

GplR1, an unusual TetR-like transcription factor in *Mycobacterium abscessus*, controls the production of cell wall glycopeptidolipids, colony morphology, and virulence

Scarlet S. Shell,¹ Michal Bar-Oz,² Junpei Xiao,¹ Manitosh Pandey,³ Juan Bellardinelli,⁴ Opeyemi I. Ibitoye,¹ Mary Jackson,⁴ Stefan H. Oehlers,³ Daniel Barkan,² Michal Meir⁵

AUTHOR AFFILIATIONS See affiliation list on p. 17.

ABSTRACT *Mycobacterium abscessus* is a major human pathogen, mostly infecting people with pre-existing lung conditions, such as cystic fibrosis. The production of glycopeptidolipids (GPL) is a major determinant of virulence of this bacterium, with clinical isolates that lack GPL generally exhibiting more aggressive clinical behavior. The current paradigm is that GPL production is abolished *in vivo* via irreversible, spontaneous mutations taking place as part of in-host evolution. Little is known about the mechanisms or extent to which GPL production may be regulated. Here, we describe an unusual TetR-like transcription factor of *M. abscessus*, *mab_1638*, that appears to be a strong positive regulator of the entire GPL biosynthesis and export gene cluster through a combination of direct and indirect mechanisms. The inactivation of *mab_1638* abolished GPL production, leading to stable rough colony morphology and increased virulence in infection models, characteristics of rough, non-GPL producers. Transcriptome analysis found that the *mab_1638* mutant had 118 differentially expressed genes, including the GPL locus and a second, recently described GPL-like locus that produces a related glycosylated lipopeptide called GP8L. Chromatin immunoprecipitation and sequencing revealed a consensus inverted-repeat DNA sequence motif, characteristic of genes regulated by *mab_1638*. Together, these findings found that *mab_1638* encodes a transcription factor required for GPL production and, therefore, has a profound effect on virulence traits. We propose naming this gene GPL regulator 1 (*gplR1*). This finding raises the important possibility that *M. abscessus* strains appearing smooth in laboratory growth conditions may nonetheless downregulate GPL-cluster genes in other conditions, including in-patient conditions, and thus acquire the phenotypic characteristics of rough strains.

IMPORTANCE *Mycobacterium abscessus* is an important human pathogen, causing disease that is difficult to treat. *M. abscessus* strains have been observed to have two distinct colony morphologies, smooth and rough, which substantially impact clinical presentation. Rough strains are associated with later-stage, more severe disease and are more virulent in animal models. Smooth morphology is conferred by a molecule called glycopeptidolipid in the outer cell envelope, and rough morphology is known to occur when mutations inactivate genes required for glycopeptidolipid biosynthesis. Little is known about the possibility that glycopeptidolipid production could be regulated. Here, we have identified a transcription factor that is required for glycopeptidolipid biosynthesis, indicating that glycopeptidolipid production is indeed a regulated process and raising the important possibility that strains exhibiting smooth morphology in the lab may downregulate GPL production in the human host, thereby acquiring the virulence properties of rough strains.

Editor Zoe A. Dyson, London School of Hygiene & Tropical Medicine, London, United Kingdom

Address correspondence to Daniel Barkan, daniel.barkan@mail.huji.ac.il, or Scarlet S. Shell, sshell@wpi.edu.

The authors declare no conflict of interest.

Received 18 June 2025

Accepted 24 July 2025

Published 28 August 2025

Copyright © 2025 Shell et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

KEYWORDS *Mycobacterium*, abscessus, transcriptional regulation, virulence regulation, glycopeptidolipid, GPL

The genus *Mycobacterium* includes highly pathogenic species, such as *M. tuberculosis* and *M. abscessus*. These bacteria, as do others, use complex mechanisms to regulate their physiology and virulence. Bacterial genomes contain a variety of transcription factors that directly bind to regulatory sequences in the chromosome, most often in the promoters or upstream regulatory elements of the affected genes, resulting in higher or lower transcription. The identification of regulatory elements and their mechanisms of activity is important for understanding bacterial stress responses and pathogenesis.

An important virulence-associated mechanism in *M. abscessus* is the production of glycopeptidolipids (GPL), which are exported to the outer cell envelope. GPL production is a variable phenotype, evidenced by some clinical isolates but not others. Isolates producing GPL have a smooth colony morphology and, in some studies, have been reported to be more fit for intracellular growth and less pro-inflammatory than non-GPL producers (1, 2). Non-GPL producers have a rough morphology, cause more pronounced inflammation, and are associated with more severe and destructive disease (2–8). It is widely accepted that GPL-negative isolates emerge *in vivo* during chronic infection and are associated with worse clinical outcomes (6, 9, 10). The majority of GPL-negative clinical isolates arise via inactivating mutations in the GPL-related genes—mostly the large non-ribosomal peptide synthase genes *mps1* and *mps2* (*mab_4099c* and *4098c*, respectively) and the transporter gene *mmpL4b* (*mab_4115c*) (10–12). However, some rough clinical isolates do not have mutations in these genes (6, 10, 12). Furthermore, strains that shift reversibly between smooth and intermediate morphologies have been observed (13), and our previous work demonstrated that a small regulatory RNA affects the expression of GPL biosynthesis and transport genes (14). Collectively, these observations suggest that in addition to all-or-nothing inactivating mutations, *M. abscessus* can regulate GPL production. We therefore opted to identify additional genes that are involved in the smooth-to-rough transition, possibly by regulating, rather than mutating, GPL-related synthetic pathways. Here, we report the identification and characterization of a transcription factor required for high-level expression of GPL genes and thus for GPL production and smooth morphology. This finding highlights the possibility that GPL production is not a constitutive phenotype for strains with WT GPL biosynthesis genes, but instead regulated in response to growth conditions.

RESULTS

Transposon disruption of *mab_1638* leads to rough colony morphology due to loss of GPL production

We previously described the creation of a comprehensive transposon (Tn) mutant library in *M. abscessus* (ATCC19977, smooth), using a zeocin-selected Himar-1 transposon (15). We now turned to use this library to identify additional genes that lead to rough colony morphology when inactivated. We plated the Tn-library on 7H9/glycerol plates with no selection, picked several rough-appearing colonies, and replated them to confirm a stable rough phenotype. In those isolates with stable rough phenotypes, we identified the Tn insertion sites as previously described (14). In one of these Tn mutants, the identified insertion site was mapped to the TA at position 175 bp in the coding sequence of *mab_1638*. Given that the whole coding sequence was 648 bp, we expected this insertion to effectively inactivate the encoded protein. The insertion site was confirmed by PCR (Fig. 1A and B), followed by Sanger sequencing and will subsequently be referred to as the *mab_1638::tn* strain. To confirm that the morphology phenotype was due to the inactivation of *mab_1638*, we complemented the *mab_1638::tn* strain with a single-copy integrating vector containing an intact *mab_1638* coding sequence with the upstream 261 bp to include the native promoter. The complemented strain regained WT-like, smooth colony morphology (Fig. 1C I–IV).

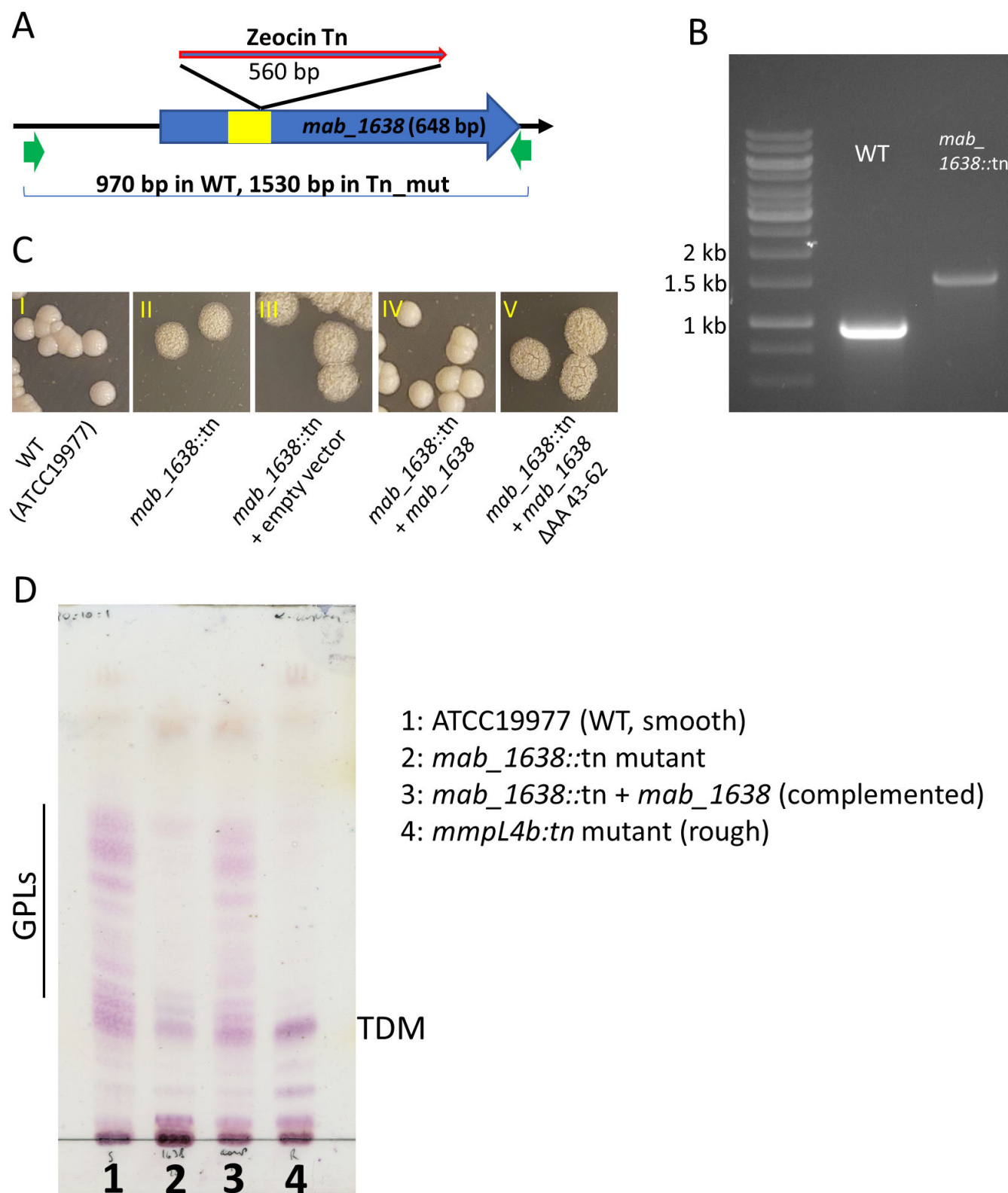


FIG 1 Disruption of *mab_1638* causes rough morphology in *M. abscessus*, due to lack of GPL in the cell envelope. (A) Schematic representation of the transposon insertion location in ATCC19977. (B) PCR confirmation of the insertion in *mab_1638*. (C) Representative colonies of the indicated strains grown on 7H10 solid media. Images III–V show the *mab_1638::tn* insertion mutant transformed with the indicated plasmids. In the plasmids harboring *mab_1638*, the gene was expressed from its native promoter on an integrating plasmid. The plasmid in V expressed *mab_1638* with an in-frame deletion of the indicated amino acids. (D) TLC analysis of lipids extracted from the cell walls of the indicated strains. Predicted GPL bands are shown. TDM, trehalose dimycolate.

