteria (Singh, Ghosh et al. 2010). Further work is needed to resolve these possibilities and should benefit from the collection of CRISPRi mutants described here.

In general, the CRISPRi-induced morphological changes appeared consistent within pathways and following exposure to the various anti-mycobacterial drugs. This was especially true of genes involved in DNA replication, cell division, cell wall metabolism, and energy metabolism – a property which allowed us to leverage the platform to assign functional predictions for a number of genes, and to infer antimicrobial MOA. While this approach cannot claim single-gene sensitivity, morphological profiling does appear to offer a rapid means of preliminary gene function assignment or compound MOA. The discovery – triggered by morphological profiling and validated by transcriptomic profiling and combinatorial CRISPRi – of a previously unreported Type II R-M system in *M. smegmatis* supports this contention. The biological function of this system is currently unclear; however, one practical implication is that, if involved as expected in intrinsic phage defense, it might inadvertently eliminate phages which could be otherwise useful for clinically relevant mycobacteria such as *M. tuberculosis* or *M. abscessus* (Dedrick, Guerrero-Bustamante et al. 2019). It is tempting to consider, too, if the differential complement of R-M systems among mycobacteria might determine the apparent species-selectivity of some mycobacteriophages – for example, DS6A (Mayer, Jain et al. 2016) – and whether the MSMEG_3213 methylase, which is conserved in *M. tuberculosis*, functions in cell-cycle regulation (Wion and Casadesús 2006) in addition to its role in self-defense.

During the course of this work, we noted a number of additional genes whose putative functional annotation was supported by our results (***Supplementary Figure 4***). For example, the *C. glumatimicum* homolog of *MSMEG_0317* is required for lipomannan maturation and lipoarabinomannan (LM/LAM) synthesis (Cashmore, Klatt et al. 2017). We observed that *MSMEG_0317* clustered closely with genes involved in arabinogalactan synthesis and localized to the cell wall (***Supplementary Figure 4B***). It was pleasing, therefore, when a separate study emerged suggesting that *MSMEG_0317* was involved in the transport of LM/LAM (Gupta, Rahlwes et al. 2019). In another example, the transcriptional regulator, *whiA*, which is involved in sporulation in *Streptomyces* (Bush, Bibb et al. 2013),

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

appears to play a key role in the mycobacterial cell cycle, clustering with other components of cell division (***Supplementary Figure 4C***). Furthermore, *MSMEG_6276* – a putative mur ligase – clusters closely with the peptidoglycan synthesis protein, *mviN* (Gee, Papavinasasundaram et al. 2012), and exhibits strong homology to *murT/gatD* from *S. pneumoniae* (Morlot, Straume et al. 2018) (***Supplementary Figure 4D***). In combination, these additional examples support the utility of image-based profiling for informing or validating hypothetical or predicted gene function – particularly those involved in cell wall metabolism or DNA metabolism and cell cycle regulation.

Cytological profiling has been extensively exploited to determine antimicrobial MOA (Nonejuie, Burkart et al. 2013, Huang 2015, Nonejuie, Trial et al. 2016). However, to our knowledge, existing approaches have consistently utilized whole-cell drug treatment to create profiles of known mechanisms. In this work, we instead applied a genetic approach, generating a database of phenoprints based on essential gene knockdown which were then compared with profiles produced by drug exposure of whole cells. Notwithstanding the fundamental difference between genetic (transcriptional) knockdown and pharmacological (small molecule) inhibition, we observed strong overlap in the resulting phenotypes, especially for compounds targeting energy metabolism (BDQ), the mycolic acid (INH) or arabinogalactan (EMB) components of the cell wall, as well as DNA metabolism (fluoroquinolones and griselimycin). In some cases, we noted, too, that different concentrations of compound could produce disparate morphological profiles, suggesting a possible dependence of MOA on drug concentration (Bernier and Surette 2013). Furthermore, not all compounds produced defining phenotypes. For example, compounds targeting peptidogylcan synthesis (DCS and VAN) led to rapid lysis, while aminoglycosides did not cluster with the majority of ribosomal knockdowns. While these exceptions expose limitations in cytological profiling (discussed below), the approach nevertheless appears useful as rapid pre-screen of compound MOA in drug discovery pipelines, prior to more in-depth investigation.

It was notable that MOXI-treated cells did not sit closely with *gyrA* and *gyrB* knockdowns and were instead associated with DNA metabolism, an observation common to other fluoroquinolones. While this separation in UMAP space might be deemed surprising, two factors are salient to the interpretation of this result: firstly, the role of mycobacterial DNA gyrase in the removal of torsional stress and the maintenance of template topology during RNA transcription (Ahmed, Sala et al. 2017) is consistent with the dominant effect of *gyrA* and *gyrB* knockdown manifesting as more similar to RNA polymerase inhibition (here, the close association with *rpoB* in the UMAP analysis is potentially telling); and, secondly, the filamentation observed on fluoroquinolone exposure is SOS-dependent (Drlica, Malik et al. 2008) and involves the formation of lethal double-strand DNA breaks following drug-mediated trapping of the DNA-gyrase complex – an effect which ATc-induced CRISPRi knockdown does not produce. Supporting this interpretation, the gyrase B inhibitor, NVB, did not trigger filamentation, and was instead closely associated with *gyrA* and *gyrB* knockdowns, a key observation given previous work noting the lack of SOS induction in NVB-treated cells (Boshoff, Myers et al. 2004). A question which commonly arises is how the apparently discrepant *gyrA*/*gyrB* and moxifloxacin phenoprints can be reconciled with recent data from a pioneering high-throughput chemical-genetic screen which reported the identification of novel gyrase inhibitors by utilizing a library which included *gyrA* and *gyrB* hypomorphs (Johnson, LaVerriere et al. 2019). In addressing this dilemma, it is important to

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

remember that hypomorphs provide a powerful means to identify hypersusceptibilities consequent on bacillary exposure to chemicals which inhibit the same (cognate) gene or pathway targeted by the knockdown system (*i.e.*, the genetic and antibiotic targets are identical); in contrast, whole-cell morphotypes report on the physiological response to (or manifestation of) lethal stress induced by essential gene silencing. Importantly, a key element distinguishing antibiotic-mediated target inhibition or corruption from transcript depletion is time: antibiotic treatment inhibits (or corrupts) active processes, triggering intracellular catastrophe (in the case of MOXI treatment, an SOS response), whereas gene silencing is much more gradual, or ordered, avoiding the impact of instantaneous loss of function.

