A *Mycobacterium tuberculosis* secreted virulence factor disrupts host snRNP biogenesis

Komal Chauhan[1], Nishat Passricha[1], Yogita Kapoor[2], Dipanwita Datta[1], Naresh Arora[1], Mrinmoy Das[1], Kavya Rao[1], Lakshyaveer Singh[1], Ravi Dutt[1], Raman Deep Sharma[1], Binayak Sarkar[3], Basanti Malakkar[3], Haroon Kalam[1], Prince Saini[4], Lalita Mehra[5], Prasenjit Das[5], Vineet Ahuja[6], Amit Singhal[7,8], Vinay Nandicoori[2, 3] and Dhiraj Kumar[1]*

[1]Cellular Immunology Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110067, India
[2]Centre for Cellular and Molecular Biology, Hyderabad, India
[3]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India
[4]Department of Biological Sciences, Indian Institute of Science Education and Research Mohali, Punjab 140306, India
[5]Department of Pathology, All India Institute of Medical Sciences, New Delhi-110012, India
[6]Department of Gastroenterology, All India Institute of Medical Sciences, New Delhi-110012, India
[7]A*STAR Infectious Diseases Labs (A*STAR ID labs), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore.
[8]Singapore Immunology Network (SIgN), A*STAR, Singapore, Singapore.

Running title: Regulation of host RNA splicing by Mtb

*Correspondence- dhiraj@icgeb.res.in

Abstract

Mycobacterium tuberculosis (Mtb) infection of human macrophages triggers extensive dysregulation of host alternative splicing (AS). Here, we present an unexpected mechanism exploited by the Mtb to achieve the same. We identify several bacterial secretory proteins that interact with the host splicing factors (SFs) and alter select RNA splicing events *in vitro*, *in cell* and during *ex vivo* infections. We show bacterial proteins disrupt AS by impairing Small nuclear ribonucleoprotein (snRNP) biogenesis, primarily by interacting with and sequestering U5snRNA and SNRPF, key spliceosome constituents. These interactions are driven, in part, by one of the identified Mtb proteins, Hsr1, whose access to the host cytosol is ESX1-dependent. The Mtb*Δhsr1* is unable to interact and sequester U5snRNA or SNRPF and fails to alter host AS. Consequently, Mtb*Δhsr1* shows compromised survival and pathogenesis in macrophages and mice. Lungs of infected mice show distinct SNRPF staining, which is hsr1 dependent. Similar distinctive SNRPF staining is also noted in human tuberculous granulomas. Moreover, monocytes from tuberculosis patients show substantial splicing dysregulation, indicating splicing defects as a key determinant behind systemic tuberculosis pathology. We propose mycobacterial exploitation of the evolutionarily conserved mechanism of snRNP biogenesis, aided its adaptation for human pathogenesis and provides a unique therapeutic opportunity.

63 **Introduction**

64 Tuberculosis (TB) is one of the leading causes of human mortality due to infectious diseases.[1]
65 *Mycobacterium tuberculosis* (Mtb), the causative organism, is known to alter a variety of host
66 cellular processes that eventually help establish the infection and cause disease. Thus, cellular
67 mechanisms like signaling, autophagy, apoptosis, inflammation etc., have been extensively
68 explored for potential host-directed therapy (HDT) against tuberculosis [2-5]. Mtb infection also
69 alters host gene expression, which determines the long-term host responses against infection.
70 In fact, the pathogen regulates host gene expression at multiple tiers, including epigenetic,
71 transcriptional and post-transcriptional stages [6-9]. Previously, reports from our group and
72 other showed that infection of human macrophages with Mtb triggers a massive alteration in
73 the pattern of host RNA splicing[10,11]. Altered RNA splicing can impact the target proteins
74 function, localization, stability and stoichimetry[12]. We reported increased instances of
75 transcripts which either do not get translated or are substrate of a process called nonsense-
76 mediated decay[13], were upregulated as a result of alternative splicing in the infected
77 macrophages[8,11,13]. Curiously, genes typically belonging to functional classes that are
78 activated upon infection are impacted by AS[8,9,12]. This targeted blunting of host response by
79 bacterial infection favours the bacteria[11]. Incidentally, infection of macrophages with
80 *Salmonella typhimurium* or *Lysteria monocytogenes* were also reported to alter isoform
81 usage, in addition to a widespread shortening of 3'UTR[14]. The macrophage transcriptome
82 also appears to get spliced differently during other inflammatory conditions [15,16]. However,
83 the mechanism determining differential isoform usage in most of these conditions is unclear.
84 Given the fundamental function of RNA splicing in shaping the host proteome and its
85 perturbation by Mtb during macrophage infections, insights into the mechanism of infection-
86 dependent alternative RNA is critical to understand molecular details of tuberculosis
87 pathogenesis[9]

88 Host RNA splicing is a complex and fundamental physiological process in all eukaryotes.
89 RNA splicing is performed by a multimolecular and dynamic spliceosome complex, which
90 employs a large number of proteins to ensure the accuracy of splicing to form mature
91 mRNAs [17-19]. One of the key components is U-rich U snRNAs, which form protein-RNA
92 complexes known as small nuclear ribonucleoproteins (snRNPs)[20-22]. The snRNPs are
93 recruited stepwise onto the pre-mRNA exon-intron boundaries to form a functional
94 spliceosome, join the adjoining exons, and splice out the intervening intron [18,23].
95
96 Like other bacterial pathogens, Mtb also secretes virulence factors to alter host physiology [24-
97 29]. Since mycobacterial secretory proteins in the infected macrophages have access to both
98 cytosol and the nucleus [24-26,30], we asked whether secreted Mtb proteins could impact
99 spliceosome complex assembly and, consequently, host RNA splicing. This study extensively
100 maps the interactions between mycobacterial secretory proteins and host proteins involved in
101 RNA splicing. We identify the Mtb proteins capable of altering the assembly of spliceosome
102 complex by exploiting an evolutionarily conserved quality control mechanism of RNA
103 splicing in the eukaryotes.
104
105

106 **Results**

107 **A targeted high-throughput yeast-2-hybrid (Y2H) screen to identify *Mtb*-human protein**
108 **interactions**

109 Eukaryotic RNA splicing is a highly conserved, multi-step process involving a large number
110 of proteins and protein-RNA complexes (herein referred to as splicing factors or SF). Our
111 hypothesis was to test whether Mtb secretory proteins could directly interact with the SFs to
112 interfere in the host RNA splicing. For this, we performed a high-throughput one-to-one
113 yeast-2-hybrid assay. The bait library (binding domain or BD) comprised of mycobacterial
114 secreted proteins while the prey library (activation domain or AD) comprised the protein-
115 coding isoforms of genes belonging to functional classes "RNA splicing" and "Spliceosome
116 function", making it a unique custom library of human transcripts (Table S1). Briefly, our BD
117 library contained 384 *Mtb* genes and also included 37 putative negative controls (Fig. S1A-B
118 Table S1 and S2)[24,27-29,31].

119 Four independent mating reactions were performed between 384 BD and 162 AD clones to
120 obtain high-quality interaction data, leading to 2,48,832 mating reactions. The schematic of
121 the Y2H strategy followed for the study is shown in Fig. 1A. In total, 123 Mtb proteins
122 showed interactions with one or more SFs, together consisting of 439 incidences of
123 interactions, including 121 autoactivation incidences (Table S3). The BD clones showing
124 autoactivation were excluded from the list of interactors. A few negative control list proteins
125 also showed interactions, including 56 autoactivation incidences (Table S3).

126 The interactome between *Mtb* secretory proteins and host SFs, as observed by the Y2H
127 screen, is shown in Fig. 1B (Table S4). The interactions spanned across all classes of host
128 genes involved in RNA splicing, including RNA polymerases, Sm proteins, SMN complex
129 proteins, U2 snRNPs, LSm proteins, HnRNPs, U4/U6.U5 snRNP tricomplex proteins and
130 non-snRNP proteins (Fig. 1C). The Venn-Diagram shows the number of Mtb proteins

131 interacting with the host SFs (Fig. 1C). Overall, host proteins from RNA Pol II subunits, Sm
132 Core, catalytic step II factors, HnRNPs and SMN complex (Fig. S1C) and Mtb proteins from
133 information pathway, virulence factors and conserved hypothetical proteins functional class
134 showed most interactions (Fig. S1D). To note, there were many instances where different
135 proteins of similar functional sub-class shared their interacting partner with *Mtb* (Fig. 1D).

136 **Interaction of Mtb proteins with the splice variants of human proteins**

137 Since the human library consisted of all protein-coding isoforms of the corresponding genes,
138 among the 162 AD clones that were screened, a total of eighteen genes were represented by
139 more than one transcript (Fig. S2A). Among them POLR2H, SYF2, SNRPD2, LSm7,
140 RBMX, LUC7L2, DCAF4, PPIH and U2AF1 showed interaction with Mtb protein in Y2H
141 (Fig. S2A). The analysis of interaction patterns observed for the different transcripts of the
142 same gene revealed unique insights (Fig. S2B). It shows the gain or loss of interacting
143 partners as a consequence of the gain or loss of exons (Fig. S2B). Most strikingly, truncated
144 isoforms for three genes showed an unusually high number of interactions with the Mtb
145 proteins (Fig. S2B). In our earlier report, we noted a significant increase in the expression of
146 truncated protein-coding isoforms of many host genes, which appeared to help Mtb. Possibly,
147 the truncated spliced variants take confirmations that allow it to interact with more partners.
148 However, this remains a preliminary observation and needs a more focused study to
149 specifically understand the underlying mechanisms and consequences.

150 ***In vitro* splicing assays point to a direct role of *Mtb* proteins in regulating host-RNA**
151 **splicing.**

152 We asked whether the mycobacterial proteins interacting with the host SFs, either through
153 Y2H or MS approach, could also alter host RNA splicing. For functional assays, we selected
154 transcripts from a handful of genes like *ACSL1*, *ACADVL*, *DDX11*, *DDX42* and *NFYA*, each
155 representing a specific exon exclusion event upon H37Rv infection (Fig 1E)[11]. These

156 transcripts were specifically upregulated in H37Rv infected macrophages but not in H37Ra
157 infected ones[11]. We designed primers to screen for the specific splicing events (Fig. S3A).
158 The selected transcripts from the above genes showed an exon exclusion event specifically
159 upregulated upon H37Rv infection (Fig. S3B). The increase was not due to an increased
160 transcriptional event since the exon inclusion primers did not show any change in expression
161 (Fig. S3C). Moreover, using primers that spanned the junction between flanking exons after
162 excluding the middle exon, we reconfirmed that the increased expression of the transcripts
163 noted above was indeed due to an increased exon exclusion event (Fig. S3D). The first set of
164 functional assays involved *in vitro* splicing assays. For *in vitro* splicing assays, we generated
165 minigene constructs for four specific isoforms for the genes *ACSL1*, *ACADVL*, *DDX11* and
166 *DDX42*. The schematic details for *in vitro* splicing assay are shown in Fig. 1E and Fig.S3E.
167 We verified the substrate RNAs produced via *in vitro* transcribing the minigene constructs,
168 undergo splicing and form the desired product in the presence of splicing-competent nuclear
169 extract from THP-1 macrophages (Fig. S3F).
170 Next, to test whether Mtb proteins could directly interfere with RNA splicing, we purified
171 several Mtb proteins identified in the Y2H screen and spliceosome complex pulldown assays
172 (Fig. 1, Fig. S1, Fig. S3G) and used them in independent in *vitro* splicing assays. Most Mtb
173 proteins did not have any impact on the inclusion or exclusion of the test exon in the *in vitro*
174 assays, a few mycobacterial proteins appeared to cause the exclusion of test exon in the case
175 of ACADVL minigene substrate (Fig. 1F). Thus, Mtb secretory proteins could indeed
176 selectively alter the choice of inclusion/exclusion of specific exons. While the *in vitro* assay
177 provided evidence for the possible role of Mtb proteins in altering the host RNA splicing, it
178 did not capture the mechanism for global impact on RNA splicing as reported earlier and
179 which was our major interest [11].

180 ***Mycobacterium smegmatis*-based knock-in system to verify the role of *Mtb* proteins in**
181 **regulating host RNA splicing**
182 As an alternate approach, we recapitulated the infection-like scenario, by generating 19 non-
183 pathogenic *M. smegmatis* (*Msmeg*) lines inducibly expressing Mtb proteins identified through
184 Y2H (*Rv1390, Rv0170, Rv0639, Rv0780, Rv0288, Rv1094, Rv3864, Rv1566c, Rv1643,*
185 *Rv3849* and *Rv1435c*) (Fig S4A). These were sub-cloned under the IVN-inducible promotor
186 to develop corresponding knock-in *Msmeg::Rv* lines (Fig. 2A). We next used each of the
187 *Msmeg::Rv* lines to infect THP-1 macrophages and analyzed specific RNA splicing events of
188 a set of target genes through RT-PCR using the exclusion primers (Fig. S3A), against
189 transcripts from *ACSL1*, *ACADVL*, *DDX11*, *DDX42* and *NFYA* (Fig. S4A). After the first
190 biological experiment, out of 11 genes screened 6 genes (Rv1566c, Rv1643, RV1435c,
191 Rv3849, Rv1390) showed IVN-induction dependent AS event in one or more of the five test
192 exons analysed (Fig. S4A). Out of these Rv1390 did not show consistent patterns during the
193 subsequent biological replicates (Fig. S4B). The *Msmeg::Rv* lines expressing *Rv1566c*,
194 *Rv1435c* or *Rv1643* showed significant exon skipping events for one or more test exons
195 analyzed across three biological replicates (Fig. 2B). Another *Msmeg::Rv* line, expressing
196 Rv3849, also showed consistent AS events across three biological replicates (Fig. S4B);
197 unlike Rv1566c, Rv1435c and Rv1643, we could not confirm its secretion from the
198 *Msmeg::Rv* line and therefore Rv3849 was not taken forward for further studies (Fig. S4C).
199 Most importantly, we observed insignificant changes in the AS for the target transcripts when
200 *Msmeg* alone or upon IVN-induction were used for infection (Fig. S4D). Thus, with the
201 limitation of using only five test exons, each Mtb protein expressed in *Msmeg* impacted the
202 splicing of a unique and overlapping subset of test exons analyzed (Fig. 2B). Despite the
203 limitations, the results from this surrogate approach suggests that select Mtb proteins were
204 able to alter specific host RNA splicing events during infection in THP-1 macrophages.

205 ***In-cell* splicing reporter assays illustrate the direct role of *Mtb* proteins in regulating**
206 **host-RNA splicing.**
207 We next developed an *in-cell* splicing reporter assay to test the three Mtb proteins identified
208 in the results in Fig. 2B. The reporters included a constitutive pRint vector and a set of test
209 pGint vector[32]; the latter was modified to incorporate the custom exon-intron boundaries from
210 our list of splicing events (Fig. 2C). In HEK293T cells, stably expressing pRint, pGint
211 constructs (pGint_ACADVL, pGint_DDX11, pGint_DDX42, pGint_ACSL1 and
212 pGint_NFYA) were co-transfected with plasmids encoding *Rv1566c, Rv1435c and Rv1643*
213 respectively, in separate experiments (Fig. 2D). We used a plasmid encoding *Rv1252c* as a
214 non-interacting control (NIC) for this experiment. We measured the %pGint$^+$ cells (GFP$^+$) in
215 the pRint$^+$ gated population (Fig. 2D). As shown in Figure 2E, transfecting NIC marginally
216 but consistently increased the splicing event for *ACADVL* and *ACSL1*. However, transfecting
217 *Rv1566c*, *Rv1435c,* and *Rv1643* further increased the exon exclusion event in *ACADVL* but
218 not in *ACSL1* (Fig. 2E). NIC transfection had no impact on the exon exclusion event for
219 *DDX11*, *DDX42* and *NFYA* (Fig. 2E). In contrast, expressing *Rv1566c*, *Rv1435c* or *Rv1643*
220 led to a significant increase in the exon exclusion event for *DDX11*, *DDX42* and *NFYA* (Fig.
221 2E). We also analyzed pRint expression in pGint$^+$-population across all the conditions, which
222 did not show any significant differences, suggesting no impact on the basal splicing event as
223 a result of ectopically expressing Mtb proteins in 293T cells (Fig. S5A). Taken together,
224 through the validation experiments in figures 1-2, *in vitro* assays, *ex vivo* infection assays,
225 and *in-cell* reporter-based assays, we demonstrate that the Mtb proteins identified in this
226 study were capable of systemically altering host RNA splicing in a targeted fashion.
227 **Verification of physical interaction between Mtb proteins and host splicing factors**
228 The three Mtb proteins identified above (Rv1435c, Rv1566c and Rv1643) were next verified
229 for their physical interaction with the human SFs in a mammalian system. Mtb proteins

230 (Flag-tagged) and human SFs (3X-HA-tagged) were co-transfected in HEK293T cells, and
231 co-immunoprecipitation was performed using anti-HA and anti-FLAG antibodies,
232 respectively. The Y2H interaction combinations thus verified included PRPF18-Rv1435c,
233 SNRPD2-Rv1435c, GEMIN3-Rv1435c, POLR2H-Rv1566c, and SNRPD2-Rv1643 (Fig. 2F).
234 The combination of SNRPD2-Rv1566c, while showed interactions in the Y2H screen, did not
235 show interaction in the co-IP experiment (Fig. S5B). We also included a few combinations
236 that did not show interaction in the Y2H screen as the additional negative control. Thus, co-
237 transfecting *SNRPF* with *Rv1566c* in HEK293T cells did not result in their co-
238 immunoprecipitation (Fig. S5B). On the other hand, co-transfecting *SNRPF* with *Rv1435c*
239 showed only limited interaction (Fig. S5B). While Rv1435c could pulldown SNRPF, the
240 reverse was not true (Fig. S5B). For a couple of Rv1435c interactions, we also verified the
241 interaction by exchanging the BD and AD clones in Y2H (Fig. S5C). Overall, from the above
242 set of results, we conclude that the Mtb proteins, which were able to alter targeted host-RNA
243 splicing in the functional assays, also showed physical interactions with the host splicing
244 factors in the mammalian system. Finally, each of these three proteins was able to alter exon
245 inclusion for ACADVL test transcript in the *in vitro* splicing assay (Fig. S5D).