To examine whether the rough phenotype of the *mab_1638::tn* strain was due to lack of GPL in the cell envelope, we performed a TLC analysis of lipids extracted from the cell envelope of the mutant, the smooth parental ATCC19977 strain, and an *mmpL4b::tn* mutant with a classic rough morphology (Fig. 1D). Whereas the smooth ATCC19977 strain had abundant GPL in the cell envelope, both the *mab_1638::tn* and *mmpL4b::tn* mutants did not have the characteristic bands. The complemented strain (lane 3) showed full reversion to the parental phenotype.

MAB_1638 protein is a TetR-like transcription factor, with a DNA binding domain essential for its effect

mab_1638 is designated in databases (such as Mycobrowser.epfl.ch) as a *tetR*-like transcription factor. By extrapolation from other TetR-like transcription factors, we identified the AA 43–62 of MAB_1638 as a helix-turn-helix structure that enables binding to DNA (16, 17). We therefore created a version of *mab_1638* with a deletion of AA 43–62 and introduced it into *mab_1638::tn*. In contrast to the intact version of the protein, MAB_1638^{ΔAA43–62} did not restore the WT-like, smooth colony morphology (Fig. 1C, V). This was not due to degradation of the truncated protein, as abundance of MAB_1638^{ΔAA43–62} was similar to that of full-length MAB_1638 (Fig. S1).

The inactivation of *mab_1638* in *M. abscessus* causes small increases in sensitivity to some antibiotics

TetR family transcriptional regulators frequently repress expression of efflux pumps involved in antibiotic resistance, as has been shown previously in *M. abscessus* for MAB_2299c, MAB_2648c, and MAB_4384 (18–21). We therefore compared the drug sensitivity of the *mab_1638::tn* strain to its WT parent (ATCC19977). No significant differences in MIC were noted for meropenem, ceftioxin, trimethoprim/sulfamethoxazole, moxifloxacin, and isoniazid (Table 1). However, the *mab_1638::tn* mutant had modestly decreased MIC values for vancomycin, teicoplanin, and linezolid compared to the WT (Table 1; Fig. S2). The directionality of this effect is the opposite of what would be expected for a transcription factor that represses efflux pump expression. Additionally, these antibiotics have two distinct mechanisms of action: vancomycin and teicoplanin block peptidoglycan biosynthesis, while linezolid blocks translation. A non-specific mechanism, such as changes in the myco-membrane composition and permeability (consistent with the colony morphology phenotype), may therefore be the cause for the subtle MIC shifts.

M. abscessus mab_1638::tn has a virulence phenotype resembling that of classic “rough” strains

To further characterize the phenotypic changes conferred by *mab_1638* inactivation, we examined the mutant for virulence. We first used the adult zebrafish infection model of persistent granulomatous infection (22). Two groups of at least eight adult zebrafish were infected with 10⁷ CFU of bacteria, and survival of the fish was followed for 10 days.

TABLE 1 The MICs (in μg/mL) of various antibiotics against *M. abscessus* strain ATCC19977 and its *mab_1638::tn* derivative

Drug	Drug target	Method	ATCC19977	<i>mab_1638::tn</i>	MIC fold change
Meropenem	Peptidoglycan (β-lactam)	E-test	>32	>32	1
Ceftioxin	Peptidoglycan (β-lactam)	E-test	32	48	1.5
Trimethoprim/sulfamethoxazole	Folate biosynthesis	E-test	>32	>32	1
Moxifloxacin	DNA gyrase	E-test	1–2	1–2	1
Vancomycin	Peptidoglycan (non-β-lactam)	E-test	16	4	0.25
Teicoplanin	Peptidoglycan (non-β-lactam)	E-test	12	3	0.25
Linezolid	Ribosome	E-test	Poorly determined		0.25
		Agar dilution	50	12.5	
Isoniazid	Cell envelope (mycolic acids synthesis)	Agar dilution	2,000	2,000	1

Whereas there was no mortality in the group infected with WT bacteria, there was 50% mortality in the *mab_1638::tn* group by three days post-infection and 100% mortality at day 9 (Fig. 2A). An additional experiment was performed to assess CFU during infection. To allow longer survival, a lower inoculum of only 10^6 CFU per fish was used (six fish for WT infection, eight fish for *mab_1638::tn* infection). However, even at this inoculum, only half of the fish infected by the *mab_1638::tn* mutant survived until the sacrifice point on day 14. In those four surviving fish, no CFU difference was seen compared to the fish infected with the WT strain (Fig. 2B).

To determine if the *mab_1638::tn* strain had infection characteristics similar to those of a classic rough *M. abscessus* strain, we turned to live imaging of the zebrafish embryo–*M. abscessus* infection model, as our adult infection experiments yielded insufficient survivors for histological comparison. We performed an infection experiment with the *mab_1638::tn* mutant, smooth ATCC19977, and a rough variant of ATCC19977 called CIP104536^T-R (2). Bacterial burden was measured on day 3 and day 6 post-infection in embryos infected by ~250 CFU/embryo. Total bacterial burden (as measured by fluorescence) was detected on days 3 and 6, recapitulating the known growth advantage of the rough control strain over the parental smooth strain (Fig. 2C). Comparing the burdens between days 3 and 6 revealed a statistically significant increase in burden in the rough control, but not the parental smooth strain. The *mab_1638::tn* mutant appeared to have an intermediate growth phenotype with non-statistically significant trends for both comparisons. Consistent with previous reports, the rough control strain produced infections in which a greater proportion of bacteria were in large clusters (abscess-like structures) compared to the smooth control strain (Fig. 2D and E). The behavior of the *mab_1638::tn* mutant was indistinguishable from that of the rough control strain and was distinct from that of the smooth parent strain (Fig. 2D and E). Altogether, these results confirm that the *mab_1638::tn* mutant behaves similarly to a GPL-null rough isolate, despite having no genetic defect in the GPL complex *per se*.

RNA sequencing reveals the transcriptional impact of *mab_1638*

To better understand the molecular mechanisms underlying the phenotypes of the *mab_1638::tn* strain, we first opted to characterize the transcriptome of the mutant compared to WT bacteria by RNAseq. We grew both strains in independent triplicates, extracted RNA at similar growth-phase stages, and performed RNA-seq as widely described. We found that 79 genes were significantly downregulated, and 39 genes were upregulated (Fig. 3A; Table S1; $\log_2$ fold change < -1 or > 1 and adjusted $P < 0.05$). Strikingly, the downregulated genes included 18 genes in the GPL complex (*mab_4097c*–*4117c*), which are responsible for GPL production and export (Fig. 3A). These downregulated genes included *mps1* and *mps2* (*mab_4099c* and *4098c*, respectively), as well as *mmpL4b*, all of which are implicated in the clinical transition of smooth to rough morphology (8, 10–12) and were downregulated by more than 8-fold. This downregulation may explain the strong rough phenotype of the *mab_1638::tn* strain, with no genetic disruption of the GPL locus itself.

Gene set enrichment analysis indicated that genes downregulated in the *mab_1638::tn* strain were enriched in several functional categories related to GPL biosynthesis (Fig. 3B; “lipid glycosylation,” “amino acid activation for non-ribosomal peptide biosynthetic process,” and “secondary metabolite biosynthetic process”). A closer examination of the differentially expressed genes revealed the presence of a second locus spanning *mab_4689c*–*4705c* that included genes with similarities to the GPL locus genes, including predicted non-ribosomal peptide synthases, methyltransferases, sugar modification genes, and transport proteins (Fig. 3C). This GPL-related locus was recently reported to produce a glycosylated lipooctapeptide termed GP8L that is important for *M. abscessus* virulence in macrophage and mouse models (23). All of the genes in this GP8L locus were downregulated in the *mab_1638::tn* strain. The genes upregulated in the *mab_1638::tn* strain showed modest enrichment for those with roles in aerobic respiration, as well as enrichment for genes involved in phenylacetate metabolism.