As for all large-scale assays, the approach detailed here inevitably carries inherent limitations which must be borne in mind when analyzing and applying the data: (i) Although motivated primarily by pragmatism, the decision to apply an 18-hour endpoint for all analyses might inadvertently have biased the "strongest" phenotypes to genes involved in DNA replication, cell division, and cell wall metabolism. One obvious complication pertains to the drug MOA analyses, and is evident in the striking frequency with which the data points for 1X, 2X and 4X concentrations map differently in UMAP space for different compounds. Intuitively, increasing drug concentration might be expected to have two principal effects – increasing the rate of cell death while retaining the primary MOA (put simply, cells treated at 4X MIC will be further along the "path to death" at 18 hours than those treated at 1X MIC) or increasing the likelihood of hitting secondary targets. This issue has been noted previously in transcriptional profiling of the *M. tuberculosis* drug-exposure response (Boshoff, Myers et al. 2004) in which optimizing the drug concentration and exposure time were critical in avoiding convergence on a common stress pathway, obscuring the informative, drug-specific transcriptional signature. Similarly, the propensity for the rate of gene-product depletion to impact qualitatively the terminal phenotype has been elegantly demonstrated in *B. subtilis* (Peters, Colavin et al. 2016). Notwithstanding the significant resources required, future work might therefore envisage the use of time-lapse imaging of all mutants to identify optimal gene-specific knockdown durations. (ii) This proof-of-concept study employed only aerobically grow mycobacteria cultivated at 37°C in media comprising Middlebrook 7H9 broth base, with glycerol as primary carbon source. There is ample precedent in the literature indicating that gene essentialities, metabolic vulnerabilities and drug susceptibilities can alter as a function of growth conditions; therefore, the phenotypes presented here are necessarily condition-specific, and could benefit significantly from expanding these analyses to other, "disease-relevant" culture systems. By extension, there is also the potential to perform equivalent analyses in different genetic backgrounds, avoiding the well-described pitfalls of "wild-type" laboratory strains. (iii) Future iterations should also aim to resolve the large cluster (Cluster 2) of genes exhibiting minimal morphological effects on knockdown into more informative sub-clusters, perhaps through the incorporation of additional features and/ or by applying alternative fluorescent reporters (possibly in addition to the ParB reporter). Some approaches to consider include the use of sentinel genes and transcription factors involved in different aspects of macromolecular biosynthesis and/ or metabolic remodeling (Naran, Moosa et al. 2016, Boot, Commandeur et al. 2018). In addition to the potential for improved MOA delineation, such refinements are necessary to address the large number of hypothetical proteins whose functions were not significantly informed by the current morphological profiling algorithm.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In conclusion, we are hopeful that the library of *M. smegmatis* mutants – and the associated online database – will offer a potentially valuable resource for the mycobacterial research community, particularly given the utilization of the ParB-mCherry reporter background which, to our knowledge, is unique in coupling whole-cell morphological information with a measure of replicative status and ploidy. While a large collection of inducible protein-degradation mutants was recently described in *M. tuberculosis* (Johnson, LaVerriere et al. 2019), there is not yet an equivalent in the faster-growing model mycobacterium, *M. smegmatis*. Moreover, in addition to the validated strains, the plasmids used to establish our library should enable relatively simple construction of either single mutants or mutant libraries in alternative genetic backgrounds and reporter strains.

METHODS

Construction of arrayed library

We limited our initial selection to 294 genes. In addition to the 288 genes found universally essential in both Tn-seq and CRISPRi-Seq (de Wet, Gobe et al. 2018), we added 6 components of the ATP synthase which were found to be non-essential in Tn-seq owing to gene duplication in *M. smegmatis* (Dragset, Ioerger et al. 2019).

While CRISPRi is a reliable technology for producing knockdown, it is subject to variation in knockdown efficacy, depending on the selected guide RNA. Moreover, when targeting a gene with CRISPRi, the efficacy of a particular guide is not known *de novo*, and activity prediction algorithms are still in their infancy, particularly for bacteria (Wang, Guo et al. 2017). As a result, it is generally best practice to target a gene with more than one sgRNA to account for differences in activity. However, in the case of arrayed mutant collections, this requirement becomes unwieldy; as such, previous arrayed CRISPRi mutant collections in bacteria have utilized one predicted sgRNA per gene (Liu, Gallay et al. 2017, Wang, Guo et al. 2017). Our previous CRISPRi-Seq data (de Wet, Gobe et al. 2018) allowed us to identify essential genes that were growth-suppressed by multiple sgRNAs, but additionally provided empirical data on sgRNA knockdown efficiency. As a result, we were able to identify the most efficient sgRNAs for any gene.

For each of the 294 selected genes, two oligonucleotides (***Supplementary Table 1***) were synthesized at 50 nmol and annealed by LGC Biosciences. Oligonucleotides were delivered in 96-well plates, and were resuspended in water to a final concentration of 100μM. Cloning was performed as previously described (Rock, Hopkins et al. 2017) but adapted for scale. Briefly, pJR962 (Rock, Hopkins et al. 2017) was digested overnight with BsmBI-FastDigest (Thermo) and gel purified. Ligations were performed in 96 well-plates with 1U T4 Ligase (NEB) and incubated at room temperate overnight. Ligation reactions were transformed into 5μl High-Efficiency DH5α cells (NEB) with heat-shock, rescued in TY Broth, and plated on LB agar prepared in 6 well plates. Following overnight growth at 37°C, single colonies were picked into 800μl of LB broth, in 2ml deep 96-well plates, and grown overnight at 37°C with vigorous shaking. All *E. coli* cultures were supplemented with Kanamycin (Roche) at a final concentration of 50μg/ml. Plasmids were extracted with a Zyppy-96 Plasmid Miniprep kit (Zymo) according to manufacturer instructions and quantified via Nanodrop.

Approximately 100ng of plasmid was electroporated into the *M. smegmatis* ParB-mCherry reporter strain (Santi and McKinney 2015) and rescued in 7H9 broth supplemented with OADC and 0.05% Tween 80, for 3 hours at 37°C. Electroporation reactions were plated

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

on 7H10 agar supplemented with OADC and Kanamycin at 20ug/ml, prepared in 6-well plates, and grown at 37°C until visible colonies formed. Single colonies were picked into 400µl 7H9 supplemented with kanamycin (20ug/ml) and grown with shaking until stationary phase. Glycerol stocks of all transformants were stored at -80°C. Any failed ligations or transformations were repeated.

Sequencing Validation

PCR Primers (***Supplementary Table 1***) designed to target the sgRNA-containing region of plasmid pJR962 were synthesized by Inqaba Biotech. PCR reactions were prepared with OneTaq Master-mix (NEB) in 96 well plates to a final volume of 20µl. Using pipette tips, scrapings of glycerol stocks were taken and mixed with the PCR reaction. The PCR was performed with the settings described (***Supplementary Methods***). PCR reactions were purified and sequenced by the Stellenbosch Central Analytic Facility using the same forward primer as the PCR reaction. Sequencing results were analyzed using a custom written R script. For PCR amplification, we opted not to use a high-fidelity polymerase owing to the number of reactions performed. Since this risked introducing artificial mismatches or ambiguities into the sequencing results, a sequence was considered correct if the mutant had a high-quality alignment with its expected sequence. We accepted ambiguities in the sequencing trace if the strain functionally validated in downstream growth analysis. PCR reactions that did not produce sequencing results were repeated; ultimately, sequence-validation was obtained for 238 mutants from the initial cloning workflow, a success rate of 83%. Of the invalidated mutants, 15 wells contained plasmid backbone, 15 were cross-contaminants which occurred at some point during the cloning process, and the remaining 20 failed to sequence. We maintained the plasmid-backbone strains in our library and downstream workflow as these represented useful empty vector sequences, so that control strains were nested within the arrayed library. Cloning was repeated for invalidated mutants.

Growth Validations

Square plates containing standard Middelbrook 7H10 OADC agar were prepared and supplemented with kanamycin (Roche,20ug/ml), with or without anhydrotetracycline (ATc, Sigma, 100ng/ml). Fresh cultures were inoculated from glycerol freezer stocks into 400µl Middelbrook 7H9 liquid medium supplemented with kanamycin (Roche, 20ug/ml) in deep 96-well plates. Cultures were grown with shaking until saturated. Stationary-phase cultures were diluted 1:10 000 into fresh 7H9 before spotting onto prepared 7H10 plates using a 96-well replicator pin (Sigma). All spotting was performed in triplicate. Plates were grown until colonies formed and photographed using a lightbox. Colony sizes were quantified using Iris (Kritikos, Banzhaf et al. 2017).

Image Sample Preparation

Fresh cultures were inoculated from glycerol stocks into 400µl 7H9 with kanamycin (Roche, 20ug/ml) and grown until saturated. Saturated cultures were diluted 1:800 into fresh 7H9 with kanamycin (Roche, 20ug/ml) and grown for 24 hours until exponential phase. Exponential phase cultures were inoculated 1:40 into 7H9 supplemented with kanamycin (Roche, 20ug/ml) and ATc (Sigma 100ng/ml) and grown for 18 hours, with shaking. Large format agarose pads were prepared with water to a final concentration of 2%. Briefly, 2ml of molten low-melt agarose was sandwiched between 2 rectangular coverslips (No. 1.5, 24

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mm x 60 mm). The bottom coverslip was marked using a laser-cut stencil to indicate sample placement. Agarose pads were left to dry, and 1µl of induced culture spotted onto the pads, using a multichannel pipette. Each pad contained 8 samples. Twenty-four strains were imaged per day, and replicate imaging was performed for 136 samples to validate the reproducibility of the imaging workflow.

