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$^{108\text{-}111}$ and RLRRPR$^{300\text{-}305}$, 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 *il1β*, 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).

22 These evolutionarily diverse effector proteins, secreted by distinct pathogens, exhibit
23 remarkable adaptability in infiltrating host nuclei and reprogramming the transcriptional and
24 epigenetic regulatory networks (Bierne et al., 2012; Bierne et al., 2020). Upon nuclear
25 localization, nucleomodulins employ diverse molecular mechanisms to subvert host cellular
26 processes, including but not limited to: (i) direct interactions with genomic DNA or nuclear
27 proteins (Rennoll-Bankert et al., 2015; Sun et al., 2016; Prokop et al., 2017), (ii) hijacking
28 post-translational modification systems (Lebreton et al., 2014; Rolando et al., 2023), (iii)
29 dysregulating signaling molecule activity or localization (Zhao et al., 2019; Chambers et al.,
30 2020; Evans et al., 2018), and (iv) destabilizing nuclear architecture (Pourpre et al., 2022; Fu
31 et al., 2021). Specific examples include AnkA, a bacterial protein from *Anaplasma*
32 *phagocytophilum,* which binds to specific A/T-rich sequences in the host cell, alters
33 chromatin architecture, and modulates gene transcription (Rennoll-Bankert et al., 2015). In
34 *Listeria monocytogenes*, LntA interacts with the proline-rich domain of BAHD1, to disrupt
35 BAHD1-mediated gene silencing and promote histone H3 acetylation (Lebreton et al., 2014).
36 Another example is OspF from *Shigella flexneri*, which exploits its phosphothreonine lyase
37 activity to irreversibly dephosphorylate MAPKs, thereby preventing the transcriptional
38 activation of NF-κB-regulated genes (Chambers et al., 2020). Unlike these effectors, *L.*
39 *monocytogenes* InlP takes a different approach by hijacking the splicing factor RBM5. This
40 interaction not only suppresses apoptosis but also reorganizes nuclear speckles through
41 SC35 redistribution and RBM5-dependent granule formation (Pourpre et al., 2022). These
42 examples illustrate the diverse and sophisticated mechanisms used by nucleomodulins to
43 reprogram host nuclear functions, ultimately manipulating the intracellular environment to

44 their advantage.

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

63 *Mycobacterium tuberculosis* (Mtb), the causative agent of tuberculosis, is a highly
64 successful pathogen because of its ability to survive within macrophages, evade the immune
65 system, and persist in the host, ultimately leading to a chronic disease state (Chandra et al.,

66 2022; Chai et al., 2020). Central to its pathogenic strategy is the secretion of a diverse array
67 of effector proteins (Chai et al., 2022; Bates et al., 2024; Qiang et al., 2023), including
68 nucleomodulins that subvert host nuclear processes. The tyrosine phosphatase PtpA
69 exemplifies this strategy by directly binding to the promoter regions of immune-related
70 genes, including *GADD45A*, to suppress innate immune responses (Wang et al., 2017). The
71 methyltransferase Rv1988 mediates an unconventional histone modification by methylating
72 histone H3 at arginine 42 (H3R42me), thereby silencing antimicrobial defense mechanisms
73 (Yaseen et al., 2015). Similarly, the acetyltransferase Rv3423.1 modulates chromatin
74 accessibility through acetylation of histone H3 at lysine 9 and 14 (H3K9/K14ac), attenuating
75 pro-inflammatory signaling pathways (Jose et al., 2016). Rv2966c represents a unique
76 bacterial DNA methyltransferase that establishes non-canonical CpG methylation patterns to
77 epigenetically silence host immune-related gene expression (Sharma et al., 2015). Despite
78 these molecular insights, the comprehensive landscape of the nuclear effector systems of
79 Mtb, particularly the mechanistic details of nucleomodulin interactions with host chromatin-
80 modifying complexes and their spatiotemporal regulation during infection, remains poorly
81 defined. This critical knowledge gap impedes our understanding of how Mtb coordinates
82 system-level transcriptional rewiring in the host cellular environment.

83 In this study, we utilized bioinformatics analysis combined with fluorescence imaging to
84 rapidly screen for potential nucleomodulins in pathogenic mycobacterial species. Using this
85 approach, the hypothetical protein Rv1075c (designated as MgdE) was identified and
86 subsequently validated. MgdE undergoes nuclear translocation mediated by two nuclear
87 localization signals, $KRIR^{108-111}$ and $RLRRPR^{300-305}$, and interacts with the histone

methyltransferase COMPASS complex subunits ASH2L and WDR5. This interaction suppresses H3K4me3-mediated inflammatory gene expression, thus dampening host immune responses and enhancing bacterial survival in macrophages and murine infection models. Our findings identify MgdE as a nucleomodulin that targets the COMPASS complex, revealing a novel epigenetic strategy exploited by mycobacteria for intracellular survival.

RESULTS

Functional screening of potential nucleomodulins in pathogenic mycobacterial species

To identify conserved nucleomodulins in pathogenic mycobacteria, we first conducted a pan-genomic comparative analysis across four relevant species, including *M. tuberculosis* H37Rv and H37Ra, *M. bovis* BCG, *M. marinum*, and *M. avium*, to categorize the core genes. Subsequently, a dual-threshold screening approach was employed to predict the classical and non-canonical secreted proteins encoded by these shared genes using SignalP 5.0 and SecretomeP 2.0, with thresholds set at a D-score ≥0.5 (Hammond et al., 2018) and a neural network (NN) score ≥0.9 (Bendtsen et al., 2005) (Table S1). This analysis yielded 135 high confidence secreted protein candidates (Table S2). Further analysis of these candidates for nuclear localization signal (NLS) motifs using cNLS Mapper (NLS score ≥2) revealed that 56 proteins harbored at least one putative NLS. Following systematic classification based on secretion machinery dependencies, the candidate proteins were categorized into four distinct pathways: 26 Sec/SPI (signal peptide-driven), 7 Tat/SPI (twin-arginine translocase pathway), 18 Sec/SPII (lipoprotein-targeted), and eight unclassified candidates lacking canonical secretion motifs (Figure 1A and B, Figure S1A and B). Functional validation of nuclear

trafficking was conducted through the heterologous expression of enhanced green fluorescent protein (EGFP)-tagged constructs in HEK293T cells, followed by quantitative profiling of the subcellular distribution using laser confocal fluorescence microscopy. This orthogonal validation confirmed the nuclear enrichment of 23 effectors, including six Tat/SPI-associated candidates. Among these, Rv2577 exhibited the most pronounced nuclear enrichment, displaying a nuclear localization capacity stronger than that of the previously described nuclear modulator protein Rv1988 (Yaseen et al., 2015). Notably, MgdE not only localized to the host cell nucleus but also exhibited a distinct punctate distribution within the nucleus (Figure 1C). Nucleomodulins are effector proteins secreted by pathogens that translocate into the host cell nucleus to modulate host functions. To assess whether MgdE is secreted by mycobacteria, recombinant *M. bovis* BCG strains expressing C-terminally 2×Flag-tagged Ag85B (positive control) and MgdE were constructed. Immunoblot analysis confirmed high levels of Ag85B-2×Flag and MgdE-2×Flag in bacterial lysates. Notably, both proteins were also detected in the culture supernatants (Figure S1C), indicating that MgdE is secreted. These findings identify MgdE as a nucleomodulin in mycobacteria and led us to further investigate the mechanisms underlying its nuclear translocation and functional role during infection.

MgdE translocates into the host nucleus through its dual NLS

MgdE is a putative GDSL-like lipase characterized by conserved esterase sequence motifs, with the GDSx motif containing a nucleophilic Ser residue as part of block I (equivalent to the classical G-x-S-x-G motif of lipases/esterases), and the consensus amino acids Gly, Asn, and His located in blocks II, III, and V, respectively (Yang et al., 2019). Phylogenetic analysis

across the mycobacterial species revealed strict evolutionary conservation of MgdE, with most strains encoding two predicted NLS: KRIR$^{108-111}$ (NLS1) and RLRRPR$^{300-305}$ (NLS2) (Figure 2A and B). To further investigate the nuclear translocation capability of MgdE, MgdE-EGFP was transfected into HEK293T cells, and its subcellular localization was analyzed at various time points post-transfection. Confocal microscopy revealed a time-dependent decrease in the cytoplasmic levels of MgdE, accompanied by its progressive accumulation in the host cell nucleus. Notably, MgdE displayed a distinct punctate distribution pattern within the nucleus, with nearly all the protein localizing to the nucleus approximately 36 h post-transfection (Figure S2). To quantitatively validate these observations, the temporal dynamics of MgdE expression in HEK293T cells were monitored, and nuclear enrichment was further assessed through cell fractionation. The results demonstrated a time-dependent increase in MgdE expression levels, which closely correlated with its gradual and sustained nuclear accumulation (Figure 2C). Collectively, these findings indicate that MgdE is an evolutionarily conserved mycobacterial nucleomodulin with nuclear translocation capacity.

To dissect the functional contributions of the predicted NLS motifs, EGFP-tagged NLS deletion mutants including MgdE$^{\Delta NLS1}$, MgdE$^{\Delta NLS2}$, and MgdE$^{\Delta NLS1-2}$ were constructed and transfected into HEK293T cells. Fluorescence intensity profiling revealed that individual deletion of either NLS1 or NLS2 partially reduced the nuclear targeting efficiency, whereas simultaneous deletion of both the NLS motifs (MgdE$^{\Delta NLS1-2}$) completely abrogated its nuclear localization (Figure 2D and E). These observations were further confirmed by immunoblot analysis of nuclear fractions, which showed that at 36 h post-transfection, MgdE-EGFP and its mutant variants were consistently detected in the cytoplasmic compartment, and nuclear

accumulation of MgdE$^{\Delta NLS1-2}$ was nearly undetectable compared to that of MgdE, MgdE$^{\Delta NLS1}$, and MgdE$^{\Delta NLS2}$ (Figure 2F). Together, these findings demonstrate that NLS1 and NLS2 are both required for efficient nuclear translocation of MgdE.

Nuclear localization of MgdE enhances mycobacterial intracellular survival within macrophages

Given that MgdE exhibits nuclear translocation ability, we sought to determine whether this ability plays a role in bacterial survival during infection. We generated various recombinant *M. bovis* BCG strains, including wild-type BCG (WT), a *mgdE* deletion mutant (ΔMgdE) (Figure S3A and B), ΔMgdE complemented with wild-type *mgdE* (Comp-MgdE), and ΔMgdE complemented with NLS-deleted variant (Comp-MgdE$^{\Delta NLS1-2}$). All strains exhibited similar growth under standard culture conditions (Figure S3C), and subsequently used them to infect THP-1 macrophages. As shown in Figure 3A, the ΔMgdE mutant strain exhibited significantly reduced survival compared to the WT strain, with accelerated bacterial clearance over time. This defect was rescued in the Comp-MgdE strain, with intracellular survival levels restored to that in the WT strain. Notably, the Comp-MgdE$^{\Delta NLS1-2}$ showed bacterial clearance equivalent to that in the ΔMgdE strains. These results were consistent with the infection experiments using RAW264.7 macrophage (Figure 3B). Furthermore, macrophages infected with the ΔMgdE mutant strains exhibited significant upregulation of inflammatory cytokines, including *il1β, il6*, *il10,* and the colony-stimulating factors *csf1-csf3* compared to WT-infected controls (Figure 3C-H). Collectively, these results demonstrated that MgdE enhances mycobacterial intracellular survival and suppresses host inflammatory responses during infection, with both NLS1 and NLS2 motifs being essential for its nuclear

trafficking-dependent function.

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

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

The conserved residues D^{224} and H^{247} are critical for mediating the binding of MgdE to WDR5

To further confirm the interactions between MgdE and WDR5/ASH2L, the key amino acids of MgdE (S^{80}, D^{244}, and H^{247}) were individually replaced with alanine (A) (Yang et al., 2019). The Y2H assays showed that single-residue substitutions of MgdE did not affect its interaction with ASH2L or WDR5. However, the $D^{244}A/H^{247}A$ double mutation specifically abrogated its interaction with WDR5 but retained its ASH2L binding ability (Figure 5A). These findings were further confirmed by Co-IP assays in HEK293T cells. Flag-tagged MgdE precipitated HA-tagged ASH2L and WDR5 but not RbBP5, with reciprocal pull-downs confirming bidirectional binding specificity. The $D^{244}A/H^{247}A$ double mutant maintained its interaction with ASH2L but failed to bind WDR5, suggesting that these interactions are mediated through distinct molecular interfaces (Figure 5B, Figure S5). To further investigate the functional consequences of these mutations, we analyzed the subcellular localization of the MgdE mutants in transfected HEK293T cells using fluorescence co-localization assays. Wild-type MgdE exhibited progressive accumulation of punctate nuclear foci over time, a phenomenon that was abrogated following transfection with the $D^{244}A/H^{247}A$ mutant (Figure 5C), suggesting that the double mutation impairs MgdE activity.