246 **Mtb infection impacts levels and distribution of splicing-associated proteins in**
247 **macrophages.**

248 The results above together show that select Mtb proteins could interact with host SFs and
249 alter the host RNA splicing. However, these results did not provide any insights into the
250 possible mechanisms. We next analyzed *in vivo* assembled spliceosome complexes from
251 Mtb-infected macrophages by label-free quantitative proteomics. The spliceosome complexes
252 were pulled down using anti-SNRPB (Y12) antibody from nuclear and cytosolic lysates of
253 the infected and uninfected macrophages to analyse the abundance and distribution of host
254 spliceosome proteins (Table S5). In the nuclear fractions of the Mtb-infected cells, there was

255 a significant depletion of SR proteins like SRSF3, SRSF7, Sm Core proteins like SNRPC,
256 SNRPB, SNRPN and SNRPD2 and some HNRNP proteins, as compared to the uninfected
257 ones (Fig. 3A). Other proteins like RPL4, SF3B2 and SYNCRIP followed a reversed pattern
258 in the infected cells, getting more enriched in the cytosol unlike the uninfected cells (Fig.
259 3A). Similarly, proteins like EIF5A and EIF5AL1 showed increased nuclear enrichment in
260 the infected macrophages, in contrast to the uninfected cells (Fig. 3A). With the limited
261 example of proteins identified in the immunoprecipitated spliceosome complexes, it was
262 evident that a diverse set of SFs were either getting depleted or mislocalized in the Mtb-
263 infected macrophages.

264 **Mtb infection dramatically alters the levels and distribution of proteins involved in**
265 **RNA splicing.**

266 To specifically investigate the levels and distribution of host SFs in Mtb-infected cells, we
267 systematically analyzed a representative set of SFs involved at different stages of RNA
268 splicing (Fig. 3B). We monitored three different parameters- MFI/cell, foci/cell and the
269 cytosolic to the nuclear ratio (Cyt/Nuc) for these proteins. (Fig. 3C-H). We also included
270 some transcription-related proteins for this analysis, since splicing begins while the RNA is
271 still being transcribed [33,34]. Both POLR2G and RBP8 (transcription related proteins) showed,
272 unexpectedly, increased nuclear retention (S6A). The U1 snRNP protein SF1 showed
273 increased MFI and increased cytosolic retention upon infection (Fig. 3C). The U5 snRNP
274 component DDX23 showed a decrease in MFI upon infection (Fig. 3D). The SMN complex
275 proteins GEMIN8 and GEMIN3 showed differential MFI and foci/cell (Fig. 3E-F). A couple
276 of SFs like PRPF4 (U4/U5.U6 associated) and SNRP200 (U5 associated) did not show
277 alteration in any of the three parameters analyzed (Fig. S6B). Among the Sm proteins, LSM7
278 showed decreased MFI/cell in the infected macrophages and SNRPF showed an increased
279 cytosolic retention (Fig. 3G-H). However, a more perplexing observation was the direct

280 interaction of SNRPF with Mtb phagosomes (Fig. 3H). It also appeared that SNRPF was
281 getting sequestered in the cytoplasm due to its strong interaction with Mtb phagosomes (Fig.
282 3I). The results in Figure 3 together indicated that the components of the RNA splicing
283 machinery, specifically snRNP complex biogenesis, could be dysregulated in the Mtb-
284 infected macrophages.

285 **U5 snRNA directly interacts with *Mtb* phagosomes**

286 The nuclear-cytosolic transport of U-rich snRNAs is fundamental to their assembly into U
287 snRNPs, which, together with different snRNPs, form the spliceosome complex[35]. The
288 snRNP complex assembly requires cytosolic transport of the snRNAs, which then interact
289 with SMN complex and Sm Core proteins to form snRNP before entering back into the
290 nucleus to form spliceosome complexes (Fig. 4A). Since SNRPF, a part of the pentameric
291 partial Sm ring, which first recognizes the Sm-sites present on snRNAs before the assembly
292 of the full heptameric ring takes place with the help of the SMN complex [20], was increasingly
293 retained in the cytoplasm of Mtb-infected macrophages, we anticipated that some snRNAs
294 could also show dysregulation in the infected macrophages. We analyzed the U2 and U5
295 snRNAs by fluorescence in situ hybridization (FISH) in the Mtb-infected macrophages. The
296 U2 FISH did not show any aberrant pattern for the parameters- MFI/cell, foci/cell and
297 Cyt/Nuc ratio, except for an increased MFI/cell at 24 hours post-infection (hpi, Fig. 4B).
298 Unlike U2 FISH, U5 FISH showed remarkable patterns in the Mtb-infected macrophages
299 (Fig. 4C). There was a decrease in MFI/cell and foci/cell at 24 hpi; however, a substantial
300 increase was noted in all three parameters at 48 hpi (Fig. 4C). At 48 hpi, there was significant
301 cytosolic retention of U5snRNA (Fig. 4C). More surprisingly, as in the case of SNRPF, we
302 noted direct interaction of U5snRNAs with Mtb in the cytosol (Fig. 4D), which quantitatively
303 increased from 24 to 48 hpi (Fig. 4D). The direct interaction of Mtb with SNRPF and
304 U5snRNA suggests binding and cytosolic sequestration of these specific splicing components

305 by Mtb. This would also mean stoichiometric disbalance among the components of snRNPs
306 in the infected macrophages, a condition which is known to cause widespread splicing defects
307 [9,36]. Interestingly, we verified that the expression of both U5 snRNA (by qPCR) and SNRPF
308 (by immunoblot) consistently increased post-infection, which could be a compensatory
309 mechanism to sustain the snRNP biogenesis in the Mtb-infected cells (Fig. 4E-F). Therefore,
310 disruption of snRNP biogenesis appears to serve as the major mechanism that triggers large-
311 scale splicing alterations in the Mtb-infected macrophages.

312 ***Rv1435c* deletion strain of Mtb fails to alter the splicing of target transcripts.**

313 We next asked whether the Mtb proteins identified in this study were directly responsible for
314 dysregulating snRNP biogenesis. To test the above, we deleted one of the three identified
315 Mtb proteins, *Rv1435c,* in H37Rv using recombineering (Fig. S6C)[37]. This protein is
316 annotated as a secretory protein and showed positive results in most of the interaction and
317 splicing assays. We confirmed the deletion of *Rv1435c* by PCR (Fig. S6D) and also generated
318 a complemented strain *ΔRv1435c::WT* using an integrative vector (see Methods). We first
319 verified the impact of *Rv1435c* deletion on the splicing of five target transcripts used in this
320 study. All five target transcripts, *ACADVL, DDX11, DDX42, ACSL1* and *NFYA*, showed
321 significantly reduced test exon exclusion events for these transcripts upon infection with the
322 H37Rv*ΔRv1435c* strain compared with the WT Mtb (Fig. 5A). Moreover, in the cells infected
323 with H37Rv*ΔRv1435c*::*WT*, the exon exclusion event was restored to the same extent or
324 higher than the WT infected cells (Fig. 5A). These results conclusively show that Rv1435c
325 can directly alter host RNA splicing during Mtb infection. We designated the Mtb protein
326 Rv1435c as the host splicing regulator 1 or Hsr1.

327 **H37Rv*Δhsr1* alters infection-induced global alternative splicing.**

328 We next wanted to see whether the loss of *hsr1* causes global recovery in the infection-
329 induced RNA splicing events. We performed RNAseq analysis from the THP-1

330 macrophages, uninfected or infected with WT or H37Rv*Δhsr1* strains and analyzed the
331 relative abundance of different splice variants of the protein-coding genes. The clustergram in
332 figure 5B shows that H37Rv infection causes increased expression of a large number of
333 transcripts, most of which were effectively restored in H37Rv*Δhsr1* infected cells (Fig. 5B).
334 Interestingly, H37Rv*Δhsr1*-infected cells show increased expression of another unique set of
335 transcripts, likely represented those splicing events, which were possibly inhibited during WT
336 H37Rv infection (Fig. 5B). Curiously, a large number of transcripts showing upregulation in
337 the Mtb-infected macrophages actually represented NMD substrates or processed transcripts
338 (Table S6), which usually don't form a functional protein. Expression of splice variants that
339 are either truncated in length or do not code for a functional protein is one of the ways
340 through which infection-induced alternative splicing helps Mtb evade host responses [11]. The
341 results here show a direct role of Mtb protein Hsr1 in manipulating host RNA splicing events
342 favoring the pathogen.

343 **Mtb *hsr1* directly perturbs snRNP biogenesis by sequestering U5snRNAs to alter host**
344 **RNA splicing.**

345 We next investigated how the loss of Mtb *hsr1* contributed to the reversal in the splicing
346 events in the macrophages. We analyzed U5snRNA interaction and distribution in Mtb*Δhsr1*
347 infected cells by FISH. In concurrence with the altered RNA splicing events reported above,
348 Mtb*Δhsr1* infected cells did not show cytosolic retention of U5snRNA (Fig. 5C). On the
349 similar line, the interaction between Mtb and U5snRNA was almost entirely lost in Mtb*Δhsr1*
350 infected cells (Fig. 5C). Like U5 snRNA, even SNRPF interaction with Mtb*Δhsr1* was
351 significantly lower than that observed with the WT Mtb (Fig. 5D). The effect of *hsr1* deletion
352 on the regulation of U5snRNA and SNRPF was specific since we did not observe any
353 difference in any of the parameters analysed for U2snRNA (Fig. S6E). These results imply

354 that Mtb Hsr1 perturbs host snRNP biogenesis by sequestering the U5snRNA and SNRPF in
355 the cytosol, leading to widespread RNA splicing defects.

356 **Hsr1-mediated impact on host RNA splicing is ESX1-dependent**

357 While we did observe secretion of Hsr1 from Mtb as well as when expressed in *Msmeg*, it
358 was not evident whether it also gets secreted inside the macrophages during infection. We
359 generated anti-sera against Hsr1 in mice to check its secretion and localization during
360 infection in macrophages. Interestingly, we could observe significant Hsr1 secretion in the
361 cytosol of THP-1 macrophage infected with the WT or *Δhsr1::WT* Mtb (Fig. 5E). The Hsr1
362 signal was almost completely lost in the cells infected with *Δhsr1* (Fig. 5E). Curiously, when
363 we used the *ΔCE* Mtb strain (lacking both *Esx1*/ESAT6 and *Esx2/CFP10*), we could not see
364 any cytosolic staining for Hsr1. The Hsr1 staining in THP-1 cytosol was restored when the
365 complemented strain *ΔCE::WT* was used for infection (Fig. 5E). Expectedly, THP-1
366 macrophages infected with Mtb strains lacking the *esx1* system failed to alter RNA splicing
367 for the target genes like ACSL1 and ACADVL (Fig. 5F). Consistent with the above
368 observation, the Mtb *ΔCE* strain could not alter SNRPF cytosolic/nuclear distribution (Fig.
369 S6F). These results suggest that the Hsr1-mediated impact on host RNA splicing reported in
370 this study is ESX1-dependent.

371 **Evolutionary perspective of hsr1 in mycobacterial virulence**

372 On UniProtKB, Rv1435c is characterized as a conserved proline, glycine, valine-rich secreted
373 protein. There are several orthologs present in diverse mycobacterial species, like *M.*
374 *marinum*, where it is classified as a conserved hypothetical protein; in opportunistic
375 pathogenic species like *M. paratuberculosis, M. avium,* and *M. helveticum,* the identified
376 orthologue codes for secreted beta-xylosidase enzyme (Fig. S7A-B). *In M. smegmatis*, the
377 orthologue, with very poor similarity, encodes a translation initiation factor (IF-2). In a few
378 pathogenic species like *M. lacus and M. heidelbergense,* it codes for a beta-xylosidase (Fig.

379 S7A-B). Thus, most *hsr1* orthologues in the genus *Mycobacterium* encode a secreted protein,
380 and their functions are better characterized in non-pathogenic or opportunistic pathogens. It is
381 likely that the ability of hsr1 to alter host splicing events was likely a gain-of-function event
382 during the course of Mtb co-evolution as a human pathogen.

383 **Mtb Hsr1 is a virulence factor.**

384 Since Mtb*Δhsr1* infection could not alter RNA splicing events as in WT Mtb-infected cells,
385 we wondered whether loss of *hsr1* also impacts bacterial pathogenicity. The Mtb*Δhsr1*
386 showed reduced intra-macrophage survival compared to the WT Mtb (Fig. 6A, S8A). The
387 Mtb*Δhsr1* strain, as such, did not have a growth defect *in vitro*. Rather, it grew faster than the
388 WT cells in liquid culture which was consistent with a decrease in the bacterial cell length
389 (Fig. S8B-C). Curiously, *Δhsr1::WT* grew inside the THP-1 macrophages even better than the
390 WT Mtb, pointing to Hsr1 being an important virulence factor (Fig. 6A, Fig. S8A).
391 Interestingly, *Δhsr1* showed increased targeting to the lysosomes and also the cell had higher
392 cathepsin D staining (Fig. S8D). Since Mtb*Δhsr1* was unable to bind and sequester SNRPF,
393 we wondered whether overexpressing SNRPF during WT Mtb infection could have an impact
394 on Mtb CFU. Overexpressing SNRPF in human macrophages caused a relative decline in
395 cytosolic retention of SNRPF and a reduced intracellular CFU of WT Mtb (Fig. S8E),
396 indicating that SNRPF sequestration by Hsr1 serves as the virulence mechanism of Hsr1. On
397 the contrary, overexpression of SNRPD2 did not have any impact on Mtb CFU (Fig. S8F).
398 Next, to assess the role of Hsr1 in bacterial virulence *in vivo*, we infected BALB/c mice with
399 WT, Mtb*Δhsr1* and Mtb*Δhsr1::WT* strains and analyzed bacterial CFU at two and four weeks
400 post-infection. At both two- and four-weeks post-infection, Mtb*Δhsr1* had fewer CFUs in the
401 lungs of the animals (Fig. 6B). Unlike in THP-1 macrophages, the complemented strain
402 Mtb*Δhsr1::WT* was not able to rescue the bacterial CFU (Fig. 6B). At both 2- and 4- weeks
403 post-infection, analysis of a panel of eight cytokines in the lung homogenates showed an

404 increase in cytokine levels in WT infected animals, which, to some extent, were diminished
405 in the Mtb*Δhsr1* infected mice (Fig. S8G). The lung homogenates from mice infected with
406 Mtb*Δhsr1::WT* did not revert the expression to the WT levels (Fig. S8G). The lung tissue
407 sections from mice that were infected for four weeks showed strong SNRPF staining in case
408 of WT and *Δhsr1::WT* Mtb infection (Fig. 6C). The SNRPF staining was strikingly lower in
409 uninfected mice or those infected with Mtb*Δhsr1* (Fig. 6C). Remarkably, SNRPF staining in
410 the WT or *Δhsr1::WT* Mtb infected mice appeared mostly cytosolic in nature. Together, the
411 *ex-vivo* assays in macrophages and *in vivo* assays corroborate Mtb Hsr1 as a virulence factor
412 which helps its survival in the animals, largely by sequestering the host SNRPF.

413 **Macrophages in human TB granulomas show altered snRNP complex component**
414 **localization.**

415 Specific SNRPF cytosolic retention in THP1 macrophage and in the lungs of infected mice
416 prompted us to explore the clinically relevant human samples. Tissue sections from the
417 confirmed human TB granulomas belonging to abdominal TB cases were stained with
418 SNRPF. To our surprise, all granuloma lesions showed strong SNRPF staining (Fig. 6D,
419 S9A). Moreover, it was evident that there were more cytosolic SNRPF than in the nucleus
420 (Fig. 6D). Lesions that were highly necrotic showed even higher SNRPF levels (Fig. 6D).
421 Multinucleated giant cells, a hallmark of TB granulomas, showed intense staining for
422 cytosolic SNRPF (Fig. S9A). Unlike the granuloma lesions, in healthy tissue sections,
423 SNRPF staining was not remarkable (Fig. S9B). Finally, we also stained for another sm Core
424 protein, i.e., SNRPD2, in the granulomatous lesions or healthy sections. In both cases,
425 SNRPD2 staining was conspicuously nuclear and indistinctive (Fig. S9C). Thus, these results
426 highlight an unusual SNRPF-dependent dysregulation in human TB lesions. No previous
427 report of such rampant dysregulation of snRNP biogenesis in human TB granulomas exists.

