1 **Host genetic background is a barrier to broadly effective vaccine protection: Relevance**

2 **to BCG and *Mycobacterium tuberculosis* Infection**

3

4 Rocky Lai[1], Diana Gong[2], Travis Williams[1], Abiola F. Ogunsola[1], Kelly Cavallo[1], Cecilia S.

5 Lindestam Arlehamn[3], Sarah Acolatse[2], Gillian L. Beamer[4], Martin T. Ferris[5], Christopher M.

6 Sassetti[1], Douglas A. Lauffenburger[2], Samuel M. Behar[1]

7

8

9 1. Department of Microbiology and Physiological Systems, University of Massachusetts

10   Medical School, Worcester, Massachusetts, USA.

11 2. Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge,

12   MA, USA

13 3. Center for infectious disease and vaccine research, La Jolla Institute for Immunology., La

14   Jolla, CA, USA

15 4. Texas Biomedical Research Institute, San Antonio, TX, USA

16 5. Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, United

17   States

18

19

20 *Correspondence:

21 Samuel M. Behar

22 E-mail address: samuel.behar@umassmed.edu

Summary

The heterogeneity of immune responses observed in humans is difficult to model in standard inbred laboratory mice. To capture the diversity inherent in mice and better understand how host variation affects BCG-induced immunity against *Mycobacterium tuberculosis*, 24 unique Collaborative Cross (CC) recombinant inbred mouse strains and the C57BL/6 reference strain were vaccinated with or without BCG, and then challenged with low-dose aerosolized virulent *M. tuberculosis*. In contrast to standard lab strains, BCG protected only half of the CC strains tested. Furthermore, BCG efficacy is dissociable from inherent susceptibility to TB. As these strains differed primarily in the genes and alleles they inherited from the CC founder strains, we conclude that the host genetic background has a major influence on whether BCG confers protection against *M. tuberculosis* infection and indicates that host genetics should be considered as an important barrier to vaccine-mediated protection. Importantly, we wished to identify the components of the immune response stimulated by BCG, which were subsequently recalled after Mtb infection and associated with protection. The T cell immune response following BCG vaccination and Mtb challenge was extensively characterized. Although considerable diversity was observed, BCG vaccination had little impact on the composition of T cells recruited and maintained in the lung after infection. Instead, the variability was largely shaped by the genetic background. We developed models to detect vaccine-induced differences, which identified immune signatures associated with BCG-elicited protection against TB. Importantly, even when categorized as susceptible vs. resistant, and protected vs. unprotected, many of the protected CC strains had unique flavors of immunity, indicating multiple paths to protection. Thus, CC mice can be used to define correlates of protection and to identify vaccine strategies that protect a larger fraction of genetically diverse individuals instead of optimizing protection for a single genotype.

Introduction

The introduction of Bacillus Calmette Guerin (BCG) as a vaccine contributed significantly to control of the tuberculosis (TB) pandemic, which has threatened global health for millennia. As the only vaccine approved for preventing TB, 4 billion doses of BCG have been administered in 150 countries (1). Although BCG prevents extrapulmonary TB in infants, it is less effective in averting pulmonary disease in adolescents and adults (2-5). BCG's variable efficacy in different regions of the world complicates its use (6). While the development of more effective TB vaccines is an important yet unmet medical need, our lack of mechanistic insights in how protective immunity against TB is mediated impedes our progress towards this goal. The identification of correlates of protective immunity to TB would accelerate this process by allowing rationale design of new vaccine candidates.

Why is identifying immune correlates of protection following vaccination difficult? TB is primarily a human disease that is transmitted person-to-person by aerosolized *Mycobacterium tuberculosis* (Mtb). No animal model recapitulates all the features of human TB, although non-human primates come close (7, 8). Although imperfect, the mouse model facilitated mechanistic studies that defined many of the critical immunological pathways required for host resistance, permitted *in vivo* antibiotic testing, and accelerated preclinical vaccine development. As all individuals within an inbred mouse strain are genetically identical, a strength of the mouse model is its "infinite" reproducibility. For experimental immunology, C57BL/6 mice are the preferred inbred strain as most genetically modified mice are produced using this genetic background. It is also the most frequently used mouse strain for TB research. However, several features of C57BL/6 mice suggest it is an outlier strain for TB research. For example, it only expresses a single class II MHC molecule (i.e., I-A^b) and is missing a class I MHC genes. Importantly, it is among the most resistant of all inbred strains to Mtb infection (9), which makes it difficult to identify conditions that increase its resistance to disease. Investigators have addressed these limitations by using F1 hybrids and susceptible strains to increase MHC and

74   genetic diversity and to identify strategies to increase host resistance (10-13). However, even

75   among a range of classical inbred strains, the genetic diversity is limited (14). Importantly, this

76   problem also affects many areas of biomedical research.

