A Conserved Mycobacterial Nucleomodulin Hijacks the Host COMPASS Complex to Reprogram Pro-Inflammatory Transcription and Promote Intracellular Survival

Liu Chen[1,2], Baojie Duan[1], Pingping Chen[1], Qiang Jiang[1], Yifan Wang[1], Lu Lu[1], Yingyu Chen[1,3,4], Changmin Hu[1,3], Lei Zhang[1,3,4]* and Aizhen Guo[1,2,3,4]*

[1] National Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China

[2] Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, China

[3] National Professional Laboratory for Animal Tuberculosis, Ministry of Agriculture and Rural Affairs, Huazhong Agricultural University, Wuhan 430070, China

[4] Hubei Jiangxia Laboratory, Wuhan 430207, China

*Address correspondence to: Lei Zhang, zhanglei2023@mail.hzau.edu.cn; Aizhen Guo, aizhen@mail.hzau.edu.cn

ABSTRACT

Nucleomodulins are a class of effector proteins secreted by bacterial pathogens that translocate into the host cell nucleus to modulate nuclear processes. However, their target proteins and underlying molecular mechanisms remain poorly understood in mycobacteria. Herein, we identified a conserved hypothetical protein Rv1075c, designated MgdE, as a nucleomodulin that enhances mycobacterial intracellular survival. MgdE undergoes nuclear translocation via two nuclear localization signals, KRIR$^{108-111}$ and RLRRPR$^{300-305}$, and interacts with ASH2L and WDR5, two subunits of the host histone methyltransferase COMPASS complex. This interaction suppresses histone H3 lysine 4 (H3K4) methylation-mediated transcription of pro-inflammatory genes, including *IL6* and *IL1B*, thereby promoting mycobacterial survival in both macrophages and mice. Our study provides the first experimental evidence that a bacterial nucleomodulin facilitates intracellular survival by directly targeting the host COMPASS complex. These findings advance our understanding of mycobacterial pathogenesis by revealing a novel mechanism that contributes to its intracellular survival strategy.

Keywords Mycobacteria; Nucleomodulin; Histone methyltransferase; Inflammatory genes; Pathogenesis.

INTRODUCTION

Nucleomodulins constitute a distinct class of bacterial proteins capable of translocating into the nuclei of eukaryotic cells, where they strategically modulate nuclear processes to promote pathogen survival and persistence (Radoshevich et al., 2018; Hanford et al., 2021). These

evolutionarily diverse effector proteins, secreted by distinct pathogens, exhibit remarkable adaptability in infiltrating host nuclei and reprogramming the transcriptional and epigenetic regulatory networks (Bierne et al., 2012; Bierne et al., 2020). Upon nuclear localization, nucleomodulins employ diverse molecular mechanisms to subvert host cellular processes, including but not limited to: (i) direct interactions with genomic DNA or nuclear proteins (Rennoll-Bankert et al., 2015; Sun et al., 2016; Prokop et al., 2017), (ii) hijacking post-translational modification systems (Lebreton et al., 2014; Rolando et al., 2023), (iii) dysregulating signaling molecule activity or localization (Zhao et al., 2019; Chambers et al., 2020; Evans et al., 2018), and (iv) destabilizing nuclear architecture (Pourpre et al., 2022; Fu et al., 2021). Specific examples include AnkA, a bacterial protein from *Anaplasma phagocytophilum,* which binds to specific A/T-rich sequences in the host cell, alters chromatin architecture, and modulates gene transcription (Rennoll-Bankert et al., 2015). In *Listeria monocytogenes*, LntA interacts with the proline-rich domain of BAHD1, to disrupt BAHD1-mediated gene silencing and promote histone H3 acetylation (Lebreton et al., 2014). Another example is OspF from *Shigella flexneri*, which exploits its phosphothreonine lyase activity to irreversibly dephosphorylate MAPKs, thereby preventing the transcriptional activation of NF-κB-regulated genes (Chambers et al., 2020). Unlike these effectors, *L. monocytogenes* InlP takes a different approach by hijacking the splicing factor RBM5. This interaction not only suppresses apoptosis but also reorganizes nuclear speckles through SC35 redistribution and RBM5-dependent granule formation (Pourpre et al., 2022). These examples illustrate the diverse and sophisticated mechanisms used by nucleomodulins to reprogram host nuclear functions, ultimately manipulating the intracellular environment to their advantage.

The evolutionary refinement of nucleomodulins is exemplified by their ability to exploit chromatin-based regulatory systems, with histone methylation emerging as a pivotal axis in host–pathogen interactions (Bierne et al., 2020; Khan et al., 2021). As a central epigenetic modification, histone methylation dynamically controls chromatin accessibility and transcriptional outcomes (Yu et al., 2024; Wang et al., 2025). By precisely targeting specific histone residues, nucleomodulins reprogram chromatin states to either suppress host immune responses or promote cellular pathways conducive to pathogen survival (Lebreton et al., 2014; Schator et al., 2023; Rolando et al.,2023; Denzer et al., 2020). A paradigm of this molecular strategy is demonstrated by the *Chlamydia trachomatis* nuclear effector NUE, which utilizes its conserved SET domain to catalyze the methylation of histones H2B, H3, and H4, establishing a repressive chromatin state that downregulates antimicrobial gene expression (Fol et al., 2020). Similarly, RomA uses its ankyrin and SET domains to bind and methylate histone H3, thereby altering the host chromatin to enhance *Legionella* intracellular survival (Schator et al., 2023; Rolando et al., 2013). *Legionella* LegAS4 catalyzes H3K4 methylation to promote an open chromatin state, enhancing ribosomal RNA transcription and facilitating bacterial intracellular replication (Denzer et al., 2020). These cases emphasize how pathogens exploit histone methylation as a molecular lever to hijack the host epigenetic machinery.

Mycobacterium tuberculosis (Mtb), the causative agent of tuberculosis, is a highly successful pathogen because of its ability to survive within macrophages, evade the immune system, and persist in the host, ultimately leading to a chronic disease state (Chandra et al., 2022; Chai et al., 2020). Central to its pathogenic strategy is the secretion of a diverse array of effector proteins (Chai et al., 2022; Bates et al., 2024; Qiang et al., 2023), including

66 nucleomodulins that subvert host nuclear processes. The tyrosine phosphatase PtpA
67 exemplifies this strategy by directly binding to the promoter regions of immune-related genes,
68 including *GADD45A*, to suppress innate immune responses (Wang et al., 2017). The
69 methyltransferase Rv1988 mediates an unconventional histone modification by methylating
70 histone H3 at arginine 42 (H3R42me), thereby silencing antimicrobial defense mechanisms
71 (Yaseen et al., 2015). Similarly, the acetyltransferase Rv3423.1 modulates chromatin
72 accessibility through acetylation of histone H3 at lysine 9 and 14 (H3K9/K14ac), attenuating
73 pro-inflammatory signaling pathways (Jose et al., 2016). Rv2966c represents a unique
74 bacterial DNA methyltransferase that establishes non-canonical CpG methylation patterns to
75 epigenetically silence host immune-related gene expression (Sharma et al., 2015). Despite
76 these molecular insights, the comprehensive landscape of the nuclear effector systems of Mtb,
77 particularly the mechanistic details of nucleomodulin interactions with host chromatin-
78 modifying complexes and their spatiotemporal regulation during infection, remains poorly
79 defined. This critical knowledge gap impedes our understanding of how Mtb coordinates
80 system-level transcriptional rewiring in the host cellular environment.

81 In this study, we utilized bioinformatics analysis combined with fluorescence imaging to
82 rapidly screen for potential nucleomodulins in pathogenic mycobacterial species. Using this
83 approach, the hypothetical protein Rv1075c (designated as MgdE) was identified and
84 subsequently validated. MgdE undergoes nuclear translocation mediated by two nuclear
85 localization signals, $KRIR^{108-111}$ and $RLRRPR^{300-305}$, and interacts with the histone
86 methyltransferase COMPASS complex subunits ASH2L and WDR5. This interaction
87 suppresses H3K4me3-mediated inflammatory gene expression, thus dampening host immune

88 responses and enhancing bacterial survival in macrophages and murine infection models. Our
89 findings identify MgdE as a nucleomodulin that targets the COMPASS complex, revealing a
90 novel epigenetic strategy exploited by mycobacteria for intracellular survival.

91 **RESULTS**

92 **Functional screening of potential nucleomodulins in pathogenic mycobacterial species**

93 To identify conserved nucleomodulins in pathogenic mycobacteria, we first conducted a
94 comparative pan-genome analysis across four relevant species, including *M. tuberculosis*
95 H37Rv and H37Ra, *M. bovis* BCG, *M. marinum*, and *M. avium*, to categorize the core genes.
96 Subsequently, a dual-threshold screening approach was employed to predict the classical and
97 non-canonical secreted proteins encoded by these shared genes using SignalP 5.0 and
98 SecretomeP 2.0, with thresholds set at a D-score ≥0.5 (Hammond et al., 2018) and a neural
99 network (NN) score ≥0.9 (Bendtsen et al., 2005) (Table S1). This analysis yielded 135 high
100 confidence secreted protein candidates. Further analysis of these candidate nuclear
101 localization signal (NLS) motifs using cNLS Mapper (NLS score ≥2) identified 56 proteins
102 containing at least one putative NLS (Table S2). Following systematic classification based on
103 secretion machinery dependencies, the candidate proteins were categorized into four distinct
104 pathways: 26 Sec/SPI (signal peptide-driven), 7 Tat/SPI (twin-arginine translocase pathway),
105 18 Sec/SPII (lipoprotein-targeted), and 5 unclassified candidates lacking canonical secretion
106 motifs (Figure 1A and B, Figure S1A and B). Functional validation of nuclear trafficking was
107 conducted through the heterologous expression of enhanced green fluorescent protein
108 (EGFP)-tagged constructs in HEK293T cells, followed by quantitative profiling of the
109 subcellular distribution using laser scanning confocal microscopy (LSCM). Among them, six

110 Tat/SPI-associated proteins including Rv0519c, Rv0846c, Rv0945, Rv1075c/MgdE, Rv1291c
111 and Rv2577 (Table S3) exhibited nuclear localization and MgdE exhibited distinct nuclear
112 aggregation (Figure 1C). Nucleomodulins are effector proteins secreted by pathogens that
113 translocate into the host cell nucleus to modulate host functions. To examine whether MgdE
114 is secreted by mycobacteria, recombinant *M. bovis* BCG strains expressing C-terminally Flag-
115 tagged MgdE were constructed. Immunoblotting showed that MgdE and the secreted antigen
116 Ag85B were present in both cell pellets and culture supernatants, while the cytoplasmic protein
117 GlpX as the negative control was only detected in the cell pellets (Figure S2A), indicating that
118 MgdE is indeed secreted by *M. bovis* BCG. Collectively, these results suggest that MgdE might
119 be a potential nucleomodulin in the pathogenic mycobacteria.

120 **MgdE translocates into the host nucleus through its dual NLS**

121 MgdE is a putative GDSL-like lipase characterized by conserved esterase sequence motifs,
122 with the GDSx motif containing a nucleophilic Ser residue as part of block I (equivalent to the
123 classical G-x-S-x-G motif of lipases/esterases), and the consensus amino acids Gly, Asn, and
124 His located in blocks II, III, and V, respectively (Yang et al., 2019). Phylogenetic analysis
125 across the mycobacterial species revealed strict evolutionary conservation of MgdE, with most
126 strains encoding two predicted NLS: KRIR$^{108\text{-}111}$ (NLS1) and RLRRPR$^{300\text{-}305}$ (NLS2) (Figure 2A
127 and B). To further investigate the nuclear translocation capability of MgdE, HEK293T cells
128 were transfected with MgdE-EGFP, and its subcellular localization was examined at multiple
129 time points post-transfection. Confocal microscopy revealed a time-dependent increase of
130 MgdE-EGFP in the nucleus by 48 hours post-transfection (Figure S2B and C). To
131 quantitatively assess this process, MgdE expression and nuclear enrichment were analyzed

132 by cell fractionation and immunoblotting. These results demonstrated that MgdE protein levels
133 increased over time and progressively accumulated in the nucleus, consistent with the
134 confocal imaging observations (Figure S2D–E). Taken together, these experiments indicate
135 that MgdE is an evolutionarily conserved mycobacterial nucleomodulin with nuclear
136 translocation capacity.

137 To dissect the functional contributions of the predicted NLS motifs, EGFP-tagged NLS
138 deletion mutants including MgdE$^{\Delta NLS1}$, MgdE$^{\Delta NLS2}$, and MgdE$^{\Delta NLS1\text{-}2}$ were constructed and
139 transfected into HEK293T cells. Fluorescence intensity profiling revealed that individual
140 deletion of either NLS1 or NLS2 partially reduced the nuclear targeting efficiency, whereas
141 simultaneous deletion of both the NLS motifs (MgdE$^{\Delta NLS1\text{-}2}$) completely abrogated its nuclear
142 localization (Figure 2C and D). These observations were further confirmed by immunoblot
143 analysis of nuclear fractions, which showed that at 36 h post-transfection, MgdE-EGFP and
144 its mutants were consistently detected in the cytoplasmic compartment, and nuclear
145 accumulation of MgdE$^{\Delta NLS1\text{-}2}$ was nearly undetectable compared to that of MgdE (Figure 2E,
146 Figure S2F). To determine whether MgdE translocates into the host nucleus during
147 mycobacterial infection, recombinant BCG strains expressing Flag-tagged MgdE mutants
148 were constructed, including single and double NLS deletion mutants (BCG/MgdE$^{\Delta NLS1}$-Flag,
149 BCG/MgdE$^{\Delta NLS2}$-Flag, and BCG/MgdE$^{\Delta NLS1\text{-}2}$-Flag). THP-1 macrophages were infected with
150 the recombinant strains, followed by subcellular fractionation. Nuclear-cytoplasmic
151 fractionation experiments showed that WT MgdE and the NLS single mutants could be
152 detected both in the cytoplasm and in the nucleus by immunoblotting, while the double mutant
153 MgdE$^{\Delta NLS1\text{-}2}$ was detectable only in the cytoplasm (Figure 2F). These results indicate that

154 MgdE translocates into the host nucleus via a dual NLS mechanism, with both NLS1 and NLS2
155 required for efficient nuclear import.

156 **Nuclear localization of MgdE enhances mycobacterial intracellular survival within**
157 **macrophages**

158 Given that MgdE exhibits nuclear translocation ability, we sought to determine whether this
159 ability plays a role in bacterial survival during infection. We generated various recombinant *M.*
160 *bovis* BCG strains, including wild-type BCG (WT), a *mgdE* deletion mutant (ΔMgdE) (Figure
161 S3A and B), ΔMgdE complemented with WT *mgdE* (Comp-MgdE), and ΔMgdE complemented
162 with NLS-deleted mutant (Comp-MgdE$^{\Delta NLS1\text{-}2}$). All strains exhibited similar growth in 7H9 liquid
163 media supplemented with OADC (Figure S3C). To assess intracellular survival, we infected
164 THP-1 macrophages with each strain and monitored bacterial burden over time. As shown in
165 Figure 3A, the ΔMgdE mutant strain exhibited significantly reduced survival compared to the
166 WT strain, with accelerated bacterial clearance over time. This defect was rescued in the
167 Comp-MgdE strain, with intracellular survival levels restored to that in the WT strain. Notably,
168 the Comp-MgdE$^{\Delta NLS1\text{-}2}$ showed bacterial clearance equivalent to that in the ΔMgdE strains.
169 These results were consistent with the infection experiments using RAW264.7 macrophage
170 (Figure 3B). Furthermore, THP-1 macrophages infected with the ΔMgdE mutant strain
171 exhibited significant upregulation of inflammatory cytokines, including *IL1B*, *IL6*, *IL10,* and the
172 colony-stimulating factors *CSF1–CSF3* compared to the macrophages infected with the WT
173 strain (Figure 3C–H). Collectively, these results demonstrated that (i) MgdE enhances
174 mycobacterial intracellular survival and suppresses host inflammatory responses during
175 infection, and (ii) both NLS1 and NLS2 motifs are essential for the nuclear trafficking-

176 dependent function.