Microscopy

Large-format agarose pads were imaged using a Zeiss Axio Observer Z1 and ZEN 2 (blue edition) software, with the ZEN Tiles and Positions and ZEN Autofocus Modules installed. Images were captured using a Zeiss Axiocam 503 with 3X analogue gain. Using a low magnification 10X objective, each sample was localized on the pad. Approximately 24 fields-of-view were selected with a 100X Phase Contrast Objective (1.4NA), before capturing images using bright-field and fluorescent imaging. Fluorescence was excited using a Colibri Green LED (555/30nm) and filtered at 590 – 650nm. Exposure times were maintained across imaging sessions for Brightfield Images, and for fluorescent images. A software autofocus regimen was used before each field-of-view was captured. In the case of low-density samples, fields-of-view were manually chosen and increased in number. Raw images were saved as CZI files prior to processing and data extraction.

Time-lapse imaging was performed on 1.5% agarose pads embedded with 7H9 OADC medium containing ATc (Sigma, 100ng/ml) and kanamycin (Roche, 20ug/ml), in glass-bottomed dishes (NEST Biotechnology). Cells were maintained in an incubated chamber at 37°C and imaged every 15 minutes with a software autofocus regimen implemented between each frame. Quantitation of microcolony growth rates was performed in ImageJ and R with a custom script. Briefly, images were thresholded in ImageJ and the microcolony area extracted per frame. Time-lapse plots were produced in R using a combination of ggplot2 (Wickham 2016) and gganimate.

Image Processing and Data Extraction

All image processing was performed in FIJI (Schindelin, Arganda-Carreras et al. 2012). CZI Images were converted to TIFF images and were manually inspected and out-of-focus fields removed. To enhance foci in the fluorescent channel, a Gaussian blur filter was applied to the fluorescent channel and subtracted from the original image. Furthermore, overall intensity of each fluorescent channel was normalized throughout the imaging dataset, to a mean pixel intensity of 15. ImageJ processing scripts are available in ***Supplementary Methods***. Pre-processed images were analyzed using MicrobeJ (Ducret, Quardokus et al. 2016) with limited constraints on cell and foci detection. MicrobeJ output (cell contours, shape descriptions and fluorescent localization) was exported as CSV files.

Dataset Curation

To curate the output of MicrobeJ, a classifier was built in R. Two fields-of-view were sampled from each mutant and replica dataset and processed and analyzed as described above. All detected objects were initially exported. To create a reference of correctly identified cells, cells were inspected with the interface built into MicrobeJ, and misidentified cells removed from the dataset. Thus, two datasets were exported: the full dataset of all identified objects, and the dataset of classified cells. In total, approximately 22,000 objects were manually classified. The classified data were imported into R and utilized to build a classification model. Briefly, 20% of the dataset was reserved as a test sample, and the remaining 80% was

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

used to train a variety of models, using the Caret package (Kuhn 2008). For model training, 5-fold cross-validation was used, and resampling performed across the default tuning parameters. ROC was used to select the optimal tuning parameters for each model. A selection of models was tested and compared based on ROC AUC. An Averaged Neural Network was empirically chosen as the best-performing model (***Supplementary Figure 8***). The trained classifier was used to classify the entire dataset and remove misidentified objects. Data were input through the curation pipeline and labelled based on sequencing results and growth validations. Reproducibility of the imaging workflow was tested by comparing length of the replica image sets. The scripts used to develop and test the models, and all input data, are available at https://osf.io/pdcw2/.

Representative Cell Generation

For each imaged mutant strain, we derived the mean values for a selection of 13 morphological features – Angularity, Area, Aspect Ratio, Circularity, Curvature, Feret, Length, Perimeter, Sinuosity, Solidity, Roundness, Width and Width Variation – and counts of ParB maxima. We utilized the approximate nearest neighbor approach in the R package, RANN, to identify a single cell that best approximated the mean values of the mutant. The cell contour, derived from MicrobeJ, was plotted using ggplot, and maxima identified by MicrobeJ were superimposed on the plotted contour. Visualizing scripts are available at https://osf.io/pdcw2/.

Data Input and Processing

For data processing, we adapted a previously described approach (Campos, Govers et al. 2018). For each mutant, we calculated the mean of each measured feature, and the coefficient of variation (mean / standard deviation). We transformed each feature into a Z-score, relative to the distribution of means derived from the empty vector strains utilizing the equation:

$$Z = (F_i - mean(F_i^{WT}))/sd(F_i^{WT})$$

The normalized Z-Score therefore represents the number of standard deviations away from the mean of the wild-type distribution (***Supplementary Figure 1***). All subsequent analyses utilized Z-Score adjusted data, unless otherwise described.

Cluster of Orthologous Groups (COG) Annotations

Each gene in our dataset was assigned to a Cluster of Orthologous Groups using annotations obtained from eggNOG (Huerta-Cepas, Szklarczyk et al. 2016). We utilized both *M. smegmatis* and *M. tuberculosis* annotations to assign COGs to as many genes as possible. In addition, we manually annotated a number of genes with recently identified function (Wu, Zhang et al. 2018).

Phenoprint Visualization

For phenoprint visualization, variables were chosen for their interpretability and potential biological relevance. Variables included angularity, area, circularity, curvature, length, sinuosity, roundness, width, width variation and number of ParB foci (maxima). For visualization, the mean and CV Z-scores for each variable were displayed in a circular coordinate system. Where visualizations represented groups of mutants (for example COGs) means of the mutant Z-scores were used for visualization. Phenoprint visualization code is available at https://osf.io/pdcw2/.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

COG Enrichment Analysis

For each feature, we utilized a cut-off of 3 standard deviations above or below the mean of the control distribution, and tested for enrichment of each COG category utilizing Fisher's Exact Test. We corrected for multiple testing using the Benjamini-Hochberg Procedure and utilized a significance cut-off of $p < 0.05$. For display of data, we opted to replicate prior work (Campos, Govers et al. 2018) in displaying all genes found above or below the threshold. COG categories that were enriched for a particular feature were highlighted to differentiate them from the background. We only displayed plots where at least one COG was found to be statistically enriched (***Supplementary Figure 2***) Scripts are available at https://osf.io/pdcw2/.

Dimensionality Reduction and Analysis

Dimensionality reduction allows visualization of multidimensional data in two-dimensional space. Utilizing our normalized Z-score data, we performed dimensionality reduction in R. We tested a number of approaches for dimensionality reduction, including a Principle Component Analysis (PCA), t-SNE (Maaten and Hinton 2008) and UMAP (Maaten and Hinton 2008) algorithms. We superimposed these dimensionality reduction techniques with a hierarchical Density-based spatial clustering of applications with noise (hdbscan) (Lel, McInnes et al. 2017). For each technique, we tested a variety of optimization parameters to produce distinct clusters of points, while still maintaining a consistent cluster of wild-type samples. On comparison, UMAP produced the best separated clusters, while maintaining a relatively uniform wild-type cluster. UMAP is characterized by a degree of stochasticity during initialization (Lel, McInnes et al. 2017); therefore, we generated 100 maps which were each clustered with hdbscan (minPts = 8). Consistently present clusters across initializations were determined with hierarchical clustering.

To explain the features that assigned points to particular clusters, we superimposed a color-gradient based on the mutant Z-Scores for a chosen set of features. Additionally, we generated heatmaps describing mean Z-scores of chosen features, for each cluster. We performed COG enrichment on each cluster, utilizing Fisher's Exact Test, and adjusted for multiple testing using the Benjamini-Hochberg Procedure. For visualizing sub-clusters within the UMAP, we adapted an approach inspired by the previously described spatial analysis of functional enrichment (SAFE) approach (Baryshnikova 2016, Campos, Govers et al. 2018), to identify regions of the UMAP output that were enriched for particular functions. Briefly, for each point on our 2-dimensional UMAP projection, we examined surrounding points, within a predefined radius (the 10th percentile of the distribution of pairwise distances between all points) and tested for enrichment of either COG or KEGG ontologies using a hypergeometric test. We combined enrichment testing with a k-nearest neighbor interpolation approach to identify spatial regions of our UMAP manifold that were enriched for particular functions.