77   The mouse genetics community responded to the limited genetic diversity of inbred mice

78   by developing a resource that captures the considerable natural genetic diversity of *Mus*

79   *musculus.* Starting with eight founder strains, including representatives from all three *Mus*

80   *musculus* subspecies, an eight-way funnel breeding scheme led to random assortment of the

81   founder genomes among the progeny (15). The resulting progeny eventually became

82   Collaborative Cross (CC) strains. Each CC strain has a genome that is a unique genetic

83   combination (i.e., genotype) of the eight founder strains. Importantly, as each CC strain is

84   inbred, they are a renewable resource, and unlimited experiments can be performed with the

85   same genotype. Their population structure is well suited to study the impact of genetic variation

86   of host traits. These attributes make the CC ideal for identifying experimental correlates and

87   defining mechanisms.

88   We previously reported that the CC strains have a large variation in primary

89   susceptibility to Mtb (16, 17). There is also variation in protection conferred by BCG against Mtb

90   in the CC founder strains, suggesting that the CC strains would be an excellent model to

91   understand how host genetics affects vaccine-induced protection (16). To understand how host

92   variation affects BCG-induced immunity against Mtb, 24 CC mouse strains were vaccinated with

93   or without BCG, and then challenged with aerosolized Mtb and ranked how well BCG protected

94   these mice against Mtb. Importantly, the host's genetic background had a major effect on

95   vaccine-induced protection. We hypothesized that CC strains protected by BCG would share

96   common immunological features. However, many of the protected CC strains had unique flavors

97   of immunity, indicating multiple paths to protection. We propose CC mice can be used as a

98   preclinical model to facilitate the development of vaccines that protect a diverse range of

99   genotypes against Mtb infection.

Results

BCG vaccination protects only a subset of CC strains against TB.

To establish how the host genetic background affects vaccine mediated protection against Mtb, 24 different CC strains were vaccinated subcutaneously with BCG or left unvaccinated and rested for 12 weeks. Then, the mice were infected with low dose aerosolized H37Rv expressing YFP (H37Rv::YFP) (18). Four weeks after infection, the lung and spleen CFU were determined. The aggregated lung CFU from the unvaccinated and BCG-vaccinated CC strains (Fig.1A, top) and C57BL/6 mice (Fig.1A, bottom) shows that as a population, CC mice are protected by BCG similarly as C57BL/6 mice. When the data is segregated by individual CC strains, we can see that the population-wide variance reflects strain-specific differences in protection conferred by BCG vaccination (Fig.1B, top). Each CC strain was classified as being protected by BCG if there was a statistically significant difference between the unvaccinated and vaccinated mice. The Δlog_{10} lung CFU was calculated for each CC strain, and they were ranked from most protected (CC037, Δlog_{10} CFU = 1.6) to least protected (CC040, Δlog_{10} CFU = −0.6) (Fig.1B, bottom). BCG vaccination exacerbated subsequent Mtb infection in the lung for two CC strains (CC003 and CC040), similar to what had been observed in the NZO founder strain (16). Importantly, BCG only led to significant reductions in the lung Mtb burden in 13 of the 24 CC strains.

We next asked whether the protective effect of BCG vaccination correlated with the inherent ability (i.e., the unvaccinated state) of each CC strain to control Mtb replication. We categorized the 24 CC strains relative to C57BL/6J mice. The lung CFU of unvaccinated mice defined the inherent susceptibility or resistance to Mtb (susceptibility, y-axis). Whether each CC strain was protected or not was based on the Δlog_{10} lung CFU (relative BCG-induced protection, x-axis). The plot was centered on C57BL/6J mice, which represented a median strain in terms of lung susceptibility and protection and divided the 24 CC strains into four different groups, (Fig.1C). Based on this plot of lung susceptibility and protection, there is no significant

5

126   correlation between intrinsic susceptibility of the mouse strain and protection conferred by BCG

127   at this time point (Pearson r=0.104, p=0.62).

128   The same analysis was applied to the spleen bacterial burden from infected mice as a

129   measure of systemic infection. As was observed in the lung, the CC population was protected

130   by BCG (Fig.1D). Protection in the spleen extended from most protected (CC026, Δlog$_{10}$ CFU =

131   2.56) to least protected (CC045, Δlog$_{10}$ CFU = 0.13) (Fig.1E). BCG significantly protected

132   against systemic infection in 17 of the 24 CC strains (Fig.1E). There was also not enough

133   evidence to suggest correlation between susceptibility and protection in the spleen (Pearson

134   r=0.089, p=0.67). Interestingly, while C57BL/6 mice were a "median" strain in the lung analysis,

135   C57BL/6 mice had among the highest spleen CFU in primary infection but were also among the

136   best protected strains following BCG vaccination (Fig.1F). The skewing (Fig.1F) illustrates that

137   vaccination studies using C57BL/6 mice are not representative of genetically diverse

138   populations. Among the CC strains, protection in the lung and the spleen were moderately

139   correlated (Pearson r=0.513, p=0.0087).