177 **MgdE physically interacts with the host COMPASS complex subunits ASH2L and WDR5**

178 To identify the potential host target proteins of MgdE, we employed AlphaFold to predict the
179 interactions between MgdE and majority of the nuclear proteins in the host. The prediction
180 results showed that MgdE binds to WDR5, a core component of the histone methyltransferase
181 COMPASS complex, with high confidence (pLDDT = 0.77) (Figure S4A and B). To
182 experimentally validate this prediction, we performed yeast two-hybrid (Y2H) assays. The
183 yeast cells co-expressing MgdE with either ASH2L or WDR5 exhibited robust growth on
184 selective SD/-Trp-Leu-His-Ade medium, whereas the cells co-expressing MgdE with DPY30
185 or RbBP5 showed no detectable growth, indicating specific interactions between MgdE and
186 ASH2L/WDR5 (Figure 4A). Further, these interactions were validated by co-
187 immunoprecipitation (Co-IP) assays. HEK293T cells were co-transfected with Flag-tagged
188 MgdE and HA-tagged ASH2L, WDR5, or RbBP5. The results revealed that HA-tagged ASH2L
189 and WDR5, but not RbBP5, co-precipitated with Flag-tagged MgdE (Figure 4B). Conversely,
190 Flag-tagged MgdE precipitated HA-tagged ASH2L and WDR5, but not RbBP5 (Figure 4C).
191 Taken together, these results demonstrate that MgdE specifically interacts with the
192 COMPASS complex subunits ASH2L and WDR5.

193 **The conserved residues D224 and H247 are critical for mediating the binding of MgdE**
194 **to WDR5**

195 To further confirm the interactions between MgdE and WDR5/ASH2L, the key amino acids of
196 MgdE (S80, D244, and H247) were individually replaced with alanine (A) (Yang et al., 2019).

197 The Y2H assays showed that single-residue substitutions of MgdE did not affect its interaction
198 with ASH2L or WDR5. However, the D244A/H247A double mutation specifically abrogated its
199 interaction with WDR5 but retained its ASH2L binding ability (Figure 5A). These findings were
200 further confirmed by Co-IP assays in HEK293T cells. Flag-tagged MgdE precipitated HA-
201 tagged ASH2L and WDR5 but not RbBP5, with reciprocal pull-downs confirming bidirectional
202 binding specificity. The D244A/H247A double mutant maintained its interaction with ASH2L
203 but failed to bind WDR5, suggesting that these interactions are mediated through distinct
204 molecular interfaces (Figure 5B, Figure S4C). To further investigate the functional
205 consequences of these mutations, we analyzed the subcellular localization of the MgdE
206 mutants in transfected HEK293T cells using fluorescence co-localization assays. Wild-type
207 MgdE exhibited distinct nuclear aggregation over time, a phenomenon that was abrogated
208 following transfection with the D244A/H247A mutant (Figure 5C), suggesting that the double
209 mutation impairs MgdE activity.

210 Given that MgdE interacts with WDR5, a subunit of COMPASS complex, which is a key
211 regulator of H3K4me3 methylation (Deng et al., 2024), we sought to determine whether this
212 interaction contributes to an epigenetic regulatory mechanism that facilitates bacterial
213 persistence within host cells. HEK293T cells were transfected with EGFP-tagged wild-type
214 MgdE (WT), the catalytic mutant MgdE-D244A/H247A, or EGFP protein. Immunoblot analysis
215 of nuclear extracts showed that cells expressing WT MgdE had ~25% lower H3K4me3 levels
216 than EGFP-expressing cells and ~40% lower levels than those expressing the D244A/H247A
217 mutant (Figure 5D). Taken together, these data indicate that the residues D244 and H247 are
218 essential for the MgdE-WDR5 interaction, which enables MgdE to attenuate H3K4me3

219 deposition in the host nuclei by subverting COMPASS complex function, a mechanism critical
220 for bacterial survival and immune evasion.

221 **MgdE suppresses host inflammatory responses probably by inhibition of COMPASS**
222 **complex-mediated H3K4 methylation**

223 As H3K4me3, a hallmark of transcriptionally active promoters, facilitates transcriptional
224 machinery assembly and RNA polymerase II recruitment (Wang et al., 2023), we further
225 investigated the transcriptional consequences of MgdE-mediated H3K4me3 suppression via
226 its interaction with the COMPASS complex. To assess the genome-wide effects of MgdE on
227 host gene expression, we performed RNA-seq profiling of THP-1 macrophages infected with
228 either WT BCG or ΔMgdE strains. Transcriptomic analysis at 24 h post-infection revealed that
229 271 genes showed significant differential expression in the cells infected with ΔMgdE strain
230 compared to those infected with the WT BCG strain, with 222 genes being transcriptionally
231 upregulated (Figure 6A). Gene ontology (GO) enrichment analysis indicated that these
232 upregulated genes were primarily associated with biological processes involved in immune
233 responses, with a notable enrichment in "Positive regulation of cytokine production (GO:
234 0001819)", a critical pathway for amplifying inflammatory signaling (Figure 6B). Furthermore,
235 the Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analysis
236 highlighted a significant activation of the "Cytokine–cytokine receptor interaction (hsa04060)"
237 and "JAK–STAT signaling pathway (hsa04630)" in THP-1 cells infected with the ΔMgdE strain
238 (Figure 6C, Figure S5A). Heatmap analysis further confirmed the significant upregulation of
239 inflammation-associated transcripts in ΔMgdE-infected cells, including inflammatory cytokines
240 (*IL6*, *IL7*, *IL10*, *IL12B*) and colony-stimulating factors (*CSF2*–*CSF3*) (Figure 6D, Figure S5B).

241 Protein–protein interaction (PPI) network analysis of the upregulated genes revealed tightly
242 interconnected clusters within inflammatory networks, particularly those orchestrated by the
243 central inflammatory mediator IL-6 (Figure S5C). To validate this result, we infected
244 macrophages with the *M. bovis* BCG strains, including WT, ΔMgdE, Comp-MgdE, and Comp-
245 MgdE$^{\Delta NLS1\text{-}2}$, under identical conditions for 24 h and examined the expression of relevant
246 inflammatory factors. Compared to cells infected with the WT and Comp-MgdE strains, those
247 infected with the ΔMgdE and Comp-MgdE$^{\Delta NLS1\text{-}2}$ strains exhibited significantly increased
248 expression levels of inflammatory factors (*IL1B*, *IL6*, *IL10*) (Figure 6E–G). Collectively, these
249 results demonstrate that MgdE effectively inhibits the production of inflammatory factors
250 during BCG infection and highlights the critical role of its NLS in this process. Through its
251 interaction with the COMPASS complex, MgdE suppresses H3K4me3 levels, thereby
252 attenuating the transcriptional activation of inflammation-related genes.

253 **Nuclear localization of MgdE is essential for mycobacterial survival in mice**

254 Our experimental data demonstrated that MgdE enhances bacterial survival in macrophages.
255 To further investigate the *in vivo* functional significance of MgdE in host–pathogen interplay,
256 we infected C57BL/6 mice with different *M. bovis* BCG strains, including WT, ΔMgdE, Comp-
257 MgdE, and Comp-MgdE$^{\Delta NLS1\text{-}2}$, and comprehensively evaluated bacterial survival, pulmonary
258 inflammation, and cytokine dynamics. Mice infected with ΔMgdE exhibited a significantly
259 reduced lung bacterial burden compared to those infected with WT or Comp-MgdE strains.
260 Moreover, mice infected with Comp-MgdE$^{\Delta NLS1\text{-}2}$ exhibited a significantly lower pulmonary
261 bacterial load than those infected with Comp-MgdE (Figure 7A). These findings collectively
262 establish that MgdE is essential for optimal bacterial survival during infection, with the NLS

263 being critical for its full functionality. Furthermore, histopathological assessment of the lung
264 tissues corroborated these observations. Hematoxylin–eosin (H&E) staining revealed more
265 pronounced inflammatory pathology in the lungs of mice infected with WT or Comp-MgdE
266 strain compared to those infected with ΔMgdE or Comp-MgdE$^{\Delta NLS1\text{-}2}$ (Figure 7B), underscoring
267 the necessity of MgdE and its NLS in sustaining inflammation during infection. To further
268 investigate the systemic immune modulation mediated by MgdE, we quantified splenic
269 bacterial colonization and the inflammatory cytokine profiles. Consistent with the attenuated
270 bacterial survival observed in the lungs, mice infected with ΔMgdE and Comp-MgdE$^{\Delta NLS1\text{-}2}$
271 strains exhibited a significantly reduced splenic bacterial load compared to those infected with
272 WT or Comp-MgdE (Figure S6). Despite diminished bacterial dissemination, infections with
273 ΔMgdE and Comp-MgdE$^{\Delta NLS1\text{-}2}$ strains triggered elevated expression of inflammatory
274 cytokines (*Il6, Il1b*) relative to infection with the WT and Comp-MgdE strains (Figure 7C and
275 D). Collectively, these findings indicate that MgdE orchestrates mycobacterial survival within
276 host tissues by suppressing inflammatory responses to evade pathogen clearance, with the
277 NLS playing a pivotal role in mediating these host–pathogen interactions.

278 **DISCUSSION**

279 **MgdE functions as a critical nucleomodulin that contributes to mycobacterial**
280 **pathogenesis.** Its evolutionary conservation across mycobacterial species, coupled with
281 nuclear translocation via dual N-terminal and C-terminal NLS, supports its classification as a
282 nucleomodulin. Previous investigations have described MgdE as a secreted effector protein
283 (Penn et al., 2018), and our experimental results confirm its secretion (Figure S2A).
284 Furthermore, our findings demonstrate that MgdE's nuclear localization plays a critical role

285 during infection. Disruption of MgdE expression or its NLS motifs significantly attenuated *M.*
286 *bovis* BCG viability within macrophages (Figure 3A and B) and reduced bacterial burden in
287 mice (Figure 7A, Figure S6), supporting the notion that nuclear targeting contributes to
288 bacterial virulence. This phenotype aligns with transcriptomic data showing upregulation of
289 MgdE during *M. tuberculosis* infection of alveolar macrophages (Pisu et al., 2020). In parallel,
290 Yang et al. (2019) characterized MgdE as a lipase involved in intracellular lipid metabolism
291 and showed that its transposon mutant exhibited impaired intracellular growth in THP-1 cells
292 and PBMC-derived macrophages, as well as a moderately reduced bacterial load in
293 C3HeB/FeJ mice at later stages of infection. These findings align with our data and highlight
294 the multifaceted role of MgdE in both metabolic adaptation and host-cell interaction. Together,
295 these findings suggest a functional duality in which MgdE contributes to both metabolic
296 adaptation and host nuclear modulation. By coupling enzymatic activity with nuclear targeting,
297 MgdE may enable Mtb to balance metabolic needs and immune evasion.

298 **MgdE directly subverts the COMPASS complex to reprogram host epigenetics.** To our
299 knowledge, MgdE is the first identified bacterial effector that directly targets the COMPASS
300 complex, a central regulator of H3K4 methylation. Nuclear-localized MgdE specifically
301 interacts with the core COMPASS subunits ASH2L and WDR5, disrupting H3K4me3
302 deposition through structural interference (Guarnaccia et al., 2021; Hsu et al., 2018; Xue et
303 al., 2019). The COMPASS complex facilitates H3K4 methylation through a conserved
304 assembly mechanism involving multiple core subunits. WDR5, a central scaffolding
305 component, interacts with RbBP5 and ASH2L to promote complex assembly and enzymatic
306 activity (Qu et al., 2018; Wysocka et al., 2005). It also recognizes the WIN motif of

307 methyltransferases such as MLL1, thereby anchoring them to the complex and stabilizing the
308 ASH2L–RbBP5 dimer (Hsu et al., 2018). ASH2L further contributes to COMPASS activation
309 by interacting with both RbBP5 and DPY30 and by stabilizing the SET domain, which is
310 essential for efficient substrate recognition and catalysis (Qu et al., 2018; Park et al., 2019).
311 Our work shows that MgdE interacts with both WDR5 and ASH2L, leading to a decrease in
312 H3K4me3 levels (Figure 4A-C, Figure 5D). Site-directed mutagenesis revealed that residues
313 D224 and H247 of MgdE are critical for WDR5 binding, as the double mutant MgdE-
314 D224A/H247A fails to interact with WDR5 and shows diminished suppression of H3K4me3
315 levels (Figure 5D).

316 Moreover, MgdE accumulates in the nucleus (Figure 1C, Figure 2C, Figure 5C, Figure S2B),
317 reminiscent of *L. monocytogenes* InlP, which modulates nuclear speckles to inhibit apoptosis
318 (Pourpre et al., 2022). This interaction silences the expression of antimicrobial effector genes,
319 thereby facilitating immune evasion. Similar to *L. pneumophila* RomA, which methylates
320 H3K14 to repress innate immune responses (Schator et al., 2023; Rolando et al., 2013), or
321 Mtb Rv2067c, a structural mimic of host DOT1L that catalyzes non-nucleosomal H3K79
322 trimethylation to subvert pro-inflammatory signaling (Singh et al., 2023), MgdE exemplifies a
323 broader pathogen strategy of hijacking the host histone methylation machinery.

324 **MgdE-mediated epigenetic silencing suppresses pro-inflammatory responses.**
325 Transcriptome analysis showed that the *mgdE*-deleted BCG strain induced hyperactivation of
326 “Cytokine–cytokine receptor interaction (hsa04060)” and “JAK–STAT signaling pathway
327 (hsa04630)” (Figure 6C), along with significantly elevated production of inflammatory
328 cytokines, including *IL6* and *IL1B*, in THP-1 macrophages compared to the WT BCG strain

329 (Figure 3C and D, Figure 6E and F). This profile is consistent with the functional restoration of
330 the COMPASS complex (Wang et al., 2025; Yu et al., 2017). These findings support the
331 hypothesis that MgdE suppresses host inflammatory responses by disrupting COMPASS-
332 mediated H3K4me3 deposition at active promoters. H3K4me3 is a well-established histone
333 mark associated with transcriptional activation, in part through the recruitment of PHD domain-
334 containing reader proteins (Hyun et al., 2017). For instance, SET1-dependent deposition of
335 H3K4me3 at NF-κB target promoters facilitates the activation of inflammatory gene expression
336 (Bhattacharya et al., 2023). In addition, WDR5 has been implicated in upregulating
337 immunosuppressive cytokines, including *IL6* and *TGFβ* (Deng et al., 2024). In transfected
338 HEK293T cells, WT MgdE significantly reduced global H3K4me3 levels, whereas the
339 catalytically inactive D224A/H247A mutant failed to do so (Figure 5D), supporting the notion
340 that MgdE-mediated suppression of H3K4me3 depends on its ability to interact with WDR5.
341 Furthermore, during infection, the ΔMgdE strain markedly activated the expression of
342 inflammatory factors compared to the WT BCG strain (Figure 6E–G). These results suggest
343 that MgdE attenuates the expression of inflammatory mediators such as *IL6* and *IL1B* by
344 reducing host H3K4me3 levels during Mycobacterium infection. This mechanism is similar to
345 those observed in other pathogens, such as spirochete-derived factors that suppress
346 inflammation by reducing H3K4me3 levels at the *TNFα* and *IL6* promoters (Chauhan et al.,
347 2015).

348 In summary, our study identified MgdE as a critical mycobacterial nucleomodulin and
349 uncovered a novel paradigm of pathogen-mediated epigenetic regulation through the "MgdE–
350 COMPASS complex–H3K4me3–cytokine" axis (Figure 7E). Mechanistically, MgdE might

351 destabilize the COMPASS assembly via the interaction with ASH2L or WDR5, resulting in
352 reduced H3K4me3 levels at the promoters of pro-inflammatory genes. This epigenetic
353 suppression downregulates cytokine transcription (e.g., *IL6 and IL1B*) and enhances bacterial
354 survival in macrophages and in mice. Our findings reveal the COMPASS complex as a
355 previously unrecognized target of bacterial effectors and offers mechanistic insights into
356 immune evasion and host–pathogen epigenetic interplay.

357 **MATERIALS AND METHODS**

358 **Bacterial strains and culture conditions**

359 The bacterial strains used in this study are detailed in Table S5. *Escherichia coli* DH5α was
360 cultured in Luria-Bertani medium under standard conditions. *Mycobacterium bovis BCG*
361 strains were grown in Middlebrook 7H9 broth (BD Biosciences), containing 1× OADC (oleic
362 acid, albumin, dextrose, catalase), 0.05% Tween-80, and 2% glycerol. Antibiotics were added
363 as required at the following final concentrations: 50 μg/mL hygromycin (for mycobacteria) and
364 50 μg/mL kanamycin (for both mycobacteria and *E. coli*). The Y2H Gold yeast strain (Takara
365 Bio) was maintained in YPDA (yeast extract, peptone, dextrose, and adenine sulfate) medium.
366 Transformants were selected on SD/-Trp medium, which lacks tryptophan, to select for
367 plasmid uptake. Protein-protein interactions were assessed in SD/-Trp/-Leu (lacking
368 tryptophan and leucine) and SD/-Trp/-Leu/-Ade/-His (lacking tryptophan, leucine, adenine,
369 and histidine) media, with growth on higher stringency media indicating positive interactions.