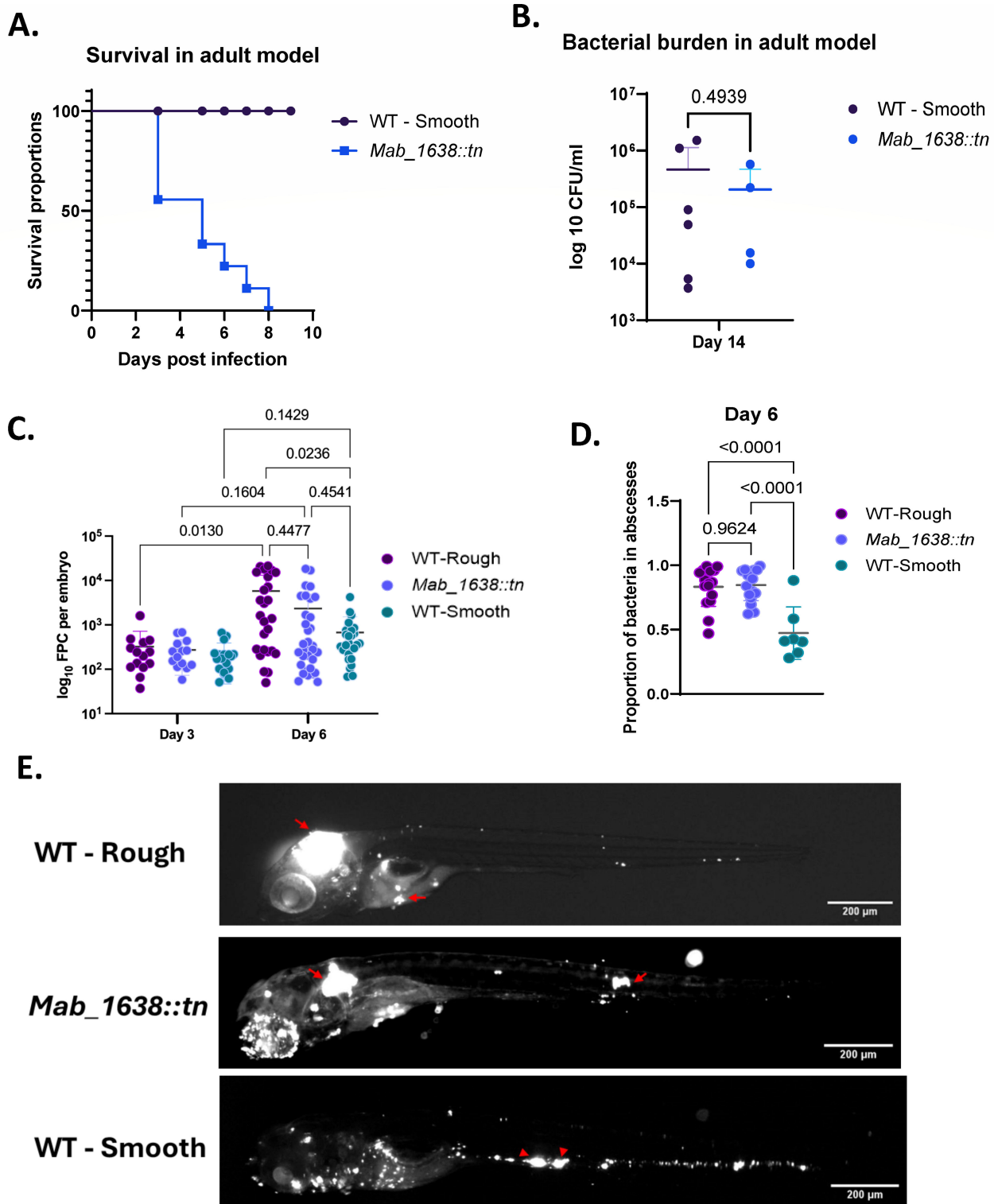


FIG 2 The *mab_1638:tn* mutant is hypervirulent in a zebrafish model and behaves similarly to a rough control strain. WT smooth is ATCC19977, and WT rough is ATCC19977/CIP104536^{T-R}. (A) Adult zebrafish were infected by 10^7 CFU/fish, and survival was followed for 10 days. $n = 8$ for WT, $n = 9$ for *mab_1638:tn*. (B) Bacterial burden was measured on day 14 post-infection by 10^6 CFU/fish and compared using unpaired *t*-test. (C) Bacterial burden, as represented by fluorescent pixel count (FPC), was measured on day 3 and day 6 post-infection in embryos infected by ~ 250 CFU/embryo. (D) The proportion of bacteria found in abscess-like structures (>500 mm², two-dimensional area) was measured on day 6 post-infection in embryos. (C and D) Ordinary one-way ANOVA with Tukey's multiple comparisons test. (E) Representative embryos from panel D. Red arrows point to clusters of fluorescent bacteria (>500 mm², two-dimensional area), referred to as "abscess-like."

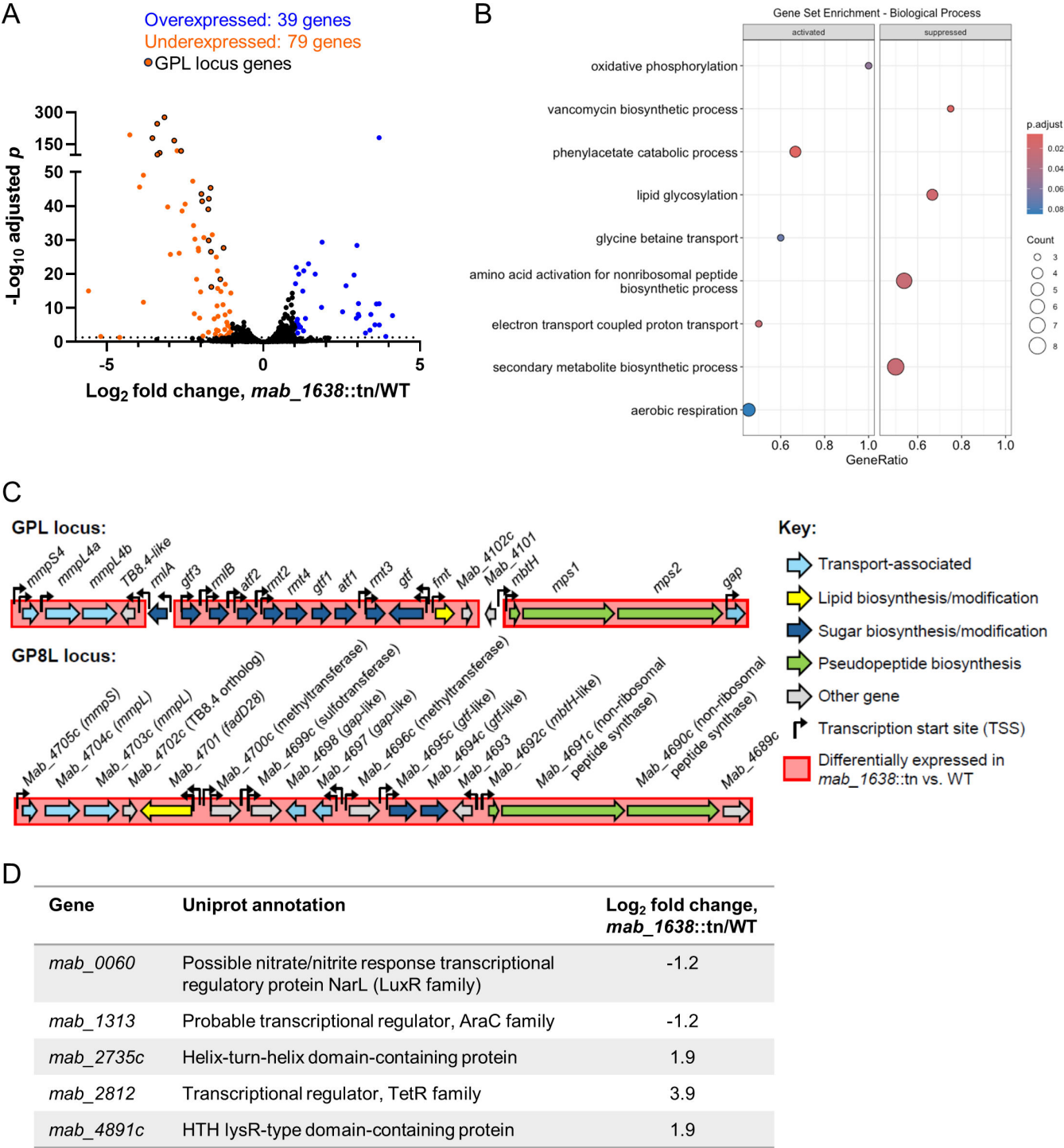


FIG 3 MAB₁₆₃₈ negatively impacts expression of GPL biosynthesis and transport genes as well as the GP8L gene cluster. (A) RNAseq revealed genes both up- and downregulated in the *mab*_{1638::tn} strain compared to the WT parental strain ATCC19977 (log₂ fold change < -1 or > 1, and adjusted *P* < 0.05). Genes known or expected to participate in GPL biosynthesis and transport are indicated. (B) Gene set enrichment analysis was used to identify Gene Ontology biological process categories that were disproportionately affected by disruption of *mab*₁₆₃₈. “Activated” genes had higher expression in the *mab*_{1638::tn} strain, while “Suppressed” genes had lower expression. (C) Diagrams of the GPL biosynthesis and transport gene cluster and a recently described GPL-like gene cluster termed the GP8L cluster, which was also downregulated in the *mab*_{1638::tn} strain. Red shadowing indicates downregulated genes. Of the 38 total genes in both clusters, only two genes did not meet the criteria for downregulation and are not shadowed. Genes are not shown to scale. Similarities between genes are indicated by naming and coloration. Bent arrows indicate reported TSSs. (D) Summary of genes predicted to encode transcription factors that were differentially expressed in the *mab*_{1638::tn} strain (log₂ fold change < -1 or > 1 and adjusted *P* < 0.05).

Finally, the differentially expressed genes included several predicted TFs (Fig. 3D), suggesting that some genes were directly controlled by MAB_1638, while others were affected indirectly. Indeed, one of these differentially expressed TFs, MAB_4891c, was previously implicated in the control of morphology in a transposon mutagenesis screen (15).

MAB_1638 recognizes an inverted-repeat consensus sequence and appears to both positively and negatively regulate its targets

While the RNAseq results demonstrated the effect of *mab_1638* on the transcriptome, it was not clear which expression changes were due to direct binding of the MAB_1638 protein to its target genes, and which were due to other transcription factors in the MAB_1638 regulon or downstream consequences of the physiological perturbations caused by changes in expression of MAB_1638's targets. We therefore performed chromatin immunoprecipitation sequencing (ChIP-seq) to identify the sequences directly bound by MAB_1638. For the ChIP-seq, we cloned both N- and C-terminally HA-tagged versions of MAB_1638 separately in a multi-copy vector and inserted the resulting plasmids into the *mab_1638::tn* mutant. The N-terminally tagged version used a strong constitutive promoter, while the C-terminally tagged version used the native promoter. Both tagged proteins were shown to have biological activity (complement the Tn-disruption) by their ability to revert the morphology back to smooth (Fig. S3). We performed ChIP-seq with the two tagged strains and two negative control strains: WT (no HA tag) and a strain expressing an unrelated HA-tagged protein not expected to bind DNA. Genome regions with sequence coverage peaks were identified, and those present in both of the MAB_1638-tagged strains but absent from both negative controls were classified as putative MAB_1638 binding sites (Table S2). Three hundred twenty-four putative binding sites were located within 500 nt upstream or downstream of a reported transcription start site (TSS) associated with a gene (24) (Table S3). We compared the frequencies of these putative binding sites among genes that were differentially expressed in the *mab_1638::tn* strain compared to the WT strain. Among genes with known TSSs, 30% of those that were downregulated and 53% of those that were upregulated had putative MAB_1638 binding sites, while only 15% of genes that were not differentially expressed had putative MAB_1638 binding sites (Fig. 4A). These statistically significant enrichments in putative binding sites among both up- and downregulated genes suggest that MAB_1638 may act as both an activator and a repressor, depending on the target gene. While TetR family TFs are often repressors, examples of activators in this family have been described (25, 26), as well as TetR family proteins that serve as both activators and repressors (27–29; see Discussion for further information).