140   In follow-up studies, 13 of the 24 CC strains were retested to confirm the initial

141   phenotype, (Table S1). Except for CC023, we were able to confirm our findings from the first

142   round of analysis. Thus, using 24 CC strains, with diverse genotypes, we find that less than

143   50% of the strains were protected by BCG. Protection did not correlate with the intrinsic

144   susceptibility of the strain to Mtb infection. Finally, the lack of correlation between the reduction

145   in lung versus spleen highlights the importance of tissue-specific immunity and the problems

146   inherent in protecting the lung against infection. We next assessed how BCG vaccination

147   altered the histological appearance of TB lung lesions.

148

149   ***BCG modifies the lung lesions caused by Mtb in a strain-specific manner.***

150   We assessed how BCG vaccination modified lung lesion development after Mtb infection

151   in 19 different CC strains plus C57BL/6 mice. A board-certified veterinarian pathologist (G.B.),

152 blinded to vaccination status of each strain, examined the lung for necrosis, neutrophilic and

153 lymphocytic infiltrates, lesion size and number. Within each strain, the vaccinated group was

154 correctly identified in 15 of the 20 strains analyzed based on reduced severity of the

155 histopathology. In some protected CC strains, increased lymphocytic infiltration, and reduced

156 necrosis were observed after vaccination (Table S2). CC025 (Fig. 2, top left) and CC037 (Fig. 2,

157 middle left) were both susceptible to Mtb infection, and the infected lung showed innate cells

158 and tissue necrosis. BCG vaccination protected these mice by promoting lymphocyte

159 recruitment, and in CC025 BCG vaccination reduced necrosis. However, lung CFU reductions

160 associated with BCG vaccination were not always accompanied with an improved histological

161 appearance. Despite a significant reduction in lung CFU after BCG, no discernable

162 histopathological differences were observed between unvaccinated and vaccinated CC072

163 mice, and both groups had non-necrotizing lymphocyte-rich granulomas (Fig. 2, bottom left).

164   Two of the 11 non-protected CC strains, CC003 (Fig. 2, top right) and CC024 (Fig. 2,

165 middle right) had no discernable histological differences in granuloma structure, content, or

166 severity between control and BCG vaccination when evaluated blindly. In these two strains,

167 lesions were generally small to medium-sized, non-necrotizing, with abundant perivascular and

168 peribronchiolar lymphocytes. Given the intermediate susceptibility of the CC003 strain, and

169 BCG's inability to alter lung CFU or histology, we suggest that BCG does not change CC003

170 innate or adaptive immunity to Mtb. Likewise, resistant CC024 mice are unprotected by BCG

171 vaccination, and its lung lesions are unaltered, suggesting that BCG vaccination fails to

172 enhance the natural resistance of CC024 to Mtb. In contrast, CC004 (Fig. 2, bottom right) had

173 discernable histopathological changes when controls and BCG vaccinees were evaluated. The

174 differences attributable to BCG vaccination were the greater lymphocytic infiltration with denser

175 perivascular and peribronchiolar lymphocytes. The altered histology shows that BCG modulated

176 the immunity to Mtb, despite being unable to control bacterial replication. Together, these data

7

177   show that the host genetic background impacts the granuloma content, structure, and quality of

178   the immune response to BCG and Mtb challenge in a manner that is not wholly predictable.

179

180   **BCG-protected CC strains have an altered lung microenvironment after Mtb infection**

181   The Mtb-infected lung is a complex environment of resident and recruited cells that

182   produce inflammatory mediators and counter-regulatory anti-inflammatory signals. We

183   hypothesized that BCG vaccination would prime an immune response that would generate a

184   lung microenvironment that would be less conducive to bacillary replication upon Mtb challenge.

185   To identify immunological features that correlated with protection across diverse genotypes,

186   cytokine levels in lung homogenates from unvaccinated and BCG-vaccinated C57BL/6 or CC

187   mice were measured 4 weeks post infection.

188   Given the critical role of IFNγ in Mtb control, its levels in control and BCG-vaccinated

189   mice were compared. In C57BL/6 mice and many of the protected CC strains, including CC037,

190   CC025, CC059, CC031, and CC011, IFNγ concentrations in BCG-vaccinated mice were

191   reduced compared to unvaccinated mice (Fig.3A). However, this pattern did not hold for all

192   protected strains. No changes in IFNγ levels were detected in the CC032, CC042, CC072, and

193   CC026 strains despite being protected by BCG. Thus, when considered alone, neither the

194   absolute level of IFNγ nor its change after vaccination can explain protection.