370 **Cell culture**

371 THP-1 cells were cultured in Roswell Park memorial institute 1640 (RPMI 1640) medium
372 supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS), 1 mM sodium

373 pyruvate, 2 mM L-glutamine, 10 mM HEPES buffer (pH 7.2–7.5), and 50 μM 2-
374 mercaptoethanol. For differentiation into macrophages, THP-1 cells were treated with 200 nM
375 phorbol 12-myristate 13-acetate (PMA) for 48 h to allow complete differentiation prior to their
376 use in experiments.
377 HEK293T and RAW264.7 monocytes were cultured in Dulbecco's modified Eagle's medium
378 (DMEM) supplemented with 10% FBS and 50 μg/mL penicillin-streptomycin. All cells were
379 maintained at 37 °C in a humidified atmosphere with 5% CO_2.

380 **Screening of conserved nucleomodulins in Mycobacteria**

381 To systematically identify evolutionarily conserved nucleomodulins in pathogenic
382 mycobacteria, a comparative pan-genome analysis was performed across four
383 phylogenetically representative mycobacterium strains: *Mycobacterium tuberculosis* H37Rv
384 (NCBI RefSeq assembly GCF_000195955.2), *Mycobacterium tuberculosis* H37Ra (NCBI
385 RefSeq assembly GCF_001938725.1), *Mycobacterium bovis* BCG (NCBI RefSeq assembly
386 GCF_000009445.1), *Mycobacterium marinum* (NCBI RefSeq assembly GCF_000018345.1),
387 and *Mycobacterium avium* subsp (NCBI RefSeq assembly GCF_009741445.1). Complete
388 genome sequences and annotations were obtained from NCBI RefSeq with strict inclusion
389 criteria requiring "Complete Genome" status and fewer than 50 contigs to ensure the assembly
390 integrity. Plasmid sequences, as well as low-complexity and repetitive regions, were filtered
391 using the NCBI Genome Data Viewer toolkit. Genome quality control was rigorously conducted
392 using CheckM with a *Mycobacteriaceae*-specific marker set, confirming that all the selected
393 genomes exhibited >99% completeness and <1% contamination. Ortholog clustering of
394 protein-coding sequences was performed using OrthoFinder with DIAMOND for sequence

395 alignment. Orthogroups were inferred using the Markov Cluster Algorithm (MCL) with an
396 inflation parameter of 1.5 to balance sensitivity and specificity. Core orthologs were defined
397 as groups present in all four strains, with additional filtering to retain only single-copy genes in
398 ≥90% of the strains, thereby minimizing paralog-related artifacts.
399 To prioritize candidate nucleomodulins, a sequential screening pipeline was applied, beginning
400 with the prediction of classical secretory proteins using SignalP v5.0 (D-score ≥0.5), followed
401 by the identification of non-classically secreted proteins via SecretomeP v2.0 (NN score ≥0.9),
402 and culminating in NLS detection using cNLS Mapper (NLS score ≥2.0), which predicts both
403 monopartite and bipartite NLS. This integrated strategy enabled the systematic identification
404 of conserved candidates that exhibit both secretory potential and nuclear targeting capacity.

405 **Immunoblot analysis of MgdE secretion**

406 The procedure for immunoblot analysis of MgdE secretion was adapted from Zhang et al.
407 (2022) with minor modifications. *M. bovis* BCG strains expressing C-terminally Flag-tagged
408 MgdE were cultured in Middlebrook 7H9 broth supplemented with 0.5% glycerol, 0.02%
409 Tyloxapol, and 50 μg/mL kanamycin. Cultures were incubated at 37 °C with gentle agitation
410 (80 rpm) until reaching an OD_{600} of ~0.6. Cells were harvested by centrifugation at 4,000 × *g*
411 for 15 min at 4 °C. The culture supernatants were filtered through 0.22 μm PES membranes
412 (Millipore) to remove residual bacteria and subsequently concentrated ~100-fold using Amicon
413 Ultra-15 centrifugal filters with a 10 kDa molecular weight cutoff (Millipore). The resulting
414 culture filtrates were used for detection of secreted MgdE. Bacterial pellets were washed twice
415 with PBS and resuspended in lysis buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl,
416 1 mM PMSF, and 1× protease inhibitor cocktail. Cells were lysed by sonication on ice, and the

417 lysates were cleared by centrifugation at 10,000 × *g* for 15 min at 4 °C to obtain whole-cell
418 lysates. Whole-cell lysates and concentrated culture filtrates were resolved on 10% SDS–
419 PAGE gels and transferred to PVDF membranes (Millipore). Membranes were blocked with
420 5% skimmed milk powder in TBST (Tris-buffered saline containing 0.1% Tween-20) for 1 h at
421 room temperature and then incubated overnight at 4 °C with a monoclonal anti-Flag antibody.
422 After washing, membranes were incubated with HRP-conjugated secondary antibodies for 1 h
423 at room temperature. Immunoreactive bands were visualized using enhanced
424 chemiluminescence (ECL) substrate and imaged with a ChemiDoc MP imaging system (Bio-
425 Rad).

426 **Phylogenetic and sequence analysis**

427 To investigate the evolutionary conservation and sequence features of MgdE homologs,
428 phylogenetic reconstruction was performed using MEGA 12.0 with a neighbor-joining
429 algorithm. The parameters included pairwise deletion for gap treatment, a Poisson substitution
430 model (selected for its suitability in analyzing closely related bacterial sequences), and 1,000
431 bootstrap replicates to assess node support. Nodes with ≥50% bootstrap values were retained,
432 and no additional pruning was performed.

433 The analysis incorporated MgdE homologs from 12 Mycobacterium strains spanning clinical,
434 environmental, and attenuated lineages. These included Rv1075c (*M. tuberculosis* H37Rv),
435 BCG_1133c (*M. bovis* BCG), MRA_1085 (*M. tuberculosis* H37Ra), MAF_10880 (*M.*
436 *africanum*), MT1105 (*M. tuberculosis* CDC1551), TBMG_02912 (*M. tuberculosis* GM041182),
437 MSHI_02760 (*M. shinjukuense*), MMAR_4392 (*M. marinum*), MULP_04607 (*M. lupini*),
438 MAV_1199 (*M. avium*), MMAN_54310 (*M. manitobense*), and MIP_01781 (*M. intracellulare*).

439 Genomic sequences were retrieved from the NCBI RefSeq database, filtered for
440 completeness, and aligned using Clustal Omega v1.2.4 with default parameters (gap opening
441 penalty = 10, gap extension = 0.2, and iterative refinement enabled). Sequence conservation
442 was visualized using ESPript 3.0, highlighting residues with ≥90% identity across all homologs
443 in red.

444 **Computational prediction of sequence and structural features**

445 AlphaFold v2.2.0 was used to predict the interaction between MgdE and COMPASS complex
446 subunits, following the methodology established in a previous study (Gómez Borrego et al.,
447 2024). The specific parameters used for the prediction were as follows: model_preset =
448 multimer; db_preset = full_dbs. A FASTA file containing the sequences of both MgdE and
449 COMPASS complex subunit proteins was input into the AlphaFold software. The resulting
450 models were evaluated and ranked based on the AF-score, which is a linear combination of
451 the interface score (ipTM) and the predicted TM-score (pTM). The top-ranked model (rank 1)
452 was selected for further analysis.

453 **Cell transfection and confocal microscopy**

454 HEK293T cells were seeded at 70% confluency in poly-L-lysine-coated 35-mm glass-bottom
455 dishes and maintained in DMEM supplemented with 10% FBS at 37 °C under 5% CO_2.
456 Transfection was performed using the Hieff Trans™ Liposomal Transfection Reagent (Yeasen)
457 according to the manufacturer's protocol. Briefly, 0.5 μg of DNA and 1.5 μL of transfection
458 reagent were mixed with 50 μL of Opti-MEM and incubated for 20 min at room temperature.
459 The mixture was added dropwise to each chamber of 35-mm six-chamber glass-bottom dishes

460 containing cells at 40–60% confluency.

461 At 4–48 h post-transfection, the cells were washed with PBS, fixed in 4% paraformaldehyde
462 for 15 min, and permeabilized with 0.1% Triton X-100 for 10 min. Nuclei were counterstained
463 with 1 μg/mL DAPI for 10 min. Confocal imaging was performed using an Olympus FV1000
464 Confocal Laser Scanning Microscope equipped with a 100×/1.40 NA oil immersion objective.
465 EGFP and DAPI signals were detected with 488 nm and 405 nm lasers, respectively. Spectral
466 unmixing was applied using the Olympus FV10-ASW software (v4.2) to minimize signal cross-
467 talk, and Z-stacks were acquired in sequential mode by applying a step size of 0.5 μm across
468 all samples to ensure comparability.

469 **Cell fractionation and Immunoblotting**

470 During infection, THP-1 cells were differentiated with 200 ng/mL PMA for 48 hours, followed
471 by infection with BCG/MgdE-Flag, BCG/MgdE$^{\Delta NLS1}$-Flag, BCG/MgdE$^{\Delta NLS2}$-Flag, or
472 BCG/MgdE$^{\Delta NLS1-2}$ -Flag strains at a multiplicity of infection (MOI) of 50 for 24 hours. Cell
473 fractionation was performed with minor modifications based on a previously described method
474 (Wang J et al., 2017). Briefly, cells were lysed in a hypotonic buffer containing 10 mM HEPES
475 (pH 7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 0.34 M sucrose, 10% glycerol, and 0.02% Tyloxapol,
476 supplemented with 1% protease inhibitor cocktail, and incubated for 10 minutes at 4 °C. The
477 lysates were then centrifuged at 1300 × *g* for 10 minutes. The pellet from uninfected THP-1
478 cells was collected as the nuclear fraction. For infected THP-1 cells, the pellet was
479 resuspended in PBS containing 0.1% SDS and centrifuged at 2500 × *g* for 15 minutes at 4 °C.
480 The resulting supernatant was collected as the nuclear fraction.

481 During transfection, HEK293T cells were cultured in DMEM supplemented with 10% FBS at

482 37 °C in a 5% CO_2 incubator. Transfection was performed as described in the "Cell transfection
483 and confocal microscopy" section. At 36 hours post-transfection, nuclear and cytoplasmic
484 protein fractionation was conducted using the Nuclear and Cytoplasmic Protein Extraction Kit
485 (Beyotime) with minor modifications. Cells were washed with PBS, detached using a cell
486 scraper, and pelleted by centrifugation at 2000 × *g* for 5 minutes. After removing the
487 supernatant, 150 μL of Cytoplasmic Extraction Reagent A (with PMSF) was added to ~20 μL
488 of packed cells, vortexed for 5 seconds, and incubated on ice for 15 minutes. Then, 10 μL of
489 Reagent B was added, followed by vortexing and ice incubation. The mixture was centrifuged
490 at 12,000 × *g* for 5 minutes at 4 °C, and the supernatant was collected as the cytoplasmic
491 fraction. The pellet was resuspended in 50 μL of Nuclear Extraction Reagent (with PMSF),
492 vortexed intermittently over 30 minutes on ice, and centrifuged at 12,000 × *g* for 10 minutes
493 at 4 °C. The resulting supernatant was collected as the nuclear protein fraction.

494 Protein concentrations were quantified using the BCA assay (Thermo Fisher Scientific). Equal
495 amounts of protein (20–40 μg) were resolved by sodium dodecylsulfate polyacrylamide gel
496 electrophoresis (SDS–PAGE) and transferred onto PVDF membranes. The membranes were
497 blocked for 1 h at room temperature with 5% non-fat milk in TBSTand incubated overnight at
498 4 °C with primary antibodies. After three washes with TBST, the membranes were incubated
499 with HRP-conjugated secondary antibodies at room temperature for 1 h. Following three
500 additional TBST washes, protein signals were detected using an ECL substrate and imaged
501 on a ChemiDoc MP system (Bio-Rad). The antibody information is listed in Table S4.

502 **Co-IP assays**

503 The Co-IP assays were performed to analyze protein-protein interactions. HEK293T cells were
504 washed twice with ice-cold PBS and lysed in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM
505 NaCl, 1% NP-40) supplemented with 1.5% protease inhibitor cocktail (Boster) for 30 min on
506 ice. Crude lysates were centrifuged at 12,000 ×*g* for 15 min at 4 °C to remove the debris. For
507 immunoprecipitation, 1 mg of total protein was incubated with 5 μg of primary antibody
508 overnight at 4 °C with gentle agitation. Protein A/G magnetic beads (MedChemExpress) were
509 added to the antibody-protein complexes and incubated for 6 h at 4 °C. The beads were
510 washed three times with PBS-T buffer (PBS containing 0.5% Tween-20) to remove nonspecific
511 interactions. Proteins were eluted by boiling the beads in 1× SDS sample buffer at 95 °C for 5
512 min. The eluted proteins were resolved by SDS-PAGE and analyzed by immunoblotting.
513 Antibody information is listed in Table S4.

514 **Y2H assay**

515 The Yeast Two-Hybrid System (Clontech) was used to assay protein-protein interactions
516 between MgdE (WT and mutants) and subunits of the COMPASS complex (ASH2L, WDR5,
517 RbBP5, and DPY30). Full-length cDNAs encoding WT MgdE and its mutants (see Table S6)
518 were cloned into the bait vector pGBKT7, whereas genes encoding the COMPASS subunits
519 were cloned into the prey vector pGADT7. The yeast strain Y2H Gold was co-transformed with
520 bait and prey plasmids using the lithium acetate/polyethylene glycol method, and the
521 transformed colonies were initially selected on SD agar plates lacking leucine and tryptophan
522 (-Leu/-Trp) to confirm plasmid retention. For interaction screening, colonies were replica-
523 plated onto selective SD agar plates lacking leucine, tryptophan, adenine, and histidine (-Leu/-
524 Trp/-Ade/-His) and supplemented with 200 ng/μL aureobasidin A (AbA, Sigma-Aldrich). The

525 plates were incubated at 30 °C, and colony growth was monitored daily, with final assessments
526 performed after 5 days of incubation. Interaction strength was qualitatively assessed by
527 comparing colony density and growth rates with those of the controls. The positive controls
528 included the co-transformation of pGBKT7-p53 with pGADT7-T, which produced robust growth
529 on -Leu/-Trp/-Ade/-His plates. Negative controls included co-transformation of pGBKT7-Lam
530 (human lamin C, a non-interacting bait) with pGADT7-T, as well as individual transformations
531 with empty pGBKT7 or pGADT7 vectors to assess the background growth. The yeast strains
532 used in this study are listed in Table S5.

533 **RNA-seq analysis**

534 THP-1 cells were seeded at a density of 1×10^6 cells/well in 6-well plates and differentiated
535 into macrophages by treatment with 200 nM PMA for 48 h. Following differentiation, the cells
536 were washed with PBS and infected with *M. bovis* BCG WT or ΔMgdE strains at a MOI of 10:1.
537 Following 24 h of infection, total RNA was extracted using TRIzol® reagent (Thermo Fisher
538 Scientific) according to the manufacturer's instructions. High-quality reads were then mapped
539 to the human reference genome (GRCh38) using HISAT2 (v2.2.1). FeatureCounts (v2.0.3)
540 was used to quantify transcript abundance, and differential gene expression analysis was
541 conducted using DESeq2 (v1.38.3) in R. Genes with an adjusted *P*-value <0.05 and an
542 absolute $\log_2$-foldchange ≥1 was considered significantly differentially expressed. RNA-seq
543 data were visualized using GraphPad Prism 8.0, and volcano plots, heatmaps, and other
544 visualizations were generated to summarize the results.

545 **qRT-PCR assay**

546 THP-1 cells were infected with the *M. bovis* BCG strain and harvested 24 h post-infection for
547 RNA isolation. Total RNA was extracted using 1 mL of TRIzol® reagent (Aidlab) according to
548 the manufacturer's instructions. To remove genomic DNA contamination, the RNA samples
549 were treated with DNase I (Thermo Fisher Scientific) for 30 min at 37 °C, followed by heat
550 inactivation at 65 °C for 10 min. RNA concentration and purity were assessed using a
551 NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). For cDNA synthesis, 1 μg of
552 total RNA was reverse transcribed using the HiScript® III cDNA Reverse Transcription Kit
553 (Vazyme) with random hexamers, according to the manufacturer's protocol. qRT-PCR was
554 performed using ChamQ SYBR® Green RT-PCR Master Mix (Vazyme). The relative mRNA
555 expression of different genes was calculated by comparing their cycle threshold (Ct) values to
556 that of a control gene *HPRT* using the $2^{-\Delta\Delta Ct}$ method. Real-time qRT-PCR experiments were
557 performed in triplicate, with at least three independent biological replicates and two technical
558 replicates per condition. Primer sequences used for qRT-PCR are listed in Table S7.