Given that MgdE interacts with WDR5, a subunit of COMPASS complex, which is a key regulator of H3K4me3 methylation (Deng et al., 2024), we sought to determine whether this interaction contributes to an epigenetic regulatory mechanism that facilitates bacterial persistence within host cells. We transfected HEK293T cells with the EGFP-tagged WT MgdE and mutated MgdE-$D^{244}A/H^{247}A$, and monitored H3K4me3 modifications over time. Strikingly, immunoblotting of nuclear extracts collected at defined time points revealed that cells expressing WT MgdE exhibited approximately 2-fold lower H3K4me3 levels compared

to those expressing the MgdE-$D^{244}A/H^{247}A$ mutant, an effect detectible within 4–12 h post-transfection (Figure 5D). These findings indicate that MgdE inhibits COMPASS-dependent methyltransferase activity via its functional interaction with WDR5.

Together, these data indicate that the residues D^{244} and H^{247} are essential for the MgdE-WDR5 interaction, which enables MgdE to attenuate H3K4me3 deposition in the host nuclei by subverting COMPASS complex function, a mechanism critical for bacterial survival and immune evasion.

MgdE suppresses inflammatory responses mediated by the COMPASS complex via inhibition of H3K4 methylation

As H3K4me3, a hallmark of transcriptionally active promoters, facilitates transcriptional machinery assembly and RNA polymerase II recruitment (Wang et al., 2023), we further investigated the transcriptional consequences of MgdE-mediated H3K4me3 suppression via its interaction with the COMPASS complex. To assess the genome-wide effects of MgdE on host gene expression, we performed RNA-seq profiling of THP-1 macrophages infected with either WT BCG (BCG/wt) or MgdE-deleted (Ko*mgdE*) strains. Transcriptomic analysis at 24 h post-infection revealed that 271 genes showed significant differential expression in the cells infected with Ko*mgdE* strain compared to that in the cells infected with WT BCG strain, with 222 genes being transcriptionally upregulated in the absence of MgdE (Figure 6A). Gene ontology (GO) enrichment analysis indicated that these upregulated genes were primarily associated with biological processes involved in immune responses, with a notable enrichment in “Positive regulation of cytokine production” (GO: 0001819), a critical pathway for amplifying inflammatory signaling (Figure 6B). Furthermore, the Kyoto encyclopedia of

genes and genomes (KEGG) pathway enrichment analysis highlighted a significant activation of the “Cytokine-cytokine receptor interaction” and “JAK-STAT signaling pathway” in THP-1 cells infected with the Ko*mgdE* strain (Figure 6C, Figure S6A). Heatmap analysis further confirmed the marked upregulation of inflammation-associated transcripts, including inflammatory cytokines (*il6*, *il10,* and *il12b*), chemokines (*ccl1-ccl5*), and colony-stimulating factors (*csf1-csf3*), in Ko*mgdE*-infected cells (Figure 6D, Figure S6B). Protein-protein interaction network analysis of the upregulated genes revealed tightly interconnected clusters within inflammatory networks, particularly those orchestrated by the central inflammatory mediator IL-6 (Figure S6C). To validate this result, we infected macrophages with the *M. bovis BCG* strains, including WT, ΔMgdE, Comp-MgdE, and Comp-MgdE$^{\Delta NLS1-2}$, under identical conditions for 24 h and examined the expression of relevant inflammatory factors. Compared to cells infected with the WT and Comp-MgdE strains, those infected with the ΔMgdE and Comp-MgdE$^{\Delta NLS1-2}$ strains exhibited significantly increased expression levels of inflammatory factors (*il6*, *il10,* and *il12b*) (Figure 6E-G). Collectively, these results demonstrate that MgdE effectively inhibits the production of inflammatory factors during BCG infection and highlights the critical role of its NLSs in this process. Through its interaction with the COMPASS complex, MgdE suppresses H3K4me3 levels, thereby attenuating the transcriptional activation of inflammation-related genes.

Nuclear localization of MgdE is essential for mycobacterial survival in mice

Our experimental data demonstrated that MgdE enhances bacterial survival in macrophages. To further investigate the *in vivo* functional significance of MgdE in host-pathogen interplay, we infected C57BL/6 mice with different *M. bovis* BCG strains, including WT, ΔMgdE, Comp-

MgdE, and Comp-MgdE$^{\Delta NLS1-2}$, and comprehensively evaluated bacterial survival, pulmonary inflammation, and cytokine dynamics. Mice infected with ΔMgdE exhibited a significantly reduced lung bacterial burden compared to those infected with WT or Comp-MgdE strains. Moreover, mice infected with Comp-MgdE$^{\Delta NLS1-2}$ exhibited a significantly lower pulmonary bacterial load than those infected with Comp-MgdE (Figure 7A). These findings collectively establish that MgdE is essential for optimal bacterial survival during infection, with the NLS being critical for its full functionality. Furthermore, histopathological assessment of the lung tissues corroborated these observations. Hematoxylin-eosin (H&E) staining revealed more pronounced inflammatory pathology in the lungs of mice infected with WT or Comp-MgdE strain compared to those infected with ΔMgdE or Comp-MgdE$^{\Delta NLS1-2}$ (Figure 7B), underscoring the necessity of MgdE and its NLSs in sustaining inflammation during infection. To further investigate the systemic immune modulation mediated by MgdE, we quantified splenic bacterial colonization and the inflammatory cytokine profiles. Consistent with the attenuated bacterial survival observed in the lungs, mice infected with ΔMgdE and Comp-MgdE$^{\Delta NLS1-2}$ exhibited a significantly reduced splenic bacterial load compared to those infected with WT or omp-MgdE (Figure S7). Despite diminished bacterial dissemination, infections with ΔMgdE and Comp-MgdE$^{\Delta NLS1-2}$ strains triggered elevated expression of inflammatory cytokines (*il-1β, il-6*) relative to infection with the WT and Comp-MgdE strains (Figure 7C and D). Collectively, these findings indicate that MgdE orchestrates mycobacterial survival within host tissues by suppressing inflammatory responses to evade pathogen clearance, with the NLS playing a pivotal role in mediating these host-pathogen interactions.

284 **DISCUSSION**

285 **MgdE functions as a critical nucleomodulin that contributes to mycobacterial**
286 **pathogenesis.** Its evolutionary conservation across mycobacterial species, coupled with
287 nuclear translocation via dual N-terminal and C-terminal NLSs, supports its classification as
288 a nucleomodulin. Previous investigations have described MgdE as a secreted effector
289 protein (Penn et al., 2018), and our experimental results independently confirm its secretion
290 (Figure S1C). Furthermore, our findings demonstrate that MgdE's nuclear localization plays
291 a critical role during infection. Disruption of MgdE expression or its NLS motifs significantly
292 attenuated *M. bovis* BCG viability within macrophages (Figure 3A and B) and reduced
293 bacterial burden in mice (Figure 7A and Figure S7), supporting the notion that nuclear
294 targeting contributes to bacterial virulence. This phenotype aligns with the observed
295 upregulation of MgdE during Mtb infection of alveolar macrophages (Pisu et al., 2020),
296 suggesting that nuclear activity of MgdE may contribute to an adaptive persistence
297 mechanism during infection. Although MgdE was previously characterized as a lipase (Yang
298 et al., 2019), our findings reveal a novel functional duality wherein nuclear targeting
299 coordinates bacterial survival through epigenetic interference while preserving its metabolic
300 roles. This functional pleiotropy positions MgdE as a master virulence regulator, exploiting
301 host-pathogen interface plasticity to sustain infection.

302 **MgdE directly subverts the COMPASS complex to reprogram host epigenetics.** To our
303 knowledge, MgdE is the first identified bacterial effector that directly targets and subverts the
304 COMPASS complex, a central regulator of H3K4 methylation. Nuclear-localized MgdE
305 specifically interacts with the core COMPASS subunits ASH2L and WDR5, disrupting

306 H3K4me3 deposition through structural interference. The COMPASS complex facilitates
307 H3K4 methylation via a conserved assembly mechanism in which WDR5 scaffolds
308 methyltransferases such as MLL1 through its WIN motif and stabilizes the ASH2L–RbBP5
309 dimer (Guarnaccia et al., 2021; Hsu et al., 2018). This dimer, in turn, positions the SET
310 domain for efficient substrate recognition (Xue et al., 2019). MgdE appears to interfere with
311 this assembly process, resembling strategies employed by host regulatory proteins.

312 Site-directed mutagenesis studies have identified critical residues (D^{224} and H^{247}) within
313 MgdE that are essential for WDR5 binding. Mutating these residues abolished both WDR5
314 interaction and H3K4me3 suppression (Figure 5C, D), confirming their functional relevance.
315 Moreover, nuclear MgdE accumulates in punctate subnuclear foci (Figure 1C, Figure 2D,
316 Figure 5C), reminiscent of *L. monocytogenes* InlP, which modulates nuclear speckles to
317 inhibit apoptosis (Pourpre et al., 2022). This interaction silences the expression of
318 antimicrobial effector genes, thereby facilitating immune evasion. Similar to *Legionella*
319 *pneumophila* RomA, which methylates H3K14 to repress innate immune responses (Schator
320 et al., 2023; Rolando et al., 2013), or *M. tuberculosis* Rv2067c, a structural mimic of host
321 DOT1L that catalyzes non-nucleosomal H3K79 trimethylation to subvert pro-inflammatory
322 signaling (Singh et al., 2023), MgdE exemplifies a broader pathogen strategy of hijacking the
323 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
326 of cytokine-cytokine receptor interactions and the JAK-STAT signaling pathway, along with
327 significantly elevated production of pro-inflammatory cytokines, including *il6* and *il1β*, in THP-

328 1 macrophages compared to the WT BCG strain (Figure 3C-F, Figure 6E-F). This profile is
329 consistent with the functional restoration of the COMPASS complex (Wang et al., 2025; Yu et
330 al., 2017). These findings support the hypothesis that MgdE suppresses host inflammatory
331 responses by disrupting COMPASS-mediated H3K4me3 deposition at active promoters.
332 H3K4me3 is a well-established histone mark associated with transcriptional activation, in
333 part through the recruitment of PHD domain-containing reader proteins (Hyun et al., 2017).
334 For instance, SET1-dependent H3K4me3 accumulation at NF-κB target promoters such as
335 *il6* and *tnfα* enhances pro-inflammatory gene expression (Bhattacharya et al., 2023). In
336 addition, WDR5 has been implicated in upregulating immunosuppressive cytokines,
337 including *il6* and *tgfβ* (Deng et al., 2024). In transfected HEK293T cells, WT MgdE
338 significantly reduced global H3K4me3 levels, whereas the catalytically inactive $D^{224}A/H^{247}A$
339 mutant failed to do so (Figure 5D), supporting the notion that MgdE-mediated suppression of
340 H3K4me3 depends on its ability to interact with WDR5. Furthermore, during infection, the
341 *MgdE*-deleted BCG strain markedly activated the expression of inflammatory factors
342 compared to the WT BCG strain (Figure 6C-G). These results suggest that MgdE attenuates
343 the expression of inflammatory mediators such as *il6* and *il1β* by reducing host H3K4me3
344 levels during Mycobacterium infection. This mechanism is similar to those observed in other
345 pathogens, such as spirochete-derived factors that suppress inflammation by reducing
346 H3K4me3 levels at the *tnfα* and *il-6* promoters (Chauhan et al., 2015).

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

350 interacts with ASH2L and WDR5 to destabilize the COMPASS assembly, resulting in
351 reduced H3K4me3 levels at the promoters of pro-inflammatory genes. This epigenetic
352 suppression downregulates cytokine transcription (e.g., *il6 and il1β*) and enhances bacterial
353 survival in macrophages and in mice. Our findings reveal the COMPASS complex as a
354 previously unrecognized target of bacterial effectors and offers mechanistic insights into
355 immune evasion and host-pathogen epigenetic interplay.

356 **MATERIALS AND METHODS**

357 **Bacterial strains and culture conditions**

358 The bacterial strains used in this study are detailed in Table S4. *Escherichia coli* DH5α was
359 cultured in Luria-Bertani medium under standard conditions. *Mycobacterium bovis BCG*
360 strains were grown in Middlebrook 7H9 broth (BD Biosciences), containing 1× OADC (oleic
361 acid, albumin, dextrose, catalase), 0.05% Tween-80, and 2% glycerol. Antibiotics were
362 added as required at the following final concentrations: 50 μg/mL hygromycin (for
363 mycobacteria) and 50 μg/mL kanamycin (for both mycobacteria and *E. coli*). The Y2HGold
364 yeast strain (Takara Bio) was maintained in YPDA (yeast extract, peptone, dextrose, and
365 adenine sulfate) medium. Transformants were selected on SD/-Trp medium, which lacks
366 tryptophan, to select for plasmid uptake. Protein-protein interactions were assessed in SD/-
367 Trp/-Leu (lacking tryptophan and leucine) and SD/-Trp/-Leu/-Ade/-His (lacking tryptophan,
368 leucine, adenine, and histidine) media, with growth on higher stringency media indicating
369 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
375 nM phorbol 12-myristate 13-acetate (PMA) for 24 h, followed by an additional 48-h
376 incubation in fresh PMA-free medium to allow complete differentiation prior to their use in
377 experiments.
378 HEK293T and RAW264.7 monocytes were cultured in Dulbecco's modified Eagle's medium
379 (DMEM) supplemented with 10% FBS and 50 μg/mL penicillin-streptomycin. All cells were
380 maintained at 37 °C in a humidified atmosphere with 5% CO_2.