195   We next asked if there was a correlation between the different cytokines in the lung

196   homogenate and CFU, in either unvaccinated or BCG-vaccinated mice, and if so, whether it

197   differed among protected CC strains, unprotected CC strains, and C57BL/6 mice. A similar

198   pattern existed for several cytokines, including IL-17, LIF, KC, IL-6, MIP-1α, IFNγ, and TNF

199   (Figure 3B and C). A significant correlation exists between the lung cytokine and CFU for

200   unvaccinated and vaccinated mice among the unprotected CC strains. Among the protected CC

201   strains, a significant correlation was found only for the unvaccinated mice. C57BL/6 mice, which

202   are protected, followed the same pattern as protected CC strains. Comparing the Pearson

203   correlation coefficients for unvaccinated and vaccinated mice within each group indicated that

204   for several cytokines, the correlation coefficients significantly differed between unvaccinated and

205   vaccinated mice in protected CC strains but not unprotected CC strains. Interestingly, the

206   differences between the unvaccinated and vaccinated correlation coefficients were not

207   significant for IFNγ and TNF (pdiff = 0.09 and 0.21, respectively). This indicates that BCG

208   vaccination changed the relationship between bulk lung cytokine and bacterial burden in

209   protected CC for IL-17, LIF, KC, IL-6, and MIP-1α (Figure 3B) more so than for IFNγ and TNF

210   (Figure 3C). Meanwhile, an unchanged cytokine-CFU relationship for these cytokines was a

211   signature of the unprotected phenotype.

212   A few cytokines had different trends (Figure 3D). LIX (CXCL2) was negatively

213   associated with lung CFU in unvaccinated but not in BCG-vaccinated C57BL/6 mice, and it was

214   the only cytokine in the panel for which vaccination significantly changed the correlation

215   coefficient in C57BL/6 mice (pdiff = 0.012). In contrast, neither unprotected nor protected CC

216   mice had a correlation between LIX and CFU. RANTES is also unique in that it strongly

217   correlated with lung CFU in unvaccinated and vaccinated C57BL/6 mice and in vaccinated

218   protected CC strains, but not in unprotected CC strains. Finally, IP-10 correlated with CFU in

219   unvaccinated and vaccinated CC mice of both protection groups, but not with C57BL/6 mice.

220   The LIX and IP-10 results further support the idea of C57BL/6 mice being an outlier.

221   Across the board, what is largely consistent is that the relationship between cytokines

222   and bacterial burden did not change in unprotected CC strains, while this relationship did

223   change in protected CC mice. This relationship was mostly unchanged statistically in C57BL/6

224   mice, though this could be due to insufficient power to detect differences in a single strain.

225   These data show how differences in the host genetic background result in varied relationships

226   between lung homogenate cytokines and lung bacterial burden.

227

228 ***Genetic background and not BCG vaccination largely determines T cell subset***

229 ***distribution in the lung after Mtb infection.***

230 We hypothesized that changes in the frequency of T cell subsets would correlate with

231 protection induced by BCG vaccination. A multi-parameter flow cytometry panel was used to

232 identify and enumerate different immune subsets (Fig.S1). We quantified CD4 and CD8 T cells,

233 and MAIT cells (Fig.4A, and Fig.S2). Memory and effector T cell populations were defined as

234 central memory (T$_{CM}$, CD44$^+$CD62L$^+$CD127$^+$), effector memory (T$_{EM}$, CD44$^+$CD62L$^-$CD127$^+$),

235 effector (T$_{Eff}$, CD44$^+$CD62L$^-$CD127$^-$), and resident memory (T$_{RM}$, CD44$^+$CD103$^+$CD69$^+$) (Fig.4B,

236 4C). CD4 T cells subsets were defined in all CC strains based on their expression of the

237 transcription factors Foxp3, Tbet, Gata3, and RORγt to identify Treg, Th1, Th2, and Th17 cells,

238 respectively (Fig.4D). Gata3 was universally lowly expressed and was not analyzed further. For

239 the other T cell subsets, we observed tremendous strain-to-strain variation. In general, BCG

240 vaccination did not modulate the distribution of T cells subsets nor was there any correlation

241 between the relative frequency of these T cell subsets and the outcome of BCG vaccination.

242 Thus, the frequency of different T cells subsets in the lungs of CC mice after Mtb infection was

243 largely determined by each strain's genetic background and not their immunization status.