559 **CFU assay**

560 For bacterial preparation, mid-log phase cultures (OD_{600} = 0.8) of *M. bovis* BCG were washed
561 with PBS and resuspended in fresh medium. THP-1 cells were seeded at a density of 5×10^5
562 cells/well in 24-well plates and subsequently infected at a MOI of 10:1. Following 4 h of
563 infection, extracellular bacteria were eliminated by treatment with penicillin-streptomycin (100
564 μg/mL) for 2 h. The cells were then washed three times with PBS and maintained in an
565 antibiotic-free medium until harvest. For bacterial recovery, the macrophages were lysed with
566 0.25% SDS for 10 min at room temperature, followed by serial dilution of the lysates in PBS.
567 Aliquots (100 μL) of each dilution were plated on Middlebrook 7H10 agar supplemented with

568 10% OADC and kanamycin (50 μg/mL). Plates were incubated at 37 °C for 14 days, and CFUs
569 were calculated by applying dilution factors to the colony counts. *M. bovis* BCG strains used
570 in this study are listed in Table S5.

571 **Animal experiments**

572 Specific pathogen-free (SPF) female C57BL/6 mice (6–8 weeks old) were obtained from
573 Chang-sheng Bio (Liaoning, China). All animals were housed under SPF conditions under
574 controlled temperature, humidity, and a 12-hour light/dark cycle, with free access to food and
575 water.

576 Mice were infected as described previously with slight modifications (Chen et al., 2022). Mid-
577 log phase *M. bovis* BCG cultures were washed twice in PBS containing 0.05% Tween 80 and
578 sonicated to disrupt clumps. Female C57BL/6 mice (6–8 weeks old, 16–18g) were
579 anesthetized using isoflurane and intranasally administered 1×10^7 CFUs of *M. bovis* BCG in
580 40 μL of PBS. At 2 days post-infection, the bacterial load in the lungs of six mice was assessed
581 to confirm the inoculation dose. The lungs were homogenized in 1 mL of PBS using a tissue
582 homogenizer, and serial dilutions of the homogenates were plated onto Middlebrook 7H10
583 agar with 10% OADC. The plates were then incubated at 37 °C with 5% CO_2 for 14 d prior to
584 colony counting.

585 **Histological analysis**

586 Lung tissues from *M. bovis* BCG-infected mice were fixed in 4% phosphate-buffered formalin
587 at room temperature for 24 h and embedded in paraffin wax. Paraffin-embedded tissue
588 samples were sectioned into 2–3 μm thick slices using a microtome. The sections were

589 deparaffinized, rehydrated with graded ethanol, and stained with hematoxylin and eosin (H&E).
590 Stained slides were visualized using an Olympus BX53 light microscope, and CaseViewer
591 version 2.0 (3DHISTECH) was used to view and annotate the scanned slides.

592 **Ethics approval**

593 All animal experiments were conducted in accordance with the protocols approved by the
594 Animal Ethics Committee of the Huazhong Agricultural University.

595 **Statistical analyses**

596 Data are presented as mean ± standard deviation (SD). Statistical significance was assessed
597 using unpaired *t*-tests or two-way ANOVA. Statistical significance was set at $P < 0.05$.
598 Significance levels are denoted in the figures as follows: $P > 0.05$ (ns, not significant), $P <$
599 0.05 (*), $P < 0.01$ (**), and $P < 0.001$ (***). All statistical analyses were performed using
600 GraphPad Prism (version 8.0). Data were collected from at least three independent biological
601 replicates, each with two or more technical replicates.

602 **ACKNOWLEDGMENTS**

603 This work was supported by the National Key Research and Development Program of China
604 (#2021YFD1800403), Fundamental Research Funds for the Central Universities
605 (#2662023DKQD001), Talent Start-up Funds of Huazhong Agricultural University
606 (#11042310008), National Natural Science Foundation of China (#32473123), China
607 Agriculture Research System of MOF and MARA (#CARS-37).

608 **AUTHOR CONTRIBUTIONS**

609 Conceptualization: L. C., A. G., and L. Z.; Methodology: L. C.; Validation: L. C., B. D., and Q.

610 J.; Formal analysis: L. C., Y. W.; Investigation: L. C., B. D., P. C., Q. J., Y. W., and L. L.;
611 Resources: A. G., L. Z., and Y. C.; Data curation: L. C., Q. J., Y. W., and L. L.; Writing - original
612 draft: L. C., L. Z., and A. G.; Writing - review & editing: L. C., L. Z., and A. G.; Visualization: L.
613 C., Q. J., Y. C., C. H., L. Z., and A. G.; Supervision: A. G., L. Z., Y. C., and C. H.; Project
614 administration: A. G., L. Z., Y. C., and C. H.; Funding acquisition: A. G., L. Z., and Y. C.

615 **DISCLOSURE AND COMPETING INTEREST STATEMENT**

616 The authors declare no competing interests.

617 **REFERENCES**

618 Bates TA, Trank-Greene M, Nguyenla X, Anastas A, Gurmessa SK, Merutka IR, Dixon SD,
619 Shumate A, Groncki AR, Parson MAH (2024) **ESAT-6 undergoes self-association at**
620 **phagosomal pH and an ESAT-6-specific nanobody restricts *M. tuberculosis* growth in**
621 **macrophages** *Elife* **12**:RP91930.

622 Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005) **Non-classical protein secretion in**
623 **bacteria** *BMC Microbiol* **5**:58.

624 Bhattacharya A, Fushimi A, Wang K, Yamashita N, Morimoto Y, Ishikawa S, Daimon T, Liu T,
625 Liu S, Long MD (2023) **MUC1-C intersects chronic inflammation with epigenetic**
626 **reprogramming by regulating the set1a compass complex in cancer progression**
627 *Commun Biol* **6**:1030.

628 Bierne H, Cossart P (2012) **When bacteria target the nucleus: the emerging family of**
629 **nucleomodulins** *Cell Microbiol* **14**:622-633.

630 Bierne H, Pourpre R (2020) **Bacterial factors targeting the nucleus: the growing family of**
631 **nucleomodulins** *Toxins* (Basel) **12**:220.

632 Chai Q, Wang L, Liu CH, Ge B (2020) **New insights into the evasion of host innate**
633 **immunity by *Mycobacterium tuberculosis*** *Cell Mol Immunol* **17**:901-913.

634 Chai Q, Yu S, Zhong Y, Lu Z, Qiu C, Yu Y, Zhang X, Zhang Y, Lei Z, Qiang L (2022) **A bacterial**
635 **phospholipid phosphatase inhibits host pyroptosis by hijacking ubiquitin** *Science*
636 **378**:eabq0132.

637 Chambers KA, Scheck RA (2020) **Bacterial virulence mediated by orthogonal post-**
638 **translational modification** *Nat Chem Biol* **16**:1043-1051.

639 Chandra P, Grigsby SJ, Philips JA (2022) **Immune evasion and provocation by**
640 ***Mycobacterium tuberculosis*** *Nat Rev Microbiol* **20**:750-766.

641 Chauhan A, Quenum FZ, Abbas A, Bradley DS, Nechaev S, Singh BB, Sharma J, Mishra BB
642 (2015) **Epigenetic modulation of microglial inflammatory gene loci in helminth-induced**
643 **immune suppression: implications for immune regulation in neurocysticercosis** *ASN*
644 *Neuro* **7**:1759091415592126.

645 Chen L, Li X, Xu P, He ZG (2022) **A Novel Zinc Exporter CtpG Enhances Resistance to**
646 **Zinc Toxicity and Survival in *Mycobacterium bovis*** *Microbiol Spectr* **10**:e0145621.

647 Deng K, Liang L, Yang Y, Wu Y, Li Y, Zhang R, Tian Y, Lu C (2024) **The Wdr5-H3K4me3**
648 **epigenetic axis regulates pancreatic tumor immunogenicity and immune suppression**
649 *Int J Mol Sci* **25**:8773.

650 Denzer L, Schroten H, Schwerk C (2020) **From gene to protein-how bacterial virulence**
651 **factors manipulate host gene expression during infection** *Int J Mol Sci* **21**:3730.

652 Evans SM, Rodino KG, Adcox HE, Carlyon JA (2018) ***Orientia tsutsugamushi* uses two Ank**
653 **effectors to modulate NF-κB p65 nuclear transport and inhibit NF-κB transcriptional**
654 **activation** *PLoS Pathog* **14**:e1007023.

655 Fol M, Włodarczyk M, Druszczyńska M (2020) **Host epigenetics in intracellular pathogen**
656 **infections** *Int J Mol Sci* **21**:4573.

657 Fu M, Liang S, Wu J, Hua Y, Chen H, Zhang Z, Liu J, Li X, Zhang B, Zhao W, Wan C (2021)
658 **An *Escherichia coli* effector protein EspF may induce host DNA damage via interaction**
659 **with SMC1** *Front Microbiol* **12**:682064.

660 Gómez Borrego J, Torrent Burgas M (2024) **Structural assembly of the bacterial essential**
661 **interactome** *Elife* **13**:e94919.

662 Guarnaccia AD, Rose KL, Wang J, Zhao B, Popay TM, Wang CE, Guerrazzi K, Hill S, Woodley
663 CM, Hansen TJ (2021) **Impact of WIN site inhibitor on the WDR5 interactome** *Cell Rep*
664 **34**:108636.

665 Hammond DE, Kumar JD, Raymond L, Simpson DM, Beynon RJ, Dockray GJ, Varro A (2018)
666 **Stable isotope dynamic labeling of secretomes (SIDLS) identifies authentic secretory**
667 **proteins released by cancer and stromal cells** *Mol Cell Proteomics* **17**:1837-1849.

668 Hanford HE, Von Dwingelo J, Abu Kwaik Y (2021) **Bacterial nucleomodulins: a**
669 **coevolutionary adaptation to the eukaryotic command center** *PLoS Pathog* **17**:e1009184.

670 Hsu PL, Li H, Lau HT, Leonen C, Dhall A, Ong SE, Chatterjee C, Zheng N (2018) **Crystal**
671 **structure of the COMPASS H3K4 methyltransferase catalytic module** *Cell* **174**:1106-
672 1116.e9.

673 Hyun K, Jeon J, Park K, Kim J (2017) **Writing, erasing and reading histone lysine**
674 **methylations** *Exp Mol Med* **49**:e324.

675 Jose L, Ramachandran R, Bhagavat R, Gomez RL, Chandran A, Raghunandanan S,
676 Omkumar RV, Chandra N, Mundayoor S, Kumar RA (2016) **Hypothetical protein Rv3423.1**
677 **of *Mycobacterium tuberculosis* is a histone acetyltransferase** *FEBS J* **283**:265-281.

678 Khan AA, Khan Z (2021) **Bacterial nucleomodulins and cancer: an unresolved enigma**
679 *Transl Oncol* **14**:100922.

680 Lebreton A, Job V, Ragon M, Le Monnier A, Dessen A, Cossart P, Bierne H (2014) **Structural**
681 **basis for the inhibition of the chromatin repressor BAHD1 by the bacterial**
682 **nucleomodulin LntA** *mBio* **5**:e00775-13.

683 Park SH, Ayoub A, Lee YT, Xu J, Kim H, Zheng W, Zhang B, Sha L, An S, Zhang Y, Cianfrocco
684 MA, Su M, Dou Y, Cho US (2019) **Cryo-EM structure of the human MLL1 core complex**
685 **bound to the nucleosome** *Nat Commun* **10**:5540.

686 Penn BH, Netter Z, Johnson JR, Von Dollen J, Jang GM, Johnson T, Ohol YM, Maher C, Bell
687 SL, Geiger K (2018) **An Mtb-human protein-protein interaction map identifies a switch**
688 **between host antiviral and antibacterial responses** *Mol Cell* **71**:637-648.e5.

689 Pisu D, Huang L, Grenier JK, Russell DG (2020) **Dual RNA-Seq of Mtb-Infected**

690 **Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions** *Cell*
691 *Rep* **30**:335-350.e4.

692 Pourpre R, Lakisic G, Desgranges E, Cossart P, Pagliuso A, Bierne H (2022) **A bacterial**
693 **virulence factor interacts with the splicing factor RBM5 and stimulates formation of**
694 **nuclear RBM5 granules** *Sci Rep* **12**:21961.

695 Prokop A, Gouin E, Villiers V, Nahori MA, Vincentelli R, Duval M, Cossart P, Dussurget O (2017)
696 **OrfX, a nucleomodulin required for *Listeria monocytogenes* virulence** *mBio* **8**:e01550-
697 17.

698 Qiang L, Zhang Y, Lei Z, Lu Z, Tan S, Ge P, Chai Q, Zhao M, Zhang X, Li B (2023) **A**
699 **mycobacterial effector promotes ferroptosis-dependent pathogenicity and**
700 **dissemination** *Nat Commun* **14**:1430.

701 Qu Q, Takahashi YH, Yang Y, Hu H, Zhang Y, Brunzelle JS, Couture JF, Shilatifard A, Skiniotis
702 G (2018) **Structure and Conformational Dynamics of a COMPASS Histone H3K4**
703 **Methyltransferase Complex** *Cell* **174**:1117-1126.e12.

704 Radoshevich L, Cossart P (2018) ***Listeria monocytogenes*: towards a complete picture of**
705 **its physiology and pathogenesis** *Nat Rev Microbiol* **16**:32-46.

706 Rennoll-Bankert KE, Garcia-Garcia JC, Sinclair SH, Dumler JS (2015) **Chromatin-bound**
707 **bacterial effector ankyrin A recruits histone deacetylase 1 and modifies host gene**
708 **expression** *Cell Microbiol* **17**:1640-1652.

709 Rolando M, Sanulli S, Rusniok C, Gomez-Valero L, Bertholet C, Sahr T, Margueron R,

710 Buchrieser C (2013) ***Legionella pneumophila* effector RomA uniquely modifies host**
711 **chromatin to repress gene expression and promote intracellular bacterial replication**
712 *Cell Host Microbe* **13**:395-405.

713 Rolando M, Wah Chung IY, Xu C, Gomez-Valero L, England P, Cygler M, Buchrieser C (2023)
714 **The SET and ankyrin domains of the secreted *Legionella pneumophila* histone**
715 **methyltransferase work together to modify host chromatin** *mBio* **14**:e0165523.

716 Schator D, Mondino S, Berthelet J, Di Silvestre C, Ben Assaya M, Rusniok C, Rodrigues-Lima
717 F, Wehenkel A, Buchrieser C, Rolando M (2023) **Legionella para-effectors target chromatin**
718 **and promote bacterial replication** *Nat Commun* **14**:2154.

719 Sharma G, Upadhyay S, Srilalitha M, Nandicoori VK, Khosla S (2015) **The interaction of**
720 **mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG**
721 **methylation and histone H3/H4 binding** *Nucleic Acids Res* **43**:3922-3937.

722 Singh PR, Dadireddy V, Udupa S, Kalladi SM, Shee S, Khosla S, Rajmani RS, Singh A,
723 Ramakumar S, Nagaraja V (2023) **The *Mycobacterium tuberculosis* methyltransferase**
724 **Rv2067c manipulates host epigenetic programming to promote its own survival** *Nat*
725 *Commun* **14**:8497.

726 Sun H, Kamanova J, Lara-Tejero M, Galán JE (2016) **A Family of Salmonella Type III**
727 **Secretion Effector Proteins Selectively Targets the NF-κB Signaling Pathway to**
728 **Preserve Host Homeostasis** *PLoS Pathog* **12**:e1005484.

729 Wang H, Fan Z, Shliaha PV, Miele M, Hendrickson RC, Jiang X, Helin K (2023) **H3K4me3**
730 **regulates RNA polymerase II promoter-proximal pause-release** *Nature* **615**:339-348.

731 Wang H, Helin K (2025) **Roles of H3K4 methylation in biology and disease** *Trends Cell*
732 *Biol* **35**:115-128.

733 Wang J, Ge P, Qiang L, Tian F, Zhao D, Chai Q, Zhu M, Zhou R, Meng G, Iwakura Y (2017)
734 **The mycobacterial phosphatase PtpA regulates the expression of host genes and**
735 **promotes cell proliferation** *Nat Commun* **8**:244.

736 Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH,
737 Allis CD (2005) **WDR5 associates with histone H3 methylated at K4 and is essential for**
738 **H3 K4 methylation and vertebrate development** *Cell* **121**:859-72.

739 Xue H, Yao T, Cao M, Zhu G, Li Y, Yuan G, Chen Y, Lei M, Huang J (2019) **Structural basis**
740 **of nucleosome recognition and modification by MLL methyltransferases** *Nature*
741 **573**:445-449.

742 Yang D, He X, Li S, Liu J, Stabenow J, Zalduondo L, White S, Kong Y (2019) **Rv1075c of**
743 ***Mycobacterium tuberculosis* is a GDSL-Like esterase and is important for intracellular**
744 **survival** *J Infect Dis* **220**:677-686.