381 **Pan-genomic nucleomodulin screening**

382 To systematically identify evolutionarily conserved nucleomodulins in pathogenic
383 mycobacteria, a pan-genomic comparative analysis was performed across four
384 phylogenetically representative mycobacterium strains: *Mycobacterium tuberculosis* H37Rv
385 (NCBI RefSeq assembly GCF_000195955.2), *Mycobacterium tuberculosis* H37Ra (NCBI
386 RefSeq assembly GCF_001938725.1), *Mycobacterium bovis* BCG (NCBI RefSeq assembly
387 GCF_000009445.1), *Mycobacterium marinum* (NCBI RefSeq assembly GCF_000018345.1),
388 and *Mycobacterium avium* subsp. (NCBI RefSeq assembly GCF_009741445.1). Complete
389 genome sequences and annotations were obtained from NCBI RefSeq with strict inclusion
390 criteria requiring "Complete Genome" status and fewer than 50 contigs to ensure the
391 assembly integrity. Plasmid sequences, as well as low-complexity and repetitive regions,
392 were filtered using the NCBI Genome Data Viewer toolkit. Genome quality control was
393 rigorously conducted using CheckM v1.2.2 with a *Mycobacteriaceae*-specific marker set

394 (including essential housekeeping genes such as *dnaA* and *rpoB*), confirming that all the
395 selected genomes exhibited >99% completeness and <1% contamination.

396 Ortholog clustering of protein-coding sequences (CDS) was executed through OrthoFinder
397 v2.5.4, employing Diamond alignment in ultra-sensitive mode (--ultra-sensitive), an E-value
398 cutoff of ≤1e-5 (--blastp_evalue 1e-5), and a minimum sequence identity threshold of ≥50%
399 (--min_percent_identity 50). The Markov cluster algorithm (MCL) was applied with an
400 inflation parameter of 1.5 (--mcl_inflation 1.5) to define orthologous groups, balancing
401 sensitivity, and specificity. Core orthologs were identified as clusters present in all four
402 strains, with additional filtering to retain only single-copy genes in ≥90% of the strains,
403 minimizing paralog-related artifacts.

404 Functional annotation of the core genes was conducted using InterProScan v5.62, which
405 integrates domain predictions from the Pfam, TIGRFAM, and SUPERFAMILY databases. To
406 prioritize candidate nucleomodulins, a sequential screening pipeline was applied, beginning
407 with the prediction of classical secretory proteins using SignalP v5.0 (D-score ≥0.5), followed
408 by the identification of non-classically secreted proteins via SecretomeP v2.0 (NN score
409 ≥0.9), and culminating in NLS detection using cNLS Mapper (NLS score ≥2.0), which
410 predicts both monopartite and bipartite NLSs. This integrated strategy enabled the
411 systematic identification of conserved candidates that exhibit both secretory potential and
412 nuclear targeting capacity.

413 **Recombinant expression and secretion of MgdE**

414 Recombinant *M. bovis* BCG strains expressing C-terminally 2×Flag-tagged Ag85B (positive

415 control) and MgdE were constructed to assess protein secretion under standard in vitro
416 culture conditions. The strains were cultured in 200 mL of Middlebrook 7H9 broth
417 supplemented with albumin-dextrose-catalase (ADC), glycerol, and 0.05% Tween-80 at
418 37 °C until reaching mid-log phase (OD_{600} = 0.6). Bacterial cultures were centrifuged at
419 4,000 × g for 60 minutes to separate cells from supernatants. Pellets were washed and lysed
420 by bead beating in PBS supplemented with protease inhibitors. Culture supernatants were
421 passed through 0.22 μm filters to remove residual bacteria and subsequently concentrated to
422 2 mL using Amicon Ultra centrifugal filters (10 kDa cutoff). Both bacterial lysates and
423 concentrated supernatants were subjected to SDS-PAGE followed by immunoblotting with
424 anti-Flag antibodies to detect the expression and secretion of 2×Flag-tagged proteins.

425 **Phylogenetic and sequence analysis**

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

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

437 *lupini*), MAV_1199 (*M. avium*), MMAN_54310 (*M. manitobense*), and MIP_01781 (*M.*
438 *intracellulare*). Genomic sequences were retrieved from the NCBI RefSeq database, filtered
439 for completeness, and aligned using Clustal Omega v1.2.4 with default parameters (gap
440 opening penalty = 10, gap extension = 0.2, and iterative refinement enabled). Sequence
441 conservation was visualized using ESPript 3.0, highlighting residues with ≥90% identity
442 across all homologs in red.

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

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

452 **Cell transfection and confocal microscopy**

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

458 temperature. The mixture was added dropwise to each chamber of 35-mm six-chamber
459 glass-bottom dishes containing cells at 40–60% confluency.
460 At 24–48 h post-transfection, the cells were washed with phosphate-buffered saline (PBS),
461 fixed in 4% paraformaldehyde for 15 min, and permeabilized with 0.1% Triton X-100 for 10
462 min. Nuclei were counterstained with 1 μg/mL DAPI (4′,6-diamidino-2-phenylindole) for 10
463 min. Confocal imaging was performed using an Olympus FV1000 Confocal Laser Scanning
464 Microscope equipped with a 100×/1.40 NA oil immersion objective. EGFP and DAPI signals
465 were detected with 488 nm and 405 nm lasers, respectively. Spectral unmixing was applied
466 using the Olympus FV10-ASW software (v4.2) to minimize signal cross-talk, and Z-stacks
467 were acquired in sequential mode by applying a step size of 0.5 μm across all samples to
468 ensure comparability.

469 **Immunoblot analysis**

470 For immunoblotting, the cells were lysed in radioimmunoprecipitation assay (RIPA) buffer
471 supplemented with 1.5% protease inhibitor cocktail (Boster). Protein concentrations were
472 quantified using the BCA assay (Thermo Fisher Scientific). Equal amounts of protein (20–40
473 μg) were resolved by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-
474 PAGE) and transferred onto polyvinylidene difluoride (PVDF) membranes (Millipore). The
475 membranes were blocked for 1 h at room temperature with 5% non-fat milk in TBST (25 mM
476 Tris, 150 mM NaCl, 2 mM KCl, and 0.2% Tween-20, pH 7.4) and incubated overnight at 4 °C
477 with primary antibodies. After three washes with TBST (10 min each), the membranes were
478 incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies at room
479 temperature for 1 h. Following three additional TBST washes, protein signals were detected

480 using an enhanced chemiluminescence (ECL) substrate and imaged on a ChemiDoc MP
481 system (Bio-Rad). β-actin, GAPDH, or histone H3 served as housekeeping proteins for
482 normalization. The antibody information is listed in Table S3.

483 Cell fractionation was performed using a Nuclear and Cytoplasmic Protein Extraction Kit
484 (Beyotime) according to the manufacturer's protocol. Briefly, after treatment, the cells were
485 washed and harvested in cold PBS. Cytoplasmic and nuclear fractions were separated
486 according to the manufacturer's instructions. Protein concentrations were quantified using a
487 BCA protein assay kit, and equal amounts of protein were subjected to immunoblotting, as
488 described above.

489 Co-IP assays were performed to analyze protein-protein interactions. HEK293T cells were
490 washed twice with ice-cold PBS and lysed in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM
491 NaCl, 1% NP-40) supplemented with 1.5% protease inhibitor cocktail (Boster) for 30 min on
492 ice. Crude lysates were centrifuged at 12,000 ×*g* for 15 min at 4 °C to remove the debris. For
493 immunoprecipitation, 1 mg of total protein was incubated with 5 μg of primary antibody
494 overnight at 4 °C with gentle agitation. Protein A/G magnetic beads (MedChemExpress)
495 were added to the antibody-protein complexes and incubated for 6 h at 4 °C. The beads
496 were washed three times with PBS-T buffer (PBS containing 0.5% Tween-20) to remove
497 nonspecific interactions. Proteins were eluted by boiling the beads in 1× SDS sample buffer
498 at 95 °C for 5 min. The eluted proteins were resolved by SDS-PAGE and analyzed by
499 immunoblotting.

500 **Y2H assay**

501 The Yeast Two-Hybrid System (Clontech) was used to assay protein-protein interactions
502 between MgdE (WT and mutants) and subunits of the COMPASS complex (ASH2L, WDR5,
503 RbBP5, and DPY30). Full-length cDNAs encoding WT MgdE and its mutants (see Table S5)
504 were cloned into the bait vector pGBKT7, whereas genes encoding the COMPASS subunits
505 were cloned into the prey vector pGADT7. The yeast strain Y2HGold was co-transformed
506 with bait and prey plasmids using the lithium acetate/polyethylene glycol method, and the
507 transformed colonies were initially selected on SD agar plates lacking leucine and
508 tryptophan (-Leu/-Trp) to confirm plasmid retention. For interaction screening, colonies were
509 replica-plated onto selective SD agar plates lacking leucine, tryptophan, adenine, and
510 histidine (-Leu/-Trp/-Ade/-His) and supplemented with 200 ng/μL aureobasidin A (AbA,
511 Sigma-Aldrich). The plates were incubated at 30 °C, and colony growth was monitored daily,
512 with final assessments performed after 5 d of incubation. Interaction strength was
513 qualitatively assessed by comparing colony density and growth rates with those of the
514 controls. The positive controls included the co-transformation of pGBKT7-p53 with pGADT7-
515 T, which produced robust growth on -Leu/-Trp/-Ade/-His plates. Negative controls included
516 co-transformation of pGBKT7-Lam (human lamin C, a non-interacting bait) with pGADT7-T,
517 as well as individual transformations with empty pGBKT7 or pGADT7 vectors to assess the
518 background growth. The yeast strains used in this study are listed in Table S4. All
519 experiments were performed in triplicate with biologically independent yeast cultures. The
520 results were consistent across replicates, confirming the reliability of the observed
521 interactions.

522 **qRT-PCR assay**

523 THP-1 cells were infected with the *M. bovis* BCG strain and harvested 12 h post-infection for
524 RNA isolation. Total RNA was extracted using 1 mL of TRIzol® reagent (Aidlab) according to
525 the manufacturer's instructions. To remove genomic DNA contamination, the RNA samples
526 were treated with DNase I (Thermo Fisher Scientific) for 30 min at 37 °C, followed by heat
527 inactivation at 65 °C for 10 min. RNA concentration and purity were assessed using a
528 NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). For cDNA synthesis, 1 μg of
529 total RNA was reverse transcribed using the HiScript® III cDNA Reverse Transcription Kit
530 (Vazyme) with random hexamers, according to the manufacturer's protocol. qRT-PCR was
531 performed using ChamQ SYBR® Green RT-PCR Master Mix (Vazyme). The relative mRNA
532 expression of different genes was calculated by comparing their cycle threshold (Ct) values
533 to that of a control gene *hprt* using the $2^{-\Delta\Delta Ct}$ method. Real-time qRT-PCR experiments were
534 performed in triplicate, with at least three independent biological replicates and two technical
535 replicates per condition. Primer sequences used for qRT-PCR are listed in Table S6.

536 **RNA-seq analysis**

537 THP-1 cells were seeded at a density of 1×10^6 cells/well in 6-well plates and differentiated
538 into macrophages by treatment with 100 nM PMA for 48 h. Following differentiation, the cells
539 were washed with PBS and infected with *M. bovis* BCG WT or *KomgdE* strains at a
540 multiplicity of infection (MOI) of 10:1. Following 12 h of infection, total RNA was extracted
541 using TRIzol® reagent (Thermo Fisher Scientific) according to the manufacturer's
542 instructions. High-quality reads were then mapped to the human reference genome
543 (GRCh38) using HISAT2 (v2.2.1). FeatureCounts (v2.0.3) was used to quantify transcript
544 abundance, and differential gene expression analysis was conducted using DESeq2

545 (v1.38.3) in R. Genes with an adjusted P-value <0.05 and an absolute log_2-foldchange ≥1
546 was considered significantly differentially expressed. RNA-seq data were visualized using
547 GraphPad Prism 8.0, and volcano plots, heatmaps, and other visualizations were generated
548 to summarize the results.

549 **CFU assay**

550 For bacterial preparation, mid-log phase cultures (OD_{600} = 0.8) of *M. bovis* BCG were
551 washed with PBS and resuspended in fresh medium. THP-1 cells were seeded at a density
552 of 5 × 10^5 cells/well in 24-well plates and subsequently infected at a multiplicity of infection
553 (MOI) of 1:10. Following 4 h of infection, extracellular bacteria were eliminated by treatment
554 with penicillin-streptomycin (100 μg/mL) for 2 h. The cells were then washed three times with
555 PBS and maintained in an antibiotic-free medium until harvest. For bacterial recovery, the
556 macrophages were lysed with 0.25% SDS for 10 min at room temperature, followed by serial
557 dilution of the lysates in PBS. Aliquots (100 μL) of each dilution were plated on Middlebrook
558 7H10 agar supplemented with 10% OADC and kanamycin (50 μg/mL). Plates were
559 incubated at 37 °C for 14 d, and CFUs were calculated by applying dilution factors to the
560 colony counts. *M. bovis* BCG strains used in this study are listed in Table S4.