Antimicrobial Profiling

For each drug, 2-fold serial dilutions were prepared in 400µl of 7H9-OADC medium in deep 96-well plates. Exponential-phase *M. smegmatis* ParB-mCherry cells were inoculated into prepared media at a 1:40 dilution and cultured with shaking for 18 hours. Prior to imaging, ODs were measured and a dose-response curve fitted to the results with the DRC package in R (Ritz, Baty et al. 2015). The MIC_{90} was derived from the fitted curve and defined as 1X MIC. We validated that the MIC_{90} was consistent with results from a fluorescent resazurin based

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

assay. To this end, 10µl of resazurin (Sigma) was added to 50µl of cell culture, incubated for one hour at 37°C and color change determined visually. For imaging, cultures were spotted on agarose pads, and imaged and analyzed as previously described.

Cell Cycle Heatmaps

A population of cells at a single timepoint contains a degree of temporal information, as the population contains cells at different stages of the cell cycle (Campos, Govers et al. 2018). To visualize movement of ParB through the cell cycle, we utilized extracted maxima-localization data and cell length from MicrobeJ and processed the data further utilizing a bespoke R script. For each mutant, we arranged cells by length, and produced consensus heatmaps of ParB localization relative to cell-length using ggplot2. Scripts are available at https://osf.io/pdcw2/.

Nanopore RNA-Seq Library Preparation

Cultures of *M. smegmatis* were grown, with shaking, to exponential phase (OD_{600} ~0.9 - 1.0) in 15ml of 7H9 OADC at 37°C prior to RNA extraction. For CRISPRi knockdowns, the strain of interest was inoculated from an exponential phase culture into 15ml of 7H9-OADC supplemented with kanamycin (Roche, 20ug/ml) and ATc (Sigma, 100ng/ml) and cultured for 18 hours at 37°C with shaking. RNA was extracted using a FastRNA Blue Kit (MP Biomedicals) and ethanol-precipitated, according to manufacturer instructions. Extracted RNA was quantified by Nanopore. Ten µg of RNA was treated with 2U DNase (NEB) at 37°C for 10 minutes, prior to cleanup with a Zymo RNA Clean and Concentrator-5 kit (Zymo Research) according to manufacturer instructions. The purified DNAse-treated RNA was poly-A tailed using *E. coli* Poly(A) Polymerase (NEB), according to manufacturer instructions, prior to clean-up with a Zymo RNA Clean and Concentrator-5 kit (Zymo Research). Ribosomal RNA was depleted with a Bacterial Ribo-Zero rRNA Removal Kit (Illumina) according to manufacturer instructions, prior to clean-up with a Zymo RNA Clean and Concentrator-5 kit (Zymo Research). Library preparation for Nanopore sequencing was performed using either a Nanopore Direct RNA Sequencing Kit (SQK-RNA001, Oxford Nanopore Technologies) or Nanopore Direct cDNA (SQK-DCS109, Oxford Nanopore Technologies) according to manufacturer protocols, and sequenced on MinION sequencers (Oxford Nanopore Technologies). Base-calling was performed by MinKNOW software (Oxford Nanopore Technologies), and data stored as FASTQ files for downstream analysis.

RNA-Seq Data Analysis

Reads were aligned to the *M. smegmatis* mc^2155 genome (NC_008596) with Minimap2 (Li 2018), utilizing the map-ont command. Produced SAM files were sorted and indexed with samtools (Li, Handsaker et al. 2009). For visualization, coverage was calculated using the bamCoverage command of deepTools2 (Ramírez, Ryan et al. 2016) with normalization by Reads Per Kilobase per Million mapped reads (RPKM). featureCounts (Liao, Smyth et al. 2013) was utilized for assigning read counts to genes, and DESeq2 (Love, Huber et al. 2014) was used for comparisons of upregulated transcripts. All visualizations were produced in R, using ggplot2.

Dual CRISPRi Knockdown

sgRNAs were chosen to target *MSMEG_3214* and cloned into pJR962 as previously described (Love, Huber et al. 2014). Golden Gate (***Supplementary Table 1***) primers were designed to amplify the promoter-sgRNA-terminator region of the plasmid, with SapI restriction sites

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

included at the 5' region of the primers, and synthesized by Inqaba Biotech. The PCR reaction was performed with Q5 High-Fidelity Polymerase Master Mix (NEB) and the product column purified with a QIAquick PCR purification kit (Qiagen) and quantified by Nanodrop. PCR primers and reaction settings are available in ***Supplementary Methods*** Golden Gate cloning was performed utilizing the purified *MSMEG_3213*-targeting plasmid cloned during establishing the library as a backbone. Briefly, a single-pot reaction was set up containing T4 ligase (10 000U) and buffer (NEB), SapI (NEB, 10Units), 75ng of backbone, and purified PCR product at a molar ratio of 2:1. The reaction was incubated at 37°C for one hour, and at 55°C for 5 minutes, prior to transformation into *E. coli* DH5α (NEB). Colonies were screened by PCR using the sequencing primers, plJR962_Seq_F and plJR962_Seq_R, and OneTaq polymerase (NEB), as described previously, and positive clones sequenced by the Stellenbosch Central Analytic Facility utilizing plJR962_Seq_F to confirm the insertion. The sequence-verified plasmid was electroporated into *M. smegmatis* mc^{2}155 and tested for growth rescue by spotting assays.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ACKNOWLEDGEMENTS

This work was supported by grants from the US National Institute of Child Health and Human Development (NICHD) U01HD085531 (to DFW), the Research Council of Norway (R&D Project 261669 "Reversing antimicrobial resistance") (to DFW), the South African Medical Research Council (to VM and DFW), the National Research Foundation of South Africa (to VM and DFW), and core funding from the Wellcome Trust [203135/Z/16/Z]. TdW is grateful to South African Medical Research Council for financial assistance during his PhD training. We are thankful to Jeremy Rock and Sarah Fortune for the CRISPRi vectors, Keith Derbyshire and Todd Gray for the Dendra-tagged translational fusions, Bree Aldridge and Trever Smith for useful discussions, and members of the MMRU for constructive critique and technical advice throughout this project.

SUPPLEMENTARY DATA

Supplementary Table 1 – Oligonucleotide Sequences Utilized in this Work

Supplementary Table 2 – Essentiality Calls, Functional Annotations and Identified Transcriptional Units

Supplementary Table 3 – Literature Review of Genes Targeted in this Work

Supplementary Video 1 – Time-lapse microscopy of *ftsZ* and *wag31* silencing by CRISPRi.

Supplementary Video 2 – Time-lapse microscopy of *guaB2* silencing by CRISPRi.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Abrahams, K. A. and G. S. Besra (2018). "Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target." Parasitology **145**(2): 116-133.

Ahmed, W., C. Sala, S. R. Hegde, R. K. Jha, S. T. Cole and V. Nagaraja (2017). "Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase." PLOS Genetics **13**(5): e1006754.

Baranowski, C., M. A. Welsh, L.-T. Sham, H. A. Eskandarian, H. C. Lim, K. J. Kieser, J. C. Wagner, J. McKinney, G. E. Fantner, T. R. Ioerger, S. Walker, T. G. Bernhardt, E. J. Rubin and E. H. Rego (2018). "Maturing Mycobacterium smegmatis peptidoglycan requires non-canonical crosslinks to maintain shape." eLife **7**: e37516.

Baryshnikova, A. (2016). "Systematic Functional Annotation and Visualization of Biological Networks." Cell Systems **2**(6): 412-421.

Becht, E., L. McInnes, J. Healy, C.-A. Dutertre, I. W. H. Kwok, L. G. Ng, F. Ginhoux and E. W. Newell (2018). "Dimensionality reduction for visualizing single-cell data using UMAP." Nature Biotechnology **37**(1).

Bernier, S. P. and M. G. Surette (2013). "Concentration-dependent activity of antibiotics in natural environments." Frontiers in Microbiology **4**: 20.