244 CXCR3 and CX$_3$CR1 expression, which has been used to discriminate between lung

245 parenchymal and circulating T cells, was analyzed in a limited number of CC strains (19). Like

246 other markers, we observed variation in CXCR3 expression between different CC strains, but

247 little or no variation that correlated with vaccination status or susceptibility. We could not detect

248 CXCR3 expression in CC003 mice (Fig.4E). CC003 has the PWK allele of the CXCR3 gene,

249 and we are also unable to detect CXCR3 expression by PWK mice (data not shown). The

250 inability of the anti-CXCR3 mAb to detect the PWK allele of CXCR3 could be arise because of a

251 haplotype-specific lack of expression, or a private mutation as we have described in other CC

252 strains (20). Finally, a reduction in CX3CR1$^+$ CD4 and CD8 T cells was observed in several of

253 the protected CC strains but not in unprotected strains (Fig.4E). This pattern mirrors the

254   reduction in IFNγ levels observed in some protected CC mice (Fig.3), suggesting that these

255   changes reflect Th1 modulation.

256   Finally, we used principal components analysis to holistically visualize the variation in

257   the T cell phenotyping data. What is striking is the small effect of BCG vaccination in altering T

258   cell phenotypes, in comparison to large effect of genetic strain differences. These data all

259   indicate that genetic differences between the CC strains determine the differences among T cell

260   subsets recruited to the lung, and not BCG vaccination. Therefore, we next wished to determine

261   whether there were functional differences among the T cell subsets in protected vs.

262   nonprotected CC strains.

263

264   ***Cytokine production by T cells differ in CC strains after BCG vaccination and Mtb***

265   ***infection.***

266   As the cytokines produced by T cells, including IFNγ, TNF, and IL-21, are important in

267   Mtb control, we hypothesized that changes in T cell function would be associated with BCG-

268   mediated protection. To measure function, we stimulated lung mononuclear cells (MNC) from

269   each mouse with anti-CD3 or the MTB300 megapool, which contains 300 peptides representing

270   epitopes from 90 Mtb proteins that are frequently recognized by human CD4 T cells and murine

271   T cells (21, 22). Twenty-six cytokines and chemokines in culture supernatants were 24 hours

272   after stimulation. Lung MNC from unvaccinated or BCG-vaccinated Mtb-infected C57BL/6 mice

273   did not differ in their secretion of cytokines or chemokines following MTB300 stimulation. In

274   contrast, IFNγ, TNF, IL-2 or IL-17, were differentially produced by unvaccinated vs. BCG-

275   vaccinated by some CC strains (Fig.5A). Interestingly, the lung MNC from BCG vaccinated

276   CC037, CC031, CC72 and CC001 strains all produced more IL-17 than their respective

277   controls. In contrast, none of the unprotected strains produced significant amounts of IL-17,

278   except for CC023, and its lung MNC produced less IL-17 after BCG vaccination.

279   In parallel, we measured intracellular levels of IFNγ, IL-2, IL-17, TNF, IL10, and IL-4, and

280   CD107 surface expression by CD4 and CD8 T cells (Fig.S3). BCG modified CD4 T cell cytokine

281   responses in C57BL/6 mice and several CC strains (Fig.5B). Among CD8 T cells, BCG

282   vaccination primarily altered the IFNγ response to MTB300 (Fig.5C). Patterns of cytokine

283   production were observed in BCG vaccinated CC mice that were not observed in C57BL/6 mice.

284   For example, while BCG vaccination did not alter the frequency of IFNγ$^+$ or TNF$^+$ CD4 T cells in

285   CC037 mice, it did increase IL-2 and IL-17 responses. BCG vaccination of CC031 mice

286   increased the frequency of CD4 T cells producing IFNγ, IL-2 and TNF but not IL-17. Finally,

287   CD4 T cells from BCG vaccinated had CC072 had increased IFNγ and IL-2 but not TNF or IL-

288   17. No discernable differences in cytokine responses were observed in most of the non-

289   protected CC mice, except for a decrease in IFNγ responses in CC023. Thus, the cytokine

290   responses in both control and BCG-vaccinated CC mice differ greatly from the classic C57BL/6

291   model. Increased cytokine responses after T cell stimulation were more likely to be detected in

292   protected CC strains. Moreover, BCG vaccination led to quantitative and qualitative changes in

293   the type of T cell responses that persisted after Mtb challenge.