428 **Monocytes from human pulmonary TB patients show pervasive regulation of**
429 **alternative splicing.**
430 Next, from an RNAseq dataset from monocytes of 20 PTB cases and 19 healthy contacts, we
431 analyzed and compared the relative abundance of different transcripts [38]. We focused our
432 analysis on four major functional classes- RNA splicing, immune response, inflammation and
433 metabolism (Fig. 6E, Table S7). As can be seen from the clustergrams, transcript level
434 expression of genes belonging to each of these functional classes was able to classify the PTB
435 and HC subjects (Fig. 6E). In each of the cases, there were two major clusters of transcripts-
436 one showing increased expression in the TB cases (Cluster 1) and the other in the control
437 cases (Cluster 2). We analyzed the biotypes of the transcripts present in these clusters across
438 all the groups. In the RNA splicing class, cluster 2 was much bigger than cluster 1, showing
439 more downregulation events in the TB cases (Fig. 6E). Moreover, the small number of
440 transcripts that were upregulated in TB subjects were either not the protein-coding isoforms
441 or encoded a truncated protein isoform (Fig. 6F). A similar analysis performed for other
442 functional classes revealed systemic splicing defects (Fig. 6F). In inflammation, immune
443 response and metabolism classes, while a large number of transcripts were upregulated in the
444 PTB subjects, a significant number of those genes did not code for a protein or encoded a
445 truncated isoform of the corresponding protein (Fig. 6F). For example, the upregulated genes
446 in the immune response class show increased expression of the transcripts of genes like
447 TLR4, THBS1, OAS, OASL3, GRN, CXCL16, etc. (Fig. 6G), however for each of them one
448 or more splice variants with highly truncated protein length were also upregulated, which
449 could significantly blunt the host responses. Comparison of monocyte expression data with
450 those from H37Rv-infected THP-1 (Fig. 5B) showed around 45 commonly upregulated
451 transcripts; however, almost a third of those genes did not code for a protein-coding variant
452 (Fig. S9D, Table S8). Taken together, these observations make it evident that by perturbing

453 host RNA splicing, the mechanism for which we report in this study, Mtb exerts a pervasive
454 influence on host physiology and contributes towards systemic pathologies witnessed in TB
455 patients.

456 **Discussions**

457 RNA splicing is a conserved eukaryotic process which impacts cellular proteome diversity
458 [10,39]. Also, its dysregulation is associated with diverse pathologies [40]. This study was designed
459 to understand the mechanism of Mtb infection-mediated alteration of host RNA splicing[9].
460 Hijacking of host RNA splicing machinery by viral pathogens, including influenza virus and
461 SARS-CoV2 is well-known [41-44]. However, very little is understood in the context of
462 bacterial pathogens. Identification of several Mtb proteins that could directly interact with
463 and alter host RNA splicing is a significant leap in our understanding of host-pathogen
464 interactions. Similarly, the observation that different spliced variants could interact
465 differently with Mtb proteins adds a new dimension to host-pathogen interactions. There are
466 only limited efforts to systematically study and map out how alternative splicing events could
467 contribute to the functional divergence in the proteome [45]. However, our results underscore
468 that more efforts are needed to understand the consequences of this fundamental aspect of
469 eukaryotic regulation.

470 The dysregulation in the abundance and distribution of SFs in the infected macrophages
471 provided crucial insights into the underlying mechanism. The U snRNPs biogenesis is
472 fundamental to eukaryotic gene expression, given their critical role in RNA splicing [46-48]. For
473 snRNPs to form, pre-snRNAs from the nucleus are transported to the cytoplasm, where they
474 interact with the multi-molecular SMN complex and Sm Core ring [46-48]. The snRNA-protein
475 complex comprising SMN complex and Sm Core ring, is transported back to the nucleus
476 where they mature into snRNPs and participate in the splicing reactions [49,50]. SMN complex,
477 including SMN1/2 dimer and Gemin proteins, charge snRNAs with the Sm core ring [51]. A
478 loss or altered stoichiometry in either the Gemins or SMN1 or 2, can lead to widespread
479 splicing defects in different tissues [20,52-54].

480 Thus, while a few Mtb proteins could alter RNA splicing of test transcripts in the *in vitro*
481 assays, the dysregulation of SMN complex proteins and Sm core proteins reported in this
482 study appears to be the major driver for the global alteration of AS in Mtb-infected
483 macrophages. To our surprise, we also noted direct interactions between SNRPF and U5
484 snRNA with Mtb, which suggests potential sequestration of these key regulatory components
485 by Mtb. While several viral proteins are reported to directly interact with the SFs [9,36,44,55],
486 there is no precedence of such interaction of SFs or snRNAs directly with a viral or bacterial
487 pathogen.
488 There could be concerns about the specificity of this regulation since targeted dysregulation
489 of splicing factors (SFs) can lead to a pervasive impact on alternative splicing. However,
490 studies suggest it does not alter the splicing of all introns in the cell [19,56,57]. Rather, the impact
491 of one splicing factor dysregulation could be very distinct and selective from another SF
492 dysregulation [19]. At the same time, there is some functional congruence, too, in that a specific
493 SF could regulate AS of functionally related genes [19,56,57].
494 Experiments using the *hsr1* deleted Mtb strain provide definite proof that Mtb Hsr1 regulates
495 host RNA splicing in an ESX1-dependent manner. More strikingly, the loss of *hsr1* impacted
496 the ability of Mtb to interact and sequester U5snRNA or SNRPF. It is pertinent to ask why
497 Mtb targets U5snRNA and no other snRNAs or at least the U2snRNAs, which we analyzed.
498 During splicing, U5snRNPs recognize the acceptor splice site [58]. Several viral proteins
499 interact with the U5 components to disrupt the normal RNA splicing [36,55]. Thus, targeting of
500 U5snRNA through interaction reveals somewhat conserved pathogenic mechanisms across
501 diverse human pathogens. The decline in Mtb*Δhsr1* survival in macrophages and in mice
502 clearly indicates that the ability to disrupt snRNP biogenesis is a desirable phenotype for
503 tuberculosis pathogenesis. To note, the orthologues of *hsr1* are secreted xylose hydrolysing
504 enzymes, which is probably required by non-pathogenic, or opportunistic pathogenic bacteria

505 in search of diverse carbon sources. Therefore, the gain of function to interact with host SFs
506 appears to be an important adaptation during the evolution of Mtb as a human pathogen.
507 Interestingly, all snRNAs, with a few exceptions, while synthesized in the nucleus, must
508 travel to the cytosol for assembly before they are transported back to the nucleus for
509 participation in the splicing reaction [49]. The cytosolic export of pre-snRNAs ensures that
510 immature snRNAs do not get incorporated into the spliceosomal complexes in the nucleus,
511 which can alter host RNA splicing catastrophically [35,59]. This evolutionarily conserved quality
512 control step appears to be co-opted by the pathogenic mycobacteria during its evolution as the
513 human pathogen.
514 Tuberculosis granulomas evolve through a complex interplay between the bacterial factors
515 and host machinery, which largely benefits the pathogen [60]. Therefore, the intense SNRPF
516 staining of human tuberculosis granulomas is of great interest. Similarly, the monocyte
517 RNAseq data from the patients emphasize the pervasiveness of infection-induced alteration in
518 host RNA splicing. Incidentally, differentially expressed transcript isoforms were recently
519 reported in an LTBI cohort and identified those associated with resistance to TST/IGRA
520 conversion [61]. There are systemic inflammation, metabolic dysregulation and muscle loss
521 reported in TB patients [62], some of which could potentially be a consequence of the large-
522 scale splicing defects observed in these individuals.
523 To conclude, with the limitation of only select splicing events analyzed, we present a new
524 paradigm of pathogen-mediated perturbation of host physiology and show the hijacking of
525 this evolutionarily conserved mechanism by Mtb to blunt host responses. We propose that
526 inhibition of molecular interactions between the host and the pathogenic partners identified in
527 this study could be exploited for developing a novel class of therapeutics.
528
529

Acknowledgements:

The work was supported by the DBT-Wellcome Trust India Alliance Senior Fellowship to DK (IA/S/17/1/503071). The work was also partly supported by grants from SERB (EMR/2016/005296 to DK) and the Department of Biotechnology, Govt of India (BT/IC-06/003/91-Flagship program to DK). All work on virulent Mtb was performed in the Tuberculosis Aerosol Challenge Facility (TACF), a national facility at ICGEB, New Delhi, supported by the Department of Biotechnology, Govt. of India.

Author Contributions:

Original conceptualization: DK; Experiments: KC, NP, YK, DD, NA, PS, BM, HK, LS, RD, RDS, LM, BS, PY; Investigation: KC, DD; Analysis and Visualization: KC, MD, PY, DD, NSB, DK; Resources: VN, AS, VA, PD; Supervision: DK; Writing-original draft: DK, KC, MD, NP; Writing—review & editing: VN, AS, NSB, PD, DK

543 **Figures and Figure Legends**

544

Figure 1: Mycobacterial secretory proteins and host spliceosome proteins interactome.

A) The schematic flow chart for high-throughput one-to-one Y2H assay in 1536-well format. Protein pairs showing interaction in 3 or 4 biological replicates were considered as positive interactions.

B) Mtb proteins -host splicing factor interactome- the positive interactions from the Y2H screen were used to generate the network shown here. The diamond and circle represent the host and Mtb proteins, respectively, and the size corresponds to the number of interactions observed. The proteins coloured yellow in the network are discussed in detail in the text.

C) The Venn diagrams show proteins belonging to the major spliceosome components and their corresponding overlap with the interacting Mtb proteins.

D) A more descriptive representation of the Venn diagrams, shown above in Fig. 1C is presented in the table as a list of interactions.

E) Transcripts that were used for making the minigene constructs to test the role of Mtb proteins in host RNA splicing. Only exons are shown, and two different transcripts for each of the five genes are shown. The red bars indicate the overlapping sequences, and the skipped exons are shown in yellow.

F) On the left, a schematic of the minigene cassettes for DDX42 and ACADVL is shown. The constitutive exons are shown in green, while the alternative exons are shown in blue. *In vitro* transcribed products of minigenes were used for in-vitro splicing assays in the presence of purified Mtb proteins Rv3867, Rv0360, Rv1566c and Rv1471.

568 **Figure 2: *Ex vivo* infection assays and reporter assays to test the role of Mtb proteins in**
569 **host RNA splicing**

570 **A)** The schematic representation of infecting THP-1 macrophages with *M. smegmatis*
571 strains that express mycobacterial proteins under IVN inducible promotor. Isoform-
572 specific qPCR for the truncated isoforms of ACADVL, DDX11, DDX42, ACSL1 and
573 NFYA was used to assess the impact on alternative splicing.

574 **B)** Expression of spliced isoforms of the five target genes (as indicated) in THP-1
575 macrophages upon infection with WT or *Rv1566c, Rv1435c* and *Rv1643* knock-in *M*
576 *smegmatis* lines. *M smegmatis* lines were induced with IVN (indicated dose) for 12
577 hours before infection. Data from three independent biological experiments. The p-
578 values shown are from One-way ANOVA.

579 **C)** Designing of the custom *in vivo* fluorescent splicing reporter. Schematic for the
580 constitutive splicing reporter pRint with a short adenovirus intron intervening the
581 DsRed gene with a strong 5' and 3' splice sites. In the pGint splicing constructs, the
582 alternative exon of ACADVL, DDX11, DDX42, ACSL1 and NFYA, along with the
583 flanking 100bp intronic sequences, were inserted within eGFP reading frame. The exon
584 disrupts the reading frame, however, skipping of the test exon reconstitutes EGFP
585 fluorescence.

586 **D)** The pRint expressing stable HEK293T cell lines were co-transfected with a pGint
587 construct and pDest-FLAG construct encoding proteins Rv1252c (NIC), Rv1566c,
588 Rv1435c, Rv1643, or empty vector control, followed by Flow cytometry. In the $DsRed^+$
589 gate, $\%FITC^+$ population was calculated to measure the extent of the test exon-skipping
590 event.

591 **E)** $\%FITC^+$ population for each of the indicated reporter constructs in the presence of
592 indicated mycobacterial protein, non-interacting control or vector controls. Each dot
593 represents an independent biological replicate (N=3), indicating p values from One-
594 way ANOVA.

595 **F)** Validation of physical interactions between indicated pairs of *Mtb* proteins (FLAG-
596 tagged) and host proteins (HA-tagged) in HEK293T cells. Data is representative of two
597 or more biologically independent transfection and pulldown experiments.

598

599

Figure 3: Dysregulation of host RNA splicing machinery in Mtb-infected macrophages

A) Heatmap shows massive dysregulation in the level and cellular distribution of human proteins involved in RNA splicing and identified in MS of the spliceosome complex pulldown from the Mtb infected THP-1 macrophages 48 hours post-infection.

B) A schematic model depicting the ordered pathway of mRNA transcription by RNAP II and splicing catalyzed by spliceosome complex. The U1snRNP binds the 5' splice site and U2 snRNP binds the branch point adenine to form complex A. U4/U5.U6 snRNP binds to form complex B and causes conformational changes to form activated complex B. This is followed by the first transesterification reaction releasing U1 and U4 snRNP and the second transesterification reaction forming complex C releasing the lariat intron and U2, U5 and U6 snRNP. The U snRNP are the major players in splicing catalysis and their biogenesis is mediated by the SMN complex and the Sm proteins in the cytoplasm.

C-H) The proteins were visualized by indirect immunofluorescence (green), the nucleus was visualized by DAPI staining (blue) and mCherry-expressing bacteria (red) was used for infection. Each dot in the bar graphs represents a field analyzed, and an average of a hundred cells were quantified. The confocal images were analyzed for three different parameters- overall MFI (MFI/cell), number of Foci (Foci/cell) and cytosolic to the nuclear ratio (Cyt/Nuc) for all major splicing-related proteins upon Mtb infection such as **C)** U1 associated proteins (SF1), **D)** U5 associated proteins (DDX23), **E)** SMN complex proteins (GEMIN8) **F)** SMN complex proteins (GEMIN3) and **G)** Sm Core proteins (LSM7) and **H)** Sm Core proteins (SNRPF)

I) Image showing SNRPF colocalizing with bacterial phagosomes. The 3D reconstruction of the nucleus, SNRPF foci and bacteria (see methods) was done to show the SNRPF Foci at <0.2 microns distance from the bacterial phagosomes. Scale bar: 2 μm and p values indicate One-way ANOVA.

628 **Figure 4: Dysregulation in SMN complex and Sm proteins impact U snRNA localization**
629 **and overall assembly of U snRNPs.**

630 **A)** Schematic representation of U snRNA nucleo-cytoplasmic trajectory and assembly of
631 Sm core ring on snRNA by SMN complex to form snRNPs

632 **B)** Representative images for RNA FISH against U2 snRNA in THP-1 macrophages upon
633 Mtb infection 24- and 48-hrs post-infection. The U2 snRNA probe was labelled with
634 Alexa 488 (green), the nucleus was stained with DAPI (blue) and mCherry-expressing
635 H37Rv (red) was used for infection. The plots corresponding to overall MFI (MFI/cell),
636 number of Foci (Foci/cell) and cytosolic to nuclear ratio (Cyt/Nuc) of U2 snRNA are
637 shown here. Each dot represents a field analyzed, and an average of a hundred cells are
638 quantified. The p-values is from one-way ANOVA.

639 **C)** Representative images for RNA FISH against U5 snRNA in THP-1 macrophages upon
640 Mtb infection 24- and 48-hrs post-infection. The U5 snRNA probe was labelled with
641 Alexa 488 (green), the nucleus was stained with DAPI (blue) and mCherry-expressing
642 H37Rv (red) was used for infection. The plots corresponding to overall MFI (MFI/cell),
643 number of Foci (Foci/cell) and cytosolic to nuclear ratio (Cyt/Nuc) of U5 snRNA are
644 shown here. Each dot represents a field analyzed, an average of a hundred cells are
645 quantified, and p values are from one-way ANOVA.

646 **D)** Direct interaction between H37Rv (red) and U5snRNA at 24- and 48-hrs post-infection.
647 For the plot at the right, the number of U5snRNA puncta within 0.2 microns from
648 H37Rv surface was counted, p values from one-way ANOVA. Representative images
649 for colocalization of Sm core protein SNRPF (green) to the H37Rv phagosome (red).
650 The 3D reconstruction and spot analysis are done as previously described.

651 E) Expression of U5 snRNA in THP-1 macrophages, 12 and 48 hpi with H37Rv. Data
652 from three independent biological experiments.

653 F) Immunoblot showing SNRPF protein expression in the whole cell lysates of uninfected
654 (UI) and H37Rv-infected THP-1 macrophages, 48 hpi. Data is representative of two
655 independent biological experiments.

656

657

Figure 5: Role of H37Rv hsr1 protein in snRNP biogenesis and altering host RNA splicing.

A) Q-PCR for the target transcripts as mentioned after infection with H37Rv, H37Rv*Δhsr1* and H37Rv*Δhsr1::WT* strain. Data from three biological replicates (N=3), *p<0.05, **p<0.01 (one-way ANOVA).

B) Clustergram of RNAseq data obtained from uninfected, or THP1 macrophages infected with H37Rv or H37Rv*Δhsr1* strains. Transcripts showing an infection-dependent increase (red box), hsr1-triggered increase (purple) or hsr1-blocked transcripts (turquoise) are shown. The distribution of different transcript biotypes across the three groups is shown in the aligned plot towards the right.