Boot, M., S. Commandeur, A. K. Subudhi, M. Bahira, T. C. Smith, A. M. Abdallah, M. van Gemert, J. Lelièvre, L. Ballell, B. B. Aldridge, A. Pain, A. Speer and W. Bitter (2018). "Accelerating Early Antituberculosis Drug Discovery by Creating Mycobacterial Indicator Strains That Predict Mode of Action." Antimicrobial Agents and Chemotherapy **62**(7): e00083-00018.

Boshoff, H. I. M., T. G. Myers, B. R. Copp, M. R. McNeil, M. A. Wilson and C. E. Barry (2004). "The Transcriptional Responses of Mycobacterium tuberculosis to Inhibitors of Metabolism: Novel Insights Into Drug Mechanism of Action." Journal of Biological Chemistry **279**(38): 40174-40184.

Boshoff, H. I. M., M. B. Reed, C. E. Barry and V. Mizrahi (2003). "DnaE2 Polymerase Contributes to In Vivo Survival and the Emergence of Drug Resistance in Mycobacterium tuberculosis." Cell **113**(2): 183 193.

Botella, H., G. Yang, O. Ouerfelli, S. Ehrt, C. F. Nathan and J. Vaubourgeix (2017). "Distinct Spatiotemporal Dynamics of Peptidoglycan Synthesis between Mycobacterium smegmatis and Mycobacterium tuberculosis." mBio **8**(5): e01183-01117.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1035 Bush, M. J., M. J. Bibb, G. Chandra, K. C. Findlay and M. J. Buttner (2013). "Genes Required for
1036 Aerial Growth, Cell Division, and Chromosome Segregation Are Targets of WhiA before
1037 Sporulation in Streptomyces venezuelae." mBio **4**(5): e00684-00613.

1038 Caicedo, J. C., S. Cooper, F. Heigwer, S. Warchal, P. Qiu, C. Molnar, A. S. Vasilevich, J. D. Barry,
1039 H. S. Bansal, O. Kraus, M. Wawer, L. Paavolainen, M. D. Herrmann, M. Rohban, J. Hung, H.
1040 Hennig, J. Concannon, I. Smith, P. A. Clemons, S. Singh, P. Rees, P. Horvath, R. G. Linington and
1041 A. E. Carpenter (2017). "Data-analysis strategies for image-based cell profiling." Nature
1042 Methods **14**(9): 849.

1043 Campos, M., S. K. Govers, I. Irnov, G. S. Dobihal, F. Cornet and C. Jacobs-Wagner (2018).
1044 "Genomewide phenotypic analysis of growth, cell morphogenesis, and cell cycle events in
1045 Escherichia coli." Molecular Systems Biology **14**(6).

1046 Camsund, D., M. J. Lawson, J. Larsson, D. Jones, S. Zikrin, D. Fange and J. Elf (2019). "Time-
1047 resolved imaging-based CRISPRi screening." Nature Methods: 1-7.

1048 Cano, D. A., C. Mouslim, J. A. Ayala, F. G.-d. Portillo and J. Casadesús (1998). "Cell division
1049 inhibition in Salmonella typhimurium histidine-constitutive strains: an ftsI-like defect in the
1050 presence of wild-type penicillin-binding protein 3 levels." Journal of bacteriology **180**(19):
1051 5231-5234.

1052 Carey, A. F., J. M. Rock, I. V. Krieger, M. R. Chase, M. Fernandez-Suarez, S. Gagneux, J. C.
1053 Sacchettini, T. R. Ioerger and S. M. Fortune (2018). "TnSeq of Mycobacterium tuberculosis
1054 clinical isolates reveals strain-specific antibiotic liabilities." PLOS Pathogens **14**(3):
1055 e1006939.

1056 Cashmore, T. J., S. Klatt, Y. Yamaryo-Botte, R. Brammananth, A. K. Rainczuk, M. J. McConville,
1057 P. K. Crellin and R. L. Coppel (2017). "Identification of a Membrane Protein Required for
1058 Lipomannan Maturation and Lipoarabinomannan Synthesis in Corynebacterineae." Journal
1059 of Biological Chemistry **292**(12): 4976-4986.

1060 Chamier, L. v., R. F. Laine and R. Henriques (2019). "Artificial intelligence for microscopy:
1061 what you should know." Biochemical Society Transactions **47**(4): BST20180391.

1062 Chatterji, M., S. Unniraman, S. Mahadevan and V. Nagaraja (2001). "Effect of different classes
1063 of inhibitors on DNA gyrase from Mycobacterium smegmatis." Journal of Antimicrobial
1064 Chemotherapy **48**(4): 479-485.

1065 Cheng, N., M. A. Porter, L. W. Frick, Y. Nguyen, J. D. Hayden, E. F. Young, M. S. Braunstein, E. A.
1066 Hull-Ryde and W. P. Janzen (2014). "Filtration Improves the Performance of a High-
1067 Throughput Screen for Anti-Mycobacterial Compounds." PLoS ONE **9**(5): e96348.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1068 Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier,
1069 S. Gas, C. E. B. Iii, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R.
1070 Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh,
1071 J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers,
1072 S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and
1073 B. G. Barrell (1998). "Deciphering the biology of Mycobacterium tuberculosis from the
1074 complete genome sequence." **393**(6685): 31159.

1075 Cong, Q., I. Anishchenko, S. Ovchinnikov and D. Baker (2019). "Protein interaction networks
1076 revealed by proteome coevolution." Science (New York, N.Y.) **365**(6449): 185-189.

1077 Davis, E. O., E. M. Dullaghan and L. Rand (2002). "Definition of the Mycobacterial SOS Box and
1078 Use To Identify LexA-Regulated Genes in Mycobacterium tuberculosis." Journal
1079 of Bacteriology **184**(12): 3287-3295.

1080 de Wet, T. J., I. Gobe, M. M. Mhlanga and D. F. Warner (2018). "CRISPRi-Seq for the
1081 Identification and Characterisation of Essential Mycobacterial Genes and Transcriptional
1082 Units." bioRxiv: 358275.

1083 Dedrick, R. M., C. A. Guerrero-Bustamante, R. A. Garlena, D. A. Russell, K. Ford, K. Harris, K. C.
1084 Gilmour, J. Soothill, D. Jacobs-Sera, R. T. Schooley, G. F. Hatfull and H. Spencer (2019).
1085 "Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant
1086 Mycobacterium abscessus." Nature Medicine **25**(5): 730-733.

1087 DeJesus, M. A., E. R. Gerrick, W. Xu, S. W. Park, J. E. Long, C. C. Boutte, E. J. Rubin, D.
1088 Schnappinger, S. Ehrt, S. M. Fortune, C. M. Sassetti and T. R. Ioerger (2017). "Comprehensive
1089 Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon
1090 Mutagenesis." mBio **8**(1): e02133-02116.

1091 Dragset, M. S., T. R. Ioerger, Y. J. Zhang, M. Mærk, Z. Ginbot, J. C. Sacchettini, T. H. Flo, E. J.
1092 Rubin and M. Steigedal (2019). "Genome-wide Phenotypic Profiling Identifies and
1093 Categorizes Genes Required for Mycobacterial Low Iron Fitness." Scientific Reports **9**(1):
1094 11394.

1095 Drlica, K., M. Malik, R. J. Kerns and X. Zhao (2008). "Quinolone-Mediated Bacterial Death."
1096 Antimicrobial Agents and Chemotherapy **52**(2): 385-392.

1097 Ducret, A., E. M. Quardokus and Y. V. Brun (2016). "MicrobeJ, a tool for high throughput
1098 bacterial cell detection and quantitative analysis." Nature Microbiology **1**(7): 16077.

1099 Due, A. V., J. Kuper, A. Geerlof, J. P. v. Kries and M. Wilmanns (2011). "Bisubstrate specificity
1100 in histidine/tryptophan biosynthesis isomerase from Mycobacterium tuberculosis by active
1101 site metamorphosis." Proceedings of the National Academy of Sciences **108**(9): 3554-3559.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1102 Dziadek, J., S. A. Rutherford, M. V. Madiraju, M. A. L. Atkinson and M. Rajagopalan (2003).
1103 "Conditional expression of Mycobacterium smegmatis ftsZ, an essential cell division gene."
1104 Microbiology **149**(6): 1593-1603.