561 **Animal experiments**

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

566 **Mice infection**

567 Mice were infected as described previously, with slight modifications. Mid-log phase *M. bovis*
568 BCG cultures were washed twice in PBS containing 0.05% Tween 80 and sonicated to
569 disrupt clumps. Female C57BL/6 mice (6–8 weeks old, 16–18g) were anesthetized using
570 isoflurane and intranasally administered 1×10^7 CFUs of *M. bovis* BCG in 40 μL of PBS. At
571 48 h post-infection, the bacterial load in the lungs of five mice was assessed to confirm the
572 inoculation dose. The lungs were homogenized in 1 mL of PBS using a tissue homogenizer,
573 and serial dilutions of the homogenates were plated onto Middlebrook 7H10 agar with 10%
574 OADC. The plates were then incubated at 37 °C with 5% CO_2 for 14 d prior to colony
575 counting.

576 **Histological analysis**

577 Lung tissues from *M. bovis* BCG-infected mice were fixed in 4% phosphate-buffered formalin
578 at room temperature for 24 h and embedded in paraffin wax. Paraffin-embedded tissue
579 samples were sectioned into 2–3 μm thick slices using a microtome. The sections were
580 deparaffinized, rehydrated with graded ethanol, and stained with hematoxylin and eosin
581 (H&E). Stained slides were visualized using an Olympus BX53 light microscope, and
582 CaseViewer version 2.0 (3DHISTECH) was used to view and annotate the scanned slides.

583 **Ethics approval**

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

586 **Statistical analyses**

587 Data are presented as mean ± standard deviation (SD). For comparisons between the two
588 groups, statistical significance was determined using two-tailed unpaired Student's *t*-tests
589 after confirming normality. Statistical significance was set at $P < 0.05$. Significance levels are
590 denoted in the figures as follows: $P > 0.05$ (ns, not significant), $P < 0.05$ (*), $P < 0.01$ (**),
591 and $P < 0.001$ (***). All statistical analyses were performed using GraphPad Prism (version
592 8.0). Data were collected from at least three independent biological replicates, each with two
593 or more technical replicates.

594 **ACKNOWLEDGMENTS**

595 This work was supported by the National Key Research and Development Program of China
596 (#2021YFD1800403), Fundamental Research Funds for the Central Universities
597 (#2662023DKQD001), Talent Start-up Funds of Huazhong Agricultural University
598 (#11042310008), National Natural Science Foundation of China (#32473123), China
599 Agriculture Research System of MOF and MARA (#CARS-37).

600 **AUTHOR CONTRIBUTIONS**

601 Conceptualization: L. C., A. G., and L. Z.; Methodology: L. C.; Validation: L. C., B. D., and Q.
602 J.; Formal analysis: L. C., Y. W.; Investigation: L. C., B. D., P. C., Q. J., Y. W., and L. L.;
603 Resources: A. G., L. Z., and Y. C.; Data curation: L. C., Q. J., Y. W., and L. L.; Writing -
604 original draft: L. C., L. Z., and A. G.; Writing - review & editing: L. C., L. Z., and A. G.;
605 Visualization: L. C., Q. J., Y. C., C. H., L. Z., and A. G.; Supervision: A. G., L. Z., Y. C., and C.
606 H.; Project administration: A. G., L. Z., Y. C., and C. H.; Funding acquisition: A. G., L. Z., and
607 Y. C.

608 **DISCLOSURE AND COMPETING INTEREST STATEMENT**

609 The authors declare no competing interests.

610 **REFERENCES**

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

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

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

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

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

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

627 Chai Q, Yu S, Zhong Y, Lu Z, Qiu C, Yu Y, Zhang X, Zhang Y, Lei Z, Qiang L (2022) **A**

628 **bacterial phospholipid phosphatase inhibits host pyroptosis by hijacking ubiquitin**
629 *Science* **378**:eabq0132.

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

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

634 Chauhan A, Quenum FZ, Abbas A, Bradley DS, Nechaev S, Singh BB, Sharma J, Mishra BB
635 (2015) **Epigenetic modulation of microglial inflammatory gene loci in helminth-**
636 **induced immune suppression: implications for immune regulation in**
637 **neurocysticercosis** *ASN Neuro* **7**:1759091415592126.

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

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

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

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

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

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

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

656 Hammond DE, Kumar JD, Raymond L, Simpson DM, Beynon RJ, Dockray GJ, Varro A
657 (2018) **Stable isotope dynamic labeling of secretomes (SIDLS) identifies authentic**
658 **secretory proteins released by cancer and stromal cells** *Mol Cell Proteomics* **17**:1837-
659 1849.

660 Hanford HE, Von Dwingelo J, Abu Kwaik Y (2021) **Bacterial nucleomodulins: a**
661 **coevolutionary adaptation to the eukaryotic command center** *PLoS Pathog*
662 **17**:e1009184.

663 Hsu PL, Li H, Lau HT, Leonen C, Dhall A, Ong SE, Chatterjee C, Zheng N (2018) **Crystal**
664 **structure of the COMPASS H3K4 methyltransferase catalytic module** *Cell* **174**:1106-
665 1116.e9.

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

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

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

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

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

679 Pisu D, Huang L, Grenier JK, Russell DG (2020) **Dual RNA-Seq of Mtb-Infected**
680 **Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions** *Cell*
681 *Rep* **30**:335-350.e4.

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

685 Prokop A, Gouin E, Villiers V, Nahori MA, Vincentelli R, Duval M, Cossart P, Dussurget O
686 (2017) **OrfX, a nucleomodulin required for *Listeria monocytogenes* virulence** *mBio*
687 **8**:e01550-17.

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

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

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

696 Rolando M, Sanulli S, Rusniok C, Gomez-Valero L, Bertholet C, Sahr T, Margueron R,
697 Buchrieser C (2013) ***Legionella pneumophila* effector RomA uniquely modifies host**
698 **chromatin to repress gene expression and promote intracellular bacterial replication**
699 *Cell Host Microbe* **13**:395-405.

700 Rolando M, Wah Chung IY, Xu C, Gomez-Valero L, England P, Cygler M, Buchrieser C
701 (2023) **The SET and ankyrin domains of the secreted *Legionella pneumophila* histone**
702 **methyltransferase work together to modify host chromatin** *mBio* **14**:e0165523.

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

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

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

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

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

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

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

723 Xue H, Yao T, Cao M, Zhu G, Li Y, Yuan G, Chen Y, Lei M, Huang J (2019) **Structural basis**
724 **of nucleosome recognition and modification by MLL methyltransferases** *Nature*
725 **573**:445-449.

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

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

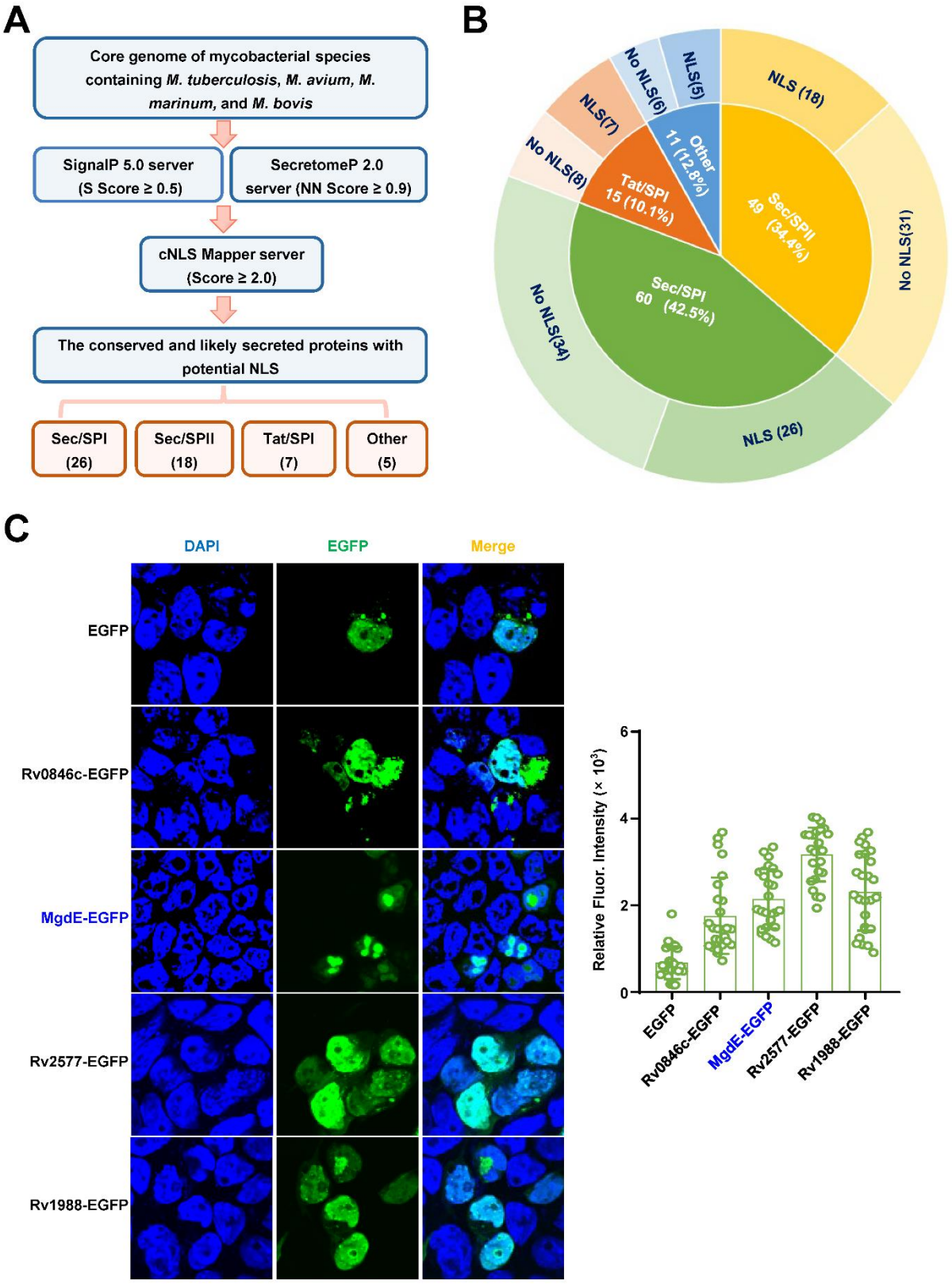
732 Yu H, Lesch BJ (2024) **Functional roles of H3K4 methylation in transcriptional**
733 **regulation** *Mol Cell Biol* **44**:505-515.

734 Yu L, Fang F, Dai X, Xu H, Qi X, Fang M, Xu Y (2017) **MKL1 defines the H3K4Me3**
735 **landscape for NF-κB dependent inflammatory response** *Sci Rep* **7**:191.

736 Zhao J, Liu Y, Lin F, Wang W, Yang S, Ge Y, Chen PR (2019) **Bioorthogonal engineering**
737 **of bacterial effectors for spatial-temporal modulation of cell signaling** *ACS Cent Sci*
738 **5**:145-152.

739 **Figures**

740 **Figure legends**



741

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

743 **functional screening.**

744 (**A**) Schematic representation of the bioinformatic pipeline for identifying conserved
745 nucleomodulins in Mycobacterium species. Genomic sequences from *M. tuberculosis*
746 *H37Rv*, *M. tuberculosis* H37Ra, *M. avium*, *M. marinum*, and *M. bovis* BCG were analyzed.
747 Signal peptides were predicted using SignalP 5.0 (D-score ≥ 0.5), non-classical secretion
748 signals were identified using SecretomeP 2.0 (NN score ≥ 0.9), and nuclear localization
749 signals (NLSs) were predicted using cNLS Mapper (score ≥ 2.0). (**B**) Classification of
750 conserved cellular nucleomodulins. Secreted proteins were categorized based on the
751 presence or absence of predicted NLS motifs. (**C**) Subcellular localization of EGFP-tagged
752 candidate nucleomodulins. (left) Confocal images of Rv0846c, MgdE, Rv2577, and Rv1988
753 fused to EGFP (green); nuclei are stained with DAPI (blue). (right) Quantification of nuclear
754 EGFP intensity. Rv1988 and pEGFP were used as positive and negative controls,
755 respectively. Scale bar, 10 μm. Data are presented as mean ± SD (n = 25 cells/group).