294

295   ***Identifying correlates of protection using multivariate approaches.***

296   We then turned to multivariate approaches to discern how BCG-induced immune

297   changes in the lung holistically lead to reductions in lung CFU in CC strains. A multivariate

298   partial least squares regression (PLSR) model successfully predicted lung CFU changes from

299   BCG-induced immune changes as validated in a five-fold cross validation framework (Fig.6A,

300   Fig.S4A). Features selected through elastic net as most correlated with lung CFU changes

301   spanned all four datasets (Fig.6B, 6C). By plotting the selected features for all CC strains,

302   ordered from most protected to least protected, it is apparent that the protected strains all had

303   BCG-induced reductions in several inflammatory mediators such as MIG, MIP-1α and -1β, IL-6,

304   RANTES, and IL-17. Their decreases could reflect resolving infection, in contrast to the

305   unprotected strains that had increased inflammatory cytokines, which could signify unimpeded

306   infection. Protected strains also had a greater vaccination-induced increase in activated CD4 T

307   cells that express RORγt and FoxP3, which could be Th17 and Treg cells, respectively.

308   Contrarily, protected strains had decreased activated CD4$^-$CD8$^-$ T cells producing T-bet.

309   Importantly, stimulated IL-2 production, unlike several other cytokines, tended to increase in

310   protected strains. Thus, BCG vaccination leads to changes in T cell immunity and in the immune

311   environment of the lung after Mtb infection in CC genotypes that are protected by BCG

312   vaccination.

313

314   ***Immune responses in CC mice differ across protection and susceptibility categories***

315   While the PLSR model highlights potential correlates of protection across all genotypes,

316   it may miss correlates that only apply to certain subsets of CC strains. We hypothesized that

317   vaccination may protect naturally resistant and susceptible strains may be protected by different

318   mechanisms. To test this, we created separate partial least squares discriminant analysis

319   (PLSDA) models for resistant and susceptible CC strains and examined the features that best

320   separated the BCG-protected and non-protected strains in each of these categories. (Fig.1C).

321   The data used for the models was processed the same way as in the PLSR models, although

322   we also examined the vaccinated and unvaccinated mice separately (Fig.7A, 7C, 7E), in

323   addition to vaccine-induced differences (Fig.7B, 7D, 7F). All models could predict the two

324   compared categories accurately and thus there are multivariate signatures of these categories

325   (Fig.S4B). Heatmaps of each model's selected features were hierarchically clustered both by

326   feature and CC strain and used to characterize the facets of the immune response

327   distinguishing the groups. Our first analysis compared resistant CC strains that were protected

328   by BCG (n=4) or not (n=8), based on lung CFU. As described earlier (Fig.4), the unvaccinated

329   and vaccinated mice of each CC strain were very similar and clustered together when the

330   model-selected features were hierarchically clustered, indicating that the genetic background

331 was an important determinant of outcome (Fig.7A). The fact that both protected and unprotected

332 mice were split into multiple clusters and there are no clear individual features distinguishing the

333 categories shows that these strains have unique immune responses. In contrast, the vaccine-

334 induced model comparing protected and non-protected resistant strains did yield more

335 univariate trends in addition to the strong multivariate signature (Fig 7B). Most prominently, CD4

336 T cells making IL-2 are increased in protected strains as are lung IL-7 and MIP-1α, while other

337 lung cytokines like TNF and IL-17 decrease (Fig.7B). This shows that while the broad immune

338 response in resistant mice is diverse, narrowing in on vaccine-induced changes reveals a more

339 uniform route to protection.

340   We next analyzed CC strains that were inherently more susceptible than B6 mice and

341 were either protected by BCG (n=4) or not (n=8) (Fig.7C). This yielded two clear subclusters

342 defining susceptible, non-protected CC strains. The CC004 and CC044 subcluster expressed

343 high levels of T_{CM}, CD4$^+$Foxp3$^+$, CD8$^+$PD-1$^+$ and polyfunctional CD4$^-$CD8$^-$ T cells. The other

344 cluster had high levels of lung MIP-1α and CD3$^-$Foxp3$^+$ T cells, and fewer polyfunctional CD4$^-$

345 CD8$^-$ T cells. The vaccination model (Fig.7D) shared several features with the model including

346 all mice (Fig.7C). However, the best distinguishing features were unique to the vaccination

347 model and included increased amounts of stimulated IL-6, IL-17A, and IL-15 in protected mice

348 compared to unprotected. Common to both the resistant and susceptible mouse models is that

349 overall lung IL-17 decreases because of BCG vaccination in protected mice, but T cell

350 populations making IL-17 increase. It is also notable that many of the best distinguishing

351 features are cytokine-based rather than T cell phenotype-based. Overall, our analysis has

352 identified features that distinguish protected from non-protected mice in both resistant and

353 susceptible backgrounds.