1105 Frandsen, N. and R. D'Ari (1993). "Excess histidine enzymes cause AICAR-independent
1106 filamentation in Escherichia coli." Molecular and General Genetics MGG **240**(3): 348-354.

1107 Gee, C. L., K. G. Papavinasasundaram, S. R. Blair, C. E. Baer, A. M. Falick, D. S. King, J. E. Griffin,
1108 H. Venghatakrishnan, A. Zukauskas, J.-R. Wei, R. K. Dhiman, D. C. Crick, E. J. Rubin, C. M.
1109 Sassetti and T. Alber (2012). "A Phosphorylated Pseudokinase Complex Controls Cell Wall
1110 Synthesis in Mycobacteria." Sci. Signal. **5**(208): ra7-ra7.

1111 Greendyke, R., M. Rajagopalan, T. Parish and M. V. V. S. Madiraju (2002). "Conditional
1112 expression of Mycobacterium smegmatis dnaA, an essential DNA replication gene."
1113 Microbiology **148**(Pt 12): 3887 3900.

1114 Gupta, K. R., K. C. Rahlwes, C. Wang, J. Liu, Y. S. Morita and E. H. Rego (2019). "Discovery of an
1115 essential mycobacterial polar growth factor involved in cell envelope integrity and
1116 lipomannan and lipoarabinomannan transport." bioRxiv: 633768.

1117 Hayashi, J. M., K. Richardson, E. S. Melzer, S. J. Sandler, B. B. Aldridge, M. S. Siegrist and Y. S.
1118 Morita (2018). "Stress-Induced Reorganization of the Mycobacterial Membrane Domain."
1119 mBio **9**(1): e01823-01817.

1120 Hildebrand, A., M. Remmert, A. Biegert and J. Söding (2009). "Fast and accurate automatic
1121 structure prediction with HHpred." Proteins **77 Suppl 9**: 128-132.

1122 Huang, K. C. (2015). "Applications of imaging for bacterial systems biology." Current Opinion
1123 in Microbiology **27**(C): 114 120.

1124 Huerta-Cepas, J., D. Szklarczyk, K. Forslund, H. Cook, D. Heller, M. C. Walter, T. Rattei, D. R.
1125 Mende, S. Sunagawa, M. Kuhn, L. J. Jensen, C. von Mering and P. Bork (2016). "eggNOG 4.5: a
1126 hierarchical orthology framework with improved functional annotations for eukaryotic,
1127 prokaryotic and viral sequences." Nucleic Acids Research **44**(D1): D286-D293.

1128 Jankute, M., J. A. G. Cox, J. Harrison and G. S. Besra (2015). "Assembly of the Mycobacterial
1129 Cell Wall." Annual Review of Microbiology **69**(1).

1130 Johnson, E. O., E. LaVerriere, E. Office, M. Stanley, E. Meyer, T. Kawate, J. E. Gomez, R. E.
1131 Audette, N. Bandyopadhyay, N. Betancourt, K. Delano, I. D. Silva, J. Davis, C. Gallo, M. Gardner,
1132 A. J. Golas, K. M. Guinn, S. Kennedy, R. Korn, J. A. McConnell, C. E. Moss, K. C. Murphy, R. M.
1133 Nietupski, K. G. Papavinasasundaram, J. T. Pinkham, P. A. Pino, M. K. Proulx, N. Ruecker, N.
1134 Song, M. Thompson, C. Trujillo, S. Wakabayashi, J. B. Wallach, C. Watson, T. R. Ioerger, E. S.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lander, B. K. Hubbard, M. H. Serrano-Wu, S. Ehrt, M. Fitzgerald, E. J. Rubin, C. M. Sassetti, D. Schnappinger and D. T. Hung (2019). "Large-scale chemical–genetics yields new M. tuberculosis inhibitor classes." Nature **571**(7763): 72-78.

Justice, S. S., D. A. Hunstad, L. Cegelski and S. J. Hultgren (2008). "Morphological plasticity as a bacterial survival strategy." Nature Reviews Microbiology **6**(2): 162-168.

Kang, C.-M., S. Nyayapathy, J.-Y. Lee, J.-W. Suh and R. N. Husson (2008). "Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria." Microbiology **154**(3): 725-735.

Kemble, H., P. Nghe and O. Tenaillon (2019). "Recent insights into the genotype–phenotype relationship from massively parallel genetic assays." Evolutionary Applications **12**(9): 1721-1742.

Kieser, K. J., C. Baranowski, M. C. Chao, J. E. Long, C. M. Sassetti, M. K. Waldor, J. C. Sacchettini, T. R. Ioerger and E. J. Rubin (2015). "Peptidoglycan synthesis in Mycobacterium tuberculosis is organized into networks with varying drug susceptibility." Proceedings of the National Academy of Sciences **112**(42): 13087-13092.

Kim, J.-H., K. M. O'Brien, R. Sharma, H. I. M. Boshoff, G. Rehren, S. Chakraborty, J. B. Wallach, M. Monteleone, D. J. Wilson, C. C. Aldrich, C. E. Barry, K. Y. Rhee, S. Ehrt and D. Schnappinger (2013). "A genetic strategy to identify targets for the development of drugs that prevent bacterial persistence." Proceedings of the National Academy of Sciences of the United States of America **110**(47): 19095 19100.

Kim, J.-H., J.-R. Wei, J. B. Wallach, R. S. Robbins, E. J. Rubin and D. Schnappinger (2011). "Protein inactivation in mycobacteria by controlled proteolysis and its application to deplete the beta subunit of RNA polymerase." Nucleic Acids Research **39**(6): 2210-2220.

Kling, A., P. Lukat, D. V. Almeida, A. Bauer, E. Fontaine, S. Sordello, N. Zaburannyi, J. Herrmann, S. C. Wenzel, C. König, N. C. Ammerman, M. B. Barrio, K. Borchers, F. Bordon-Pallier, M. Brönstrup, G. Courtemanche, M. Gerlitz, M. Geslin, P. Hammann, D. W. Heinz, H. Hoffmann, S. Klieber, M. Kohlmann, M. Kurz, C. Lair, H. Matter, E. Nuermberger, S. Tyagi, L. Fraisse, J. H. Grosset, S. Lagrange and R. Müller (2015). "Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins." Science (New York, N.Y.) **348**(6239): 1106-1112.

Koch, A., V. Mizrahi and D. F. Warner (2014). "The impact of drug resistance on Mycobacterium tuberculosis physiology: what can we learn from rifampicin?" Emerging Microbes & Infections **3**(3): emi201417.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1167 Kritikos, G., M. Banzhaf, L. Herrera-Dominguez, A. Koumoutsi, M. Wartel, M. Zietek and A.
1168 Typas (2017). "A tool named Iris for versatile high-throughput phenotyping in
1169 microorganisms." Nature Microbiology **2**(5): 17014.

1170 Kuhn, M. (2008). "Building Predictive Models in R Using the caret Package." Journal of
1171 Statistical Software, Articles **28**(5): 1--26.

1172 Kumar, R., B. S. Madhumathi and V. Nagaraja (2014). "Molecular Basis for the Differential
1173 Quinolone Susceptibility of Mycobacterial DNA Gyrase." Antimicrobial Agents and
1174 Chemotherapy **58**(4): 2013-2020.

1175 Leach, K. L., S. J. Brickner, M. C. Noe and P. F. Miller (2011). "Linezolid, the first oxazolidinone
1176 antibacterial agent." Annals of the New York Academy of Sciences **1222**(1): 49-54.

1177 Lee, H. H., N. Ostrov, B. G. Wong, M. A. Gold, A. S. Khalil and G. M. Church (2019). "Functional
1178 genomics of the rapidly replicating bacterium Vibrio natriegens by CRISPRi." Nature
1179 Microbiology **4**(7): 1105-1113.

1180 Lel, McInnes, J. Healy and S. Astels (2017). "hdbscan: Hierarchical density based clustering."
1181 The Journal of Open Source Software **2**(11): 205.