745 Yaseen I, Kaur P, Nandicoori VK, Khosla S (2015) **Mycobacteria modulate host epigenetic**
746 **machinery by Rv1988 methylation of a non-tail arginine of histone H3** *Nat Commun*
747 **6**:8922.

748 Yu H, Lesch BJ (2024) **Functional roles of H3K4 methylation in transcriptional regulation**
749 *Mol Cell Biol* **44**:505-515.

750 Yu L, Fang F, Dai X, Xu H, Qi X, Fang M, Xu Y (2017) **MKL1 defines the H3K4Me3 landscape**

751 **for NF-κB dependent inflammatory response** *Sci Rep* **7**:191.

752 Zhang L, Kent JE, Whitaker M, Young DC, Herrmann D, Aleshin AE, Ko YH, Cingolani G, Saad
753 JS, Moody DB, Marassi FM, Ehrt S, Niederweis M (2022) **A periplasmic cinched protein is**
754 **required for siderophore secretion and virulence of Mycobacterium tuberculosis** *Nat*
755 *Commun* **13**:2255.

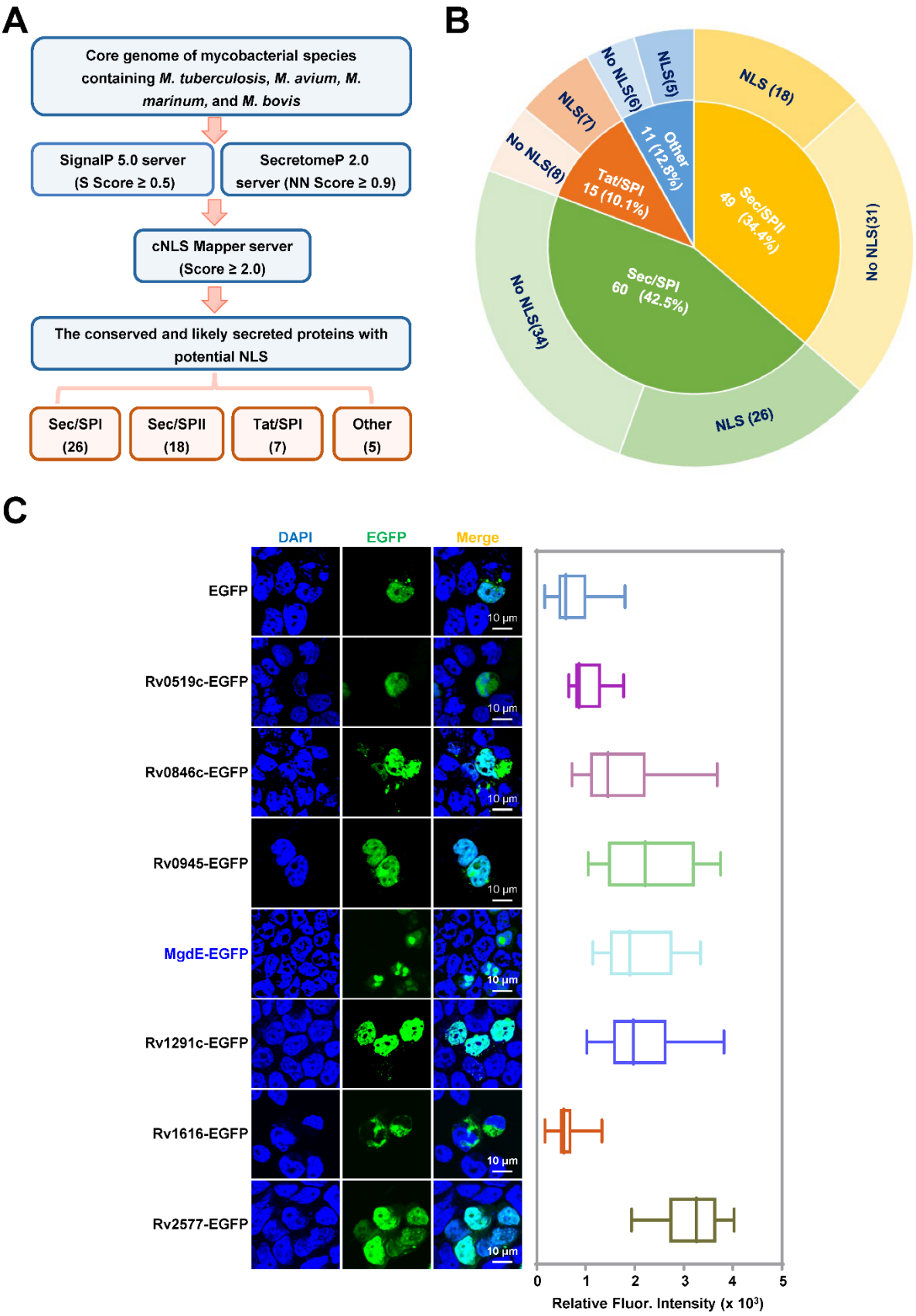
756 Zhang YJ, Reddy MC, Ioerger TR, Rothchild AC, Dartois V, Schuster BM, Trauner A, Wallis D,
757 Galaviz S, Huttenhower C, Sacchettini JC, Behar SM, Rubin EJ (2013) **Tryptophan**
758 **biosynthesis protects mycobacteria from CD4 T-cell-mediated killing** *Cell* **155**:1296-308.

759 Zhao J, Liu Y, Lin F, Wang W, Yang S, Ge Y, Chen PR (2019) **Bioorthogonal engineering of**
760 **bacterial effectors for spatial-temporal modulation of cell signaling** *ACS Cent Sci* **5**:145-
761 152.

762 **Figures**

763 **Figure legends**

764

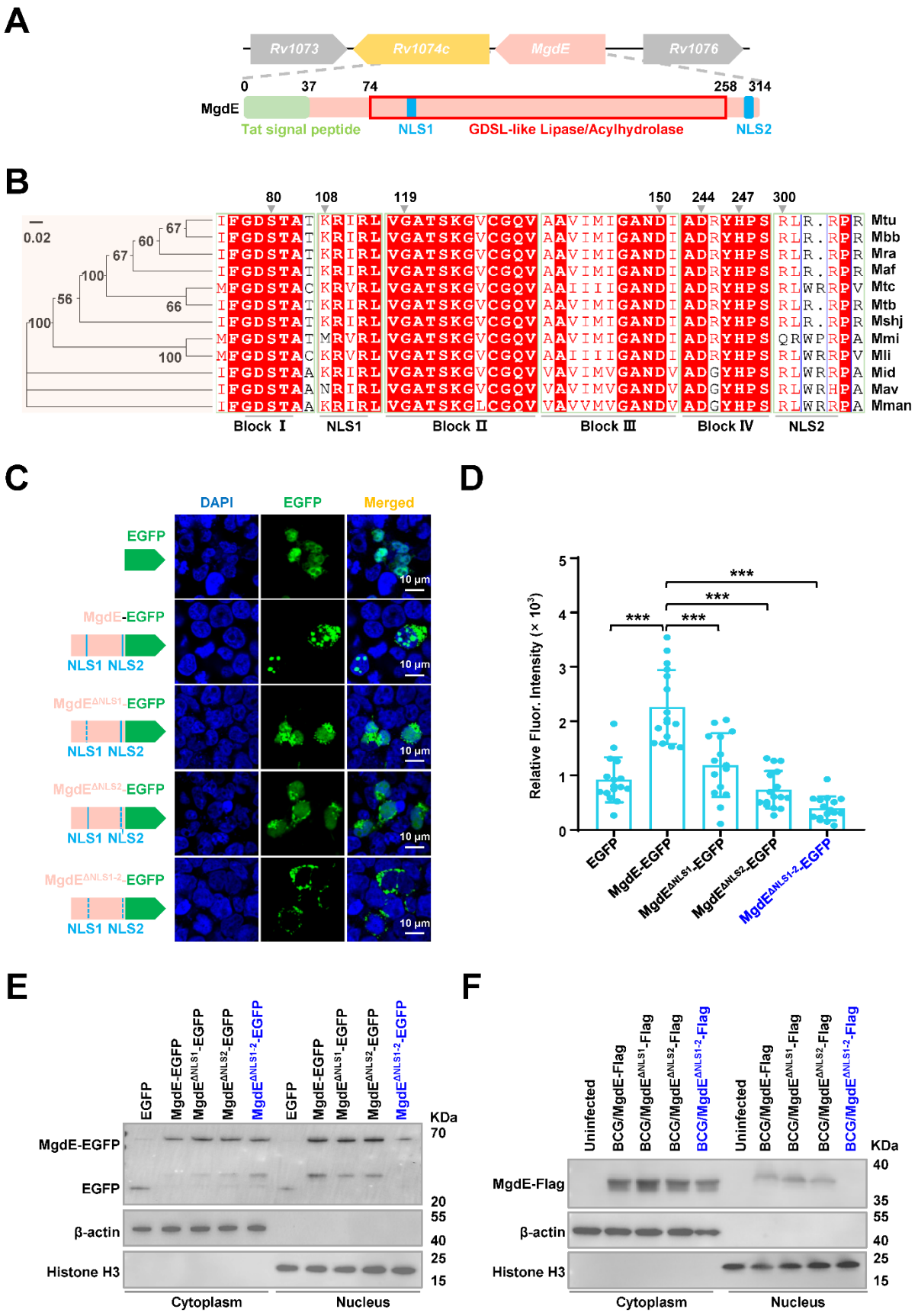


765

766 **Figure 1. Identification of conserved nucleomodulins in mycobacteria through**

767 **functional screening.**

768 (**A**) Schematic representation of the bioinformatic pipeline for identifying conserved
769 nucleomodulins in Mycobacterium species. Genomic sequences from *M. tuberculosis* H37Rv,
770 *M. tuberculosis* H37Ra, *M. avium*, *M. marinum*, and *M. bovis* BCG were analyzed. Signal
771 peptides were predicted using SignalP 5.0 (D-score ≥ 0.5), non-classical secretion signals
772 were identified using SecretomeP 2.0 (NN score ≥ 0.9), and nuclear localization signals (NLS)
773 were predicted using cNLS Mapper (score ≥ 2.0). (**B**) Classification of conserved cellular
774 nucleomodulins. Secreted proteins were categorized based on the presence or absence of
775 predicted NLS motifs. (**C**) Subcellular localization of EGFP-tagged candidate nucleomodulins.
776 (left) Confocal microscopy images of seven Tat/SPI proteins fused to EGFP (green), nuclei
777 were stained with DAPI (blue). (right) Quantification of nuclear EGFP fluorescence intensity.
778 Scale bar, 10 μm. Data are presented as mean ± SD (n = 25 cells/group).


A
Rv1073
Rv1074c
MgdE
Rv1076
0
37
74
258
314
MgdE
Tat signal peptide
NLS1
GDSL-like Lipase/Acylhydrolase
NLS2
B
80
108
119
150
244
247
300
0.02
Mtu
Mbb
Mra
Maf
Mtc
Mtb
Mshj
Mmi
Mli
Mid
Mav
Mman
Block Ⅰ
NLS1
Block Ⅱ
Block Ⅲ
Block Ⅳ
NLS2
C
DAPI
EGFP
Merged
EGFP
MgdE-EGFP
NLS1 NLS2
MgdE^ΔNLS1-EGFP
MgdE^ΔNLS2-EGFP
MgdE^ΔNLS1-2-EGFP
10 μm
D
Relative Fluor. Intensity (× 10³)
***
EGFP
MgdE-EGFP
MgdE^ΔNLS1-EGFP
MgdE^ΔNLS2-EGFP
MgdE^ΔNLS1-2-EGFP
E
EGFP
MgdE-EGFP
MgdE^ΔNLS1-EGFP
MgdE^ΔNLS2-EGFP
MgdE^ΔNLS1-2-EGFP
KDa
MgdE-EGFP
EGFP
β-actin
Histone H3
70
20
55
40
25
15
Cytoplasm
Nucleus
F
Uninfected
BCG/MgdE-Flag
BCG/MgdE^ΔNLS1-Flag
BCG/MgdE^ΔNLS2-Flag
BCG/MgdE^ΔNLS1-2-Flag
KDa
MgdE-Flag
β-actin
Histone H3
40
35
55
40
25
15
Cytoplasm
Nucleus

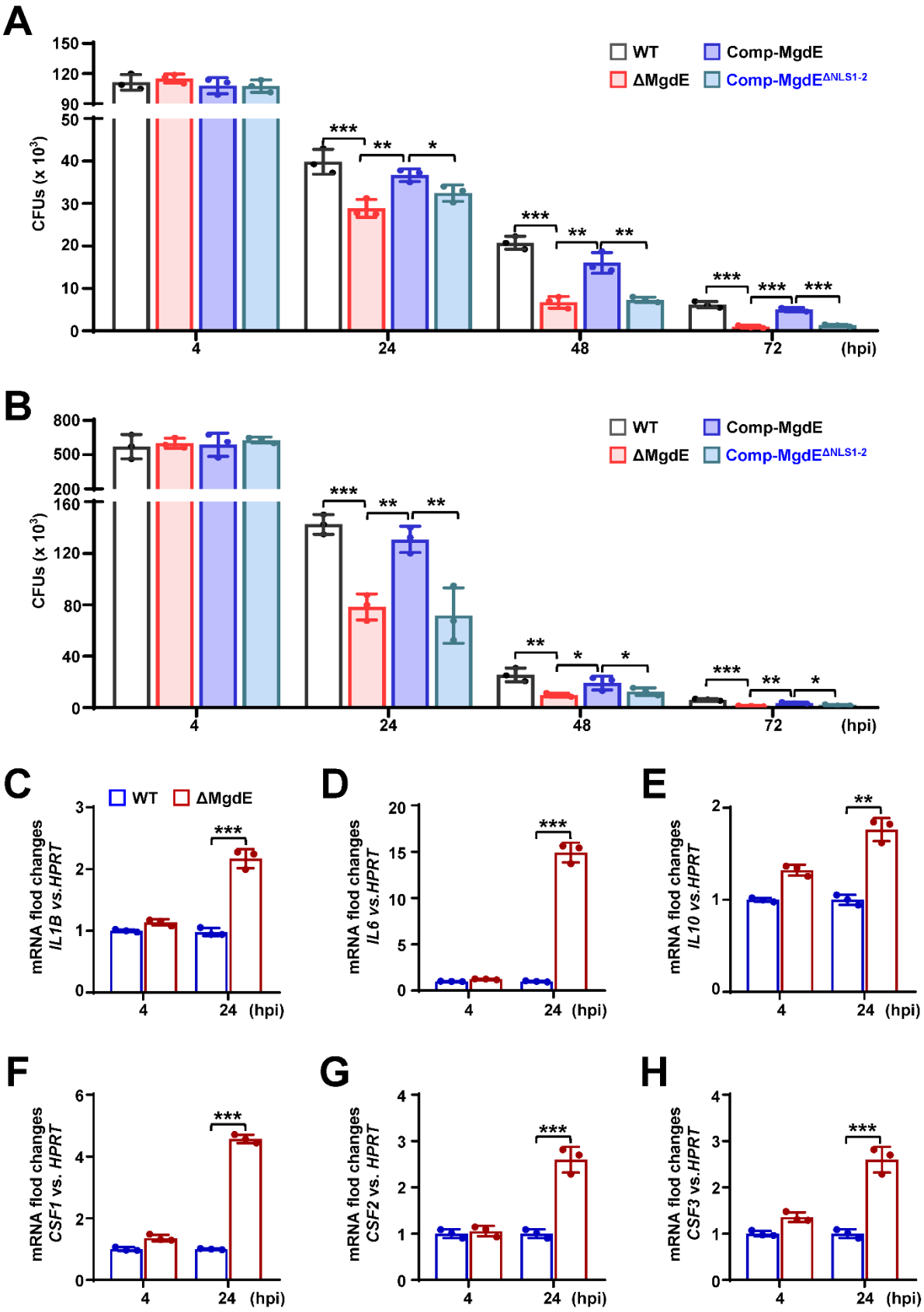
779

780

781 **Figure 2. Hypothetical protein MgdE enters the host nucleus via dual NLS.**