C) Representative image for the U5snRNA FISH in THP-1 macrophages infected with WT H37Rv or H37Rv*Δhsr1* strains. The plots at the right show the cytosolic to nuclear ratio and the ratio of interacting to non-interacting U5snRNA with WT H37Rv or H37Rv*Δhsr1*. **p<0.01 (one-way ANOVA).

D) Representative images of SNRPF staining in THP-1 macrophages infected with WT H37Rv or H37Rv*Δhsr1* strains. For the plot at the right, the number of SNRPF puncta within 0.2 microns from the H37Rv surface was counted as interacting and shows the ratio of interacting to non-interacting SNRPF with WT H37Rv or H37Rv*Δhsr1*. *p<0.05 (one-way ANOVA).

E) Representative images of THP1-macrophages stained with DAPI (nucleus) and Rv1435c antisera (green) upon infection with WT H37Rv, H37Rv*Δhsr1,* H37Rv*ΔCE* and H37Rv*ΔCE::WT* and H37Rv*Δhsr1::WT.* Images are representative of two independent biological experiments. Scale bar-2 μm.

F) Expression of truncated isoform as indicated upon infection with WT H37Rv, H37Rv*ΔCE* and H37Rv*ΔCE::WT* at 48hpi from three biologically independent experiments. The p values are from One-way ANOVA.

686 **Figure 6: Regulation of host alternative splicing as a virulence mechanism**

687 A) THP-1 macrophages were infected with WT H37Rv, H37Rv*Δhsr1,* H37Rv*Δhsr1::WT*
688 strains. At 48 hours post-infection, bacterial load was examined using CFU plating.
689 Data represent fold change in CFU at 48 hrs compared to 0 hr post-infection. (N=3), p
690 values are from one-way ANOVA.

691 B) BALB/C mice (6-8 weeks old) were infected with 200 CFU of WT H37Rv,
692 H37Rv*Δhsr1,* H37Rv*Δhsr1::WT* in separate experiments. At 2- and 4 weeks post-
693 infection, bacterial load in the lungs of the animals were determined by CFU plating. P
694 values are from one-way ANOVA.

695 C) The 4 week post-infection lung tissue sections of WT H37Rv, H37Rv*Δhsr1,*
696 H37Rv*Δhsr1::WT* were stained with DAPI (nucleus) and SNRPF. Scale bar: 20 µm

697 D) Human intestinal tuberculosis granulomatous lesions were stained with SNRPF. The
698 arrows indicate multinucleated giant cells with intense staining for cytosolic SNRPF.
699 Black arrows show areas of intense SNRPF staining in the granulomas and in the giant
700 cells.

701 **E)** RNAseq data from the monocytes of pulmonary TB cases (N=20) and healthy contacts
702 (N=19) was analyzed for expression of differential alternatively spliced transcripts.
703 Genes belonging to the indicated functional classes were selected for the cluster
704 analysis. Cluster 1 in each case shows transcripts upregulated in PTB cases.

705 F) For each of the functional classes shown in Fig. 6D, the relative distribution of transcript
706 biotypes (as indicated) is plotted and binned separately for those that were up- (cluster
707 1, red) or down- (cluster 2, blue) regulated in PTB cases.

708 G) The transcripts from the list of genes in the immune response class (Fig. 6D) were
709 further analyzed for their length with respect to the longest encoded isoform of the
710 corresponding gene. Only genes with more than one transcript in the list were chosen
711 for this analysis. The red dotted line separates cluster 1 and cluster 2 transcripts.
712 Transcripts of DYSF and RNASET2 are present in both clusters.

713

714

715

716 **References**
717

718 1 in *WHO consolidated guidelines on tuberculosis: Module 3: diagnosis - rapid*
719 *diagnostics for tuberculosis detection WHO Guidelines Approved by the Guidelines*
720 *Review Committee* (2021).
721 2 Pisu, D., Huang, L., Grenier, J. K. & Russell, D. G. Dual RNA-Seq of Mtb-Infected
722 Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions.
723 *Cell Rep* **30**, 335-350 e334 (2020). https://doi.org/10.1016/j.celrep.2019.12.033
724 3 Matteucci, K. C., Correa, A. A. S. & Costa, D. L. Recent Advances in Host-Directed
725 Therapies for Tuberculosis and Malaria. *Front Cell Infect Microbiol* **12**, 905278
726 (2022). https://doi.org/10.3389/fcimb.2022.905278
727 4 Young, C., Walzl, G. & Du Plessis, N. Therapeutic host-directed strategies to improve
728 outcome in tuberculosis. *Mucosal Immunol* **13**, 190-204 (2020).
729 https://doi.org/10.1038/s41385-019-0226-5
730 5 Kumar, D. & Rao, K. V. Regulation between survival, persistence, and elimination of
731 intracellular mycobacteria: a nested equilibrium of delicate balances. *Microbes Infect*
732 **13**, 121-133 (2011). https://doi.org/10.1016/j.micinf.2010.10.009
733 6 Karim, A. F. *et al.* Express path analysis identifies a tyrosine kinase Src-centric
734 network regulating divergent host responses to Mycobacterium tuberculosis infection.
735 *J Biol Chem* **286**, 40307-40319 (2011). https://doi.org/10.1074/jbc.M111.266239
736 7 Khan, N. *et al.* M. tuberculosis Reprograms Hematopoietic Stem Cells to Limit
737 Myelopoiesis and Impair Trained Immunity. *Cell* **183**, 752-770 e722 (2020).
738 https://doi.org/10.1016/j.cell.2020.09.062
739 8 Kalam, H., Singh, K., Chauhan, K., Fontana, M. F. & Kumar, D. Alternate splicing of
740 transcripts upon Mycobacterium tuberculosis infection impacts the expression of
741 functional protein domains. *IUBMB Life* **70**, 845-854 (2018).
742 https://doi.org/10.1002/iub.1887
743 9 Chauhan, K., Kalam, H., Dutt, R. & Kumar, D. RNA Splicing: A New Paradigm in
744 Host-Pathogen Interactions. *J Mol Biol* **431**, 1565-1575 (2019).
745 https://doi.org/10.1016/j.jmb.2019.03.001
746 10 Liu, Y. *et al.* Impact of Alternative Splicing on the Human Proteome. *Cell Rep* **20**,
747 1229-1241 (2017). https://doi.org/10.1016/j.celrep.2017.07.025
748 11 Kalam, H., Fontana, M. F. & Kumar, D. Alternate splicing of transcripts shape
749 macrophage response to Mycobacterium tuberculosis infection. *PLoS Pathog* **13**,
750 e1006236 (2017). https://doi.org/10.1371/journal.ppat.1006236
751 12 Zhang, X. O., Fu, Y., Mou, H., Xue, W. & Weng, Z. The temporal landscape of
752 recursive splicing during Pol II transcription elongation in human cells. *PLoS Genet*
753 **14**, e1007579 (2018). https://doi.org/10.1371/journal.pgen.1007579
754 13 Hug, N., Longman, D. & Caceres, J. F. Mechanism and regulation of the nonsense-
755 mediated decay pathway. *Nucleic Acids Res* **44**, 1483-1495 (2016).
756 https://doi.org/10.1093/nar/gkw010
757 14 Pai, A. A. *et al.* Widespread Shortening of 3' Untranslated Regions and Increased
758 Exon Inclusion Are Evolutionarily Conserved Features of Innate Immune Responses
759 to Infection. *PLoS Genet* **12**, e1006338 (2016).
760 https://doi.org/10.1371/journal.pgen.1006338
761 15 Bhatt, D. M. *et al.* Transcript dynamics of proinflammatory genes revealed by
762 sequence analysis of subcellular RNA fractions. *Cell* **150**, 279-290 (2012).
763 https://doi.org/10.1016/j.cell.2012.05.043

16 Pandya-Jones, A. *et al.* Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA* **19**, 811-827 (2013). https://doi.org/10.1261/rna.039081.113

17 Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718 (2009). https://doi.org/10.1016/j.cell.2009.02.009

18 Wilkinson, M. E., Charenton, C. & Nagai, K. RNA Splicing by the Spliceosome. *Annu Rev Biochem* **89**, 359-388 (2020). https://doi.org/10.1146/annurev-biochem-091719-064225

19 Papasaikas, P., Tejedor, J. R., Vigevani, L. & Valcarcel, J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* **57**, 7-22 (2015). https://doi.org/10.1016/j.molcel.2014.10.030

20 Zhang, R. *et al.* Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**, 384-395 (2011). https://doi.org/10.1016/j.cell.2011.06.043

21 Karijolich, J. & Yu, Y. T. Spliceosomal snRNA modifications and their function. *RNA Biol* **7**, 192-204 (2010). https://doi.org/10.4161/rna.7.2.11207

22 Strasser, A., Dickmanns, A., Luhrmann, R. & Ficner, R. Structural basis for m3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1. *EMBO J* **24**, 2235-2243 (2005). https://doi.org/10.1038/sj.emboj.7600701

23 Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem* **84**, 291-323 (2015). https://doi.org/10.1146/annurev-biochem-060614-034316

24 Penn, B. H. *et al.* An Mtb-Human Protein-Protein Interaction Map Identifies a Switch between Host Antiviral and Antibacterial Responses. *Mol Cell* **71**, 637-648 e635 (2018). https://doi.org/10.1016/j.molcel.2018.07.010

25 Portal-Celhay, C. *et al.* Mycobacterium tuberculosis EsxH inhibits ESCRT-dependent CD4(+) T-cell activation. *Nat Microbiol* **2**, 16232 (2016). https://doi.org/10.1038/nmicrobiol.2016.232

26 Sharma, G., Upadhyay, S., Srilalitha, M., Nandicoori, V. K. & Khosla, S. The interaction of mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG methylation and histone H3/H4 binding. *Nucleic Acids Res* **43**, 3922-3937 (2015). https://doi.org/10.1093/nar/gkv261

27 Mehra, A. *et al.* Mycobacterium tuberculosis type VII secreted effector EsxH targets host ESCRT to impair trafficking. *PLoS Pathog* **9**, e1003734 (2013). https://doi.org/10.1371/journal.ppat.1003734

28 de Souza, G. A., Leversen, N. A., Malen, H. & Wiker, H. G. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics* **75**, 502-510 (2011). https://doi.org/10.1016/j.jprot.2011.08.016

29 Malen, H., Berven, F. S., Fladmark, K. E. & Wiker, H. G. Comprehensive analysis of exported proteins from Mycobacterium tuberculosis H37Rv. *Proteomics* **7**, 1702-1718 (2007). https://doi.org/10.1002/pmic.200600853

30 Yaseen, I., Kaur, P., Nandicoori, V. K. & Khosla, S. Mycobacteria modulate host epigenetic machinery by Rv1988 methylation of a non-tail arginine of histone H3. *Nat Commun* **6**, 8922 (2015). https://doi.org/10.1038/ncomms9922

31 Malakar, B. *et al.* Phosphorylation of CFP10 modulates Mycobacterium tuberculosis virulence. *mBio* **14**, e0123223 (2023). https://doi.org/10.1128/mbio.01232-23

32 Bonano, V. I., Oltean, S. & Garcia-Blanco, M. A. A protocol for imaging alternative splicing regulation in vivo using fluorescence reporters in transgenic mice. *Nat Protoc* **2**, 2166-2181 (2007). https://doi.org/10.1038/nprot.2007.292

33 Herzel, L., Ottoz, D. S. M., Alpert, T. & Neugebauer, K. M. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol* **18**, 637-650 (2017). https://doi.org/10.1038/nrm.2017.63

34 Oesterreich, F. C. *et al.* Splicing of Nascent RNA Coincides with Intron Exit from RNA Polymerase II. *Cell* **165**, 372-381 (2016). https://doi.org/10.1016/j.cell.2016.02.045

35 Becker, D. *et al.* Nuclear Pre-snRNA Export Is an Essential Quality Assurance Mechanism for Functional Spliceosomes. *Cell Rep* **27**, 3199-3214 e3193 (2019). https://doi.org/10.1016/j.celrep.2019.05.031

36 Boudreault, S. *et al.* Reovirus mu2 protein modulates host cell alternative splicing by reducing protein levels of U5 snRNP core components. *Nucleic Acids Res* **50**, 5263-5281 (2022). https://doi.org/10.1093/nar/gkac272

37 van Kessel, J. C. & Hatfull, G. F. Recombineering in Mycobacterium tuberculosis. *Nat Methods* **4**, 147-152 (2007). https://doi.org/10.1038/nmeth996

38 Del Rosario, R. C. H. *et al.* Histone acetylome-wide associations in immune cells from individuals with active Mycobacterium tuberculosis infection. *Nat Microbiol* **7**, 312-326 (2022). https://doi.org/10.1038/s41564-021-01049-w

39 Kornblihtt, A. R. *et al.* Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol* **14**, 153-165 (2013). https://doi.org/10.1038/nrm3525

40 Bradley, R. K. & Anczukow, O. RNA splicing dysregulation and the hallmarks of cancer. *Nat Rev Cancer* (2023). https://doi.org/10.1038/s41568-022-00541-7

41 Liao, K. C. & Garcia-Blanco, M. A. Role of Alternative Splicing in Regulating Host Response to Viral Infection. *Cells* **10** (2021). https://doi.org/10.3390/cells10071720

42 Wang, C. *et al.* Abnormal global alternative RNA splicing in COVID-19 patients. *PLoS Genet* **18**, e1010137 (2022). https://doi.org/10.1371/journal.pgen.1010137

43 Manet, E. *et al.* Modulation of alternative splicing during early infection of human primary B lymphocytes with Epstein-Barr virus (EBV): a novel function for the viral EBNA-LP protein. *Nucleic Acids Res* **49**, 10657-10676 (2021). https://doi.org/10.1093/nar/gkab787

44 Thompson, M. G. *et al.* Viral-induced alternative splicing of host genes promotes influenza replication. *Elife* **9** (2020). https://doi.org/10.7554/eLife.55500

45 Louadi, Z. *et al.* DIGGER: exploring the functional role of alternative splicing in protein interactions. *Nucleic Acids Res* **49**, D309-D318 (2021). https://doi.org/10.1093/nar/gkaa768

46 Neuenkirchen, N. *et al.* Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. *EMBO J* **34**, 1925-1941 (2015). https://doi.org/10.15252/embj.201490350

47 Shaw, D. J., Eggleton, P. & Young, P. J. Joining the dots: production, processing and targeting of U snRNP to nuclear bodies. *Biochim Biophys Acta* **1783**, 2137-2144 (2008). https://doi.org/10.1016/j.bbamcr.2008.07.025

48 Massenet, S., Pellizzoni, L., Paushkin, S., Mattaj, I. W. & Dreyfuss, G. The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* **22**, 6533-6541 (2002). https://doi.org/10.1128/MCB.22.18.6533-6541.2002

49 Gubitz, A. K., Feng, W. & Dreyfuss, G. The SMN complex. *Exp Cell Res* **296**, 51-56 (2004). https://doi.org/10.1016/j.yexcr.2004.03.022

50 Pellizzoni, L. Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* **8**, 340-345 (2007). https://doi.org/10.1038/sj.embor.7400941

51 Yong, J., Kasim, M., Bachorik, J. L., Wan, L. & Dreyfuss, G. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38**, 551-562 (2010). https://doi.org/10.1016/j.molcel.2010.03.014

52 Schlaen, R. G. *et al.* The spliceosome assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms. *Proc Natl Acad Sci U S A* **112**, 9382-9387 (2015). https://doi.org/10.1073/pnas.1504541112

53 Zhang, Z. *et al.* SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**, 585-600 (2008). https://doi.org/10.1016/j.cell.2008.03.031

54 Cartegni, L., Hastings, M. L., Calarco, J. A., de Stanchina, E. & Krainer, A. R. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am J Hum Genet* **78**, 63-77 (2006). https://doi.org/10.1086/498853

55 De Maio, F. A. *et al.* The Dengue Virus NS5 Protein Intrudes in the Cellular Spliceosome and Modulates Splicing. *PLoS Pathog* **12**, e1005841 (2016). https://doi.org/10.1371/journal.ppat.1005841

56 Wang, Z. & Burge, C. B. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802-813 (2008). https://doi.org/10.1261/rna.876308

57 Grosso, A. R. *et al.* Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res* **36**, 4823-4832 (2008). https://doi.org/10.1093/nar/gkn463

58 Artemyeva-Isman, O. V. & Porter, A. C. G. U5 snRNA Interactions With Exons Ensure Splicing Precision. *Front Genet* **12**, 676971 (2021). https://doi.org/10.3389/fgene.2021.676971

59 Jia, Y., Mu, J. C. & Ackerman, S. L. Mutation of a U2 snRNA gene causes global disruption of alternative splicing and neurodegeneration. *Cell* **148**, 296-308 (2012). https://doi.org/10.1016/j.cell.2011.11.057