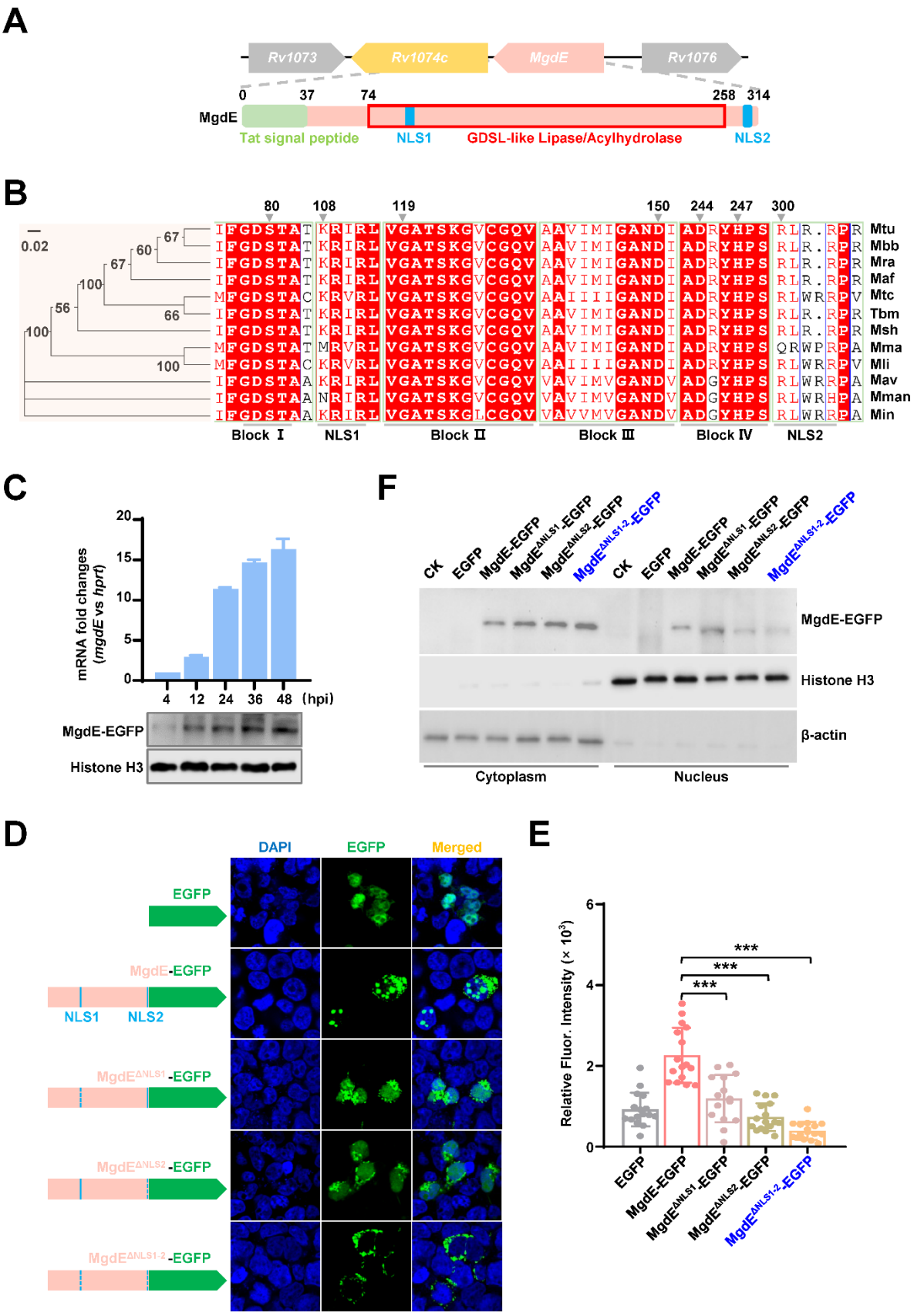
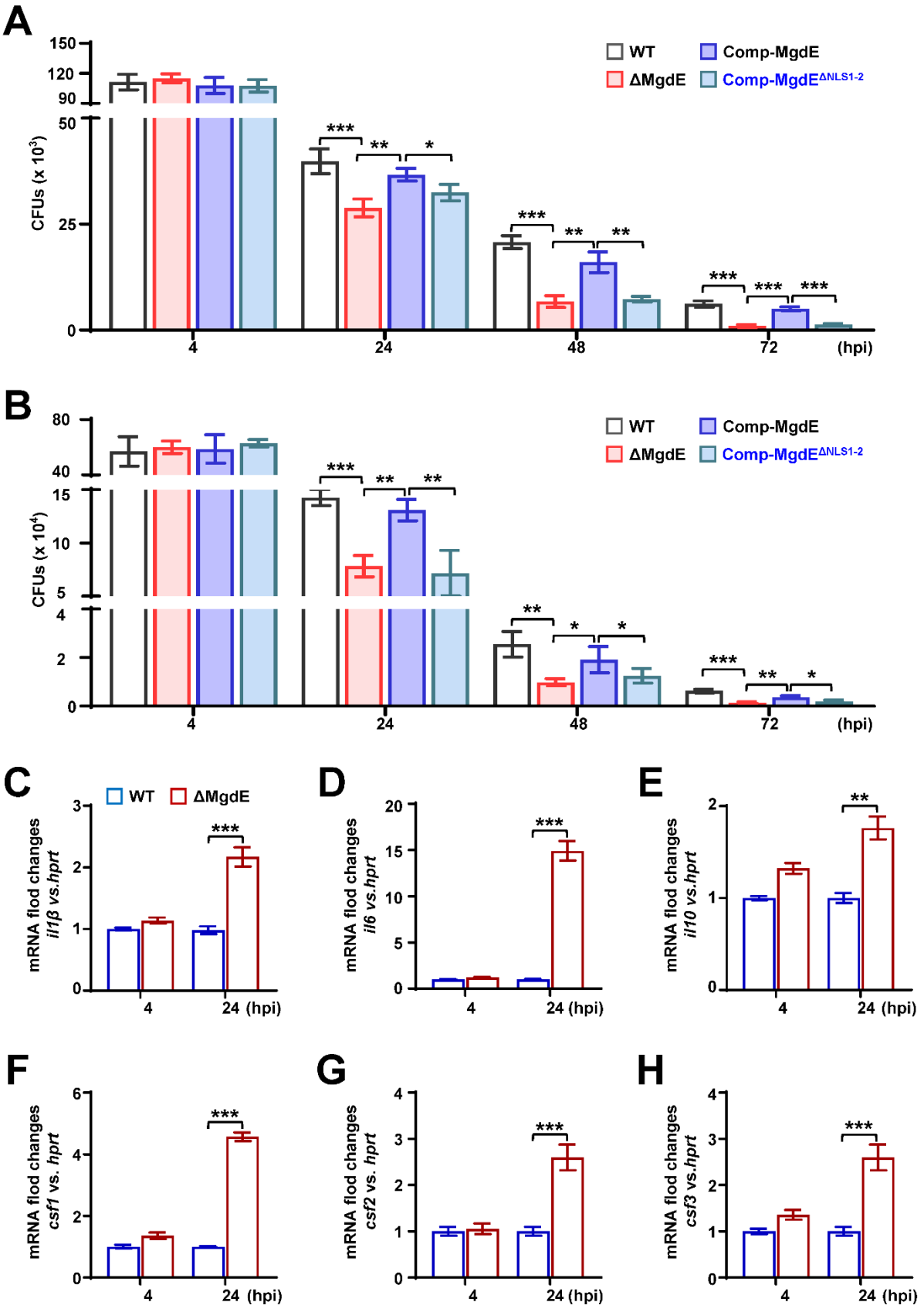


Figure 2. Hypothetical protein MgdE enters the host nucleus via dual NLS.

758 (**A**) Domain architecture of MgdE. Schematic representation of MgdE with the annotated
759 functional domains, including a Tat signal peptide (1–37 aa, twin-arginine translocation
760 motif), a GDSL-like lipase/acylhydrolase catalytic domain (74–258 aa), and two nuclear
761 localization signals (NLS1: 108–111 aa; NLS2: 300–305 aa). (**B**) Phylogenetic and structural
762 conservation of MgdE. (left) Neighbor-joining phylogenetic tree of MgdE homologs across
763 Mycobacterium species (1,000 bootstrap replicates; values ≥50% shown). (right) Clustal
764 Omega sequence alignment highlighting conserved residues (≥90% identity, red). Species
765 abbreviations: *M. tuberculosis* (Mtu), *M. bovis* BCG (Mbb), *M. tuberculosis* H37Ra (Mra), *M.*
766 *africanum* (Maf), *M. tuberculosis* CDC1551 (Mtc), *M. tuberculosis* GM041182 (Tbm), *M.*
767 *shinjukuense* (Msh), *M. marinum* (Mma), *M. lupini* (Mli), *M. avium* (Mav), *M. manitobense*
768 (Mman), *M. intracellulare* (Min). (**C**) Expression and nuclear localization of MgdE. (upper)
769 qRT-PCR analysis of mgdE mRNA in HEK293T cells over 48 h post-transfection. (lower)
770 Western blot analysis of nuclear fractions showing time-dependent nuclear accumulation of
771 MgdE-EGFP. Histone H3 and β-actin served as nuclear and cytoplasmic markers,
772 respectively. **(D)** Subcellular localization of EGFP-tagged wild-type and NLS-deleted MgdE.
773 (left) Schematic representation of EGFP-tagged constructs, including wild-type MgdE and its
774 mutants ($MgdE^{\Delta NLS1}$, $MgdE^{\Delta NLS2}$, and $MgdE^{\Delta NLS1\text{-}2}$). (right) Confocal microscopy images of
775 HEK293T cells transfected with the indicated constructs for 36 h. EGFP fluorescence
776 (green) and nuclear staining with DAPI (blue) were visualized using an FLUOVIEW software
777 (v5.0). Scale bar, 10 μm. Images were acquired with a ×100 oil immersion objective (NA =
778 1.4). (**E**) Nuclear EGFP intensity of wild-type and mutant constructs in **(D)**. Data are shown
779 as mean ± SD (n = 15 cells). (**F**) Western blot analysis of nuclear and cytoplasmic fractions

780 from HEK293T cells transfected with wild-type and mutant MgdE-EGFP confirmed their
781 nuclear localization. MgdE-EGFP was detected using an anti-GFP antibody, and histone H3
782 and β-actin served as nuclear and cytoplasmic markers, respectively. All experiments were
783 performed in three biologically independent replicates. Statistical significance was assessed
784 by two-tailed unpaired Student's *t*-tests, with $^{*}P < 0.05$ considered statistically significant.

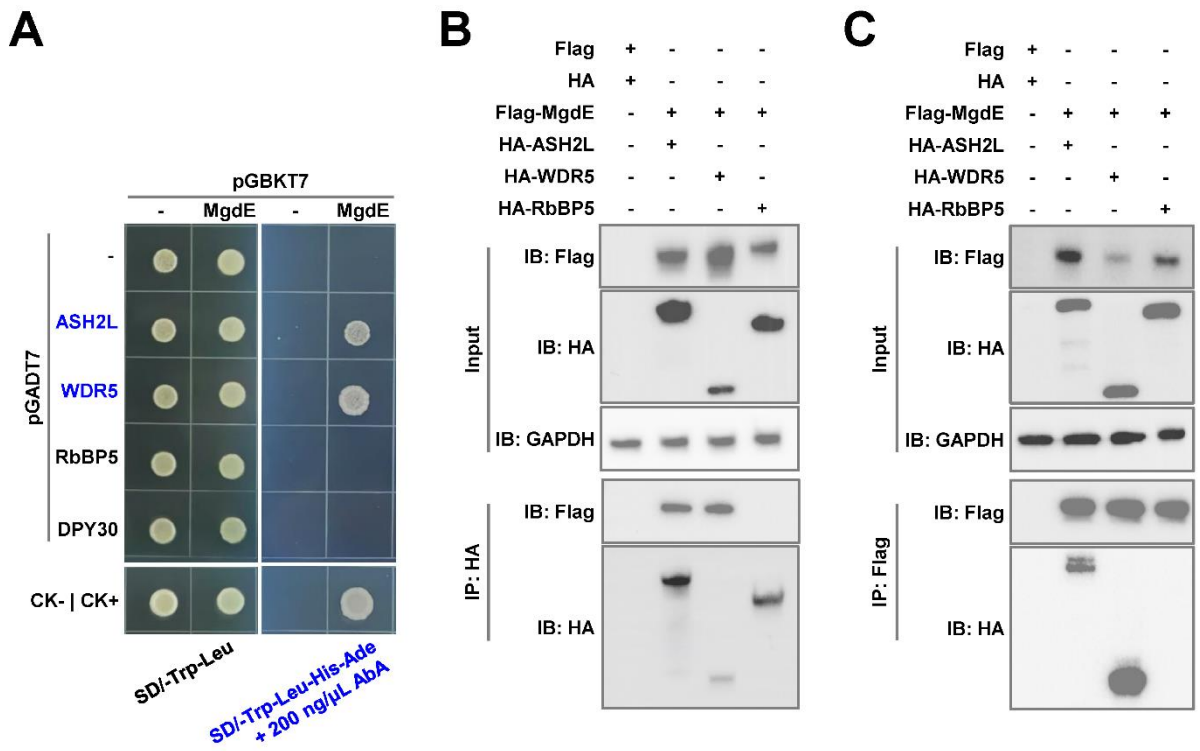


785

786 **Figure 3. Nuclear localization of MgdE facilitates mycobacterial intracellular survival**