Motif analysis with MEME and FIMO (30) revealed an inverted-repeat motif in 266 (82%) of the 324 putative MAB_1638 binding sites associated with TSSs (Fig. 4B). The motif was 23 nt in length, consisting of 11 nt inverted-repeat arms separated by a single nt. The outer 5 nt of each repeat had a more conserved sequence (5'-TCTGG-3') than the inner portion. This motif was disproportionately frequent in the 100 nt upstream of TSSs (Fig. 4C), consistent with expectations for a typical bacterial TF.

Although most of the GPL biosynthesis and transport genes were downregulated in the *mab_1638::tn* strain, there were only two MAB_1638 binding sites detected within the GPL gene cluster, both in the coding sequence of *mab_4110c* (*atf2*) and upstream of the TSS of *mab_4109c* (*rmt2*) (Fig. 4D). Like most GPL genes, both *atf2* and *rmt2* were downregulated in the *mab_1638::tn* strain. It is unlikely that a transcription factor binding at these two sites could directly impact expression of the entire locus. It therefore seems likely that MAB_1638 controls expression of a small number of GPL genes directly and controls the rest indirectly via the activity of one or more of the transcription factors listed in Fig. 3D. Notably, two of the transcription factors upregulated in the *mab_1638::tn* strain, *mab_2812* and *mab_4891c*, have MAB_1638 binding sites with inverted-repeat motifs, suggesting that they are direct targets of MAB_1638. They could

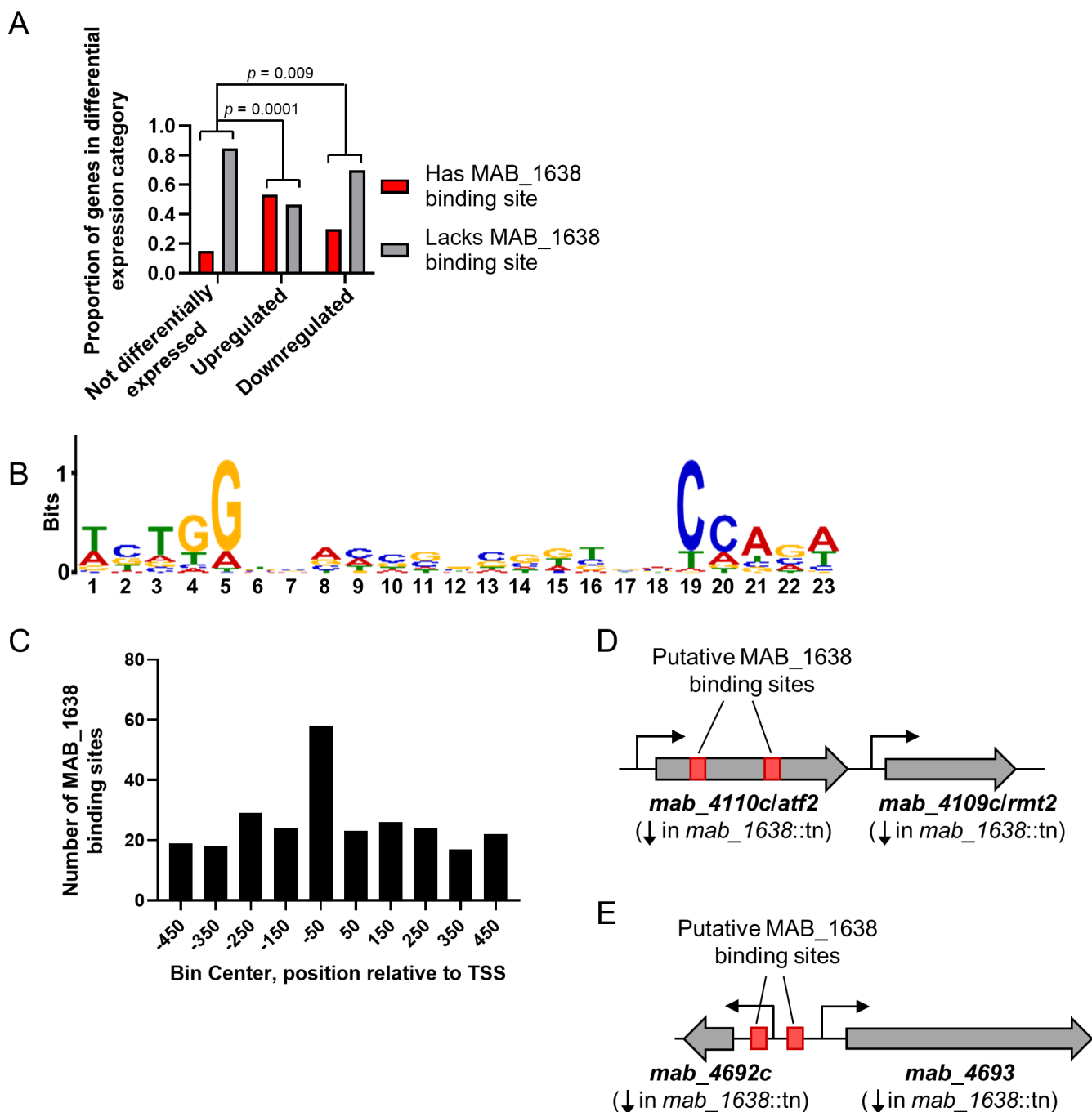


FIG 4 MAB_1638 binding sites are enriched in both up- and downregulated genes and feature a consensus inverted-repeat motif. (A) ChIP-seq was used to identify genomic sites bound by MAB_1638. Genes with annotated TSSs were then stratified based on the presence or absence of MAB_1638 binding sites within 500 nt of their TSSs and how their expression was affected by disruption of *mab_1638*. Genes that were up- or downregulated in the *mab_1638::tn* strain were significantly more likely to have MAB_1638 binding sites in their putative regulatory regions than genes that were not differentially expressed (Fisher's exact test). (B) MEME analysis revealed an inverted-repeat motif present in ~80% of the MAB_1638 binding sites in putative regulatory regions. (C) MAB_1638 binding sites motifs in putative regulatory regions (motif within 500 nt up- or downstream of an annotated TSS) were disproportionately located in the regions ~100 bp upstream of TSSs. (D) Two inverted-repeat MAB_1638 binding site motifs were identified within the GPL biosynthesis locus. Their approximate locations are indicated. (E) Two inverted-repeat MAB_1638 binding site motifs were identified within the GP8L biosynthesis locus. Their approximate locations are indicated. Schematics are not to scale.

conceivably be direct repressors of GPL locus genes and/or could regulate other transcription factors that directly impact the GPL locus.

TABLE 2 The putative regulon of MAB_1638 direct targets^d

Gene	Annotated TSS ^a	Plausibly in operon with MAB_1638 target ^b	MAB_1638 binding peak	Log ₂ fold change <i>mab_1638::tn</i> /WT ^c	<i>P</i> adjusted ^c	Uniprot annotation
MAB_0852	Yes	NA	Yes	−2.3	4.5×10^{-48}	Possible conserved polyketide synthase associated protein PapA2
MAB_0853	No	Yes	No	−2.2	5.7×10^{-35}	Probable esterase/lipase LipP
MAB_0854	No	Yes	No	−2.2	5.1×10^{-31}	Fatty acyl-AMP ligase FadD28 and polyketide synthase
MAB_0855	No	Yes	No	−1.9	1.8×10^{-31}	Membrane protein, MmpL family
MAB_0884c	No	Yes	No	−2.5	2.6×10^{-41}	Mycobacterium 19 kDa lipoprotein antigen
MAB_0885c	Yes	NA	Yes	−4.3	1.7×10^{-194}	Hypothetical lipoprotein lpqH
MAB_0939	Yes	NA	Yes	1.7	4.9×10^{-75}	Probable polyketide synthase
MAB_1013	Yes	NA	Yes	−2.8	1.2×10^{-73}	Ribonucleases G and E
MAB_1042c	Yes	NA	Yes	3.9	0.024	Cytochrome c oxidase subunit 1
MAB_1312	Yes	NA	Yes	−1.2	1.0×10^{-17}	Beta-lactamase-like
MAB_1313	No	Yes	No	−1.2	1.0×10^{-08}	Probable transcriptional regulator, AraC family
MAB_1565c	Yes	NA	Yes	1.1	1.1×10^{-20}	Alpha-amylase family protein
MAB_1674c	Yes	NA	Yes	−2.1	3.5×10^{-19}	Amidase AmiA2
MAB_1827c	Yes	NA	Yes	−1.2	0.00024	Bacteriophage protein
MAB_2464	Yes	NA	Yes	−1.5	0.0053	Chitin-binding type 2 domain-containing protein
MAB_2734c	Yes	NA	Yes	1.3	1.1×10^{-15}	Lipoprotein
MAB_2735c	Yes	NA	Yes	1.9	7.2×10^{-11}	Helix-turn-helix domain-containing protein
MAB_2741	Yes	NA	Yes	1.4	1.0×10^{-23}	DUF5666 domain-containing protein
MAB_3448	Yes	NA	Yes	−1.4	0.00050	Acetyltransferase
MAB_4109c	Yes	NA	Yes	−1.4	3.5×10^{-19}	Methyltransferase
MAB_4110c	Yes	NA	Yes	−1.7	1.3×10^{-30}	Probable acetyltransferase Atf2
MAB_4438	Yes	NA	Yes	1.7	1.1×10^{-20}	Serine aminopeptidase S33 domain-containing protein
MAB_4469c	Yes	NA	Yes	−1.0	4.3×10^{-15}	Sulfatase family protein
MAB_4689c	No	Yes	No	−3.1	1.8×10^{-40}	Uncharacterized protein
MAB_4690c	No	Yes	No	−2.9	3.5×10^{-67}	Probable non-ribosomal peptide synthetase PstA
MAB_4691c	No	Yes	No	−3.4	6.1×10^{-94}	Probable non-ribosomal peptide synthetase PstA
MAB_4692c	Yes	NA	Yes	−5.6	1.0×10^{-15}	MbtH-like protein
MAB_4693	Yes	NA	Yes	−2.1	1.5×10^{-27}	Probable cytochrome P450
MAB_4775c	Yes	NA	Yes	−1.5	1.2×10^{-21}	L,D-transpeptidase
MAB_4778c	Yes	NA	Yes	−1.3	1.2×10^{-11}	Uncharacterized protein
MAB_4891c	Yes	NA	Yes	1.9	4.1×10^{-30}	HTH lysR-type domain-containing protein

^aAs reported in reference 24.^bBased on absence of a reported TSS and examination of coverage in our RNAseq data.^cDESeq2.^dGenes differentially expressed in the *mab_1638::tn* strain were included if they had MAB_1638 binding peaks or were likely to be co-transcribed with genes with MAB_1638 binding peaks.