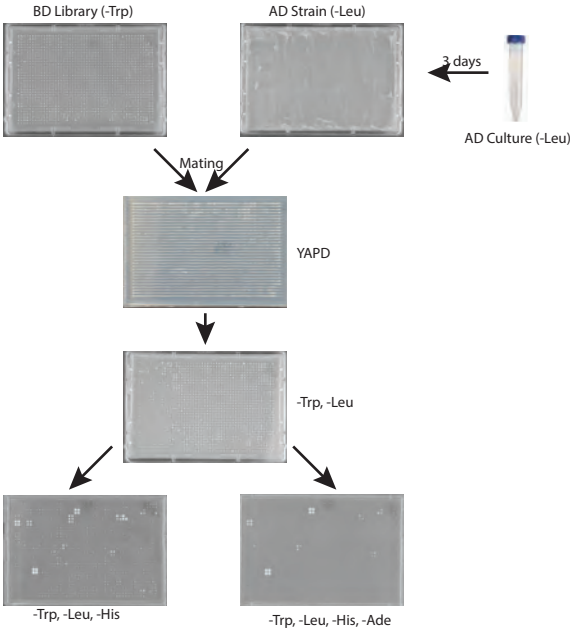
60 Ramakrishnan, L. Revisiting the role of the granuloma in tuberculosis. *Nat Rev Immunol* **12**, 352-366 (2012). https://doi.org/10.1038/nri3211

61 Simmons, J. D. *et al.* Differentially expressed transcript isoforms associate with resistance to tuberculin skin test and interferon gamma release assay conversion. *PLoS One* **18**, e0284498 (2023). https://doi.org/10.1371/journal.pone.0284498

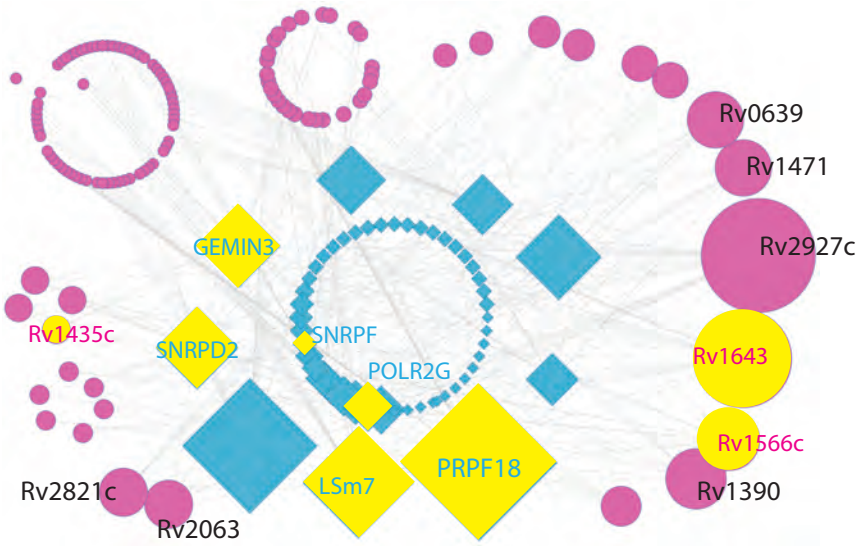
62 Mehdiratta, K. *et al.* Kupyaphores are zinc homeostatic metallophores required for colonization of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **119** (2022). https://doi.org/10.1073/pnas.2110293119

Figure 1

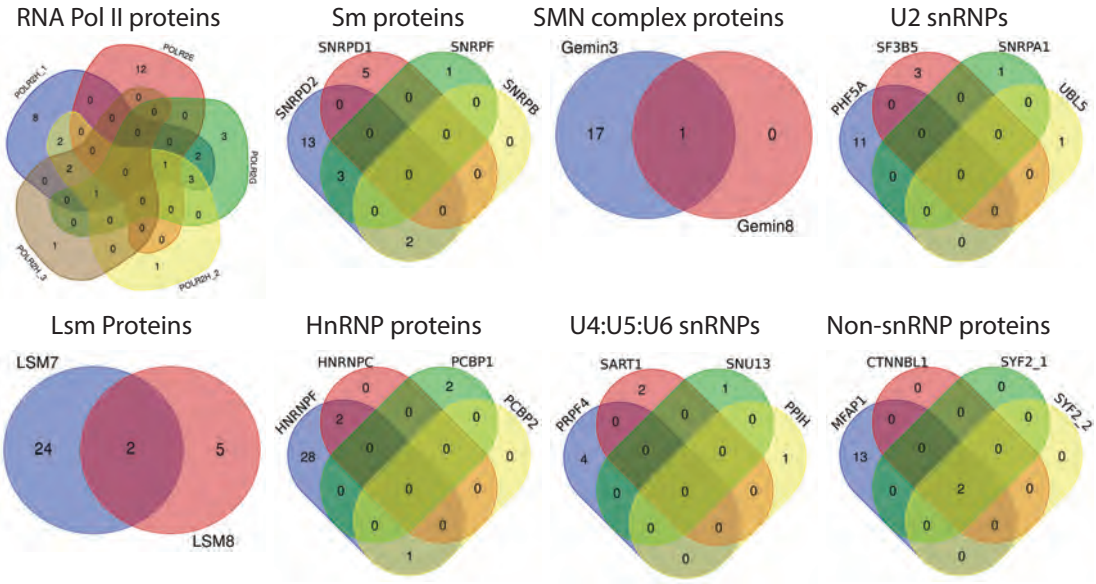
A



B



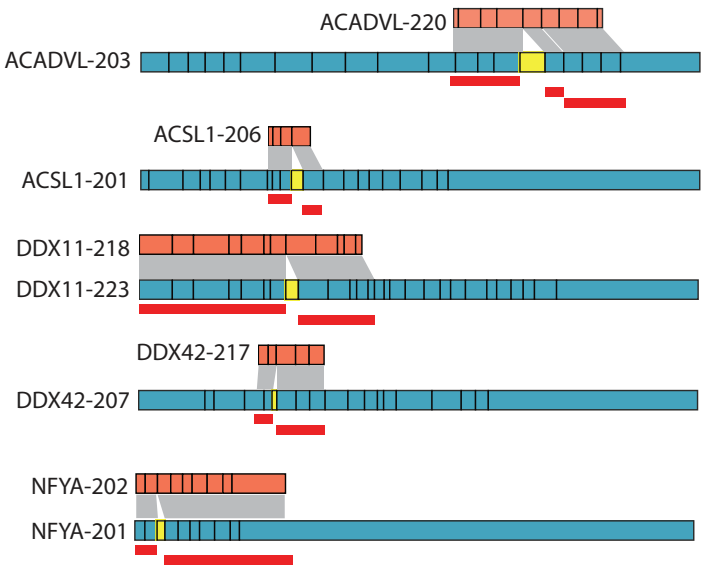
C



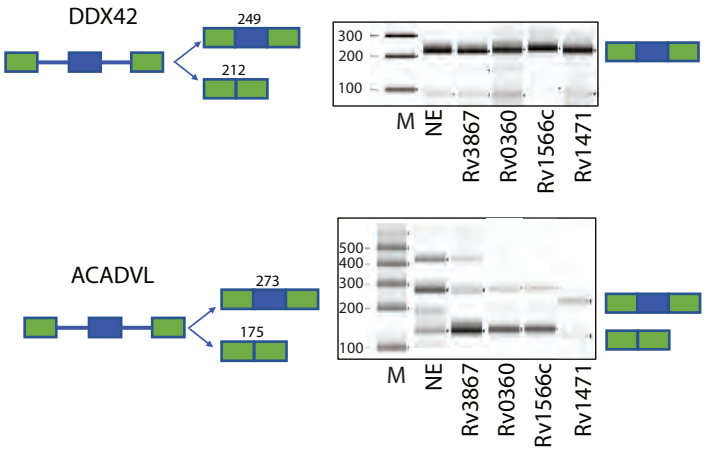
D

Functional Class	Host proteins	Interacting Mtb protein
Catalytic Step II factors	PRPF18, DHX38	Rv2927c, Rv1643
hnRNP	hnRNPF, hnRNPC	Rv1566c, Rv1471
	hnRNPF, PCBP2	Rv2927c
Non-SnRNP	MFAP1, SYF2, CTNNBL1	Rv2927c, Rv1643
RNA Pol II	POLR2E, POLR2G, POLR2H (1,2)	Rv1566c
	POLR2G, POLR2H (1,2,3)	Rv0639
	POLR2G, POLR2H (1,2)	Rv2063, Rv1390, Rv2927c
	POLR2H (1,2,3)	Rv1558, Rv1638A
	POLR2G POLR2H (1)	Rv1643, Rv0164
	POLR2H (1,2)	Rv1094, Rv2223c
	POLR2E POLR2H (4)	Rv2246
Sm Core and Sm like	LSm7 Lsm8	Rv1471,Rv1643
	SNRPD2 SNRPF	Rv1390, Rv2927c, Rv0170
	SNRPB, SNRPD2	Rv2891, Rv2821c
	LSm7 SNRPF	Rv0639
	LSm7 SNRPD2 SNRPF	Rv1390 Rv2927c
	LSm7 SNRPD2	Rv1566c
SMN Complex	Gemin3, Gemin8	Rv2927c

E



F



542 **Figure 1: Mycobacterial secretory proteins and host spliceosome proteins interactome.**

543 **A)** The schematic flow chart for high-throughput one-to-one Y2H assay in 1536-well
544 format. Protein pairs showing interaction in 3 or 4 biological replicates were considered
545 as positive interactions.

546 **B)** Mtb proteins -host splicing factor interactome- the positive interactions from the Y2H
547 screen were used to generate the network shown here. The diamond and circle represent
548 the host and Mtb proteins, respectively, and the size corresponds to the number of
549 interactions observed. The proteins coloured yellow in the network are discussed in
550 detail in the text.

551 **C)** The Venn diagrams show proteins belonging to the major spliceosome components and
552 their corresponding overlap with the interacting Mtb proteins.

553 **D)** A more descriptive representation of the Venn diagrams, shown above in Fig. 1C is
554 presented in the table as a list of interactions.

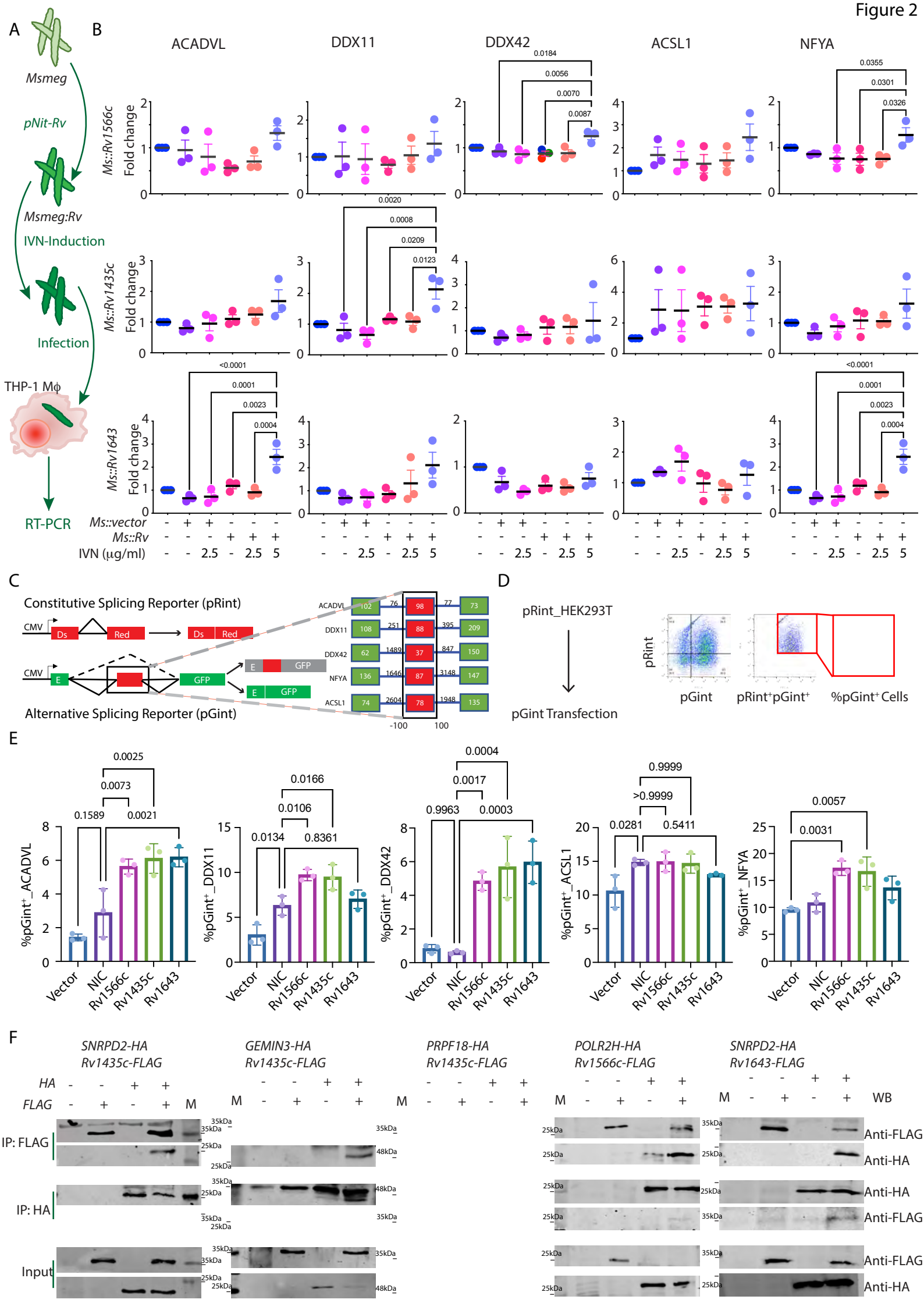
555 **E)** Transcripts that were used for making the minigene constructs to test the role of Mtb
556 proteins in host RNA splicing. Only exons are shown, and two different transcripts for
557 each of the five genes are shown. The red bars indicate the overlapping sequences, and
558 the skipped exons are shown in yellow.

559 **F)** On the left, a schematic of the minigene cassettes for DDX42 and ACADVL is shown.
560 The constitutive exons are shown in green, while the alternative exons are shown in
561 blue. *In vitro* transcribed products of minigenes were used for in-vitro splicing assays
562 in the presence of purified Mtb proteins Rv3867, Rv0360, Rv1566c and Rv1471.

563

564

Figure 2


A
Msmeg
pNit-Rv
Msmeg:Rv
IVN-Induction
Infection
THP-1 Mϕ
RT-PCR
B
ACADVL
DDX11
DDX42
ACSL1
NFYA
Ms::Rv1566c
Ms::Rv1435c
Ms::Rv1643
Fold change
Ms::vector
Ms::Rv
IVN (µg/ml)
C
Constitutive Splicing Reporter (pRint)
CMV
Ds
Red
E
GFP
Alternative Splicing Reporter (pGint)
ACADVL
DDX11
DDX42
NFYA
ACSL1
D
pRint_HEK293T
pGint Transfection
pRint
pGint
pRint+pGint+
%pGint+ Cells
E
%pGint+_ACADVL
%pGint+_DDX11
%pGint+_DDX42
%pGint+_ACSL1
%pGint+_NFYA
Vector
NIC
Rv1566c
Rv1435c
Rv1643
F
SNRPD2-HA
Rv1435c-FLAG
GEMIN3-HA
Rv1435c-FLAG
PRPF18-HA
Rv1435c-FLAG
POLR2H-HA
Rv1566c-FLAG
SNRPD2-HA
Rv1643-FLAG
HA
FLAG
M
WB
IP: FLAG
IP: HA
Input
Anti-FLAG
Anti-HA

565 **Figure 2: *Ex vivo* infection assays and reporter assays to test the role of Mtb proteins in**
566 **host RNA splicing**

567 **A)** The schematic representation of infecting THP-1 macrophages with *M. smegmatis*
568 strains that express mycobacterial proteins under IVN inducible promotor. Isoform-
569 specific qPCR for the truncated isoforms of ACADVL, DDX11, DDX42, ACSL1 and
570 NFYA was used to assess the impact on alternative splicing.

571 **B)** Expression of spliced isoforms of the five target genes (as indicated) in THP-1
572 macrophages upon infection with WT or *Rv1566c, Rv1435c* and *Rv1643* knock-in *M*
573 *smegmatis* lines. *M smegmatis* lines were induced with IVN (indicated dose) for 12
574 hours before infection. Data from three independent biological experiments. The p-
575 values shown are from One-way ANOVA.

576 **C)** Designing of the custom *in vivo* fluorescent splicing reporter. Schematic for the
577 constitutive splicing reporter pRint with a short adenovirus intron intervening the
578 DsRed gene with a strong 5' and 3' splice sites. In the pGint splicing constructs, the
579 alternative exon of ACADVL, DDX11, DDX42, ACSL1 and NFYA, along with the
580 flanking 100bp intronic sequences, were inserted within eGFP reading frame. The exon
581 disrupts the reading frame, however, skipping of the test exon reconstitutes EGFP
582 fluorescence.

583 **D)** The pRint expressing stable HEK293T cell lines were co-transfected with a pGint
584 construct and pDest-FLAG construct encoding proteins Rv1252c (NIC), Rv1566c,
585 Rv1435c, Rv1643, or empty vector control, followed by Flow cytometry. In the $DsRed^+$
586 gate, $\%FITC^+$ population was calculated to measure the extent of the test exon-skipping
587 event.

588 **E)** $\%FITC^+$ population for each of the indicated reporter constructs in the presence of
589 indicated mycobacterial protein, non-interacting control or vector controls. Each dot
590 represents an independent biological replicate (N=3), indicating p values from One-
591 way ANOVA.