787 **in macrophages.**

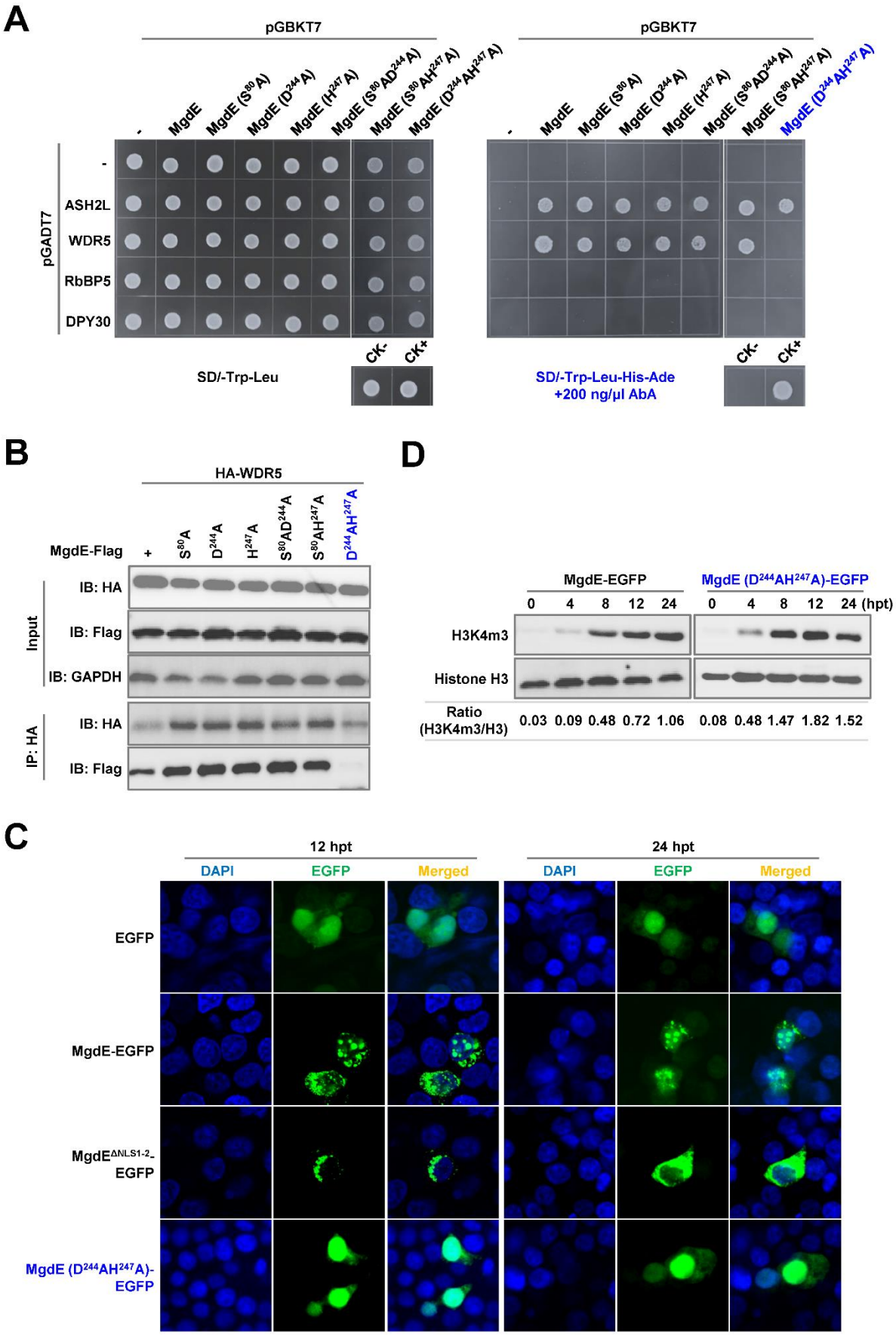
788 (**A**-**B**) Intracellular survival of *M. bovis* BCG strains. THP-1 human macrophages (A) and
789 RAW264.7 murine macrophages (B) were infected (MOI = 10) with wild-type BCG (WT),
790 MgdE deletion mutant (ΔMgdE), complemented strain (Comp-MgdE), or NLS-deficient
791 complement (Comp-MgdE$^{\Delta NLS1\text{-}2}$). Bacterial survival was assessed by CFU enumeration at 2,
792 24, 48, and 72 h post-infection. (**C-H**) Cytokine expression in the infected THP-1 cells. qRT-
793 PCR analysis of cytokine mRNA levels in THP-1 cells infected with WT or ΔMgdE strains for
794 4–24 h. Inflammatory cytokine expression in infected THP-1 cells. qRT-PCR analysis of
795 intracellular cytokine mRNA levels in THP-1 cells infected with WT or ΔMgdE strains for 4–
796 12 h. Target genes include *il1b* (C), *il6* (D), *il10* (E), *csf1* (F), *csf2* (G), and *csf3* (H). Data are
797 presented as mean ±SD of three biologically independent experiments, analyzed using two-
798 tailed unpaired Student's *t*-tests (**P* < 0.05, ***P* < 0.01, and ****P* < 0.001).



799

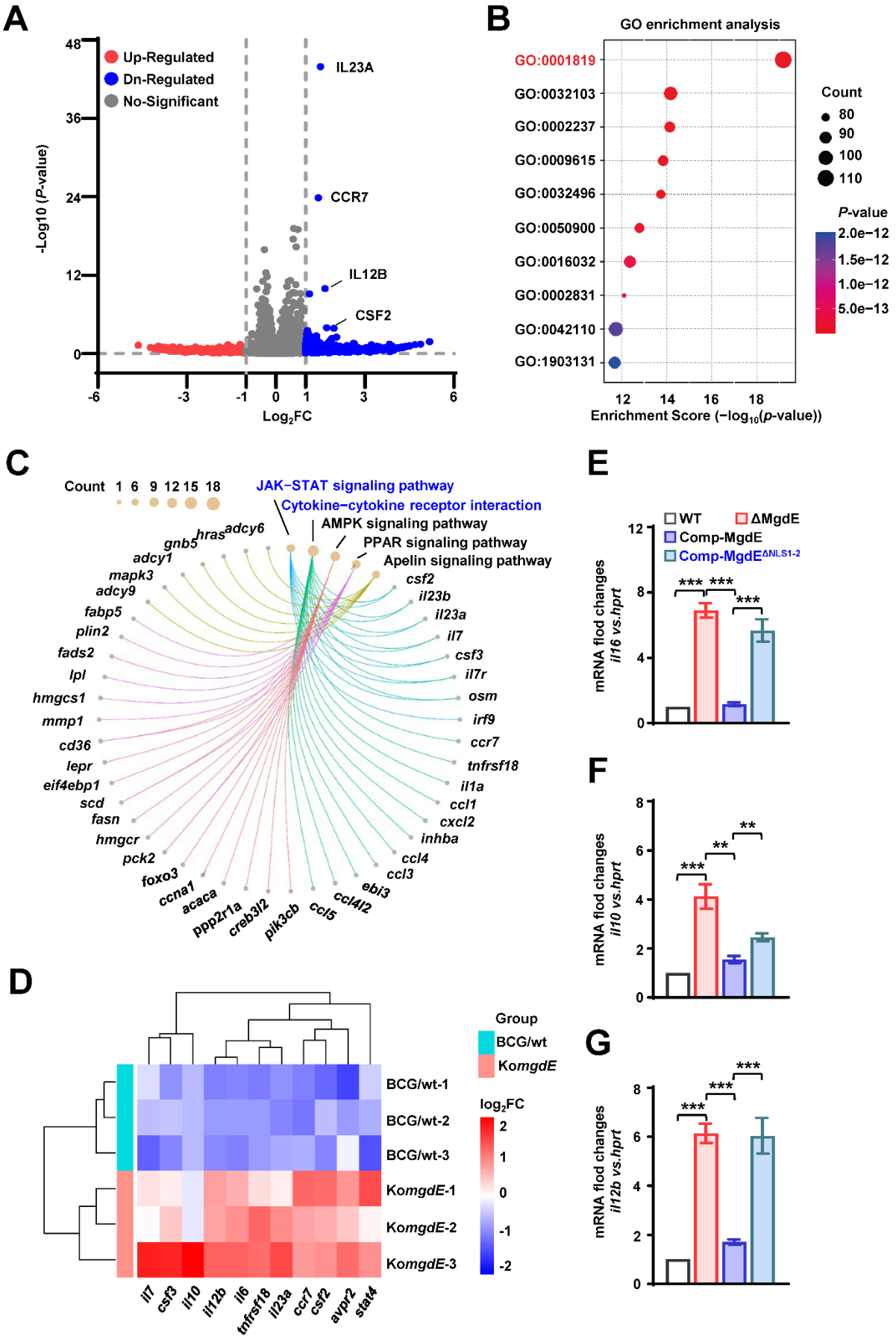
800 **Figure 4. MgdE directly interacts with ASH2L and WDR5, core components of the host**
801 **COMPASS complex.**

802 (**A**) Yeast cells were co-transformed with bait (pGBKT7) and prey (pGADT7) plasmids
803 expressing wild-type MgdE and human COMPASS components (ASH2L, WDR5, RbBP5,
804 and DPY30). Growth was monitored on non-selective (-Leu/-Trp, left) and selective (-Leu/-
805 Trp/-Ade/-His + 200 ng/μL aureobasidin A, right) media. Controls: CK− (pGBKT7-*lam* +
806 pGADT7-T, negative); CK+ (pGBKT7-*p53* + pGADT7-T, positive). (**B**-**C**) Cells were co-
807 transfected with Flag-MgdE and HA-tagged ASH2L, WDR5, or RbBP5 (1:1 molar ratio). At
808 36 h post-transfection, the lysates were immunoprecipitated using (**B**) anti-HA or (**C**) anti-
809 Flag antibodies, followed by immunoblotting with anti-HA, anti-Flag, and anti-GAPDH
810 (loading control). The input lanes represent 5% of the total lysate. Data are representative of
811 three independent experiments.


A
pGBKT7
pGBKT7
-
MgdE
MgdE (S80 A)
MgdE (D244 A)
MgdE (H247 A)
MgdE (S80AD244 A)
MgdE (S80AH247 A)
MgdE (D244AH247 A)
pGADT7
-
ASH2L
WDR5
RbBP5
DPY30
CK-
CK+
SD/-Trp-Leu
SD/-Trp-Leu-His-Ade
+200 ng/μl AbA
B
HA-WDR5
MgdE-Flag
+
S80A
D244A
H247A
S80AD244A
S80AH247A
D244AH247A
Input
IB: HA
IB: Flag
IB: GAPDH
IP: HA
IB: HA
IB: Flag
D
MgdE-EGFP
MgdE (D244AH247A)-EGFP
0 4 8 12 24
0 4 8 12 24 (hpt)
H3K4m3
Histone H3
Ratio
(H3K4m3/H3)
0.03 0.09 0.48 0.72 1.06
0.08 0.48 1.47 1.82 1.52
C
12 hpt
24 hpt
DAPI
EGFP
Merged
EGFP
MgdE-EGFP
MgdEΔNLS1-2-EGFP
MgdE (D244AH247A)-EGFP

813 **Figure 5. The conserved residues D^{224} and H^{247} mediate the binding ability of MgdE to**
814 **WDR5.**

815 (**A**) Y2H assay identifying interactions between MgdE mutants and COMPASS complex
816 subunits. Yeast cells were co-transformed with bait (pGBKT7) and prey (pGADT7) plasmids
817 expressing wild-type or mutant MgdE and human COMPASS subunits (ASH2L, WDR5,
818 RbBP5, and DPY30). Growth was assessed on non-selective (-Leu/-Trp, left) and selective
819 (-Leu/-Trp/-Ade/-His + 200 ng/μL aureobasidin A, right) media. Controls: CK− (pGBKT7-*lam*
820 + pGADT7-T, negative) and CK+ (pGBKT7-*p53* + pGADT7-T, positive). (**B**) Co-IP analysis of
821 MgdE mutants with WDR5. HEK293T cells were co-transfected with Flag-tagged MgdE
822 mutants and HA-tagged WDR5 (1:1 molar ratio). Complexes were immunoprecipitated using
823 anti-HA antibody and protein A/G beads, followed by immunoblotting with anti-Flag and anti-
824 HA antibodies. (**C**) Nuclear distribution of wild-type and mutants MgdE. Confocal microscopy
825 of HEK293T cells expressing wild-type or $D^{224}A/H^{247}A$ MgdE-EGFP at 12 and 24 h post-
826 transfection (hpt). Nuclear foci were visualized by EGFP (green) and DAPI (blue) staining.
827 Scale bar, 10 μm. Images were acquired with a ×100 oil immersion objective (NA = 1.4). (**D**)
828 Immunoblot analysis of H3K4me3 levels. HEK293T cells expressing wild-type or $D^{224}A/H^{247}A$
829 mutant MgdE were analyzed for changes in H3K4me3 levels over 0–24 h post-transfection.
830 Histone H3 was used as a loading control. The data represent two independent biological
831 replicates.