1182 Lel, McInnes, J. Healy and J. Melville (2018). "UMAP: Uniform Manifold Approximation and
1183 Projection for Dimension Reduction."

1184 Li, H. (2018). "Minimap2: pairwise alignment for nucleotide sequences." 1708.01492
1185 **34**(18): 3094-3100.

1186 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin
1187 and G. P. D. P. Subgroup (2009). "The Sequence Alignment/Map format and SAMtools."
1188 Bioinformatics **25**(16): 2078-2079.

1189 Li, X.-h., Y.-l. Zeng, Y. Gao, X.-c. Zheng, Q.-f. Zhang, S.-n. Zhou and Y.-j. Lu (2010). "The ClpP
1190 protease homologue is required for the transmission traits and cell division of the pathogen
1191 Legionella pneumophila." BMC Microbiology **10**(1): 54.

1192 Liao, Y., G. K. Smyth and W. Shi (2013). "featureCounts: an efficient general purpose program
1193 for assigning sequence reads to genomic features." 1305.3347 **30**(7): 923-930.

1194 Liu, X., C. Gallay, M. Kjos, A. Domenech, J. Slager, S. P. Kessel, K. Knoops, R. A. Sorg, J. R. Zhang
1195 and J. W. Veening (2017). "High-throughput CRISPRi phenotyping identifies new essential
1196 genes in Streptococcus pneumoniae." Molecular Systems Biology **13**(5): 931.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1197 Loenen, W. A. M., D. T. F. Dryden, E. A. Raleigh, G. G. Wilson and N. E. Murray (2014).
1198 "Highlights of the DNA cutters: a short history of the restriction enzymes." Nucleic Acids
1199 Research **42**(1): 3-19.

1200 Logsdon, M. M., P.-Y. Ho, K. Papavinasasundaram, K. Richardson, M. Cokol, C. M. Sassetti, A.
1201 Amir and B. B. Aldridge (2017). "A Parallel Adder Coordinates Mycobacterial Cell-Cycle
1202 Progression and Cell-Size Homeostasis in the Context of Asymmetric Growth and
1203 Organization." Current Biology **27**(21): 1 16.

1204 Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and
1205 dispersion for RNA-seq data with DESeq2." Genome Biology **15**(12): 550.

1206 Lunardi, J., J. Nunes, C. Bizarro, L. Basso, D. Santos and P. Machado (2013). "Targeting the
1207 Histidine Pathway in Mycobacterium tuberculosis." Current Topics in Medicinal Chemistry
1208 **13**(22): 2866-2884.

1209 Maaten, L. v. d. and G. Hinton (2008). "Visualizing Data using t-SNE." Journal of Machine
1210 Learning Research **9**: 2579--2605.

1211 Martini, M. C., Y. Zhou, H. Sun and S. S. Shell (2019). "Defining the Transcriptional and Post-
1212 transcriptional Landscapes of Mycobacterium smegmatis in Aerobic Growth and Hypoxia."
1213 Frontiers in Microbiology **10**: 591.

1214 Mayer, O., P. Jain, T. R. Weisbrod, D. Biro, L. Ho, D. Jacobs-Sera, G. F. Hatfull and W. R. Jacobs
1215 (2016). "Fluorescent Reporter DS6A Mycobacteriophages Reveal Unique Variations in
1216 Infectibility and Phage Production in Mycobacteria." Journal of Bacteriology **198**(23): 3220-
1217 3232.

1218 Mazandu, G. K. and N. J. Mulder (2012). "Function Prediction and Analysis of Mycobacterium
1219 tuberculosis Hypothetical Proteins." International Journal of Molecular Sciences **13**(6):
1220 7283-7302.

1221 Morlot, C., D. Straume, K. Peters, O. A. Hegnar, N. Simon, A.-M. Villard, C. Contreras-Martel, F.
1222 Leisico, E. Breukink, C. Gravier-Pelletier, L. L. Corre, W. Vollmer, N. Pietrancosta, L. S.
1223 Håvarstein and A. Zapun (2018). "Structure of the essential peptidoglycan amidotransferase
1224 MurT/GatD complex from Streptococcus pneumoniae." Nature Communications **9**(1): 3180.

1225 Müller, A. U., F. Imkamp and E. Weber-Ban (2018). "The Mycobacterial LexA/RecA-
1226 Independent DNA Damage Response Is Controlled by PafBC and the Pup-Proteasome
1227 System." Cell Reports **23**(12): 3551-3564.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1228 Murray, M. L. and P. E. Hartman (1972). "Overproduction of hisH and hisF gene products
1229 leads to inhibition of cell division in Salmonella." Canadian Journal of Microbiology **18**(5):
1230 671-681.

1231 Naran, K., A. Moosa, C. E. Barry, H. I. M. Boshoff, V. Mizrahi and D. F. Warner (2016).
1232 "Bioluminescent Reporters for Rapid Mechanism of Action Assessment in Tuberculosis Drug
1233 Discovery." Antimicrobial Agents and Chemotherapy **60**(11): 6748-6757.

1234 Nonejuie, P., M. Burkart, K. Pogliano and J. Pogliano (2013). "Bacterial cytological profiling
1235 rapidly identifies the cellular pathways targeted by antibacterial molecules." Proceedings of
1236 the National Academy of Sciences **110**(40): 16169-16174.

1237 Nonejuie, P., R. M. Trial, G. L. Newton, A. Lamsa, V. R. Perera, J. Aguilar, W.-T. Liu, P. C.
1238 Dorrestein, J. Pogliano and K. Pogliano (2016). "Application of bacterial cytological profiling
1239 to crude natural product extracts reveals the antibacterial arsenal of Bacillus subtilis." The
1240 Journal of Antibiotics **69**(5): 353.

1241 Opijnen, T. v. and A. Camilli (2013). "Transposon insertion sequencing: a new tool for
1242 systems-level analysis of microorganisms." Nature Reviews Microbiology **11**(7): 435-442.

1243 Paintdakhi, A., B. Parry, M. Campos, I. Irnov, J. Elf, I. Surovtsev and C. J. Wagner (2016). "Oufti:
1244 an integrated software package for high-accuracy, high-throughput quantitative microscopy
1245 analysis." Molecular Microbiology **99**(4): 767 777.

1246 Painter, R. E., G. C. Adam, M. Arocho, E. DiNunzio, R. G. K. Donald, K. Dorso, O. Genilloud, C.
1247 Gill, M. Goetz, N. N. Hairston, N. Murgolo, B. Nare, D. B. Olsen, M. Powles, F. Racine, J. Su, F.
1248 Vicente, D. Wisniewski, L. Xiao, M. Hammond and K. Young (2015). "Elucidation of DnaE as
1249 the Antibacterial Target of the Natural Product, Nargenicin." Chemistry & Biology **22**(10):
1250 1362-1373.

1251 Park, Y., A. Pacitto, T. Bayliss, L. A. Cleghorn, Z. Wang, T. Hartman, K. Arora, T. R. Ioerger, J.
1252 Sacchettini, M. Rizzi, S. Donini, T. L. Blundell, D. B. Ascher, K. Rhee, A. Breda, N. Zhou, V.
1253 Dartois, S. R. Jonnala, L. E. Via, V. Mizrahi, O. Epemolu, L. Stojanovski, F. Simeons, M. Osuna-
1254 Cabello, L. Ellis, C. J. MacKenzie, A. R. Smith, S. H. Davis, D. Murugesan, K. I. Buchanan, P. A.
1255 Turner, M. Huggett, F. Zuccotto, M. J. Rebollo-Lopez, M. J. Lafuente-Monasterio, O. Sanz, G. S.
1256 Diaz, J. Lelievre, L. Ballell, C. Selenski, M. Axtman, S. Ghidelli-Disse, H. Pflaumer, M. Bosche, G.
1257 Drewes, G. M. Freiberg, M. D. Kurnick, M. Srikumaran, D. J. Kempf, S. R. Green, P. C. Ray, K.
1258 Read, P. Wyatt, r. C. E. Barry and H. I. Boshoff (2017). "Essential but Not Vulnerable: Indazole
1259 Sulfonamides Targeting Inosine Monophosphate Dehydrogenase as Potential Leads against
1260 Mycobacterium tuberculosis." **3**(1): 18-33.