592 **F)** Validation of physical interactions between indicated pairs of *Mtb* proteins (FLAG-
593 tagged) and host proteins (HA-tagged) in HEK293T cells. Data is representative of two
594 or more biologically independent transfection and pulldown experiments.

595

596

Figure 3

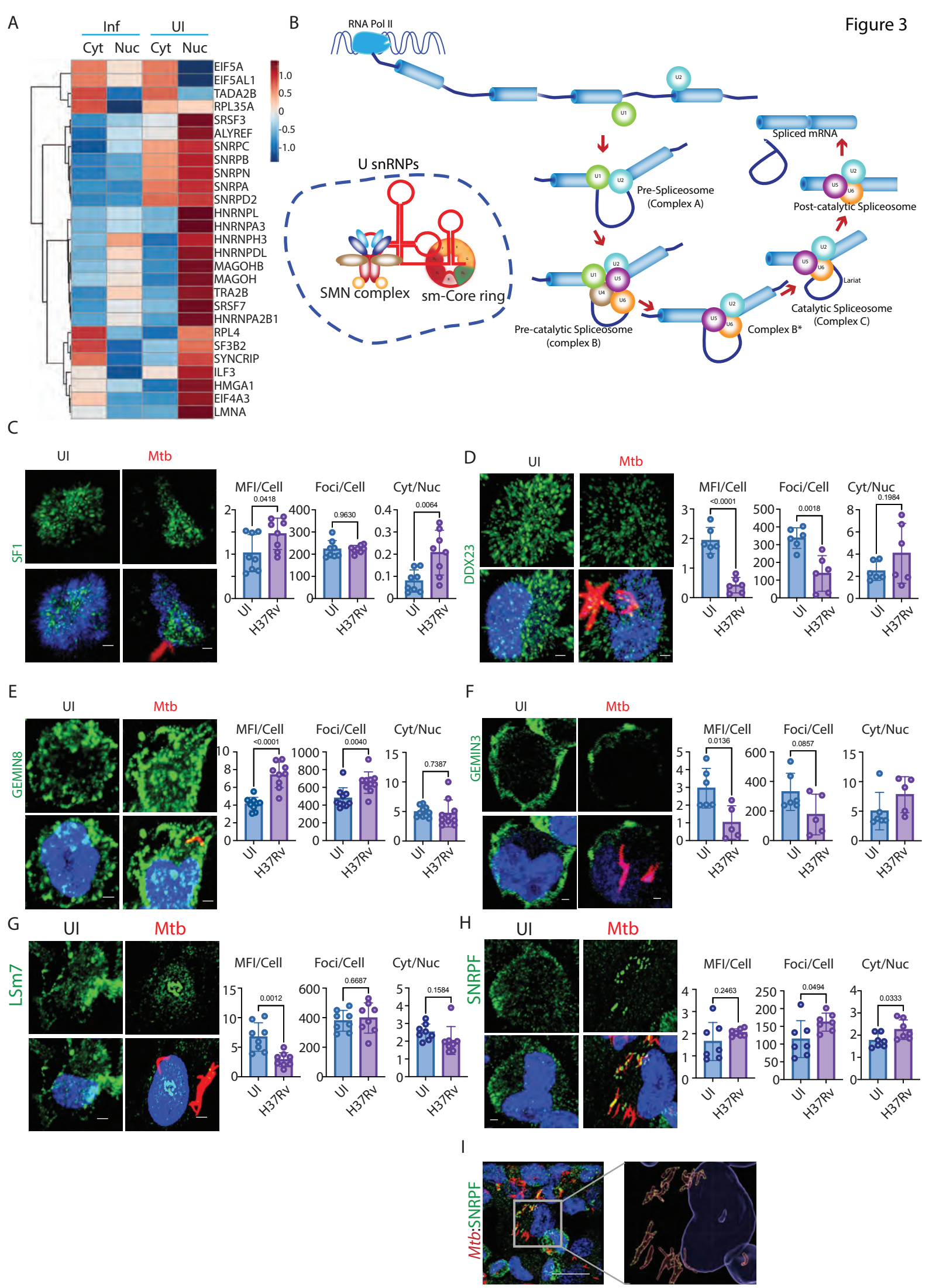

A
Inf
UI
Cyt
Nuc
Cyt
Nuc
EIF5A
EIF5AL1
TADA2B
RPL35A
SRSF3
ALYREF
SNRPC
SNRPB
SNRPN
SNRPA
SNRPD2
HNRNPL
HNRNPA3
HNRNPH3
HNRNPDL
MAGOHB
MAGOH
TRA2B
SRSF7
HNRNPA2B1
RPL4
SF3B2
SYNCRIP
ILF3
HMGA1
EIF4A3
LMNA
1.0
0.5
0
-0.5
-1.0
B
RNA Pol II
U snRNPs
SMN complex
sm-Core ring
Pre-Spliceosome
(Complex A)
Pre-catalytic Spliceosome
(complex B)
Complex B*
Catalytic Spliceosome
(Complex C)
Lariat
Post-catalytic Spliceosome
Spliced mRNA
C
UI
Mtb
SF1
MFI/Cell
Foci/Cell
Cyt/Nuc
0.0418
0.9630
0.0064
D
DDX23
<0.0001
0.0018
0.1984
E
GEMIN8
<0.0001
0.0040
0.7387
F
GEMIN3
0.0136
0.0857
G
LSm7
0.0012
0.6687
0.1584
H
SNRPF
0.2463
0.0494
0.0333
I
Mtb:SNRPF

Figure 3: Dysregulation of host RNA splicing machinery in Mtb-infected macrophages

A) Heatmap shows massive dysregulation in the level and cellular distribution of human proteins involved in RNA splicing and identified in MS of the spliceosome complex pulldown from the Mtb infected THP-1 macrophages 48 hours post-infection.

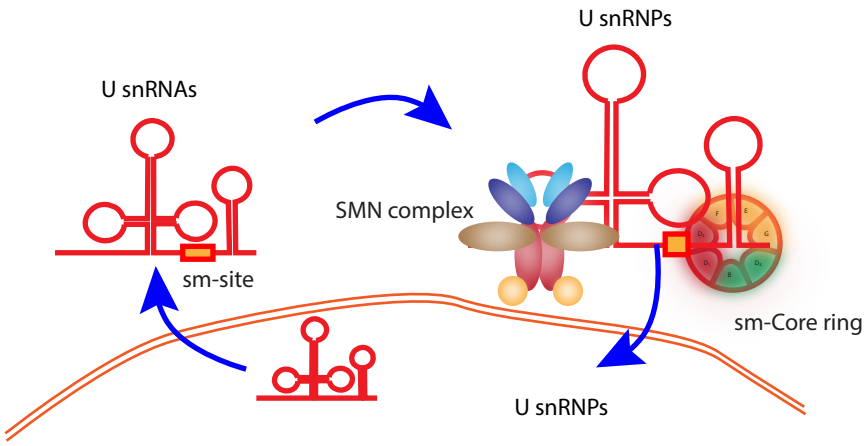
B) A schematic model depicting the ordered pathway of mRNA transcription by RNAP II and splicing catalyzed by spliceosome complex. The U1snRNP binds the 5' splice site and U2 snRNP binds the branch point adenine to form complex A. U4/U5.U6 snRNP binds to form complex B and causes conformational changes to form activated complex B. This is followed by the first transesterification reaction releasing U1 and U4 snRNP and the second transesterification reaction forming complex C releasing the lariat intron and U2, U5 and U6 snRNP. The U snRNP are the major players in splicing catalysis and their biogenesis is mediated by the SMN complex and the Sm proteins in the cytoplasm.

C-H) The proteins were visualized by indirect immunofluorescence (green), the nucleus was visualized by DAPI staining (blue) and mCherry-expressing bacteria (red) was used for infection. Each dot in the bar graphs represents a field analyzed, and an average of a hundred cells were quantified. The confocal images were analyzed for three different parameters- overall MFI (MFI/cell), number of Foci (Foci/cell) and cytosolic to the nuclear ratio (Cyt/Nuc) for all major splicing-related proteins upon Mtb infection such as **C)** U1 associated proteins (SF1), **D)** U5 associated proteins (DDX23), **E)** SMN complex proteins (GEMIN8) **F)** SMN complex proteins (GEMIN3) and **G)** Sm Core proteins (LSM7) and **H)** Sm Core proteins (SNRPF)

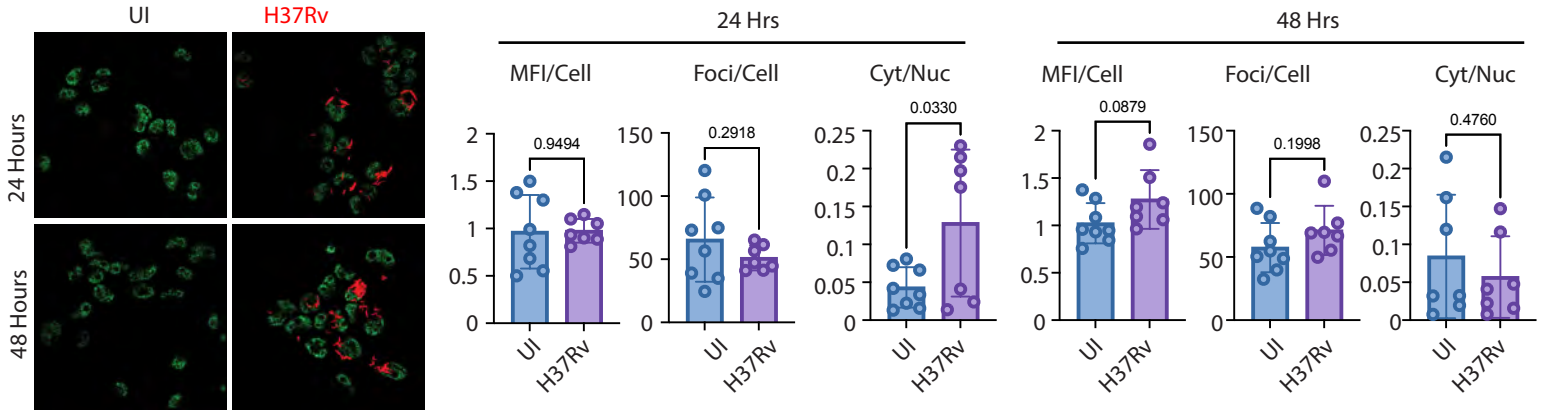
I) Image showing SNRPF colocalizing with bacterial phagosomes. The 3D reconstruction of the nucleus, SNRPF foci and bacteria (see methods) was done to show the SNRPF Foci at <0.2 microns distance from the bacterial phagosomes. Scale bar: 2 μm and p values indicate One-way ANOVA.

Figure 4

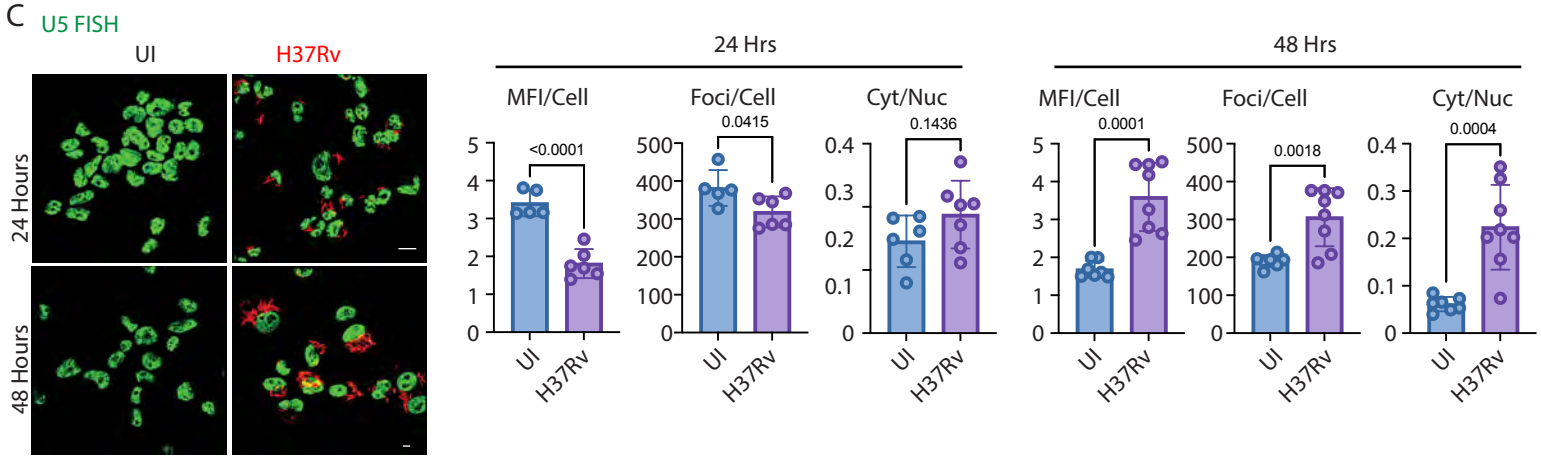
A



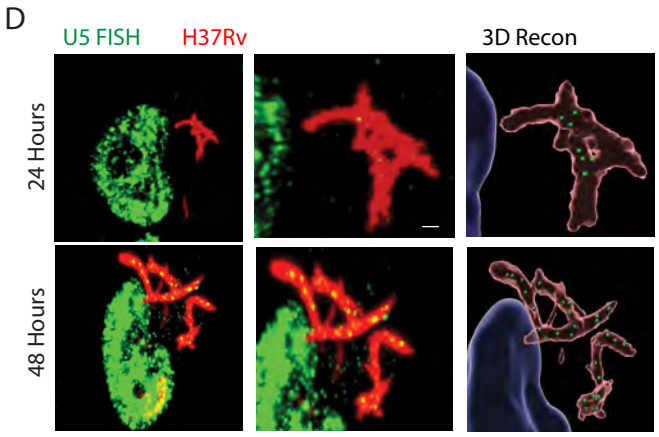
B



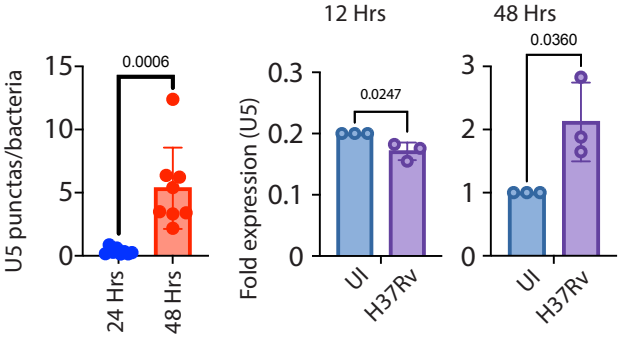
C



D



E



F

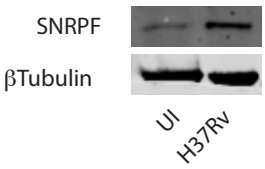


Figure 4: Dysregulation in SMN complex and Sm proteins impact U snRNA localization and overall assembly of U snRNPs.

A) Schematic representation of U snRNA nucleo-cytoplasmic trajectory and assembly of Sm core ring on snRNA by SMN complex to form snRNPs

B) Representative images for RNA FISH against U2 snRNA in THP-1 macrophages upon Mtb infection 24- and 48-hrs post-infection. The U2 snRNA probe was labelled with Alexa 488 (green), the nucleus was stained with DAPI (blue) and mCherry-expressing H37Rv (red) was used for infection. The plots corresponding to overall MFI (MFI/cell), number of Foci (Foci/cell) and cytosolic to nuclear ratio (Cyt/Nuc) of U2 snRNA are shown here. Each dot represents a field analyzed, and an average of a hundred cells are quantified. The p-values is from one-way ANOVA.