782 (**A**) Domain architecture of MgdE. Schematic representation of MgdE with the annotated
783 functional domains, including a Tat signal peptide (1–37 aa, twin-arginine translocation motif),
784 a GDSL-like lipase/acylhydrolase catalytic domain (74–258 aa), and two nuclear localization
785 signals (NLS1: 108–111 aa, NLS2: 300–305 aa). (**B**) Phylogenetic and structural conservation
786 of MgdE. (left) Neighbor–joining phylogenetic tree of MgdE homologs across Mycobacterium
787 species (1,000 bootstrap replicates, values ≥50% shown). (right) Clustal Omega sequence
788 alignment highlighting conserved residues (≥90% identity, red). Species abbreviations: *M.*
789 *tuberculosis* (Mtu), *M. bovis* BCG (Mbb), *M. tuberculosis* H37Ra (Mra), *M. africanum* (Maf), *M.*
790 *marinum* M (Mmi), *M. tuberculosis* CDC1551 (Mtc), *M. tuberculosis* KZN 1435 (Mtb), *M.*
791 *shinjukuense* (Mshj), *M. avium* (Mav), *M. liflandii* (Mli), *M. mantenii* (Mman), *M. intracellulare*
792 *subsp* (Mid). (**C**) Subcellular localization of EGFP-tagged wild-type MgdE and its NLS-deletion
793 mutants ($MgdE^{\Delta NLS1}$, $MgdE^{\Delta NLS2}$, and $MgdE^{\Delta NLS1\text{-}2}$). (left) Schematic representation of EGFP-
794 tagged constructs. (right) Representative confocal microscopy images of HEK293T cells
795 transfected with the indicated constructs for 36 h. EGFP fluorescence (green) and nuclear
796 staining with DAPI (blue) were visualized using FLUOVIEW software. Scale bar, 10 μm.
797 Images were acquired using a 100× oil immersion objective (NA = 1.4). (**D**) Quantification of
798 nuclear EGFP intensity in cells expressing wild-type or mutant MgdE constructs. Data are
799 presented as mean ± SD (n = 15 cells). Statistical significance was assessed by two-tailed
800 unpaired Student's *t*-tests (***$P < 0.001$). (**E**) Western blot analysis of nuclear and cytoplasmic
801 fractions from HEK293T cells transfected with wild-type MgdE and its NLS-deletion mutants.
802 The empty EGFP vector was used as a negative control. EGFP and MgdE-EGFP fusion

803 proteins were detected using an anti-GFP antibody. Histone H3 and β-actin served as nuclear
804 and cytoplasmic markers, respectively. (**F**) Nuclear localization of MgdE during infection. THP-
805 1 macrophages were infected with recombinant *M. bovis* BCG strains expressing Flag-tagged
806 wild-type MgdE or NLS-deletion mutants ($MgdE^{\Delta NLS1}$, $MgdE^{\Delta NLS2}$, and $MgdE^{\Delta NLS1\text{-}2}$) for 24
807 hours. Subcellular fractionation was performed, and cytoplasmic and nuclear fractions were
808 analyzed by immunoblotting using anti-Flag antibodies. Histone H3 and β-actin were used as
809 nuclear and cytoplasmic markers, respectively.
810

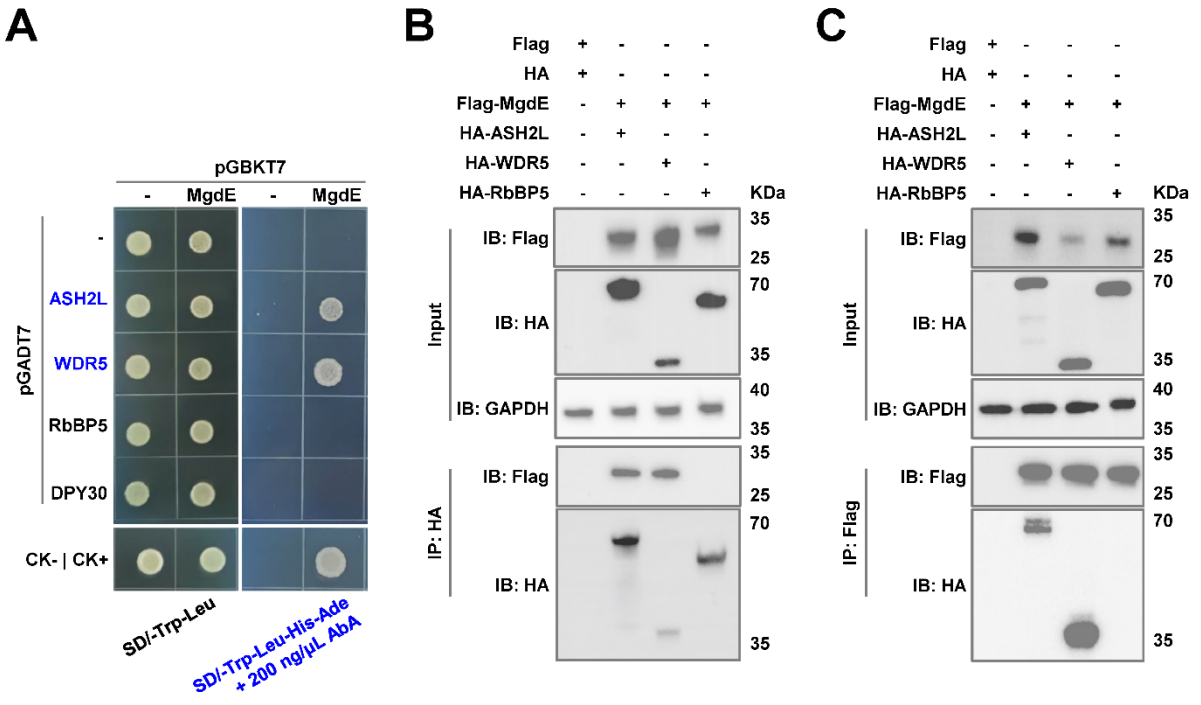

A
WT
ΔMgdE
Comp-MgdE
Comp-MgdEΔNLS1-2
CFUs (x 10³)
150
120
90
50
40
30
20
10
0
***
**
*
4
24
48
72
(hpi)
B
800
500
200
160
120
80
40
C
mRNA flod changes IL1B vs.HPRT
D
mRNA flod changes IL6 vs.HPRT
E
mRNA flod changes IL10 vs.HPRT
F
mRNA flod changes CSF1 vs. HPRT
G
mRNA flod changes CSF2 vs. HPRT
H
mRNA flod changes CSF3 vs.HPRT

811

812

Figure 3. Nuclear localization of MgdE facilitates mycobacterial intracellular survival in macrophages.

(**A**–**B**) Intracellular survival of *M. bovis* BCG strains. THP-1 human macrophages (**A**) and RAW264.7 murine macrophages (**B**) were infected (MOI = 10) with wild-type BCG (WT), MgdE deletion mutant (ΔMgdE), complemented strain (Comp-MgdE), or NLS-deficient complement (Comp-MgdE$^{\Delta NLS1\text{-}2}$). Bacterial survival was assessed by CFU enumeration at 2, 24, 48, and 72 h post-infection. (**C**–**H**) Cytokine expression in the infected THP-1 cells. qRT-PCR analysis of cytokine mRNA levels in THP-1 cells infected with WT or ΔMgdE strains for 4–24 h. Target genes include *IL1B* (**C**), *IL6* (**D**), *IL10* (**E**), *CSF1* (**F**), *CSF2* (**G**), and *CSF3* (**H**). Data represent mean ± SD of three independent biological replicates. Statistical significance determined using two-way ANOVA or two-tailed unpaired Student's *t*-tests, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.



827 **Figure 4. MgdE directly interacts with ASH2L and WDR5, core components of the host**

828 **COMPASS complex.**

829 (**A**) Yeast cells were co-transformed with bait (pGBKT7) and prey (pGADT7) plasmids

830 expressing wild-type MgdE and human COMPASS components (ASH2L, WDR5, RbBP5, and

831 DPY30). Growth was monitored on non-selective (-Leu/-Trp, left) and selective (-Leu/-Trp/-

832 Ade/-His + 200 ng/μL aureobasidin A, right) media. Controls: CK− (pGBKT7-*lam* + pGADT7-

833 T, negative), CK+ (pGBKT7-*p53* + pGADT7-T, positive). (**B**–**C**) Cells were co-transfected with

834 Flag-MgdE and HA-tagged ASH2L, WDR5, or RbBP5 (1:1 molar ratio). At 36 h post-

835 transfection, the lysates were immunoprecipitated using (**B**) anti-HA or (**C**) anti-Flag antibodies,

836 followed by immunoblotting with anti-HA, anti-Flag, and anti-GAPDH (loading control). The

837 input lanes represent 5% of the total lysate.

838

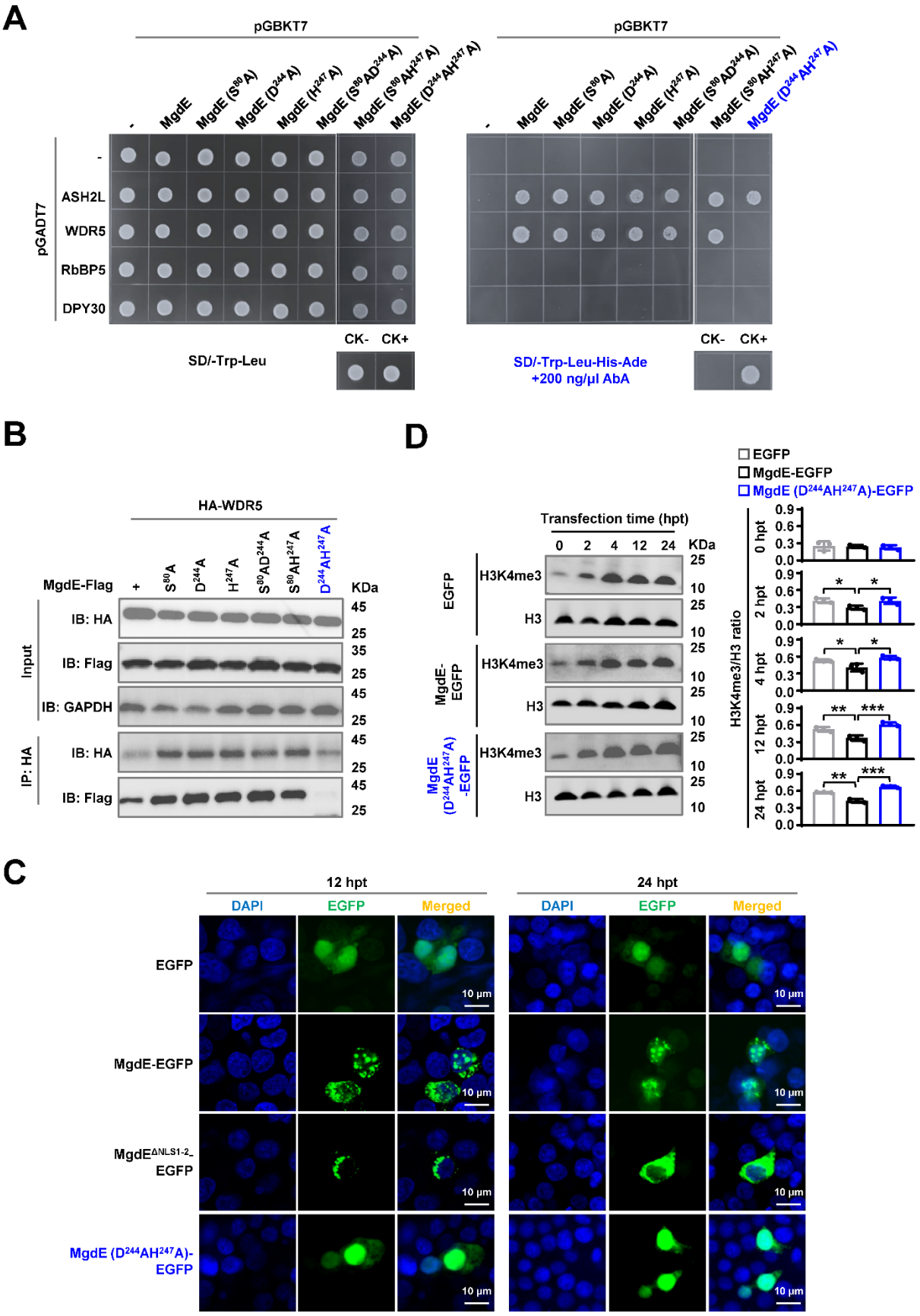
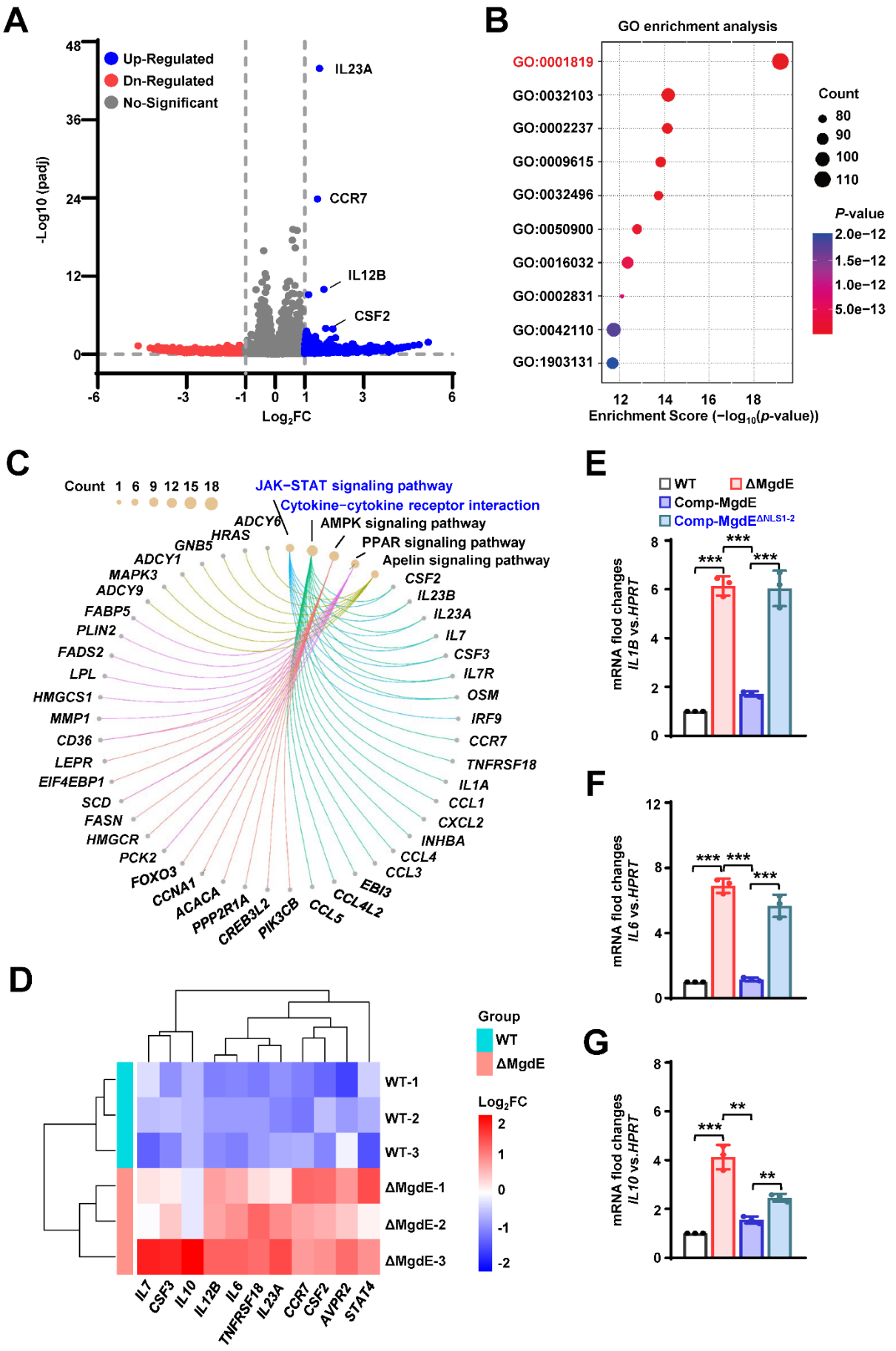


Figure 5. The conserved residues D224 and H247 mediate the binding ability of MgdE

841 **to WDR5.**

842 (**A**) Y2H assay identifying interactions between MgdE mutants and COMPASS complex
843 subunits. Yeast cells were co-transformed with bait (pGBKT7) and prey (pGADT7) plasmids
844 expressing wild-type or mutant MgdE and human COMPASS subunits (ASH2L, WDR5,
845 RbBP5, and DPY30). Growth was assessed on non-selective (-Leu/-Trp, left) and selective (-
846 Leu/-Trp/-Ade/-His + 200 ng/μL aureobasidin A, right) media. Controls: CK− (pGBKT7-*lam* +
847 pGADT7-T, negative) and CK+ (pGBKT7-*p53* + pGADT7-T, positive). (**B**) Co-IP analysis of
848 MgdE mutants with WDR5. HEK293T cells were co-transfected with Flag-tagged MgdE
849 mutants and HA-tagged WDR5 (1:1 molar ratio). Complexes were immunoprecipitated using
850 anti-HA antibody and protein A/G beads, followed by immunoblotting with anti-Flag and anti-
851 HA antibodies. (**C**) Nuclear distribution of wild-type and mutants MgdE. Confocal microscopy
852 of HEK293T cells expressing wild-type or D224A/H247A MgdE-EGFP at 12 and 24 h post-
853 transfection (hpt). Nuclear foci were visualized by EGFP (green) and DAPI (blue) staining.
854 Scale bar, 10 μm. Images were acquired with a ×100 oil immersion objective (NA = 1.4). (**D**)
855 Immunoblot analysis of H3K4me3 levels. HEK293T cells expressing wild-type or
856 D224A/H247A mutant MgdE were analyzed for changes in H3K4me3 levels over 0–24 h post-
857 transfection. Histone H3 was used as a loading control. Data represent mean ± SD of three
858 independent biological replicates. Statistical significance determined using two-tailed unpaired
859 Student's *t*-tests, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.