The GP8L locus also had two instances of the MAB_1638 inverted-repeat recognition motif within a MAB_1638 binding peak (Fig. 4E). These were located in putative regulatory regions for *mab_4692c* (an *mbtH* homolog) and *mab_4693*. As noted for the GPL locus, these binding sites are unlikely to be responsible for downregulation of the entire GP8L locus. This suggests that one or more of the transcription factors listed in Fig. 3D are also direct regulators of the GP8L locus.

Collectively, these results indicate that MAB_1638 is a bona fide transcription factor that binds an inverted-repeat motif to regulate expression of various genes, including some involved in GPL GP8L biosynthesis. MAB_1638-binding motifs were detected for

genes that were both up- and downregulated in the *mab_1638::tn* strain, consistent with the idea that MAB_1638 acts as both an activator and as a repressor.

Defining the MAB_1638/GplR1 regulon

While 118 genes were differentially expressed in the *mab_1638::tn* strain according to the cutoffs described above (Table S1), many of these genes are likely to be indirectly affected and therefore not technically part of the MAB_1638 regulon. Those differentially expressed genes for which ChIP-seq revealed MAB_1638 binding (23 genes, defined by having a binding peak within 500 nt of an annotated TSS) are likely to be direct targets. However, this gene list is limited to those genes with annotated TSSs (65/118 differentially expressed genes). Genes lacking TSSs may plausibly be encoded in polycistronic operons together with upstream genes that have annotated TSSs. We therefore considered differentially expressed genes lacking TSSs to be likely members of the MAB_1638 regulon if they were encoded immediately downstream of differentially expressed genes with TSSs and MAB_1638 binding sites. This added eight additional genes, expanding the putative MAB_1638 regulon to include 31 genes in total (Table 2).

Given the substantial impact of *mab_1638* on GPL gene expression and the likelihood that additional transcription factors regulating GPL gene expression will be discovered, we propose naming it GPL regulator 1 (*gplR1*).

DISCUSSION

Bacteria regulate their genetic and metabolic networks, and subsequently virulence, through a complex network of DNA-binding proteins known as transcription factors, which are generally responsive to various environmental cues. In *M. abscessus*, a major virulence factor (or arguably, an anti-virulence factor) is the glycopeptidolipids (GPL), which, when present in the outer cell envelope, give the bacterial colony its smooth dome shape. Lack of GPL in the cell envelope confers a rough colony morphology and leads to more destructive inflammatory responses, extracellular growth, resistance, and persistence to some antibiotics—collectively causing a more virulent phenotype with a considerably worse prognosis (6–10, 31, 32). It is thought that patients are usually infected by smooth strains and, via *in vivo* evolution, acquire rough isolates that drive the pathological clinical picture. Some rough variants arise via inactivating mutations in genes involved in GPL production or transport (the cluster *mab_4096* through *mab_4117*). However, many rough clinical isolates have what appear to be intact GPL-cluster genotypes (6, 10, 13), suggesting there are regulatory pathways that can up- or downregulate GPL production and transport, possibly in response to environmental cues. We recently identified and published one such regulatory pathway involving the sRNA B11 (14), where inactivation of B11 was sufficient to cause a rough phenotype. We therefore postulated that additional factors may regulate, rather than simply inactivate, GPL production/export, and here we have described one of these factors.

We used a transposon-mutant library to identify a mutant with a transposon insertion in the previously unstudied gene *mab_1638*, which exhibited a stable rough morphology on agar plates. This mutant recapitulated the hypervirulent phenotype associated with “classic” rough isolates when tested in the zebrafish model (adult and embryo). Although the increase in bacterial proliferation (a phenotype sometimes observed in R strains) was only mild, there was a substantial increase in mortality. However, the idea that enhanced virulence (and killing) can stem from stronger, more destructive inflammatory response (with little or no change in the bacterial burden) is well established in mycobacterial models. This could be due to enhanced immune-mediated tissue damage, as has been previously described both in *Mtb* and in *M. abscessus*. For example, in *Mtb*, a hypervirulent strain (Δ *cmaA2*) caused more death and lung pathology only due to stronger inflammation (33). In *M. abscessus*, a mutant defective in the small regulatory RNA B11 was shown to be hypervirulent and behave as a rough mutant in terms of virulence, with no substantial difference in CFU number (14). Classic rough strains with genetic

disruptions in GPL biosynthesis genes have also been shown to be more inflammatory than smooth strains (2, 8).

The protein encoded by *mab_1638* is predicted to have the classic structure of a TetR-like transcription factor—a large family of TFs present in most bacteria and involved in multiple physiologic and pathogenesis processes. Indeed, when we complemented the Tn mutant with a version of the gene coding for a protein unable to bind DNA, smooth colony morphology was not restored, supporting the hypothesis that DNA binding is required for the protein to exert its effect.

The combination of transcriptomic and ChIP-seq analyses provided the landscape of the *mab_1638* regulon. Thirty-one genes were differentially expressed and had putative GpIR1/MAB_1638 binding sites within their likely regulatory sequences, suggesting they are direct targets. This putative regulon likely includes some false positives, as GpIR1/MAB_1638 appeared to bind the regulatory regions of ~15% of non-differentially expressed genes (Fig. 4A). There are also likely additional regulon members that we did not detect because their TSSs have not yet been annotated. None of the *gplR1/mab_1638* regulon genes are located near *gplR1/mab_1638* on the *M. abscessus* chromosome. This makes GpIR1/MAB_1638 a type III TetR family transcriptional regulator, according to a scheme published by Ahn and colleagues (34). Unlike more conventional TetR family regulators which regulate a small number of adjacently transcribed genes, type III regulators control disparately located sets of genes.

The closest homolog of GpIR1/MAB_1638 in *M. tuberculosis* is Rv0275c, which was previously described as mediating isoniazid resistance in *M. bovis* (35). However, *M. abscessus* encodes another predicted TetR regulator, MAB_2594, that has more similarity and identity to Rv0275c than GpIR1/MAB_1638 does. It is therefore unsurprising that disruption of *gplR1/mab_1638* did not affect isoniazid resistance (Fig. S1).

The putative GpIR1/MAB_1638 regulon defined here includes eight genes that appear to be negatively regulated by GpIR1/MAB_1638 and 23 genes that appear to be positively regulated (Table 2). This contrasts with previously described TetR family transcription factors in *M. abscessus*, MAB_2299c, MAB_2648c, and MAB_4384, which all negatively regulate drug efflux pumps (18–21). Additional mycobacterial TetR family regulators have been described that negatively regulate cholesterol metabolism genes (36, 37), CRISPR genes (38), and oxidative stress response genes (39, 40), among others. There is a precedent in *M. smegmatis* for TetR family regulators that directly activate expression of their targets; *msmeg_6564* was reported to activate expression of DNA repair genes (41), and *msmeg_4022* was reported to activate expression of drug transporters (42). To our knowledge, the mechanisms of positive regulation in those cases are not known. There is also precedent for the idea that a TetR family transcription factor may both positively and negatively regulate genes. In *E. coli*, RutR represses transcription at several loci, including its own promoter and a pyrimidine degradation pathway promoter, while activating transcription at other loci, including a pyrimidine biosynthesis operon (29). Its activator activity was proposed to result from competing with a transcriptional repressor that could otherwise bind to a partially overlapping site (29). However, to our knowledge, this hypothesis has not been tested. In *Streptomyces autolyticus* CGMCC0516, deletion of the transcription factor GdmRIII led to upregulation of some genes and downregulation of others (27). *In vitro*, purified GdmRIII bound to the promoter regions of two genes that it appeared to upregulate and one gene that it appeared to downregulate (27). It is unknown how GdmRIII might have opposite effects on different promoters. A mechanism for both positive and negative regulation has been elucidated for AraC, an *E. coli* transcription factor that is not a member of the TetR family but, remarkably, can activate or repress the same set of genes depending on arabinose availability (reviewed in [43, 44]). AraC binds DNA in both its apo and arabinose-bound forms, but its mode of dimerization differs in these two states. This leads to the apo form binding to a different combination of sites than the arabinose-bound form. In the former case, AraC binding blocks access of RNA polymerase and CRP to the promoter, repressing transcription. In the latter case, AraC binds adjacent to RNA polymerase and

CRP, apparently helping to recruit RNA polymerase. Given the various mechanisms that could potentially be at play, further investigation will be needed to determine how GplR1/MAB_1638 carries out both positive and negative regulation. Alternatively, it is possible that some of the genes we have designated as direct targets are in fact indirect targets; further work is needed to determine this with high confidence.

One of the largest groups of genes transcriptionally affected by *gplR1/mab_1638* disruption was a set of 18 genes from the GPL locus, half of which were downregulated at least 8-fold, thus explaining the rough colony morphotype in the absence of genetic mutations in the GPL locus itself. However, only two GplR1/MAB_1638 binding sites were identified in the GPL locus, both within the *atf2* coding sequence. One of these binding sites is located ~484 nt upstream of the annotated TSS of the downstream gene, *rmt2*, and could plausibly contribute to the regulation of its expression. The other binding site is located >1.2 kb upstream of the annotated *rmt2* TSS and therefore seems unlikely to regulate its expression. Further work will be needed to determine if GplR1/MAB_1638 binding at these sites impacts the expression of *rmt2* or other genes in the GPL locus. It appears likely that at least some of the impact of *gplR1/mab_1638* on GPL gene expression is indirect, given the absence of GplR1/MAB_1638 binding sites elsewhere in the locus. Notably, several predicted transcription factors were differentially expressed in the *mab_1638::tn* strain, and two of these had GplR1/MAB_1638 binding sites revealed in our ChIP-seq data. Additional transcription factors under the control of GplR1/MAB_1638 may therefore be direct regulators of GPL locus genes.

Interestingly, a second large group of genes with similarity to the GPL locus genes was also downregulated. These were recently reported to encode proteins that produce GP8L, a glycosylated lipooctapeptide (23). GP8L is clearly not sufficient to independently confer smooth morphology in laboratory conditions, since various works have shown that single mutations in key canonical GPL locus genes cause rough morphology. Like canonical GPL, the role of GP8L in virulence appears to be complex. A strain in which the GP8L locus was disrupted in a rough background showed reduced uptake by macrophage-like THP-1 cells and possibly also a growth defect following uptake (23). In zebrafish embryo infections, a rough strain with GP8L disruption produced a bacterial burden lower than the rough parental strain and similar to a smooth strain (23). Finally, the rough GP8L disruption strain caused less mortality in mice than its rough parent (23). The impact of GP8L on strains that also produce GPL (smooth strains) is unknown, as is the significance of the apparent co-regulation of GPL and GP8L production by GplR1/MAB_1638. These questions should be investigated further.