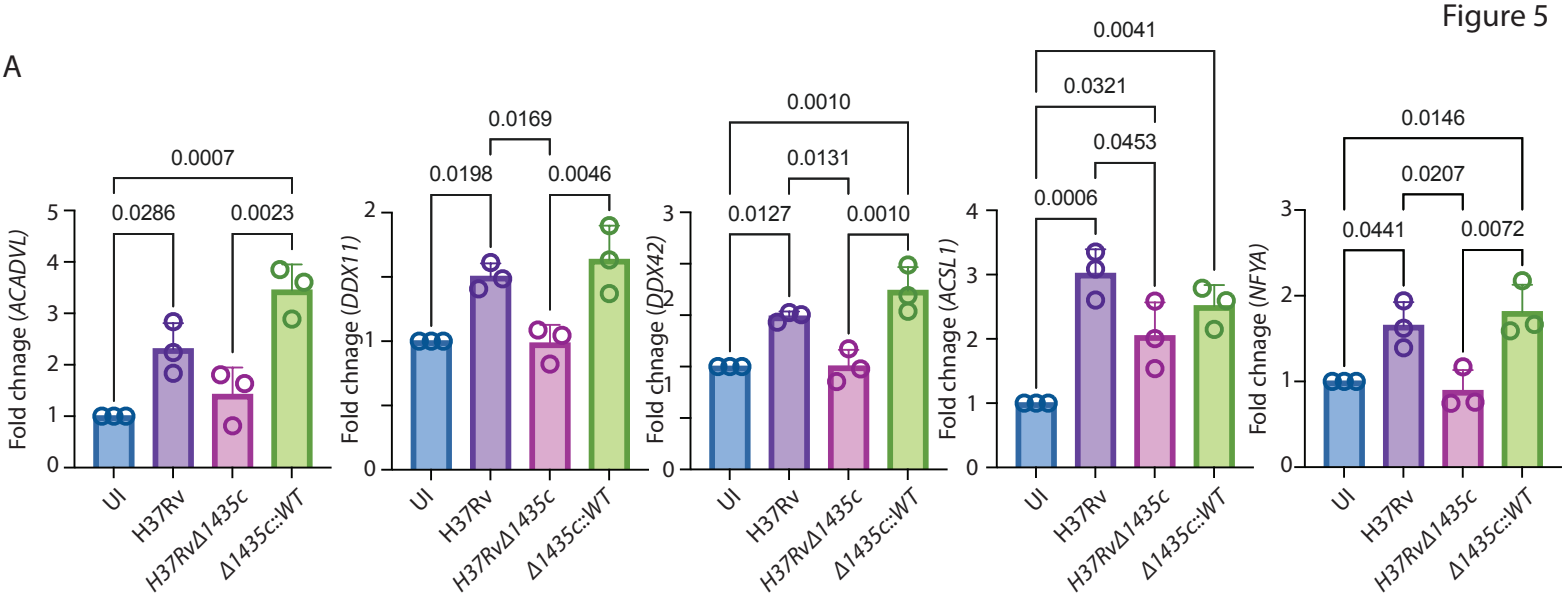
C) Representative images for RNA FISH against U5 snRNA in THP-1 macrophages upon Mtb infection 24- and 48-hrs post-infection. The U5 snRNA probe was labelled with Alexa 488 (green), the nucleus was stained with DAPI (blue) and mCherry-expressing H37Rv (red) was used for infection. The plots corresponding to overall MFI (MFI/cell), number of Foci (Foci/cell) and cytosolic to nuclear ratio (Cyt/Nuc) of U5 snRNA are shown here. Each dot represents a field analyzed, an average of a hundred cells are quantified, and p values are from one-way ANOVA.

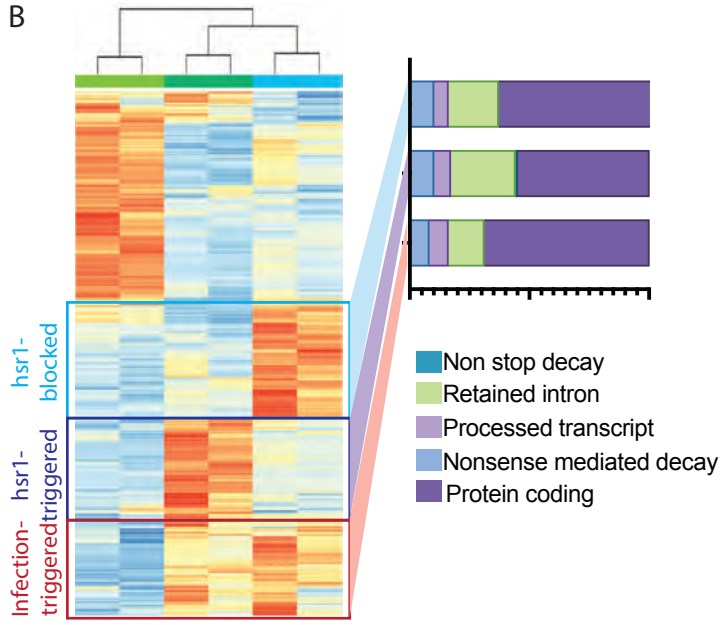
D) Direct interaction between H37Rv (red) and U5snRNA at 24- and 48-hrs post-infection. For the plot at the right, the number of U5snRNA puncta within 0.2 microns from H37Rv surface was counted, p values from one-way ANOVA. Representative images for colocalization of Sm core protein SNRPF (green) to the H37Rv phagosome (red). The 3D reconstruction and spot analysis are done as previously described.

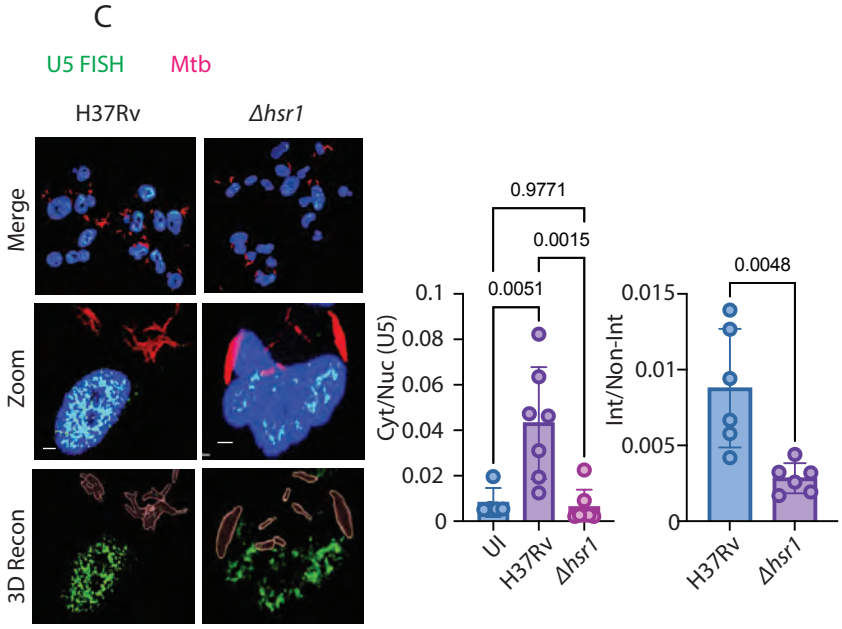
E) Expression of U5 snRNA in THP-1 macrophages, 12 and 48 hpi with H37Rv. Data from three independent biological experiments.

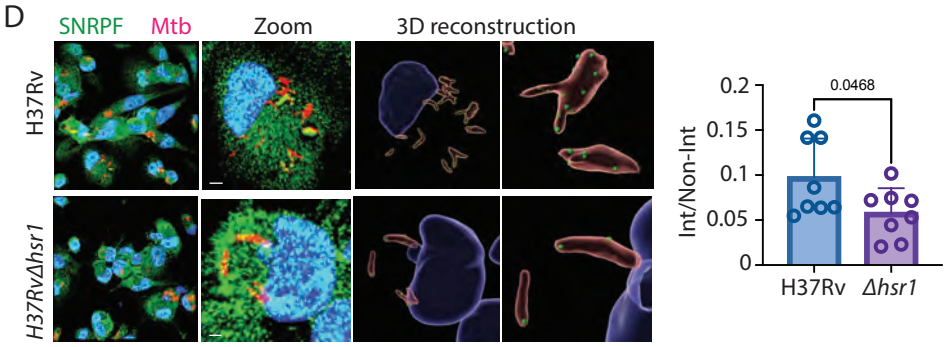
F) Immunoblot showing SNRPF protein expression in the whole cell lysates of uninfected (UI) and H37Rv-infected THP-1 macrophages, 48 hpi. Data is representative of two independent biological experiments.

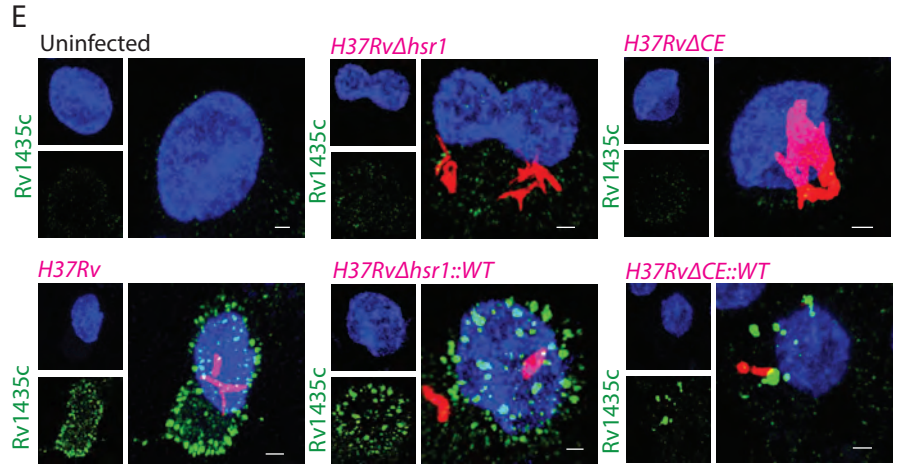
Figure 5


A
Fold chnage (ACADVL)
0.0007
0.0286
0.0023
Fold chnage (DDX11)
0.0169
0.0198
0.0046
Fold chnage (DDX42)
0.0010
0.0131
0.0127
0.0010
Fold chnage (ACSL1)
0.0041
0.0321
0.0453
0.0006
Fold chnage (NFYA)
0.0146
0.0207
0.0441
0.0072
UI
H37Rv
H37RvΔ1435c
Δ1435c::WT


B
hsr1-blocked
hsr1-triggered
Infection-triggered
Non stop decay
Retained intron
Processed transcript
Nonsense mediated decay
Protein coding


C
U5 FISH
Mtb
H37Rv
Δhsr1
Merge
Zoom
3D Recon
Cyt/Nuc (U5)
0.9771
0.0015
0.0051
Int/Non-Int
0.0048
UI
H37Rv
Δhsr1


D
SNRPF
Mtb
Zoom
3D reconstruction
H37Rv
H37RvΔhsr1
Int/Non-Int
0.0468
H37Rv
Δhsr1


E
Uninfected
H37RvΔhsr1
H37RvΔCE
H37Rv
H37RvΔhsr1::WT
H37RvΔCE::WT
Rv1435c

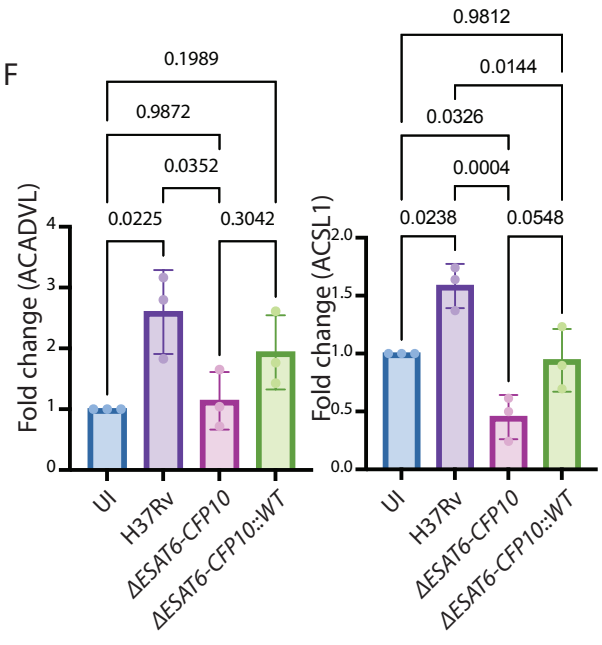

F
Fold change (ACADVL)
0.1989
0.9872
0.0352
0.0225
0.3042
Fold change (ACSL1)
0.9812
0.0144
0.0326
0.0004
0.0238
0.0548
UI
H37Rv
ΔESAT6-CFP10
ΔESAT6-CFP10::WT

Figure 5: Role of H37Rv hsr1 protein in snRNP biogenesis and altering host RNA splicing.

A) Q-PCR for the target transcripts as mentioned after infection with H37Rv, H37Rv*Δhsr1* and H37Rv*Δhsr1::WT* strain. Data from three biological replicates (N=3), *p<0.05, **p<0.01 (one-way ANOVA).

B) Clustergram of RNAseq data obtained from uninfected, or THP1 macrophages infected with H37Rv or H37Rv*Δhsr1* strains. Transcripts showing an infection-dependent increase (red box), hsr1-triggered increase (purple) or hsr1-blocked transcripts (turquoise) are shown. The distribution of different transcript biotypes across the three groups is shown in the aligned plot towards the right.

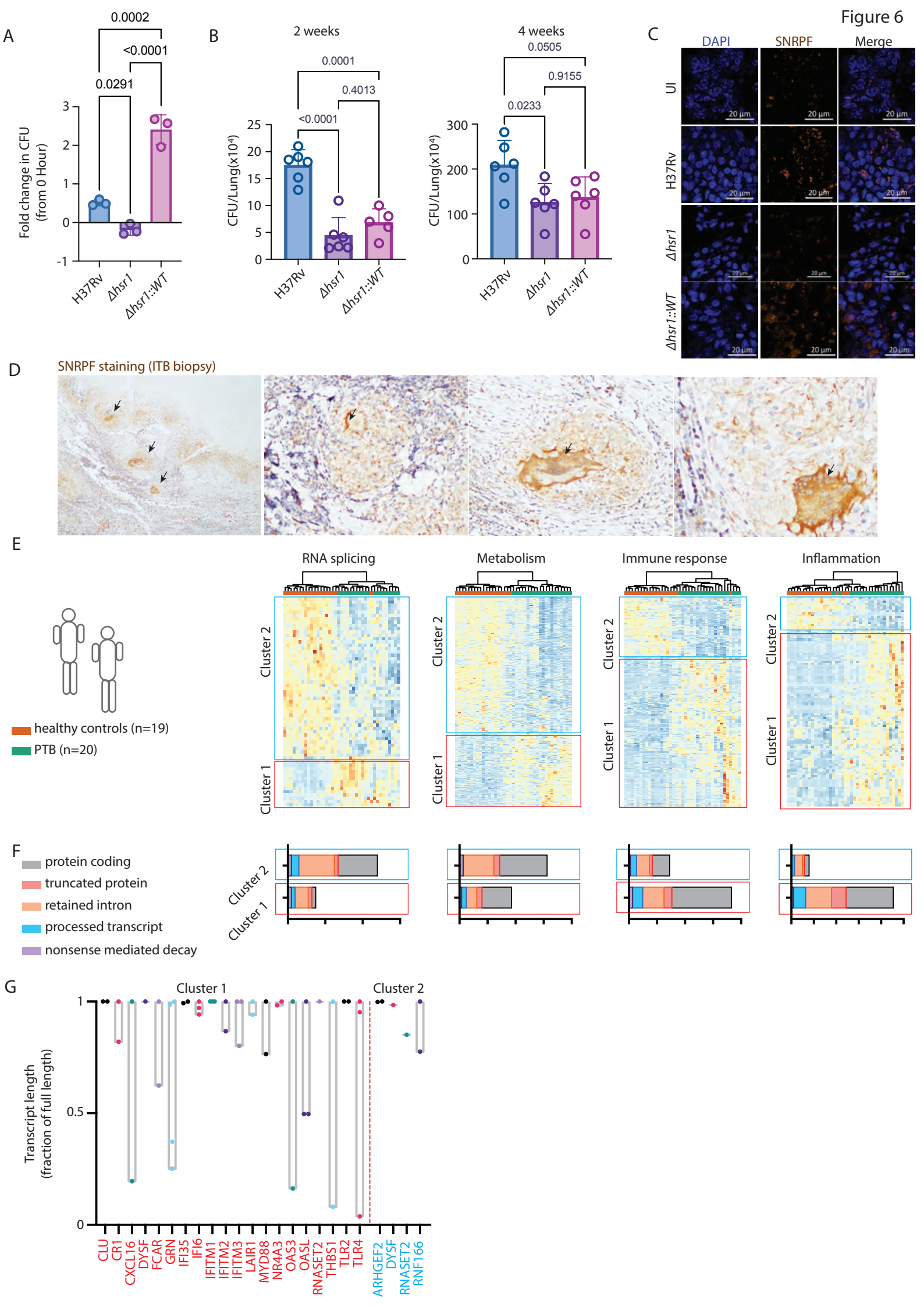
C) Representative image for the U5snRNA FISH in THP-1 macrophages infected with WT H37Rv or H37Rv*Δhsr1* strains. The plots at the right show the cytosolic to nuclear ratio and the ratio of interacting to non-interacting U5snRNA with WT H37Rv or H37Rv*Δhsr1*. **p<0.01 (one-way ANOVA).

D) Representative images of SNRPF staining in THP-1 macrophages infected with WT H37Rv or H37Rv*Δhsr1* strains. For the plot at the right, the number of SNRPF puncta within 0.2 microns from the H37Rv surface was counted as interacting and shows the ratio of interacting to non-interacting SNRPF with WT H37Rv or H37Rv*Δhsr1*. *p<0.05 (one-way ANOVA).

E) Representative images of THP1-macrophages stained with DAPI (nucleus) and Rv1435c antisera (green) upon infection with WT H37Rv, H37Rv*Δhsr1,* H37Rv*ΔCE* and H37Rv*ΔCE::WT* and H37Rv*Δhsr1::WT.* Images are representative of two independent biological experiments. Scale bar-2 μm.

F) Expression of truncated isoform as indicated upon infection with WT H37Rv, H37Rv*ΔCE* and H37Rv*ΔCE::WT* at 48hpi from three biologically independent experiments. The p values are from One-way ANOVA.