A
Up-Regulated
Dn-Regulated
No-Significant
IL23A
CCR7
IL12B
CSF2
-Log10 (padj)
Log2FC
B
GO enrichment analysis
GO:0001819
GO:0032103
GO:0002237
GO:0009615
GO:0032496
GO:0050900
GO:0016032
GO:0002831
GO:0042110
GO:1903131
Count
P-value
Enrichment Score (−log10(p-value))
C
Count 1 6 9 12 15 18
JAK−STAT signaling pathway
Cytokine−cytokine receptor interaction
AMPK signaling pathway
PPAR signaling pathway
Apelin signaling pathway
D
Group
WT
ΔMgdE
Log2FC
WT-1
WT-2
WT-3
ΔMgdE-1
ΔMgdE-2
ΔMgdE-3
IL7
CSF3
IL10
IL12B
IL6
TNFRSF18
IL23A
CCR7
CSF2
AVPR2
STAT4
E
WT
ΔMgdE
Comp-MgdE
Comp-MgdEΔNLS1-2
mRNA flod changes IL1B vs.HPRT
F
mRNA flod changes IL6 vs.HPRT
G
mRNA flod changes IL10 vs.HPRT

861 **Figure 6. MgdE suppresses host inflammatory responses probably by inhibition of**
862 **COMPASS complex-mediated H3K4 methylation.**

863 (**A**) Volcano plot of DEGs. DEGs between MgdE-deleted strain (ΔMgdE) and wild-type BCG
864 (WT) were visualized in a volcano plot. Genes with $|\log_2\text{-fold change}| \geq 1$ and $P < 0.05$ were
865 considered significant. The *x*-axis represents $\log_2$fold change, and the *y*-axis shows $-\log_{10}(P$-
866 value). (**B**) GO enrichment analysis of DEGs. GO analysis revealed the significant enrichment
867 of immune and inflammatory processes in ΔMgdE-infected macrophages compared to that in
868 WT-infected cells, including positive regulation of response to external stimulus (GO:0032103),
869 response to molecule of bacterial origin (GO:0002237), response to virus (GO:0009615),
870 response to lipopolysaccharide (GO:0032496), leukocyte migration (GO:0050900), viral
871 process (GO:0016032), regulation of response to biotic stimulus (GO:0002831), T cell
872 activation (GO:0042110), and mononuclear cell differentiation (GO:1903131). (**C**) KEGG
873 pathway enrichment analysis of DEGs. Chord diagram of KEGG pathway enrichment analysis
874 showing signaling pathways that are significantly enriched in ΔMgdE. (**D**) Heatmap of the
875 inflammatory gene expression. Heatmap depicting $\log_2$-foldchange levels of inflammatory
876 genes involved in the JAK–STAT and cytokine signaling pathways. Upregulated and
877 downregulated genes in ΔMgdE are shown in green and red, respectively. The data were *Z*-
878 score normalized. (**E**–**G**) qRT-PCR analysis of cytokine mRNA levels in THP-1 cells infected
879 with WT, ΔMgdE, MgdE-complemented (Comp-MgdE), and NLS-deleted complement (Comp-
880 $\text{MgdE}^{\Delta\text{NLS1-2}}$) at 24 h post-infection. Cytokines analyzed include *IL1B* (**E**), *IL6* (**F**), and *IL10* (**G**).
881 Data represent mean ± SD of three independent biological replicates. Statistical significance
882 determined using two-tailed unpaired Student's *t*-tests, $^{**}P < 0.01$, $^{***}P < 0.001$.

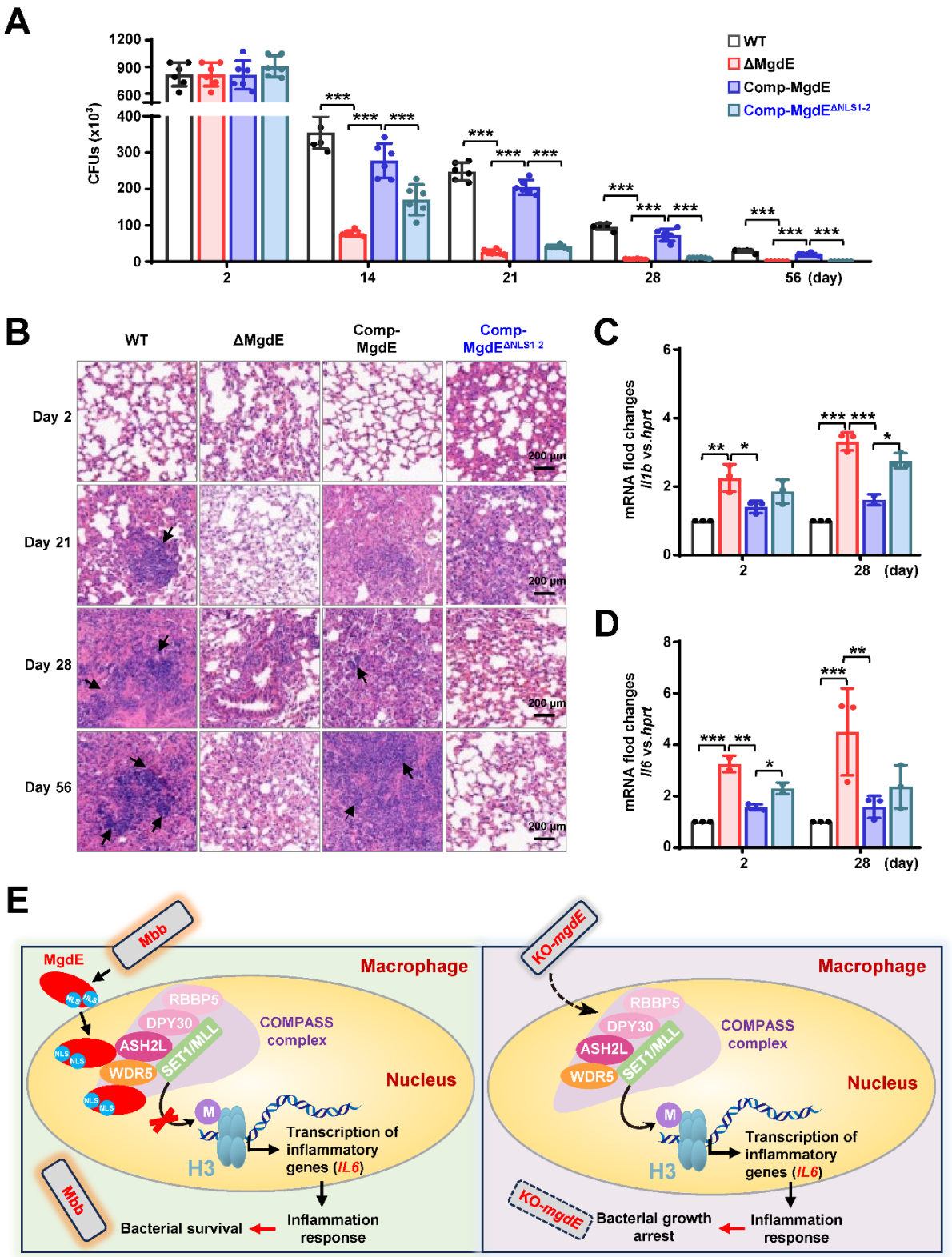
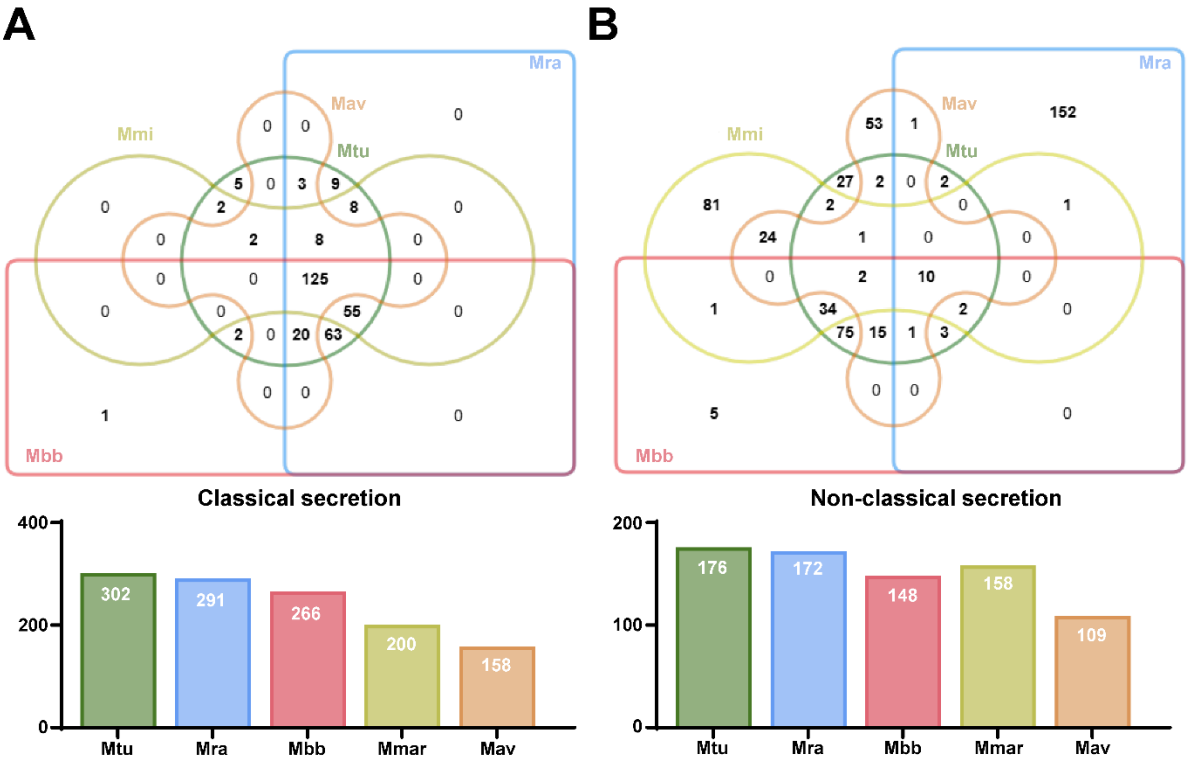


Figure 7. Nuclear localization of MgdE is essential for mycobacterial survival in mice. (**A**) Bacterial burden in the lungs of the infected mice. C57BL/6 mice (n = 6/group) maintained under SPF conditions were intratracheally infected with 1.0 × 10^7 colony-forming units (CFU)

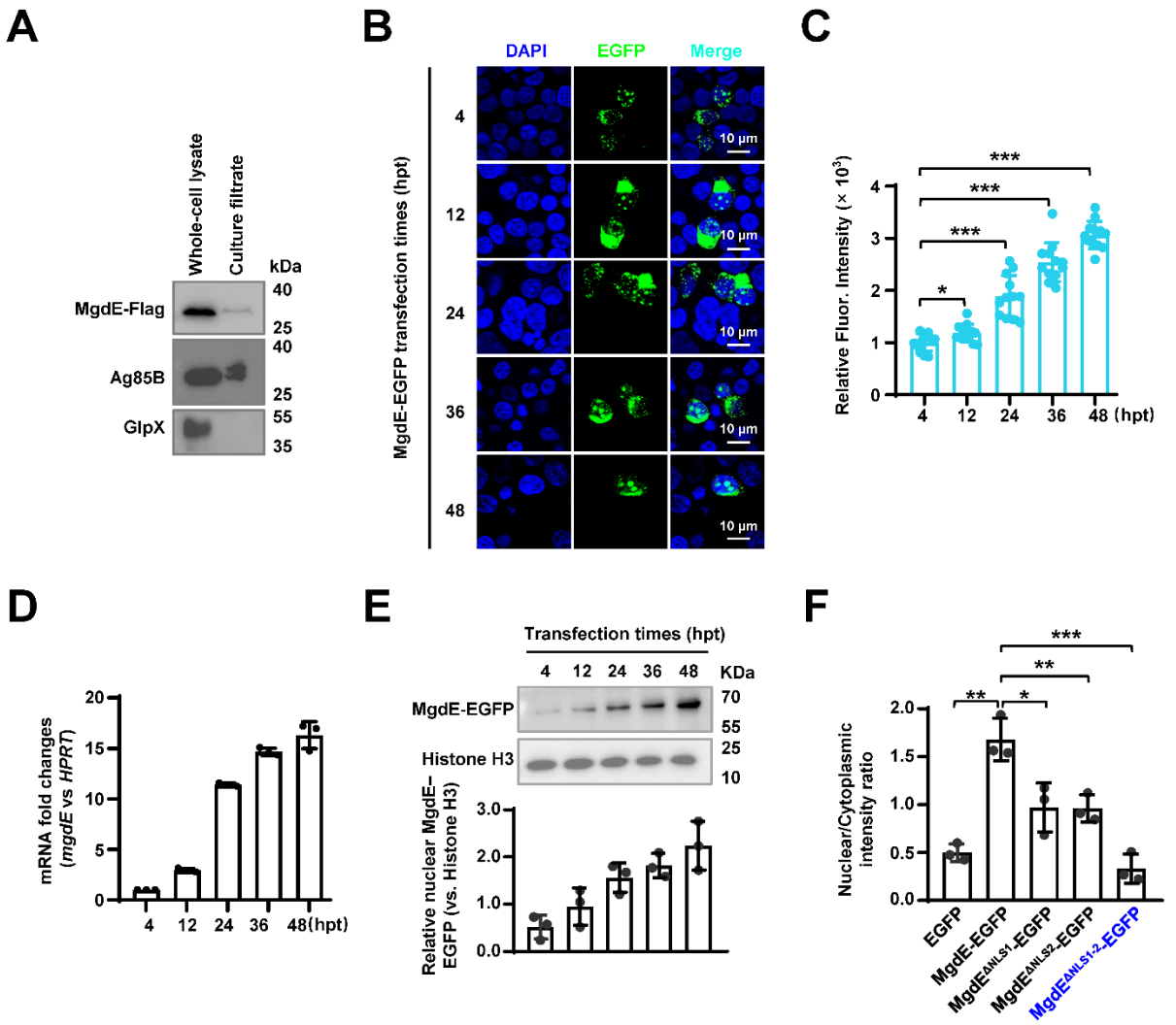
887 of *M. bovis* BCG strains, including wild-type (WT), MgdE-deleted (ΔMgdE), MgdE-
888 complemented (Comp-MgdE), and NLS-deleted complement (Comp-MgdE$^{\Delta NLS1\text{-}2}$). Lung
889 bacterial loads were quantified using CFU assays at 0, 14, 21, 28, and 56 d post-infection.
890 The data were obtained from a single experiment. (**B**) H&E-stained lung sections from infected
891 mice (as in **A**) revealed granulomatous inflammation. Scale bars: 200 μm. (**C**–**D**) Pro-
892 inflammatory cytokine expression in mice spleen. qRT-PCR analysis of cytokine mRNA levels
893 of *Il1b* (**C**) and *Il6* (**D**) in spleen tissues from infected mice (n = 6/group) at 2 and 28 days post-
894 infection. Data are presented as mean ± SD from six biologically independent experiments.
895 Statistical significance determined using two-way ANOVA, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.
896 (**E**) Mechanistic model showing how mycobacterial nucleomodulin MgdE hijacks the
897 COMPASS complex to suppress H3K4me3 and promote immune evasion. Upon *M. bovis*
898 BCG infection, the nucleomodulin MgdE is delivered into the host nucleus via its NLS and
899 directly binds to the COMPASS complex subunits, ASH2L or WDR5. This interaction disrupts
900 H3K4 trimethylation (H3K4me3) deposition, leading to the epigenetic suppression of pro-
901 inflammatory cytokine transcription (e.g., *IL6*), thereby facilitating the intracellular survival of
902 the pathogen.
903
904
905
906
907
908

Supplementary figures



Supplementary figure 1. Comparative analysis of classical and non-classical secreted proteins in mycobacterial species.