Mounting evidence—both from examining clinical isolates (6, 12, 13) and from targeted studies like this one and our previous study on the B11 sRNA (14)—points to the conclusion that “S to R” transitions are not limited to random mutagenesis events selected by *in vivo* evolutionary pressures, but that the bacteria has active mechanisms to regulate GPL production and the associated phenotypes. This is consistent with previous work in *M. smegmatis* showing that outer cell envelope composition is subject to transcriptional regulation (45). Thus far, we have identified the *gplR1/mab_1638* gene and the B11 sRNA, both of which appear to maintain GPL production under certain conditions (including those in the lab), but may, in response to some as-yet unidentified cue, downregulate GPL production and export to produce the more virulent phenotype associated with the “rough morphology.” When these bacteria are isolated from patients or from infection models and re-grown on agar plates, this phenotype would logically reverse to “smooth colony” in response to the unknown signals in the lab growth environment that lead to expression of positive regulators of GPL, such as *gplR1/mab_1638* and B11. This raises the question of whether, in clinical practice, many of the so-called “smooth isolates” may actually behave in a “rough” isolate manner in the patient, and that only a tip of the iceberg of these isolates is identified as “rough” when re-grown in the clinical microbiology lab. This means research should focus on the regulatory elements governing the GPL formation and the metabolic cues that may

signal the bacteria to downregulate it. Interfering with these signals may provide a yet-unexplored therapeutic pathway.

MATERIALS AND METHODS

Strains and growth conditions

All strains were derived from *M. abscessus* ATCC19977 WT (smooth). Unless otherwise specified, bacteria were grown in Middlebrook 7H9 supplemented with ADS (final concentrations 5 g/liter bovine serum albumin fraction V, 2 g/liter dextrose, 0.85 g/liter NaCl, and 3 mg/liter catalase), 0.2% glycerol, and 0.05% Tween 80.

Isolation, identification, and subsequent complementation of a *mab_1638::tn* mutant

The creation of a zeocin-selected transposon-mutant library in *M. abscessus* was previously described (15). The identification of the insertion site was performed the same way as the library sequencing, but only on a single colony. The primers used for corroboration of the insertion site (Fig. 1A) were 1638 F: 5'-CAGAGCCATGCCTGGA TCCAG-3' and 1638 R: 5'-AGCGTGCTCAGTGCTCGACGG-3'. The *mab_1638::tn* strain was also referred to as mDB292. For complementation, the same primers were used to PCR-amplify a 970 bp fragment from *M. abscessus* gDNA (containing the gene and the 261 upstream bases), and cloned into the *Swal* site of the attB-integrating vector pDB213, to create pDB426. To create the Δ AA43-62 mutant, PCR was used to create two separate fragments: one ending at AA 42 and the other starting at AA 63. These fragments were then fused using 1638 F and 1638 R primers described above to obtain a 910 bp fragment with the codons for AA 43-62 deleted. This fragment was again cloned into the *Swal* site of pDB213 to create pDB448.

Total lipid extraction and TLC analysis

M. abscessus ATCC19977 WT (smooth), *mab_1638::tn*, *mab_1638::tn + mab_1638*, and *mmpL4b::tn* (rough) were grown in 7H9 with ADS, glycerol, and Tween-80 0.05% until OD = 0.4. Cells were collected, washed with PBS, and lipids extracted with CHCl₃:CH₃OH (1:2 vol/vol) once, followed by CHCl₃:CH₃OH (2:1 vol/vol) twice. After Folch wash, lipids were weighed, and the same amount of lipids was loaded onto a silica gel 60 F254-coated aluminum sheet. Samples were developed in CHCl₃:CH₃OH:H₂O (90:10:1) and revealed with α -naphthol.

Antibiotic sensitivity and MIC determination

For E-tests, equal amounts of the tested bacteria were plated on 7H9/glycerol plates in top (soft) agar to ensure equal spread of bacteria, and an E-test strip (BD BBL Sensi-Disc, Becton Dickinson and Company) was placed. Plates were incubated at 37°C, and results were read after 96–120 hours. Isoniazid sensitivity was performed using agar dilution: 7H9/glycerol/ADS plates were supplemented by the determined concentrations of isoniazid, and the tested bacteria were inoculated using a 10 μ L loop from a 0.02 OD₆₀₀ culture.

Zebrafish infection models

Fluorescent labeling was done with the pTEC15 (pMSP12:Wasabi) plasmid (Addgene plasmid # 30174; <http://n2t.net/addgene:30174>; RRID:Addgene_30174) (46). Zebrafish infection experiments were approved by A*STAR IACUC in protocol 221694. *M. abscessus* WT and *mab_1638::tn* fluorescently labeled strains were grown in 7H9 broth supplemented with 0.5% glycerol, 10% OADC, and 0.05% Tween 80 in the presence of 250 μ g/mL kanamycin. Single-cell suspension stocks for bacterial infection were prepared according to a previously reported protocol (22). Briefly, 0.8 OD₆₀₀ cells were centrifuged and

washed with 1× PBS and resuspended in 7H9 broth. Cells were declumped by passaging through a 27G needle three times and stored at −80°C in aliquots.

Adult zebrafish from AB strain were infected intraperitoneally with the fluorescently labeled *M. abscessus* WT and *mab_1638::tn* strains using a 31G needle. Infected zebrafish were housed in a recirculating aquarium system (28°C, 750 µS conductivity, pH 7.4) with a 14/10-hour light/dark cycle and observed as part of their daily feeding routine (22).

CFU recovery was performed on euthanized adult zebrafish by homogenization using a bead beater and plating on LB agar for 4–5 days at 37°C.

Zebrafish embryos were infected by microinjection (47). Embryos were anesthetized, mounted in 3% methyl cellulose, and imaged at six dpi on a Nikon SMZ25 equipped with a Nikon Digital Camera DS-10. Image analysis was performed with FIJI (ImageJ) to enumerate bacterial fluorescent pixel count, and abscesses were defined as having an area of 200 px corresponding to approximately 500 mm² (1.60 mm/px) (3).

Strains used as controls were the ATCC19977 (smooth) and the ATC19977/CIP104536^T-R (rough).

RNA sequencing and analysis

M. abscessus WT and *mab_1638::tn* were grown in triplicates (each starting from a different single colony) to an OD₆₀₀ of 0.5 in 7H9 supplemented with 0.5% glycerol and 0.05% Tween 80. Bacteria were collected, and RNA was extracted as previously described (14). Briefly, frozen cultures were thawed on ice and centrifuged at 4,000 rpm for 5 minutes at 4°C. The pellets were resuspended in 1 mL Trizol (Life Technologies) and placed in tubes containing Lysing Matrix B (MP Bio). Cells were lysed by bead-beating twice for 40 seconds at 9 m/sec in a FastPrep-5G instrument (MP Bio). RNA was purified using Direct Zol RNA miniprep kit (Zymo) according to the manufacturer's instructions. rRNA depletion, library preparation, and RNA sequencing were performed in the Technion Genomics Core Facility (Haifa, Israel). Data are available on GEO, accession number [GSE287636](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287636).

Reads were aligned to the NC_010397 reference genome using the Burrows-Wheeler Aligner (48). The FeatureCounts tool (49) was used to assign mapped reads to genomic features, and DESeq2 (50) was used to assess changes in gene expression. Genes with adjusted *P*-values < 0.05 and log₂ fold changes > 1 or < −1 were considered to be differentially expressed.

The R package clusterProfiler 4.0 (51) was used to perform pathway enrichment analysis using KEGG annotation data. clusterProfiler was also used to conduct gene set enrichment analysis (GSEA) using Gene Ontology (GO) annotations from QuickGO (<https://www.ebi.ac.uk/QuickGO/annotations>) (52).

Chromatin immune-precipitation (ChIP) sequencing and analysis

For N-terminal tagging, an HA-tag sequence was added to the 5′ end of the *mab_1638* coding sequence, and this fragment was cloned under the strong constitutive MOP promoter. For C-terminal tagging, the HA sequence was simply added to the 3′ end of *mab_1638* driven by its native promoter. Both constructs were cloned into the episomal multi-copy vector pDB32 to create pDB458 and pDB459, respectively. These were introduced into the *mab_1638::tn* mutant to create strains mDB353 (MOP promoter + N-terminal tag) and mDB352 (native promoter + C-terminal tag). Two negative control strains were assayed in parallel: WT ATCC19977, and a strain expressing an HA-tagged cyclopropane synthase (MmaA2) from *M. tuberculosis*, which is expected not to bind DNA. Immunoprecipitation was performed based on the method described in (53). Bacteria were cultured overnight in 50 mL of 7H9 medium supplemented with 0.5% glycerol and 0.05% Tween 80 to an OD₆₀₀ of ~0.4–0.6 and equilibrated to room temperature. Formaldehyde (1% final concentration) was added, followed by a 20-minute shaking-incubation at room temperature. Cross-linking was quenched by adding glycine (final concentration of 0.25 M) for 5 minutes at RT. Fixed cultures were collected by centrifugation at 4,500 × *g* for 10 minutes. Pellets were washed twice with cold PBS

and resuspended in 600 μL of ChIP lysis buffer (20 mM HEPES-KOH, pH 7.9; 50 mM KCl; 0.5 mM DTT; and 10% glycerol), supplemented with a protease inhibitor cocktail (Roche cOmplete #11836153001). The resuspended cells were then transferred to tubes containing 500–700 mg of 0.1 mM glass beads (Sigma-Aldrich G8772). Bead beating was performed through three cycles, each consisting of 30 seconds of beating followed by a 5-minute pause on ice, at 4,260 RPM. Samples were centrifuged at $100 \times g$ for 2 minutes at 4°C , and supernatants were collected. DNA shearing was performed using a Covaris M220 instrument with the following parameters: 10% duty cycle, 200 cycles per burst, and 10 minutes duration. Insoluble materials were removed by centrifugation at $11,000 \times g$ for 10 minutes at 4°C , and the supernatant was collected. The efficiency of shearing was confirmed by electrophoresis on a 2% agarose gel. Anti-HA antibodies (3 μg per sample, Abcam, Ab#9110) were added to 25 μg of sheared DNA from each sample, followed by a 1-hour incubation at 4°C with rotation. Subsequently, 25 μL of prewashed magnetic beads (NEB-S1425S) were added to the samples, and incubation was continued overnight with rotation at 4°C . The next day, the samples were subjected to magnetic separation to remove liquids. The beads were washed five times with 900 μL of IPP50 buffer (10 mM Tris-HCl buffer, pH 8; 150 mM NaCl; and 0.1% NP-40), followed by two washes with TE buffer (50 mM Tris-HCl, pH 8; 10 mM EDTA). Beads–antibody–DNA complexes were collected by magnetic separation and then resuspended in 70 μL of elution buffer (50 mM Tris-HCl, pH 8; 1% SDS). This suspension was incubated at 65°C for 15 minutes. The elution step was repeated with an additional 70 μL , and the eluents were combined. RNA was removed from the eluted DNA using RNase A (NEB T3018-2, 20 mg/mL) at 37°C for 1 hour. Subsequently, reverse crosslinking was carried out by adding 3 μL of proteinase K (20 mg/mL, MAN-740506) and incubating for 2 hours at 37°C , followed by an overnight incubation at 65°C . The next morning, DNA was extracted using phenol:chloroform. Finally, the extracted DNA was resuspended in 500 μL of 10 mM Tris-HCl, pH 8.