354   Finally, to identify features that differed between restrictive and permissive immune

355 responses to Mtb, we compared susceptible CC strains that were unprotected by BCG to

356 resistant CC strains that were protected by BCG. Unsurprisingly, there is clear separation of

14

357 categories in the model with all mice. "Poor" immunity CC strains tended to have high levels of

358 inflammatory monokines (Fig.7E). "Good" immunity CC strains had high levels of RANTES after

359 anti-CD3 stimulation. A subset of "good" immunity strains including CC031 and CC011 had high

360 frequencies of IFNγ-producing CD8 T cells and various subsets of CD4⁻CD8⁻ T cells that

361 produced IFNγ or expressed CD11a. In contrast, a subset of "poor" immunity strains had high

362 levels of polyfunctional CD4⁻CD8⁻ T cells and ILCregs. The protection model revealed similar

363 features with the addition of increased activated CD4 T_{RM} cells after BCG vaccination in "poor"

364 immunity but not "good" immunity CC strains (Fig.7F). We find that immune features can be

365 identified using PLSDA that can categorize the inherent susceptibility and response to BCG of

366 CC strains following Mtb infection. Looking at more refined subgroups of CC strains allowed us

367 to find features that define subclasses of protected and non-protected strains and highlight

368 features that did not emerge in the PLSR model such as non-T_{RM} effector and T_{CM} populations.

369

370 *Th1/17 cells detected in a subset of CC mice that are protected by BCG vaccination*

371   T cells expressing IL-17 and RORγt emerged as important correlates in several of the

372 multivariate protection models. Vaccination significantly enhanced IL-17 responses univariately

373 in the CC037, CC001, CC072 and CC031 strains, which were all protected by BCG (Fig.5A,

374 5B). As the CC strains that produced IL-17 after antigen stimulation also had strong IFNγ

375 responses, we considered whether these T cells produced both IFNγ and IL-17. After

376 stimulation with MTB300, some CD4 T cells in BCG vaccinated CC037 mice produced both

377 IFNγ and IL-17 (Fig.8A). Such IFNγ/IL-17-producing CD4 T cells were not detected in C57BL/6

378 mice, regardless of vaccination status. BCG vaccination significantly increased the frequency of

379 IFNγ/IL-17-producing CD4 T cells in CC037, CC072 and CC001 mice (Fig.8A, 8B). Further

380 analysis identified BCG primed CD4 T cells from CC037 and CC001 to co-express both Tbet

381 and RORγt after Mtb infection (Fig.8C, 8D), suggesting that these are Th1/17 cells.

382 We next measured T cell polyfunctionality. In addition to IFNγ and IL-17, an increase in

383 the proportion of CD4 T cells producing IL-2 or TNF was detected (Fig.8E), indicating that CD4

384 T cell polyfunctionality was increased in BCG vaccinated mice. A detailed analysis of the

385 different polyfunctional populations revealed different distributions of Th1/17 cells within each

386 CC strain. IFNγ$^+$IL-2$^+$TNF$^+$IL-17$^+$ cells were significantly increased in CC037 following BCG

387 vaccination, while frequency of IFNγ "single positive" and IFNγ$^+$TNF$^+$ CD4 T cells were

388 diminished (Fig.S5, red and purple boxes). Similarly, IFNγ$^+$TNF$^+$IL-17$^+$ cells were increased

389 following BCG vaccination in CC072 and CC001 (Fig.S5, blue box). CC037 (Fig. S5, purple

390 box). This contrasts with C57BL/6 mice where BCG vaccination had a minimal effect on CD4 T

391 cell polyfunctionality. Significant decreases were observed for IFNγ$^+$TNF$^+$ dual producers and

392 IFNγ single positive cells in Mtb-infected BCG-vaccinated C57BL/6 mice, consistent with the

393 observation that Th1 dominant response is associated with decreased bacterial burden in this

394 background.

395 Finally, we sought to confirm that increased Th1/Th17 polyfunctionality was among the

396 most significant correlates with BCG-mediated protection in the CC strains with this phenotype.

397 We performed a correlation analysis between the measured immune features and the

398 normalized lung CFU from unvaccinated control and BCG vaccinated mice. As a point of

399 comparison, for C57BL/6 mice, IFNγ-producing and polyfunctional CD4 T cells correlated with

400 lung CFU. In CC037 mice, IFNγ-producing CD8 T cells correlated with lung CFU. Importantly,

401 many features were negatively correlated with lung bacterial burden in CC037 mice (Fig.8F).

402 The frequency of CD4$^+$RORγt$^+$, and antigen-stimulated IL-17 secretion and IFNγ/IL-17 dual-

403 producing CD4 T cells were significant even after correction for multiple testing, suggesting that

404 these features were associated with protection following BCG. There was a strong correlation

405 between the frequency of IFNγ/IL-17 dual-producing CD4 T cells and lung CFU in CC037 mice

406 (Spearman r = -0.8947) (Fig.8G).