A
Up-Regulated
Dn-Regulated
No-Significant
IL23A
CCR7
IL12B
CSF2
-Log10 (P-value)
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
BCG/wt
KomgdE
log2FC
BCG/wt-1
BCG/wt-2
BCG/wt-3
KomgdE-1
KomgdE-2
KomgdE-3
E
WT
ΔMgdE
Comp-MgdE
Comp-MgdEΔNLS1-2
mRNA flod changes il16 vs.hprt
F
mRNA flod changes il10 vs.hprt
G
mRNA flod changes il12b vs.hprt

833 **Figure 6. MgdE suppresses COMPASS complex-mediated inflammatory responses by**

834 **inhibiting H3K4 methylation.**

835 (**A**) Volcano plot of DEGs. DEGs between MgdE-deleted strain (Ko*mgdE*) and wild-type

836 BCG (WT) were visualized in a volcano plot. Genes with $|\log_2\text{-fold change}| \geq 1$ and $P < 0.05$

837 were considered significant. The *x*-axis represents $\log_2$fold change, and the *y*-axis shows -

838 $\log_{10}$(*P*-value). (**B**) GO enrichment analysis of DEGs. GO analysis revealed the significant

839 enrichment of immune and inflammatory processes in Ko*mgdE*-infected macrophages

840 compared to that in WT-infected cells, including positive regulation of response to external

841 stimulus (GO:0032103), response to molecule of bacterial origin (GO:0002237), response to

842 virus (GO:0009615), response to lipopolysaccharide (GO:0032496), leukocyte migration

843 (GO:0050900), viral process (GO:0016032), regulation of response to biotic stimulus

844 (GO:0002831), T cell activation (GO:0042110), and mononuclear cell differentiation

845 (GO:1903131). (**C**) KEGG pathway enrichment analysis of DEGs. Chord diagram of KEGG

846 pathway enrichment analysis showing signaling pathways that are significantly enriched in

847 Ko*mgdE*, including cytokine-cytokine receptor interaction and JAK-STAT signaling as the top

848 enriched pathways. (**D**) Heatmap of the inflammatory gene expression. Heatmap depicting

849 $\log_2$-foldchange levels of inflammatory genes involved in the JAK-STAT and cytokine

850 signaling pathways. Upregulated and downregulated genes in Ko*mgdE* are shown in green

851 and red, respectively. The data were *Z*-score normalized. (**E-G**) qRT-PCR analysis of

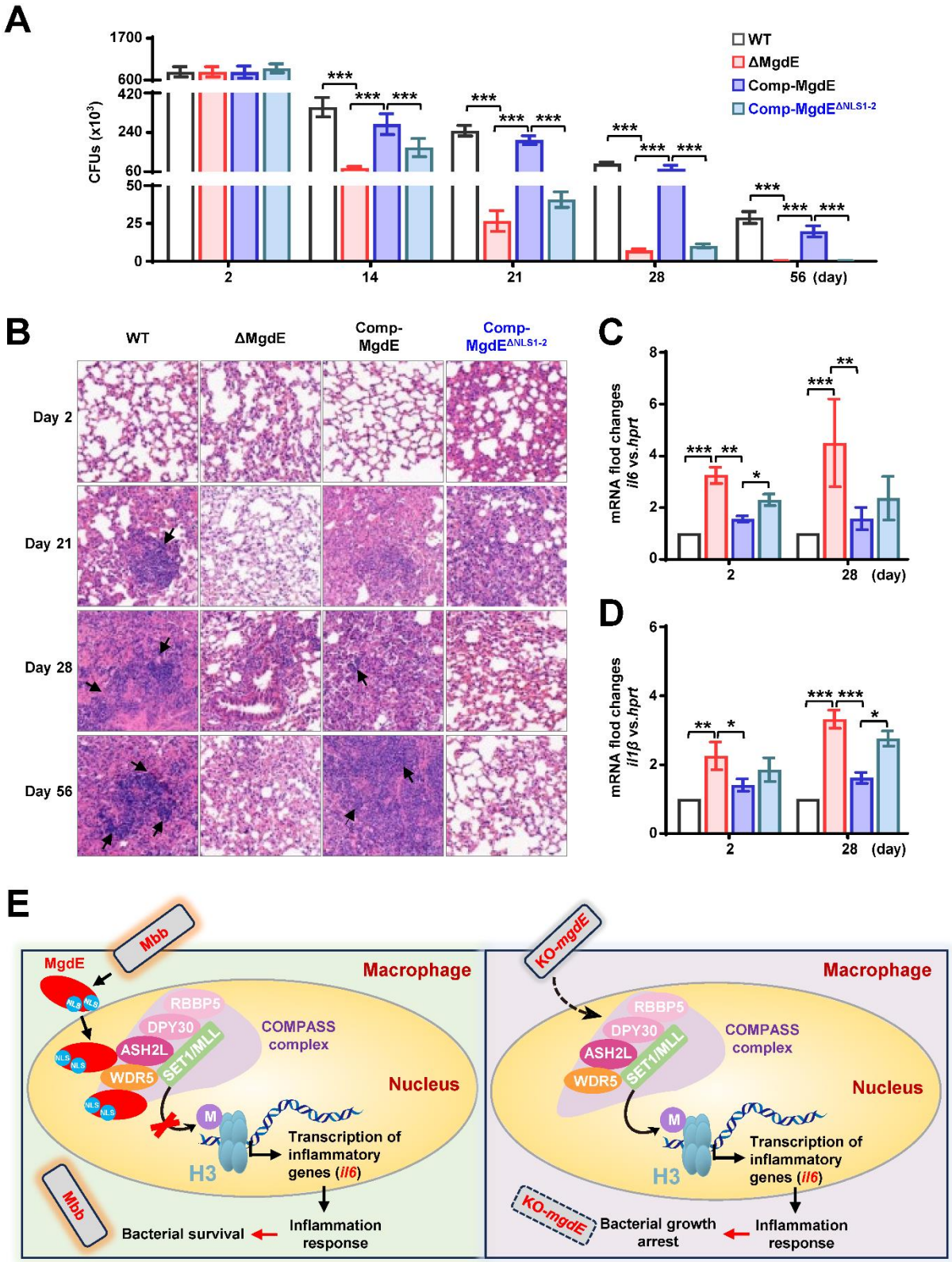
852 cytokine mRNA levels in THP-1 cells infected with WT, ΔMgdE, MgdE-complemented

853 (Comp-MgdE), and NLS-deleted complement (Comp-MgdE$^{\Delta NLS1\text{-}2}$) at 24 h post-infection.

854 Cytokines analyzed include *il6* (**E**), *il10* (**F**), and *il12b* (**G**). Data are presented as mean ± SD

855 of three biologically independent experiments, analyzed using two-tailed unpaired Student's

856 *t*-tests (**P* < 0.05, ***P* < 0.01, and ****P* < 0.001).

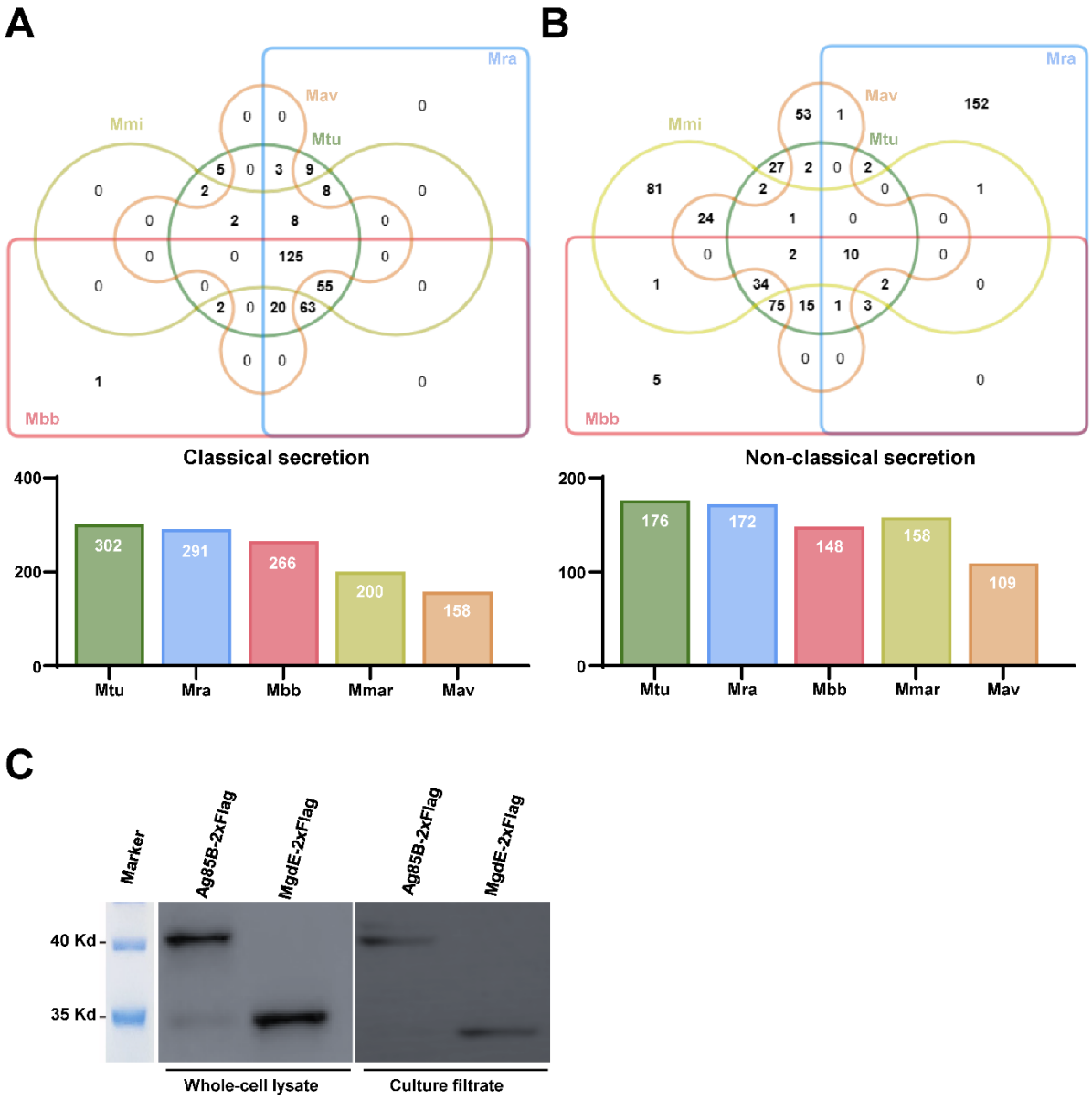


857

858 **Figure 7. Nuclear localization of MgdE is essential for mycobacterial survival in mice.**

859 (**A**) Bacterial burden in the lungs of the infected mice. C57BL/6 mice (n = 6/group)
860 maintained under SPF conditions were intratracheally infected with 1.0×10^7 colony-forming
861 units (CFU) of *M. bovis* BCG strains, including wild-type (WT), MgdE-deleted (ΔMgdE),
862 MgdE-complemented (Comp-MgdE), and NLS-deleted complement (Comp-MgdE$^{\Delta NLS1\text{-}2}$).
863 Lung bacterial loads were quantified using CFU assays at 0, 14, 28, and 56 d post-
864 infection. (**B**) H&E-stained lung sections from infected mice (as in **A**) revealed
865 granulomatous inflammation. Scale bars: 200 μm. (**C-D**) Pro-inflammatory cytokine
866 expression in mice spleen. qRT-PCR analysis of cytokine mRNA levels of *il6* (**C**) and *il1β* (**D**)
867 in spleen tissues from infected mice (n = 6/group) at 2 and 28 d post-infection. (**E**)
868 Mechanistic model showing how mycobacterial nucleomodulin MgdE hijacks the COMPASS
869 complex to suppress H3K4me3 and promote immune evasion. Upon *M. bovis* BCG infection,
870 the nucleomodulin MgdE is delivered into the host nucleus via its nuclear localization signal
871 (NLS) and directly binds to the COMPASS complex subunits, ASH2L or WDR5. This
872 interaction disrupts H3K4 trimethylation (H3K4me3) deposition, leading to the epigenetic
873 suppression of pro-inflammatory cytokine transcription (e.g., *il6*), thereby facilitating the
874 intracellular survival of the pathogen. Data are presented as mean ± SD from six biologically
875 independent experiments. Statistical significance was determined by a two-tailed unpaired
876 Student's *t*-test (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$).