Library preparation, sequencing, quality control, and peak analyses were conducted by the Technion Genomics Center, Haifa, Israel. 100 bp single-end reads were aligned to *Mycobacteroides abscessus* reference genome (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/069/185/GCA_000069185.1_ASM6918v1/GCA_000069185.1_ASM6918v1_genomic.fna.gz) and annotation file (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/069/185/GCA_000069185.1_ASM6918v1/GCA_000069185.1_ASM6918v1_genomic.gff.gz) using Bowtie2 v2.2.6 with default parameters, filtering reads for unique alignment using samtools v1.3. Peaks detection was done using MACS2 v2.1.1. Data are available on GEO, accession number [GSE287715](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287715).

Peaks were filtered to identify those that were present in samples from both strains expressing tagged MAB_1638 and absent in the two negative control strain samples. Each peak had start and end coordinates as well as a peak summit coordinate. The start and end coordinates were used to define peaks as being present in both MAB_1638-tagged strains, with 90% overlap required. Peaks that passed this filter were considered to be putative MAB_1638 binding sites (Table S2). Putative MAB_1638 binding sites were classified as associated with a gene if the peak summit was within 500 nt upstream or downstream of a TSS annotated as gene-associated in (24) (Table S3). For gene-associated putative MAB_1638 binding sites, we extracted 25 nucleotides upstream and 25 nucleotides downstream of the peak summit position. These sequences were input into MEME to discover *de novo* motifs (30) using the *M. abscessus* genome composition as the background. The discovered motif was then used as input for FIMO (46) to identify other potential motifs within the whole peak region. This step helps locate additional instances of the discovered motif, since the peak summit may not always represent the exact location of the MAB_1638 binding site.

ACKNOWLEDGMENTS

D.B. is supported by the Israeli Science Foundation IPMP grant (1349/20) and an individual grant (608/22). S.H.O. and M.P. were supported by the Singapore NMRC, under its OF-IRG program, Award OFIRG22jul-0081 to S.H.O. S.S.S. and J.X. were supported by NIH/NIAID grant 1P01AI143575-01A1. M.M. is supported by the Israeli Science Foundation 1730/22, and the CF Foundation (<https://www.cff.org/>) award, number 0055961223.

AUTHOR AFFILIATIONS

¹Department of Biology and Biotechnology and Program in Bioinformatics and Computational Biology, Worcester Polytechnic Institute, Worcester, Massachusetts, USA

²Koret School of Veterinary Medicine, The Robert E. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel

³A*STAR Infectious Diseases Labs (A*STAR ID Labs), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

⁴Mycobacteria Research Laboratories, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado, USA

⁵Ruth Rappaport Children's Hospital, Rambam Medical Center, Haifa, Israel

AUTHOR ORCIDS

Scarlet S. Shell  <http://orcid.org/0000-0003-1136-1728>

Juan Bellardinelli  <http://orcid.org/0000-0003-1336-9141>

Mary Jackson  <http://orcid.org/0000-0002-9212-0258>

Stefan H. Oehlers  <http://orcid.org/0000-0003-0260-672X>

Daniel Barkan  <http://orcid.org/0000-0002-5691-1757>

Michal Meir  <http://orcid.org/0000-0001-9979-6552>

AUTHOR CONTRIBUTIONS

Scarlet S. Shell, Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Michal Bar-Oz, Data curation, Investigation | Junpei Xiao, Data curation, Formal analysis, Investigation | Manitosh Pandey, Data curation, Investigation | Juan Bellardinelli, Data curation | Opeyemi I. Ibitoye, Data curation | Mary Jackson, Data curation | Stefan H. Oehlers, Data curation, Investigation | Daniel Barkan, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Writing – original draft, Writing – review and editing | Michal Meir, Data curation, Funding acquisition, Investigation, Supervision

DATA AVAILABILITY

RNAseq data and ChIPseq data are available on NCBI GEO, accession numbers [GSE287636](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287636) and [GSE287715](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287715), respectively.

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental Figures (mSystems00872-25-s0001.pdf). Fig. S1 to S3.

Table S1 (mSystems00872-25-s0002.xlsx). Transcriptomics comparison of *M. abscessus* ATCC19977 and *gplR1/mab_1638::tn*.

Table S2 (mSystems00872-25-s0003.xlsx). Putative GplR1/MAB_1638 binding sites identified by ChIP-seq.

Table S3 (mSystems00872-25-s0004.xlsx). Putative GplR1/MAB_1638 binding sites associated with genes and their sequence motifs.