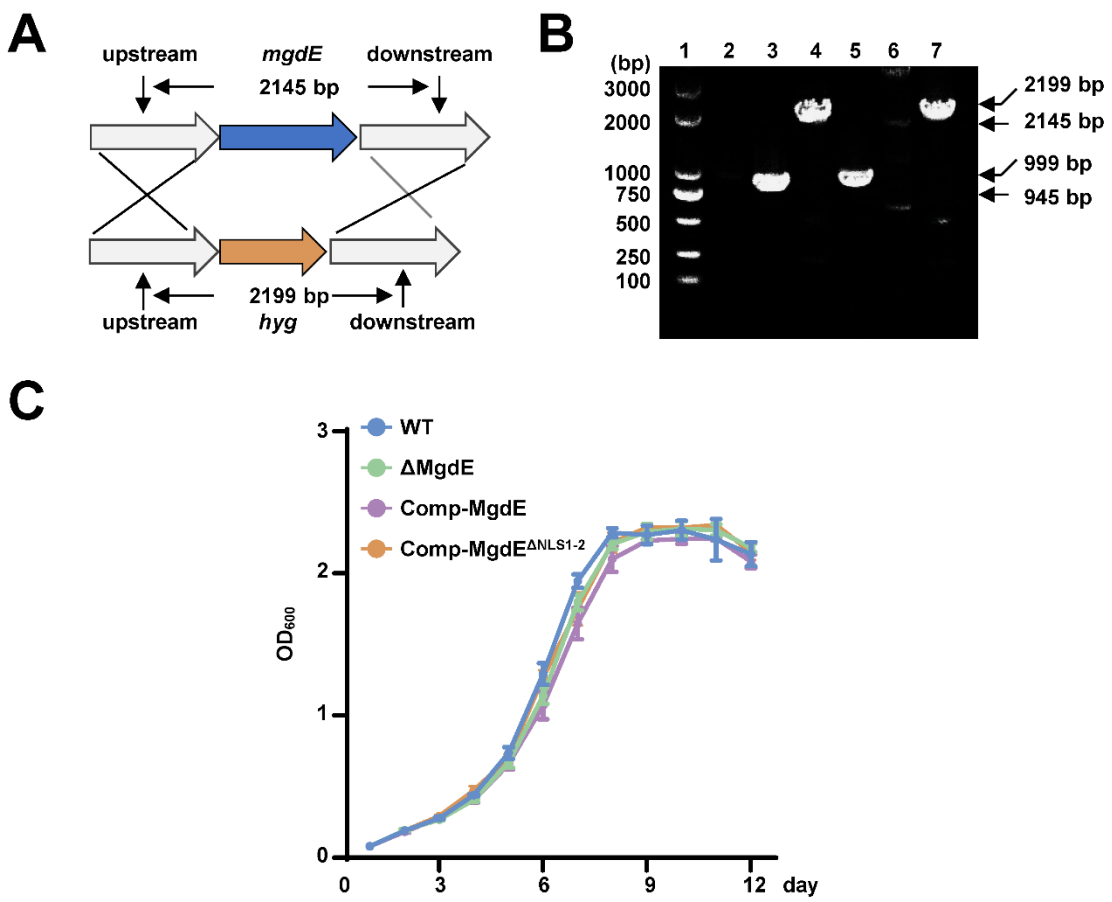
(**A**) Venn diagram showing the distribution of predicted classical secreted proteins in four Mycobacterium species: *M. tuberculosis* H37Rv (Mtu) and *M. tuberculosis* H37Ra (Mra), *M. bovis* BCG (Mbb), *M. marinum* (Mmar), and *M. avium* (Mav). Proteins were predicted using SignalP 5.0, with a signal peptide score (D-score) ≥ 0.5. A total of 125 proteins were conserved across all species. (**B**) Venn diagram showing the distribution of predicted non-classical secreted proteins across the same Mycobacterium species, predicted using SecretomeP 2.0 with a neural network (NN) score ≥ 0.9. Ten proteins were conserved across all species. Bar graphs (left: classical, right: non-classical) summarize the total number of predicted secreted proteins per species.



Supplementary Figure 2. Subnuclear localization of MgdE-EGFP.

(**A**) Immunoblot analysis of bacterial lysates and culture supernatants from *M. bovis* BCG strains expressing C-terminally Flag-tagged MgdE. Ag85B and GlpX were detected using anti-Ag85B and anti-GlpX antibodies, serving as positive and negative controls for protein secretion, respectively. MgdE-Flag was detected using an anti-Flag antibody. (**B**) Confocal microscopy was used to assess the nuclear distribution of MgdE-EGFP at various time points post-transfection. Nuclei were stained with DAPI (blue), MgdE-EGFP is shown in green. Scale bar: 10 μm. Images were acquired using a 100× oil immersion objective (NA = 1.4). (**C**)

Quantification of nuclear EGFP intensity in cells expressing wild-type or mutant MgdE constructs. Data are presented as mean ± SD (n = 12 cells). (**D**) Quantitative RT-PCR analysis of *mgdE* mRNA expression in HEK293T cells at different time points post-transfection (4–48 h). Data represent the transcriptional level of *mgdE* relative to *HPRT*. (**E**) Western blot analysis of nuclear fractions at different time points post-transfection (4–48 h), showing time-dependent nuclear accumulation of MgdE-EGFP. Histone H3 was used as a loading control for nuclear proteins. The lower panel shows quantification of nuclear MgdE-EGFP levels normalized to Histone H3. (**F**) Quantification of the nuclear and cytoplasmic distribution of EGFP, wild-type MgdE, and its NLS-deletion mutants based on the Western blot results shown in (**Figure 2E**). Data represent mean ± SD of three independent biological replicates. Statistical significance determined using two-tailed unpaired Student's *t*-tests, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



Supplementary Figure 3. Deletion of the nuclear localization signal of MgdE does not affect the growth of *M. bovis* BCG strains.

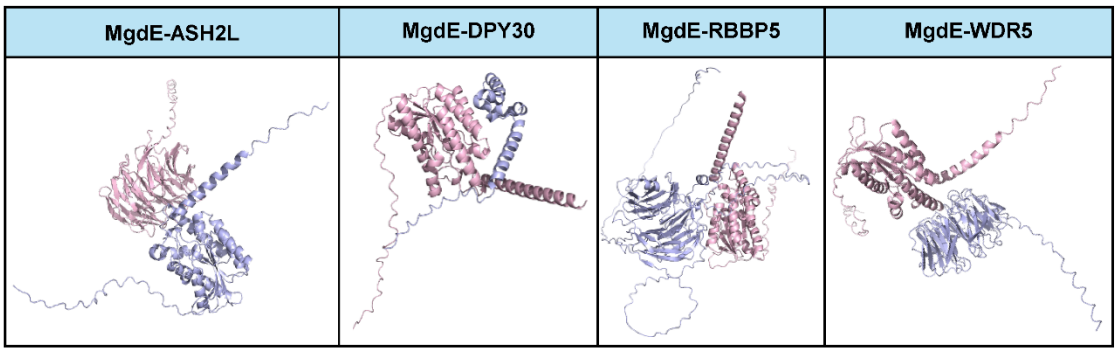
(**A-B**) Construction and validation of the MgdE-deleted strain of *M. bovis* BCG. (**A**) Schematic diagram of the homologous recombination strategy used to delete *mgdE* from the *M. bovis* BCG genome. (**B**) Wild-type and mutant strains were used as templates to amplify the *mgdE* gene (600 bp upstream–600 bp downstream) by PCR. Lanes 1 and 3: wild-type genomic DNA, lanes 2 and 4: Δ*mgdE* genomic DNA. (**C**) Growth curve analysis of *M. bovis* BCG strains. Growth of BCG strains, including wild-type BCG (WT), MgdE-deleted (ΔMgdE), MgdE-complemented (Comp-MgdE), and NLS-deleted complemented (Comp-MgdE$^{\Delta NLS1\text{-}2}$) strains, was measured in 7H9 medium. Data represent mean ± SD of three independent biological replicates.

956

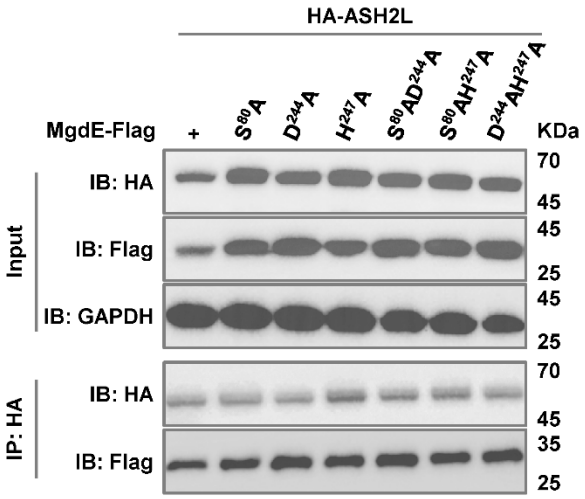
A

Function	COMPASS complexes	Predicted binding confidence (pLDDT)
The catalytic subunit	SET1A/B	-
Required for H3K4me3	ASH2L	0.47
Required for assembly	RBBP5	0.30
Required for assembly	WDR5	0.77
	DPY30	0.62
Components of the Set1 complexes	CXXC1	0.51
	WDR82	-
Components of the Set1 complexes	HCF1	0.37

B



C

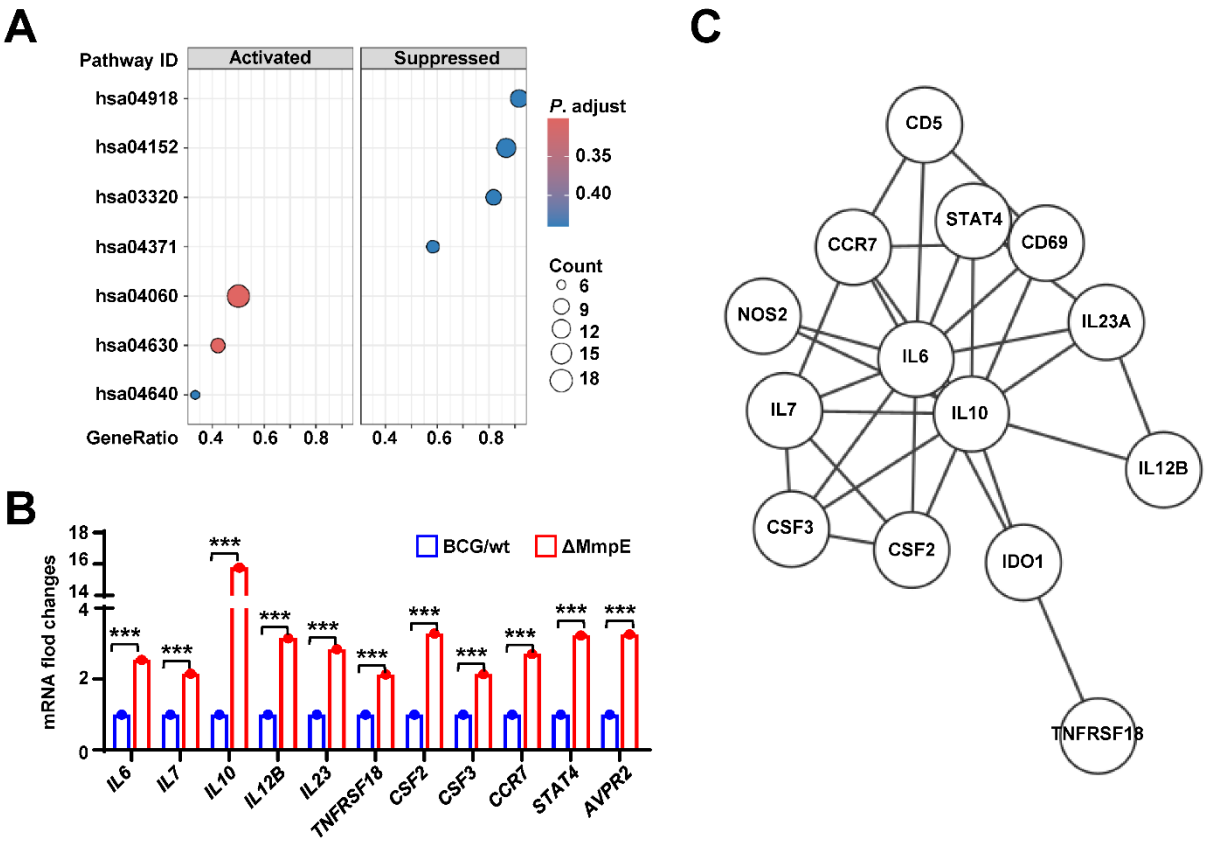


957

958 **Supplementary Figure 4. MgdE interacts with COMPASS complex subunits.**

959 (**A**) Predicted binding affinities between MgdE and COMPASS core subunits. The predicted

local distance difference test (pLDDT) scores calculated using AlphaFold 2.2.0 for the interactions between MgdE and the COMPASS subunits were as follows: ASH2L (pLDDT = 0.47), RbBP5 (pLDDT = 0.30), WDR5 (pLDDT = 0.77), and DPY30 (pLDDT = 0.62). Confidence levels are categorized as follows: High confidence: pLDDT ≥ 0.7 (strong predicted binding), Medium confidence: 0.5 ≤ pLDDT < 0.7 (moderate binding), Low confidence: pLDDT < 0.5 (weak predicted binding). (**B**) Structural modeling of MgdE-COMPASS interactions. AlphaFold generated models of the simulated binding interfaces are shown, with MgdE highlighted in red and COMPASS subunits (ASH2L, WDR5, RbBP5, and DPY30) shown in gray. (**C**) Co-IP analysis of MgdE mutants with ASH2L. HEK293T cells were co-transfected with Flag-tagged MgdE mutants and HA-tagged ASH2L (1:1 molar ratio). Co-IP was performed using anti-HA antibody-bound protein A/G beads, and immunoprecipitated complexes were analyzed by immunoblotting with anti-Flag and anti-HA antibodies.

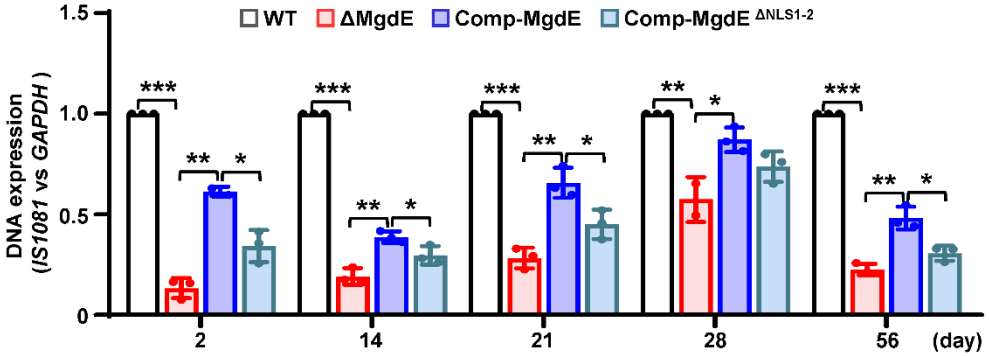


Supplementary Figure 5. MgdE suppresses cellular inflammatory responses during *M. bovis* BCG infection.

(**A**) KEGG pathway enrichment analysis. KEGG pathway analysis identified significantly enriched pathways in THP-1 cells infected with the ΔMgdE strain compared to those infected with wild-type BCG (WT). DEGs were predominantly associated with immune and signaling pathways, including thyroid hormone synthesis (hsa04918), AMPK signaling pathway (hsa04152), PPAR signaling pathway (hsa03320), apelin signaling pathway (hsa04371), Cytokine–cytokine receptor interaction (hsa04060), JAK–STAT signaling pathway (hsa04630), and hematopoietic cell lineage (hsa04640). (**B**) Elevated inflammatory gene expression in ΔMgdE-infected cells. Key upregulated inflammatory genes in ΔMgdE- *vs*. WT-infected cells are highlighted. (**C**) Functional enrichment analysis of upregulated genes. Analysis using

985 STRING and Cytoscape revealed regulation of inflammatory responses as a top enriched

986 biological process ($P < 0.05$).



987

988 **Supplementary Figure 6. MgdE facilitates bacterial colonization in the spleens of**

989 **infected mice.**

990 Bacterial colonization was assessed in splenic homogenates from infected mice (as described

991 in **Figure 7A**) by quantifying bacterial DNA using quantitative PCR at 2, 14, 21, 28, and 56

992 days post-infection. Data are presented as mean ± SD (n = 3). Statistical significance

993 determined using two-way ANOVA, *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.