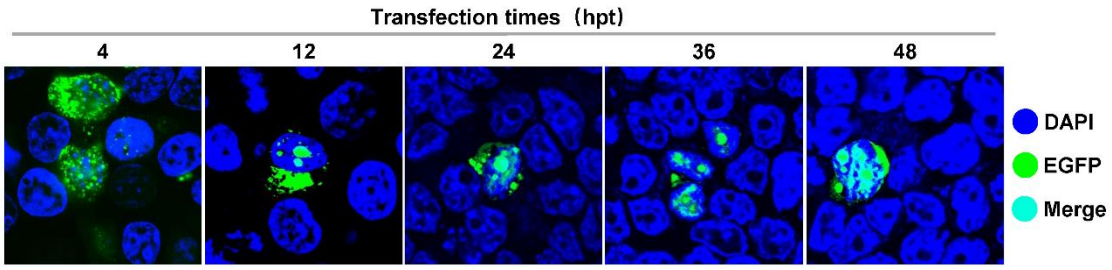
877 **Supplementary figures**



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

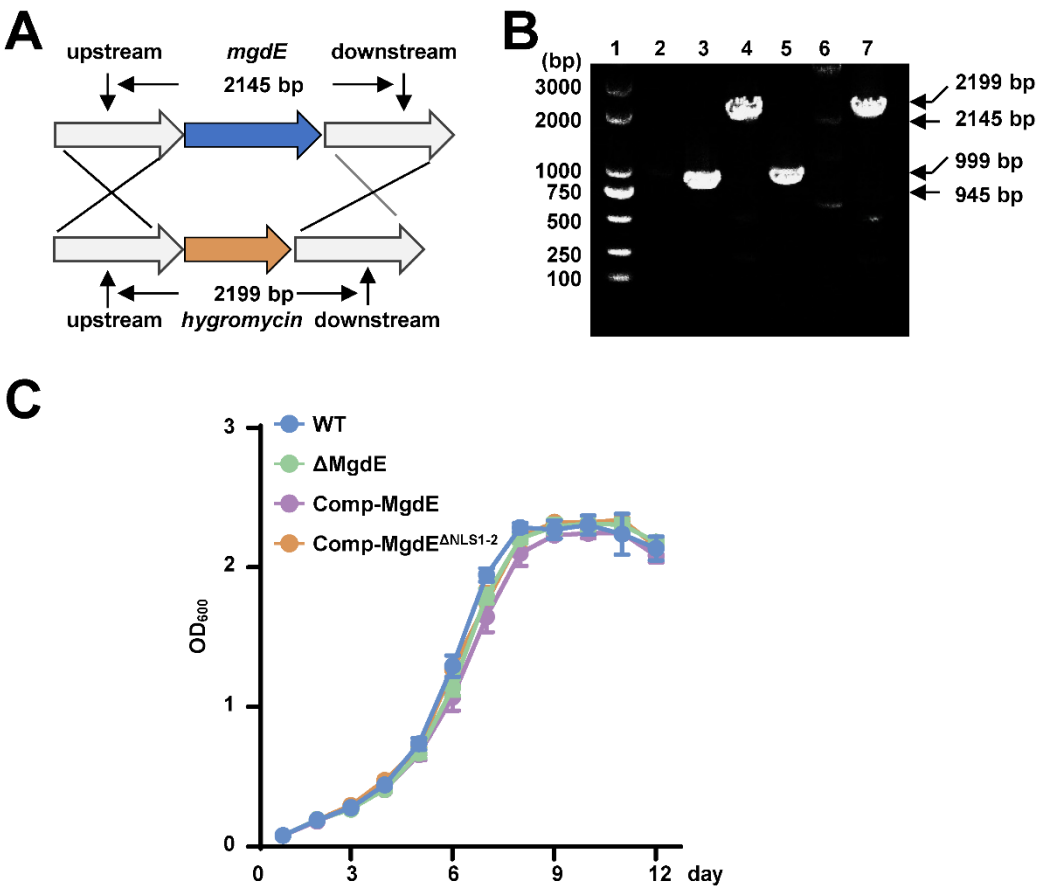
(**A**) Venn diagram showing the distribution of predicted classical secreted proteins in four Mycobacterium species: *M. tuberculosis* H37Rv (Mtu) and *M. tuberculosis* H37Ra (Mra), *M. bovis* BCG (Mbb), *M. marinum* (Mmar), and *M. avium* (Mav). Proteins were predicted using SignalP 5.0, with a signal 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. **(C)** Detection of secreted MgdE in culture supernatants. C-terminally Flag-tagged Ag85B (positive control) and MgdE were expressed in *M. bovis* BCG strains. Immunoblot analysis of bacterial lysates and culture supernatants showing the expression and secretion of Flag-tagged Ag85B and MgdE.



Supplementary Figure 2. Subnuclear localization of MgdE-EGFP.

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.



898

899 **Supplementary Figure 3. Deletion of the nuclear localization signal of MgdE does not**

900 **affect the growth of *M. bovis* BCG strains.**

901 (**A-B**) Construction and validation of the MgdE-deleted strain of *M. bovis* BCG. (**A**)

902 Schematic diagram of the homologous recombination strategy used to delete *mgdE* from the

903 *M. bovis* BCG genome. (**B**) Wild-type and mutant strains were used as templates to amplify

904 the *mgdE* gene (600 bp upstream–600 bp downstream) by PCR. Lanes 1 and 3: wild-type

905 genomic DNA; lanes 2 and 4: ΔmgdE genomic DNA. (**C**) Growth curve analysis of *M. bovis*

906 BCG strains. Growth of BCG strains, including wild-type BCG (WT), MgdE-deleted (ΔMgdE),

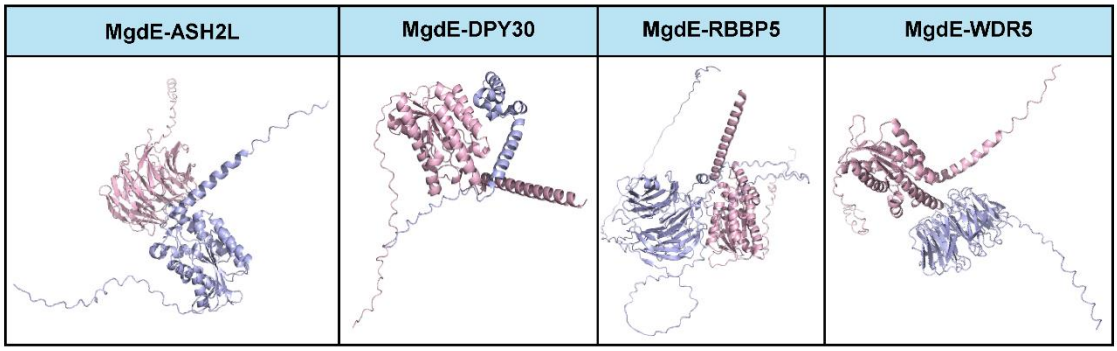
907 MgdE-complemented (Comp-MgdE), and NLS-deleted complemented (Comp-MgdE$^{\Delta NLS1\text{-}2}$)

908 strains, was measured in 7H9 medium.

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

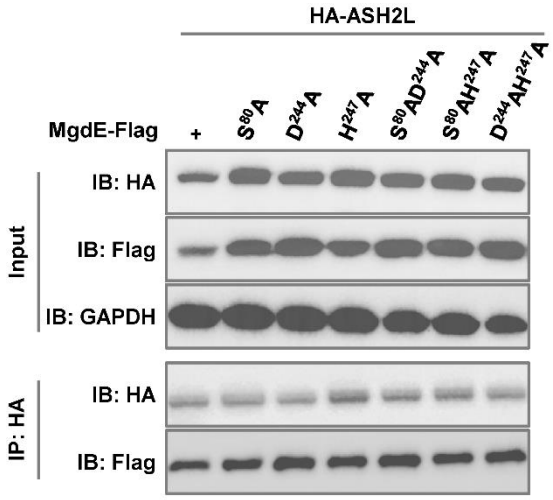
B



Supplementary Figure 4. MgdE interacts with COMPASS complex subunits.

(**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

920 DPY30) shown in gray.



921

922 **Supplementary Figure 5. Conserved residue mutations do not affect the binding of**

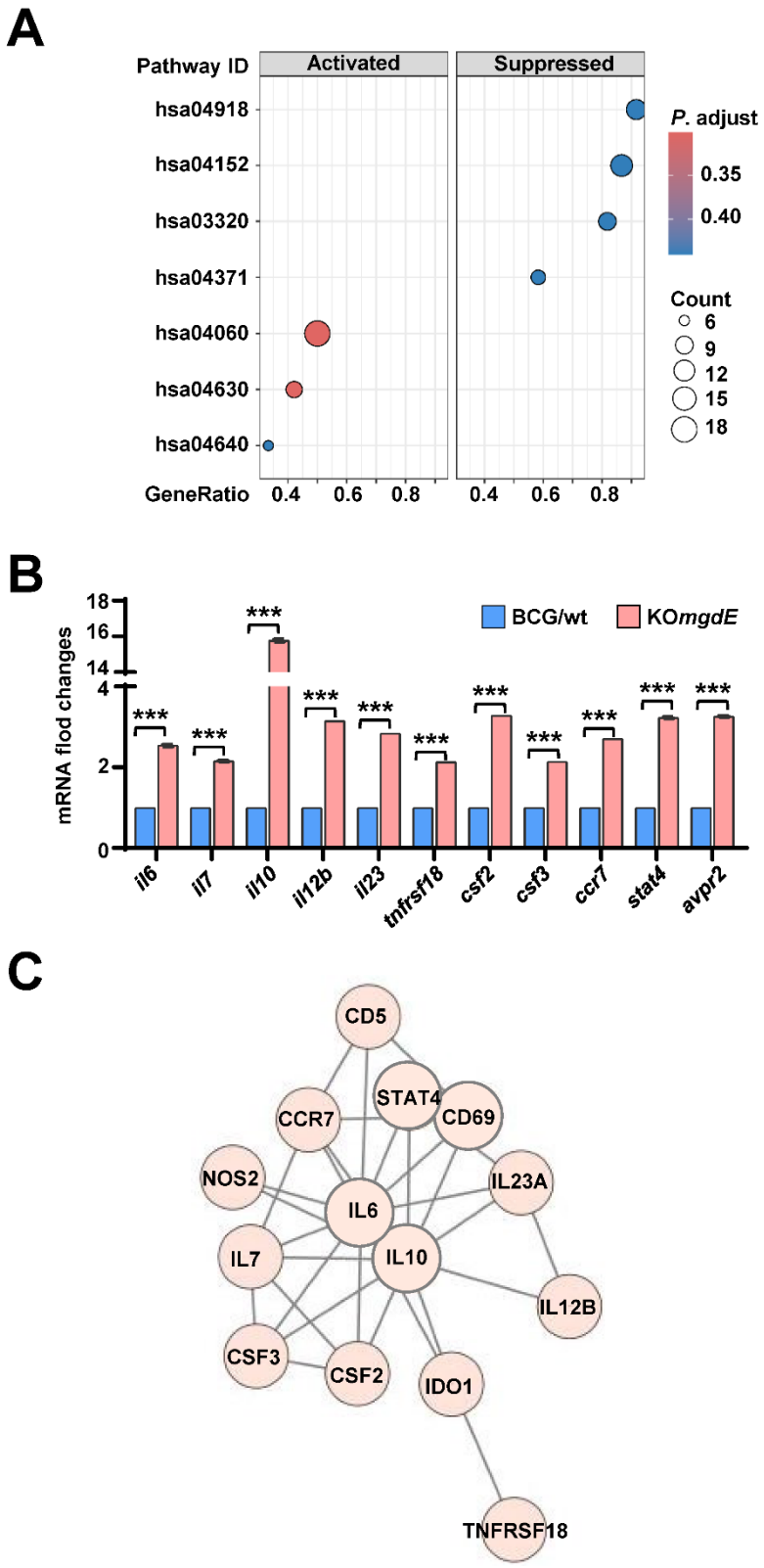
923 **MgdE to ASH2L.**

924 Co-IP analysis of MgdE mutants with ASH2L. HEK293T cells were co-transfected with Flag-

925 tagged MgdE mutants and HA-tagged ASH2L (1:1 molar ratio). Co-IP was performed using

926 anti-HA antibody-bound protein A/G beads, and immunoprecipitated complexes were

927 analyzed by immunoblotting with anti-Flag and anti-HA antibodies.

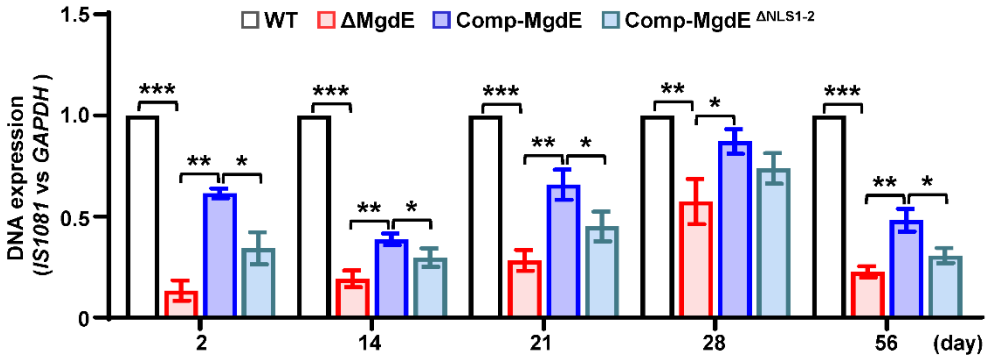


928

929 **Supplementary Figure 6. MgdE suppresses cellular inflammatory responses during *M.***

930 ***bovis* BCG infection.**

931 (**A**) KEGG pathway enrichment analysis. KEGG pathway analysis identified significantly
932 enriched pathways in THP-1 cells infected with the Ko*mgdE* strain compared to those
933 infected with wild-type BCG (WT). DEGs were predominantly associated with immune and
934 signaling pathways, including thyroid hormone synthesis (hsa04918), AMPK signaling
935 pathway (hsa04152), PPAR signaling pathway (hsa03320), apelin signaling pathway
936 (hsa04371), cytokine-cytokine receptor interaction (hsa04060), JAK-STAT signaling pathway
937 (hsa04630), and hematopoietic cell lineage (hsa04640). (**B**) Elevated inflammatory gene
938 expression in Ko*mgdE*-infected cells. Key upregulated inflammatory genes in Ko*mgdE*- *vs*.
939 WT-infected cells are highlighted. (**C**) Functional enrichment analysis of upregulated genes.
940 Analysis using STRING and Cytoscape revealed regulation of inflammatory responses as a
941 top enriched biological process ($P < 0.05$).



942

943 **Supplementary Figure 7. MgdE modulates inflammatory responses and bacterial**
944 **colonization in the spleens of infected mice.**

945 Bacterial colonization in the spleen. CFU assays were performed on splenic homogenates
946 from mice infected as described in **Figure 7A**. Bacterial loads (CFU/g tissue) were quantified
947 at 2, 14, 21, 28, and 56 d post-infection. Data are presented as mean ± SD (n = 3).
948 Statistical significance was determined using a two-tailed unpaired Student's *t*-test ($P < 0.05$,

949 **$P < 0.01$, and ***$P < 0.001$).

950