Figure 6


A
0.0002
<0.0001
0.0291
Fold change in CFU (from 0 Hour)
H37Rv
Δhsr1
Δhsr1::WT
B
2 weeks
0.0001
0.4013
<0.0001
CFU/Lung(x10⁴)
4 weeks
0.0505
0.9155
0.0233
C
DAPI
SNRPF
Merge
UI
H37Rv
Δhsr1
Δhsr1::WT
20 µm
D
SNRPF staining (ITB biopsy)
E
RNA splicing
Metabolism
Immune response
Inflammation
Cluster 2
Cluster 1
healthy controls (n=19)
PTB (n=20)
F
protein coding
truncated protein
retained intron
processed transcript
nonsense mediated decay
G
Cluster 1
Cluster 2
Transcript length (fraction of full length)
CLU
CR1
CXCL16
DYSF
FCAR
GRN
IFI35
IFI6
IFITM1
IFITM2
IFITM3
LAIR1
MYD88
NR4A3
OAS3
OASL
RNASET2
THBS1
TLR2
TLR4
ARHGEF2
DYSF
RNASET2
RNF166

683 **Figure 6: Regulation of host alternative splicing as a virulence mechanism**

684 A) THP-1 macrophages were infected with WT H37Rv, H37Rv*Δhsr1,* H37Rv*Δhsr1::WT*
685 strains. At 48 hours post-infection, bacterial load was examined using CFU plating.
686 Data represent fold change in CFU at 48 hrs compared to 0 hr post-infection. (N=3), p
687 values are from one-way ANOVA.

688 B) BALB/C mice (6-8 weeks old) were infected with 200 CFU of WT H37Rv,
689 H37Rv*Δhsr1,* H37Rv*Δhsr1::WT* in separate experiments. At 2- and 4 weeks post-
690 infection, bacterial load in the lungs of the animals were determined by CFU plating. P
691 values are from one-way ANOVA.

692 C) The 4 week post-infection lung tissue sections of WT H37Rv, H37Rv*Δhsr1,*
693 H37Rv*Δhsr1::WT* were stained with DAPI (nucleus) and SNRPF. Scale bar: 20 µm

694 D) Human intestinal tuberculosis granulomatous lesions were stained with SNRPF. The
695 arrows indicate multinucleated giant cells with intense staining for cytosolic SNRPF.
696 Black arrows show areas of intense SNRPF staining in the granulomas and in the giant
697 cells.

698 **E)** RNAseq data from the monocytes of pulmonary TB cases (N=20) and healthy contacts
699 (N=19) was analyzed for expression of differential alternatively spliced transcripts.
700 Genes belonging to the indicated functional classes were selected for the cluster
701 analysis. Cluster 1 in each case shows transcripts upregulated in PTB cases.

702 F) For each of the functional classes shown in Fig. 6D, the relative distribution of transcript
703 biotypes (as indicated) is plotted and binned separately for those that were up- (cluster
704 1, red) or down- (cluster 2, blue) regulated in PTB cases.

705 G) The transcripts from the list of genes in the immune response class (Fig. 6D) were
706 further analyzed for their length with respect to the longest encoded isoform of the
707 corresponding gene. Only genes with more than one transcript in the list were chosen
708 for this analysis. The red dotted line separates cluster 1 and cluster 2 transcripts.
709 Transcripts of DYSF and RNASET2 are present in both clusters.

710

711

712

713 **References**

714

715 1 in *WHO consolidated guidelines on tuberculosis: Module 3: diagnosis - rapid*
716 *diagnostics for tuberculosis detection WHO Guidelines Approved by the Guidelines*
717 *Review Committee* (2021).
718 2 Pisu, D., Huang, L., Grenier, J. K. & Russell, D. G. Dual RNA-Seq of Mtb-Infected
719 Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions.
720 *Cell Rep* **30**, 335-350 e334 (2020). https://doi.org/10.1016/j.celrep.2019.12.033
721 3 Matteucci, K. C., Correa, A. A. S. & Costa, D. L. Recent Advances in Host-Directed
722 Therapies for Tuberculosis and Malaria. *Front Cell Infect Microbiol* **12**, 905278
723 (2022). https://doi.org/10.3389/fcimb.2022.905278
724 4 Young, C., Walzl, G. & Du Plessis, N. Therapeutic host-directed strategies to improve
725 outcome in tuberculosis. *Mucosal Immunol* **13**, 190-204 (2020).
726 https://doi.org/10.1038/s41385-019-0226-5
727 5 Kumar, D. & Rao, K. V. Regulation between survival, persistence, and elimination of
728 intracellular mycobacteria: a nested equilibrium of delicate balances. *Microbes Infect*
729 **13**, 121-133 (2011). https://doi.org/10.1016/j.micinf.2010.10.009
730 6 Karim, A. F. *et al.* Express path analysis identifies a tyrosine kinase Src-centric
731 network regulating divergent host responses to Mycobacterium tuberculosis infection.
732 *J Biol Chem* **286**, 40307-40319 (2011). https://doi.org/10.1074/jbc.M111.266239
733 7 Khan, N. *et al.* M. tuberculosis Reprograms Hematopoietic Stem Cells to Limit
734 Myelopoiesis and Impair Trained Immunity. *Cell* **183**, 752-770 e722 (2020).
735 https://doi.org/10.1016/j.cell.2020.09.062
736 8 Kalam, H., Singh, K., Chauhan, K., Fontana, M. F. & Kumar, D. Alternate splicing of
737 transcripts upon Mycobacterium tuberculosis infection impacts the expression of
738 functional protein domains. *IUBMB Life* **70**, 845-854 (2018).
739 https://doi.org/10.1002/iub.1887
740 9 Chauhan, K., Kalam, H., Dutt, R. & Kumar, D. RNA Splicing: A New Paradigm in
741 Host-Pathogen Interactions. *J Mol Biol* **431**, 1565-1575 (2019).
742 https://doi.org/10.1016/j.jmb.2019.03.001
743 10 Liu, Y. *et al.* Impact of Alternative Splicing on the Human Proteome. *Cell Rep* **20**,
744 1229-1241 (2017). https://doi.org/10.1016/j.celrep.2017.07.025
745 11 Kalam, H., Fontana, M. F. & Kumar, D. Alternate splicing of transcripts shape
746 macrophage response to Mycobacterium tuberculosis infection. *PLoS Pathog* **13**,
747 e1006236 (2017). https://doi.org/10.1371/journal.ppat.1006236
748 12 Zhang, X. O., Fu, Y., Mou, H., Xue, W. & Weng, Z. The temporal landscape of
749 recursive splicing during Pol II transcription elongation in human cells. *PLoS Genet*
750 **14**, e1007579 (2018). https://doi.org/10.1371/journal.pgen.1007579
751 13 Hug, N., Longman, D. & Caceres, J. F. Mechanism and regulation of the nonsense-
752 mediated decay pathway. *Nucleic Acids Res* **44**, 1483-1495 (2016).
753 https://doi.org/10.1093/nar/gkw010
754 14 Pai, A. A. *et al.* Widespread Shortening of 3' Untranslated Regions and Increased
755 Exon Inclusion Are Evolutionarily Conserved Features of Innate Immune Responses
756 to Infection. *PLoS Genet* **12**, e1006338 (2016).
757 https://doi.org/10.1371/journal.pgen.1006338
758 15 Bhatt, D. M. *et al.* Transcript dynamics of proinflammatory genes revealed by
759 sequence analysis of subcellular RNA fractions. *Cell* **150**, 279-290 (2012).
760 https://doi.org/10.1016/j.cell.2012.05.043

16 Pandya-Jones, A. *et al.* Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA* **19**, 811-827 (2013). https://doi.org/10.1261/rna.039081.113

17 Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718 (2009). https://doi.org/10.1016/j.cell.2009.02.009

18 Wilkinson, M. E., Charenton, C. & Nagai, K. RNA Splicing by the Spliceosome. *Annu Rev Biochem* **89**, 359-388 (2020). https://doi.org/10.1146/annurev-biochem-091719-064225

19 Papasaikas, P., Tejedor, J. R., Vigevani, L. & Valcarcel, J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* **57**, 7-22 (2015). https://doi.org/10.1016/j.molcel.2014.10.030

20 Zhang, R. *et al.* Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**, 384-395 (2011). https://doi.org/10.1016/j.cell.2011.06.043

21 Karijolich, J. & Yu, Y. T. Spliceosomal snRNA modifications and their function. *RNA Biol* **7**, 192-204 (2010). https://doi.org/10.4161/rna.7.2.11207

22 Strasser, A., Dickmanns, A., Luhrmann, R. & Ficner, R. Structural basis for m3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1. *EMBO J* **24**, 2235-2243 (2005). https://doi.org/10.1038/sj.emboj.7600701

23 Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem* **84**, 291-323 (2015). https://doi.org/10.1146/annurev-biochem-060614-034316

24 Penn, B. H. *et al.* An Mtb-Human Protein-Protein Interaction Map Identifies a Switch between Host Antiviral and Antibacterial Responses. *Mol Cell* **71**, 637-648 e635 (2018). https://doi.org/10.1016/j.molcel.2018.07.010

25 Portal-Celhay, C. *et al.* Mycobacterium tuberculosis EsxH inhibits ESCRT-dependent CD4(+) T-cell activation. *Nat Microbiol* **2**, 16232 (2016). https://doi.org/10.1038/nmicrobiol.2016.232

26 Sharma, G., Upadhyay, S., Srilalitha, M., Nandicoori, V. K. & Khosla, S. The interaction of mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG methylation and histone H3/H4 binding. *Nucleic Acids Res* **43**, 3922-3937 (2015). https://doi.org/10.1093/nar/gkv261

27 Mehra, A. *et al.* Mycobacterium tuberculosis type VII secreted effector EsxH targets host ESCRT to impair trafficking. *PLoS Pathog* **9**, e1003734 (2013). https://doi.org/10.1371/journal.ppat.1003734

28 de Souza, G. A., Leversen, N. A., Malen, H. & Wiker, H. G. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics* **75**, 502-510 (2011). https://doi.org/10.1016/j.jprot.2011.08.016

29 Malen, H., Berven, F. S., Fladmark, K. E. & Wiker, H. G. Comprehensive analysis of exported proteins from Mycobacterium tuberculosis H37Rv. *Proteomics* **7**, 1702-1718 (2007). https://doi.org/10.1002/pmic.200600853

30 Yaseen, I., Kaur, P., Nandicoori, V. K. & Khosla, S. Mycobacteria modulate host epigenetic machinery by Rv1988 methylation of a non-tail arginine of histone H3. *Nat Commun* **6**, 8922 (2015). https://doi.org/10.1038/ncomms9922

31 Malakar, B. *et al.* Phosphorylation of CFP10 modulates Mycobacterium tuberculosis virulence. *mBio* **14**, e0123223 (2023). https://doi.org/10.1128/mbio.01232-23

32 Bonano, V. I., Oltean, S. & Garcia-Blanco, M. A. A protocol for imaging alternative splicing regulation in vivo using fluorescence reporters in transgenic mice. *Nat Protoc* **2**, 2166-2181 (2007). https://doi.org/10.1038/nprot.2007.292

811 33 Herzel, L., Ottoz, D. S. M., Alpert, T. & Neugebauer, K. M. Splicing and
812 transcription touch base: co-transcriptional spliceosome assembly and function. *Nat*
813 *Rev Mol Cell Biol* **18**, 637-650 (2017). https://doi.org/10.1038/nrm.2017.63
814 34 Oesterreich, F. C. *et al.* Splicing of Nascent RNA Coincides with Intron Exit from
815 RNA Polymerase II. *Cell* **165**, 372-381 (2016).
816 https://doi.org/10.1016/j.cell.2016.02.045
817 35 Becker, D. *et al.* Nuclear Pre-snRNA Export Is an Essential Quality Assurance
818 Mechanism for Functional Spliceosomes. *Cell Rep* **27**, 3199-3214 e3193 (2019).
819 https://doi.org/10.1016/j.celrep.2019.05.031
820 36 Boudreault, S. *et al.* Reovirus mu2 protein modulates host cell alternative splicing by
821 reducing protein levels of U5 snRNP core components. *Nucleic Acids Res* **50**, 5263-
822 5281 (2022). https://doi.org/10.1093/nar/gkac272
823 37 van Kessel, J. C. & Hatfull, G. F. Recombineering in Mycobacterium tuberculosis.
824 *Nat Methods* **4**, 147-152 (2007). https://doi.org/10.1038/nmeth996
825 38 Del Rosario, R. C. H. *et al.* Histone acetylome-wide associations in immune cells
826 from individuals with active Mycobacterium tuberculosis infection. *Nat Microbiol* **7**,
827 312-326 (2022). https://doi.org/10.1038/s41564-021-01049-w
828 39 Kornblihtt, A. R. *et al.* Alternative splicing: a pivotal step between eukaryotic
829 transcription and translation. *Nat Rev Mol Cell Biol* **14**, 153-165 (2013).
830 https://doi.org/10.1038/nrm3525
831 40 Bradley, R. K. & Anczukow, O. RNA splicing dysregulation and the hallmarks of
832 cancer. *Nat Rev Cancer* (2023). https://doi.org/10.1038/s41568-022-00541-7
833 41 Liao, K. C. & Garcia-Blanco, M. A. Role of Alternative Splicing in Regulating Host
834 Response to Viral Infection. *Cells* **10** (2021). https://doi.org/10.3390/cells10071720
835 42 Wang, C. *et al.* Abnormal global alternative RNA splicing in COVID-19 patients.
836 *PLoS Genet* **18**, e1010137 (2022). https://doi.org/10.1371/journal.pgen.1010137
837 43 Manet, E. *et al.* Modulation of alternative splicing during early infection of human
838 primary B lymphocytes with Epstein-Barr virus (EBV): a novel function for the viral
839 EBNA-LP protein. *Nucleic Acids Res* **49**, 10657-10676 (2021).
840 https://doi.org/10.1093/nar/gkab787
841 44 Thompson, M. G. *et al.* Viral-induced alternative splicing of host genes promotes
842 influenza replication. *Elife* **9** (2020). https://doi.org/10.7554/eLife.55500
843 45 Louadi, Z. *et al.* DIGGER: exploring the functional role of alternative splicing in
844 protein interactions. *Nucleic Acids Res* **49**, D309-D318 (2021).
845 https://doi.org/10.1093/nar/gkaa768
846 46 Neuenkirchen, N. *et al.* Reconstitution of the human U snRNP assembly machinery
847 reveals stepwise Sm protein organization. *EMBO J* **34**, 1925-1941 (2015).
848 https://doi.org/10.15252/embj.201490350
849 47 Shaw, D. J., Eggleton, P. & Young, P. J. Joining the dots: production, processing and
850 targeting of U snRNP to nuclear bodies. *Biochim Biophys Acta* **1783**, 2137-2144
851 (2008). https://doi.org/10.1016/j.bbamcr.2008.07.025
852 48 Massenet, S., Pellizzoni, L., Paushkin, S., Mattaj, I. W. & Dreyfuss, G. The SMN
853 complex is associated with snRNPs throughout their cytoplasmic assembly pathway.
854 *Mol Cell Biol* **22**, 6533-6541 (2002). https://doi.org/10.1128/MCB.22.18.6533-
855 6541.2002
856 49 Gubitz, A. K., Feng, W. & Dreyfuss, G. The SMN complex. *Exp Cell Res* **296**, 51-56
857 (2004). https://doi.org/10.1016/j.yexcr.2004.03.022
858 50 Pellizzoni, L. Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO*
859 *Rep* **8**, 340-345 (2007). https://doi.org/10.1038/sj.embor.7400941

51 Yong, J., Kasim, M., Bachorik, J. L., Wan, L. & Dreyfuss, G. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38**, 551-562 (2010). https://doi.org/10.1016/j.molcel.2010.03.014

52 Schlaen, R. G. *et al.* The spliceosome assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms. *Proc Natl Acad Sci U S A* **112**, 9382-9387 (2015). https://doi.org/10.1073/pnas.1504541112

53 Zhang, Z. *et al.* SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**, 585-600 (2008). https://doi.org/10.1016/j.cell.2008.03.031

54 Cartegni, L., Hastings, M. L., Calarco, J. A., de Stanchina, E. & Krainer, A. R. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am J Hum Genet* **78**, 63-77 (2006). https://doi.org/10.1086/498853

55 De Maio, F. A. *et al.* The Dengue Virus NS5 Protein Intrudes in the Cellular Spliceosome and Modulates Splicing. *PLoS Pathog* **12**, e1005841 (2016). https://doi.org/10.1371/journal.ppat.1005841

56 Wang, Z. & Burge, C. B. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802-813 (2008). https://doi.org/10.1261/rna.876308

57 Grosso, A. R. *et al.* Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res* **36**, 4823-4832 (2008). https://doi.org/10.1093/nar/gkn463

58 Artemyeva-Isman, O. V. & Porter, A. C. G. U5 snRNA Interactions With Exons Ensure Splicing Precision. *Front Genet* **12**, 676971 (2021). https://doi.org/10.3389/fgene.2021.676971

59 Jia, Y., Mu, J. C. & Ackerman, S. L. Mutation of a U2 snRNA gene causes global disruption of alternative splicing and neurodegeneration. *Cell* **148**, 296-308 (2012). https://doi.org/10.1016/j.cell.2011.11.057

60 Ramakrishnan, L. Revisiting the role of the granuloma in tuberculosis. *Nat Rev Immunol* **12**, 352-366 (2012). https://doi.org/10.1038/nri3211

61 Simmons, J. D. *et al.* Differentially expressed transcript isoforms associate with resistance to tuberculin skin test and interferon gamma release assay conversion. *PLoS One* **18**, e0284498 (2023). https://doi.org/10.1371/journal.pone.0284498

62 Mehdiratta, K. *et al.* Kupyaphores are zinc homeostatic metallophores required for colonization of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **119** (2022). https://doi.org/10.1073/pnas.2110293119