1261 Peters, J. M., A. Colavin, H. Shi, T. L. Czarny, M. H. Larson, S. Wong, J. S. Hawkins, C. H. S. Lu, B.-
1262 M. Koo, E. Marta, A. L. Shiver, E. H. Whitehead, J. S. Weissman, E. D. Brown, L. S. Qi, K. C. Huang

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and C. A. Gross (2016). "A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria." Cell **165**(6): 1 15.

Raina, M. and M. Ibba (2014). "tRNAs as regulators of biological processes." Frontiers in Genetics **5**: 171.

Ramírez, F., D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar and T. Manke (2016). "deepTools2: a next generation web server for deep-sequencing data analysis." Nucleic Acids Research **44**(W1): W160-W165.

Rego, E. H., R. E. Audette and E. J. Rubin (2017). "Deletion of a mycobacterial divisome factor collapses single-cell phenotypic heterogeneity." Nature **546**(7656): 153-157.

Rengarajan, J., B. R. Bloom and E. J. Rubin (2005). "Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages." Proceedings of the National Academy of Sciences of the United States of America **102**(23): 8327 8332.

Riska, P. F., W. R. Jacobs, Jr. and D. Alland (2000). "Molecular determinants of drug resistance in tuberculosis." The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease **4**(2 Suppl 1): S4-S10.

Ritz, C., F. Baty, J. C. Streibig and D. Gerhard (2015). "Dose-Response Analysis Using R." PLOS ONE **10**(e0146021).

Roberts, R. J., T. Vincze, J. Posfai and D. Macelis (2014). "REBASE--a database for DNA restriction and modification: enzymes, genes and genomes." Nucleic acids research **43**(Database issue): D298-299.

Rock, J. M., F. F. Hopkins, A. Chavez, M. Diallo, M. R. Chase, E. R. Gerrick, J. R. Pritchard, G. M. Church, E. J. Rubin, C. M. Sassetti, D. Schnappinger and S. M. Fortune (2017). "Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform." **2**: 16274.

Santi, I. and J. D. McKinney (2015). "Chromosome Organization and Replisome Dynamics in Mycobacterium smegmatis." mBio **6**(1): e01999-01914.

Sarathy, P. J., G. Gruber and T. Dick (2019). "Re-Understanding the Mechanisms of Action of the Anti-Mycobacterial Drug Bedaquiline." Antibiotics **8**(4).

Sassetti, C. M., D. H. Boyd and E. J. Rubin (2003). "Genes required for mycobacterial growth defined by high density mutagenesis." Molecular Microbiology **48**(1): 77 84.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1294 Sassetti, C. M. and E. J. Rubin (2003). "Genetic requirements for mycobacterial survival
1295 during infection." Proceedings of the National Academy of Sciences **100**(22): 12989-12994.

1296 Satta, G., M. Lipman, G. P. Smith, C. Arnold, O. M. Kon and T. D. McHugh (2018).
1297 "Mycobacterium tuberculosis and whole-genome sequencing: how close are we to
1298 unleashing its full potential?" Clinical Microbiology and Infection **24**(6): 604-609.

1299 Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C.
1300 Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P.
1301 Tomancak and A. Cardona (2012). "Fiji: an open-source platform for biological-image
1302 analysis." Nature Methods **9**(7): 676-682.

1303 Schnappinger, D. and S. Ehrt (2014). "Regulated Expression Systems for Mycobacteria and
1304 Their Applications." Microbiology Spectrum **2**(1): MGM2-0018-2013.

1305 Shell, S. S., E. G. Prestwich, S.-H. Baek, R. R. Shah, C. M. Sassetti, P. C. Dedon and S. M. Fortune
1306 (2013). "DNA Methylation Impacts Gene Expression and Ensures Hypoxic Survival of
1307 Mycobacterium tuberculosis." PLoS Pathogens **9**(7): e1003419.

1308 Shi, H., A. Colavin, T. K. Lee and K. C. Huang (2017). "Strain Library Imaging Protocol for high-
1309 throughput, automated single-cell microscopy of large bacterial collections arrayed on
1310 multiwell plates." Nature Protocols **12**(2): 429 438.

1311 Singh, B., J. Ghosh, N. M. Islam, S. Dasgupta and L. A. Kirsebom (2010). "Growth, cell division
1312 and sporulation in mycobacteria." Antonie van Leeuwenhoek **98**(2): 165-177.

1313 Singh, V., S. Donini, A. Pacitto, C. Sala, R. C. Hartkoorn, N. Dhar, G. Keri, D. B. Ascher, G.
1314 Mondesert, A. Vocat, A. Lupien, R. Sommer, H. Vermet, S. Lagrange, J. Buechler, D. F. Warner,
1315 J. D. McKinney, J. Pato, S. T. Cole, T. L. Blundell, M. Rizzi and V. Mizrahi (2013). "The Inosine
1316 Monophosphate Dehydrogenase, GuaB2, Is a Vulnerable New Bactericidal Drug Target for
1317 Tuberculosis." **3**(1): 5-17.

1318 Smith, T. C., K. M. Pullen, M. C. Olson, M. E. McNellis, I. Richardson, S. Hu, J. Larkins-Ford, X.
1319 Wang, J. S. Freundlich, D. M. Ando and B. B. Aldridge (2020). "Morphological profiling of
1320 tubercule bacilli identifies drug pathways of action." bioRxiv: 2020.2003.2011.987545.

1321 Su, C.-C., P. A. Klenotic, J. R. Bolla, G. E. Purdy, C. V. Robinson and E. W. Yu (2019). "MmpL3 is
1322 a lipid transporter that binds trehalose monomycolate and phosphatidylethanolamine."
1323 Proceedings of the National Academy of Sciences **116**(23): 201901346.

1324 Timmins, G. S. and V. Deretic (2006). "Mechanisms of action of isoniazid." Molecular
1325 Microbiology **62**(5): 1220-1227.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.20.000372; this version posted March 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Vilchèze, C. and W. R. J. Jacobs (2007). "The Mechanism of Isoniazid Killing: Clarity Through the Scope of Genetics." Annual Review of Microbiology **61**(1): 35-50.

Vilchèze, C., H. R. Morbidoni, T. R. Weisbrod, H. Iwamoto, M. Kuo, J. C. Sacchettini and J. W. R. Jacobs (2000). "Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis." Journal of bacteriology **182**(14): 4059-4067.

Wang, T., C. Guan, J. Guo, B. Liu, Y. Wu, Z. Xie, C. Zhang and X.-H. Xing (2018). "Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance." Nature Communications **9**(1): 2475.

Wang, T., J. Guo, C. Guan, Y. Wu, B. Liu, Z. Xie, C. Zhang and X.-H. Xing (2017). "Pooled CRISPR interference screens enable high-throughput functional genomics study and elucidate new rules for guide RNA library design in Escherichia coli." bioRxiv: 129668.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York, Springer-Verlag.

Wion, D. and J. Casadesús (2006). "N6-methyl-adenine: an epigenetic signal for DNA–protein interactions." Nature Reviews Microbiology **4**(3): 183-192.

World Health Organization (2019). Global Tuberculosis Report.

Wu, K. J., J. Zhang, C. Baranowski, V. Leung, E. H. Rego, Y. S. Morita, E. J. Rubin and C. C. Boutte (2018). "Characterization of Conserved and Novel Septal Factors in Mycobacterium smegmatis." Journal of Bacteriology **200**(6): e00649-00617.

Xu, W., M. A. DeJesus, N. Rücker, C. A. Engelhart, M. G. Wright, C. Healy, K. Lin, R. Wang, S. W. Park, T. R. Ioerger, D. Schnappinger and S. Ehrt (2017). "Chemical Genetic Interaction Profiling Reveals Determinants of Intrinsic Antibiotic Resistance in Mycobacterium tuberculosis." Antimicrobial Agents and Chemotherapy **61**(12): e01334-01317.