REFERENCES

- Roux A-L, Viljoen A, Bah A, Simeone R, Bernut A, Laencina L, Deramaut T, Rottman M, Gaillard J-L, Majlessi L, Brosch R, Girard-Misguich F, Vergne I, de Chastellier C, Kremer L, Herrmann J-L. 2016. The distinct fate of smooth and rough *Mycobacterium abscessus* variants inside macrophages. *Open Biol* 6:160185. <https://doi.org/10.1098/rsob.160185>
- Catherinot E, Clarissou J, Etienne G, Ripoll F, Emile J-F, Daffé M, Perronne C, Soudais C, Gaillard J-L, Rottman M. 2007. Hypervirulence of a rough variant of the *Mycobacterium abscessus* type strain. *Infect Immun* 75:1055–1058. <https://doi.org/10.1128/IAI.00835-06>
- Bernut A, Herrmann JL, Kissa K, Dubremetz JF, Gaillard JL, Lutfalla G, Kremer L. 2014. *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess formation. *Proc Natl Acad Sci USA* 111:E943–E952. <https://doi.org/10.1073/pnas.1321390111>
- Howard ST, Rhoades E, Recht J, Pang X, Alsop A, Kolter R, Lyons CR, Byrd TF. 2006. Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. *Microbiology (Reading)* 152:1581–1590. <https://doi.org/10.1099/mic.0.28625-0>
- Byrd TF, Lyons CR. 1999. Preliminary characterization of a *Mycobacterium abscessus* mutant in human and murine models of infection. *Infect Immun* 67:4700–4707. <https://doi.org/10.1128/IAI.67.9.4700-4707.1999>
- Li B, Ye M, Zhao L, Guo Q, Chen J, Xu B, Zhan M, Zhang Y, Zhang Z, Chu H. 2020. Glycopeptidolipid genotype correlates with the severity of *Mycobacterium abscessus* lung disease. *J Infect Dis* 221:S257–S262. <https://doi.org/10.1093/infdis/jiz475>
- Jönsson BE, Gilljam M, Lindblad A, Ridell M, Wold AE, Welinder-Olsson C. 2007. Molecular epidemiology of *Mycobacterium abscessus*, with focus on cystic fibrosis. *J Clin Microbiol* 45:1497–1504. <https://doi.org/10.1128/JCM.02592-06>
- Nessar R, Reyat JM, Davidson LB, Byrd TF. 2011. Deletion of the *mmpL4b* gene in the *Mycobacterium abscessus* glycopeptidolipid biosynthetic pathway results in loss of surface colonization capability, but enhanced ability to replicate in human macrophages and stimulate their innate immune response. *Microbiology (Reading)* 157:1187–1195. <https://doi.org/10.1099/mic.0.046557-0>
- Kreutzfeldt KM, McAdam PR, Claxton P, Holmes A, Seagar AL, Laurenson IF, Fitzgerald JR. 2013. Molecular longitudinal tracking of *Mycobacterium abscessus* spp. during chronic infection of the human lung. *PLoS One* 8:e63237. <https://doi.org/10.1371/journal.pone.0063237>
- Park IK, Hsu AP, Tettelin H, Shallom SJ, Drake SK, Ding L, Wu UI, Adamo N, Prevots DR, Olivier KN, Holland SM, Sampaio EP, Zelazny AM. 2015. Clonal diversification and changes in lipid traits and colony morphology in *Mycobacterium abscessus* clinical isolates. *J Clin Microbiol* 53:3438–3447. <https://doi.org/10.1128/JCM.02015-15>
- Pawlik A, Garnier G, Orgeur M, Tong P, Lohan A, Le Chevalier F, Sapriel G, Roux A-L, Conlon K, Honoré N, Dillies M-A, Ma L, Bouchier C, Coppée J-Y, Gaillard J-L, Gordon SV, Loftus B, Brosch R, Herrmann JL. 2013. Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent *Mycobacterium abscessus*. *Mol Microbiol* 90:612–629. <https://doi.org/10.1111/mmi.12387>
- Steindor M, Nkwouano V, Stefanski A, Stuehler K, Ioegeer TR, Bogumil D, Jacobsen M, Mackenzie CR, Kalscheuer R. 2019. A proteomics approach for the identification of species-specific immunogenic proteins in the *Mycobacterium abscessus* complex. *Microbes Infect* 21:154–162. <https://doi.org/10.1016/j.micinf.2018.10.006>
- Pichler V, Dalkilic L, Shoab G, Shapira T, Rankine-Wilson L, Boudehen YM, Chao JD, Sexton D, Prieto M, Quon BS, Tocheva EI, Kremer L, Hsiao W, Av-Gay Y. 2024. The diversity of clinical *Mycobacterium abscessus* isolates in morphology, glycopeptidolipids and infection rates in a macrophage model. *J Med Microbiol* 73. <https://doi.org/10.1099/jmm.0.001869>
- Bar-Oz M, Martini MC, Alonso MN, Meir M, Lore NI, Miotto P, Riva C, Angala SK, Xiao J, Masiello CS, Misiakou MA, Sun H, Moy JK, Jackson M, Johansen HK, Cirillo DM, Shell SS, Barkan D. 2023. The small non-coding RNA B11 regulates multiple facets of *Mycobacterium abscessus* virulence. *PLoS Pathog* 19:e1011575. <https://doi.org/10.1371/journal.ppat.1011575>
- Foreman M, Gershoni M, Barkan D. 2020. A simplified and efficient method for Himar-1 transposon sequencing in bacteria, demonstrated by creation and analysis of a saturated transposon-mutant library in *Mycobacterium abscessus*. *mSystems* 5:e00976–20. <https://doi.org/10.1128/mSystems.00976-20>
- Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* 7:215–219. <https://doi.org/10.1038/73324>
- Aramaki H, Yagi N, Suzuki M. 1995. Residues important for the function of a multihelical DNA binding domain in the new transcription factor family of Cam and Tet repressors. *Protein Eng* 8:1259–1266. <https://doi.org/10.1093/protein/8.12.1259>
- Gutiérrez AV, Richard M, Roquet-Banères F, Viljoen A, Kremer L. 2019. The TetR family transcription factor MAB_2299c regulates the expression of two distinct MmpS-MmpL efflux pumps involved in cross-resistance to clofazimine and bedaquiline in *Mycobacterium abscessus*. *Antimicrob Agents Chemother* 63:e01000–19. <https://doi.org/10.1128/AAC.01000-19>
- Richard M, Gutiérrez AV, Viljoen A, Rodríguez-Rincon D, Roquet-Baneres F, Blaise M, Everall I, Parkhill J, Floto RA, Kremer L. 2019. Mutations in the MAB_2299c TetR regulator confer cross-resistance to clofazimine and bedaquiline in *Mycobacterium abscessus*. *Antimicrob Agents Chemother* 63:e01316–18. <https://doi.org/10.1128/AAC.01316-18>
- Rodríguez R, Campbell-Kruger N, Gonzalez Camba J, Berude J, Fetterman R, Stanley S. 2023. MarR-dependent transcriptional regulation of *mmpSL5* induces ethionamide resistance in *Mycobacterium abscessus*. *Antimicrob Agents Chemother* 67:e01350–22. <https://doi.org/10.1128/aac.01350-22>
- Richard M, Gutiérrez AV, Viljoen AJ, Ghigo E, Blaise M, Kremer L. 2018. Mechanistic and structural insights into the unique TetR-dependent regulation of a drug efflux pump in *Mycobacterium abscessus*. *Front Microbiol* 9:649. <https://doi.org/10.3389/fmicb.2018.00649>
- Kam JY, Horte E, Krogman E, Warner SE, Wright K, Luo K, Cheng T, Manuneechi Cholan P, Kikuchi K, Triccas JA, Britton WJ, Johansen MD, Kremer L, Oehlers SH. 2022. Rough and smooth variants of *Mycobacterium abscessus* are differentially controlled by host immunity during chronic infection of adult zebrafish. *Nat Commun* 13:952. <https://doi.org/10.1038/s41467-022-28638-5>
- Leclercq LD, Le Moigne V, Daher W, Cortes M, Viljoen B, Tasrini Y, Trivelli X, Lavanant H, Schmitz-Afonso I, Durand N, Biet F, Guérardel Y, Kremer L, Herrmann J-L. 2025. A glycosylated lipopeptide promotes uptake and growth of *Mycobacterium abscessus* in the host. *Nat Commun* 16:3326. <https://doi.org/10.1038/s41467-025-58455-5>
- Miranda-Casoluengo AA, Staunton PM, Dinan AM, Lohan AJ, Loftus BJ. 2016. Functional characterization of the *Mycobacterium abscessus* genome coupled with condition specific transcriptomics reveals conserved molecular strategies for host adaptation and persistence. *BMC Genomics* 17:553. <https://doi.org/10.1186/s12864-016-2868-y>
- Ball AS, van Kessel JC. 2019. The master quorum-sensing regulators LuxR/HapR directly interact with the alpha subunit of RNA polymerase to drive transcription activation in *Vibrio harveyi* and *Vibrio cholerae*. *Mol Microbiol* 111:1317–1334. <https://doi.org/10.1111/mmi.14223>
- Christen S, Srinivas A, Bähler P, Zeller A, Primrose D, Bieniossek C, Baumann U, Erni B. 2006. Regulation of the Dha operon of *Lactococcus lactis*: a deviation from the rule followed by the TetR family of transcription regulators. *J Biol Chem* 281:23129–23137. <https://doi.org/10.1074/jbc.M603486200>
- Jiang M, Yin M, Wu S, Han X, Ji K, Wen M, Lu T. 2017. GdmR11, a TetR family transcriptional regulator, controls geldanamycin and elaiophyllin biosynthesis in *Streptomyces autolyticus* CGMCC0516. *Sci Rep* 7:4803. <https://doi.org/10.1038/s41598-017-05073-x>
- Nguyen Le Minh P, de Cima S, Bervoets I, Maes D, Rubio V, Charlier D. 2015. Ligand binding specificity of RutR, a member of the TetR family of transcription regulators in *Escherichia coli*. *FEBS Open Bio* 5:76–84. <https://doi.org/10.1016/j.fob.2015.01.002>
- Shimada T, Hirao K, Kori A, Yamamoto K, Ishihama A. 2007. RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines. *Mol Microbiol* 66:744–757. <https://doi.org/10.1111/j.1365-2958.2007.05954.x>
- Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. *Nucleic Acids Res* 43:W39–W49. <https://doi.org/10.1093/nar/gkv416>
- Hedin W, Fröberg G, Fredman K, Chrysanthou E, Selmerlyd I, Gillman A, Orsini L, Runold M, Jönsson B, Schön T, Davies Forsman L. 2023. A rough colony morphology of *Mycobacterium abscessus* is associated with

- cavitary pulmonary disease and poor clinical outcome. *J Infect Dis* 227:820–827. <https://doi.org/10.1093/infdis/jiad007>
32. Rhoades ER, Archambault AS, Greendyke R, Hsu FF, Streeter C, Byrd TF. 2009. *Mycobacterium abscessus* glycopeptidolipids mask underlying cell wall phosphatidyl-myo-inositol mannosides blocking induction of human macrophage TNF- α by preventing interaction with TLR2. *J Immunol* 183:1997–2007. <https://doi.org/10.4049/jimmunol.0802181>
 33. Rao V, Gao F, Chen B, Jacobs WR, Glickman MS. 2006. *Trans*-cyclopropanation of mycolic acids on trehalose dimycolate suppresses *Mycobacterium tuberculosis* –induced inflammation and virulence. *J Clin Invest* 116:1660–1667. <https://doi.org/10.1172/JCI27335>
 34. Ahn SK, Cuthbertson L, Nodwell JR. 2012. Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators. *PLoS One* 7:e50562. <https://doi.org/10.1371/journal.pone.0050562>
 35. Yang M, Gao CH, Hu J, Zhao L, Huang Q, He ZG. 2015. InbR, a TetR family regulator, binds with isoniazid and influences multidrug resistance in *Mycobacterium bovis* BCG. *Sci Rep* 5:13969. <https://doi.org/10.1038/srep13969>
 36. Kendall SL, Withers M, Soffair CN, Moreland NJ, Gurcha S, Sidders B, Frita R, Ten Bokum A, Besra GS, Lott JS, Stoker NG. 2007. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol Microbiol* 65:684–699. <https://doi.org/10.1111/j.1365-2958.2007.05827.x>
 37. Kendall SL, Burgess P, Balhana R, Withers M, Ten Bokum A, Lott JS, Gao C, Uhia-Castro I, Stoker NG. 2010. Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: *kstR* and *kstR2*. *Microbiology (Reading)* 156:1362–1371. <https://doi.org/10.1099/mic.0.034538-0>
 38. Wei W, Jiang X, Zhang L, Yan Y, Yan J, Xu L, Gao CH, Yang M. 2023. Regulation of CRISPR-associated genes by Rv1776c (CasR) in *Mycobacterium tuberculosis*. *Biomolecules* 13:400. <https://doi.org/10.3390/biom13020400>
 39. Jiang Q, Hu R, Liu F, Huang F, Zhang L, Zhang H. 2024. Characterization of a novel oxidative stress responsive transcription regulator in *Mycobacterium bovis*. *Biomedicines* 12:1872. <https://doi.org/10.3390/biomedicines12081872>
 40. Li W, Li M, Hu L, Zhu J, Xie Z, Chen J, He ZG. 2018. HpoR, a novel c-di-GMP effective transcription factor, links the second messenger's regulatory function to the mycobacterial antioxidant defense. *Nucleic Acids Res* 46:3595–3611. <https://doi.org/10.1093/nar/gky146>
 41. Yang M, Gao C, Cui T, An J, He ZG. 2012. A TetR-like regulator broadly affects the expressions of diverse genes in *Mycobacterium smegmatis*. *Nucleic Acids Res* 40:1009–1020. <https://doi.org/10.1093/nar/gkr830>
 42. Liu H, Yang M, He ZG. 2016. Novel TetR family transcriptional factor regulates expression of multiple transport-related genes and affects rifampicin resistance in *Mycobacterium smegmatis*. *Sci Rep* 6:27489. <https://doi.org/10.1038/srep27489>
 43. Schleif R. 2000. Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet* 16:559–565. [https://doi.org/10.1016/s0168-9525\(00\)02153-3](https://doi.org/10.1016/s0168-9525(00)02153-3)
 44. Schleif R. 2010. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol Rev* 34:779–796. <https://doi.org/10.1111/j.1574-6976.2010.00226.x>
 45. Kołodziej M, Łebkowski T, Płociński P, Hołówka J, Paściak M, Wojtaś B, Bury K, Konieczny I, Dziadek J, Zakrzewska-Czerwińska J. 2021. Lsr2 and its novel paralogue mediate the adjustment of *Mycobacterium smegmatis* to unfavorable environmental conditions. *mSphere* 6:e00290-21. <https://doi.org/10.1128/mSphere.00290-21>
 46. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27:1017–1018. <https://doi.org/10.1093/bioinformatics/btr064>
 47. Matty MA, Oehlers SH, Tobin DM. 2016. Live imaging of host-pathogen interactions in zebrafish larvae. *Methods Mol Biol* 1451:207–223. https://doi.org/10.1007/978-1-4939-3771-4_14
 48. Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
 49. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>
 50. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
 51. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *The Innovation* 2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>
 52. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. 2009. QuickGO: a web-based tool for gene ontology searching. *Bioinformatics* 25:3045–3046. <https://doi.org/10.1093/bioinformatics/btp536>
 53. Minch KJ, Rustad TR, Peterson EJ, Winkler J, Reiss DJ, Ma S, Hickey M, Brabant W, Morrison B, Turkarslan S, Mawhinney C, Galagan JE, Price ND, Baliga NS, Sherman DR. 2015. The DNA-binding network of *Mycobacterium tuberculosis*. *Nat Commun* 6:5829. <https://doi.org/10.1038/ncomms6829>