407

Discussion

We sought to determine how pre-existing immunity to BCG modifies the immune response to Mtb infection using CC mice in a genetically diverse population. An important outcome is our ranking of how BCG induced protection relative to control unvaccinated mice within each of these 24 CC strains expanded the range of responses as compared to the reference strain C57BL/6J. We kept constant the BCG strain, growth conditions, dose, and vaccination route. All mice were rested for 12 weeks after vaccination, challenged with the same Mtb strain, and analyzed four weeks after infection. The different CC strains were categorized as being either protected or unprotected by BCG against Mtb infection, based on whether vaccination led to a statistically significant reduction in lung bacillary burden measured four weeks after low dose aerosol infection. After controlling for these variables, the surprising result was that BCG protected only half of the CC strains tested. As these strains differed primarily in the genes and alleles they inherited from the CC founder strains, we conclude that the host genetic background has a major influence on whether BCG confers protection against Mtb infection. These results extend our analysis of the eight CC founder strains, which also vary in their ability to be protected by BCG vaccination (16). Not only do environmental mycobacteria and genetic variations in BCG strains used around the world interfere with the ability of BCG-induced protection, but we suggest that host genetics should be considered a third important barrier to vaccine-mediated protection to TB.

Why doesn't BCG induce protection in all CC strains? Protection was largely defined based on reductions in lung CFU four weeks after aerosol Mtb infection, which is a standard time point in murine vaccine studies. Interestingly, the CC strains that were not protected by BCG can be subdivided based on their lung pathology. In half of the unprotected CC strains (e.g., CC004), the BCG-vaccinated mice had changes in lung histology (e.g., more lymphocytic infiltrates) that differed from the unvaccinated controls. We conclude that while immunization generated a memory recall response in these 'unprotected' CC strains, it was insufficient to

17

434   control lung CFU. Histological changes couldn't be identified in other unprotected strains (e.g.,

435   CC024), and we suggest that BCG generated a suboptimal immune response. A failure in the

436   overall immune response to BCG is also supported by the finding that the lung homogenate

437   cytokine patterns in the unprotected vaccinated CC strains are similar to the unvaccinated

438   strains.

439   Importantly, we wished to identify the components of the immune response stimulated

440   by BCG, which were subsequently recalled after Mtb infection and associated with protection.

441   We hypothesized that CC strains protected by BCG would share common immunological

442   features. We postulated that comparing protected to unprotected CC strains would allow us to

443   filter out components of the immune response that were elicited by vaccination or infection but

444   not associated with protection. The T cell immune response following BCG vaccination and Mtb

445   challenge was extensively characterized as T cells are essential for protection against Mtb (23-

446   27). Although considerable diversity was observed, BCG vaccination had little impact on the

447   composition of T cells recruited and maintained in the lung after infection. Instead, the variability

448   was largely shaped by the genetic background. In contrast, functional skewing of the T cell

449   cytokine production by BCG was evident in several of the protected CC strains. We developed a

450   series of multivariate models to identify immune signatures associated with BCG-elicited

451   protection against TB. These revealed that the correlates of protection common among CC

452   strains are the decrease of several lung homogenate cytokines and increase in CD4 T cells

453   expressing Foxp3 or RoRγT. When creating models specific to either resistant or susceptible

454   CC mice that predicted protected vs. unprotected mice, further correlates of protection emerge

455   that could not be captured by the broader model. Importantly, there is substantial heterogeneity

456   in the immune responses of protected mice, indicating multiple paths to protection.

457   Lung homogenate levels of IL-17 strongly correlated with lung CFU. However, the

458   correlation between IL-17 and lung CFU was lost in BCG protected CC strains after vaccination

459   and challenge. Interestingly, lung IL-17 concentrations co-vary with IL-1β and IL-6, two

18

460   cytokines that are involved in skewing of Th17 cells, and with TNF, KC, MIP2, and G-CSF, all

461   which are induced by IL-17 (Fig.3 and data not shown). IL-17 could be more generally important

462   in driving lung inflammation than predicted by the C57BL/6 model. In contrast, the CD4 T cells

463   from four of the 13 protected strains by BCG co-produced IFNγ and IL-17 after Mtb infection.

464   None of the 11 unprotected strains had this phenotype, and these dual producers were detected

465   only in BCG vaccinated mice and not in sham controls, nor were they detected in the C57BL/6

466   reference strain. Thus, IFNγ/IL-17 dual producing CD4 T cells are an example of an immune

467   response that BCG primes but is not manifested until the CD4 T cell response is recalled by Mtb

468   infection. These cells are similar to Th1/17 cells, or Th1*, that are identified following TB

469   infection, and are associated with late-forming, low burden granulomas in a NHP TB infection

470   model (28-30). However, an important difference is that IL-17 is generally not produced by Th1*

471   cells after stimulation with Mtb antigens. Importantly, we find that dual IFNγ/IL-17 producing

472   CD4 T cells correlated with protection in several strains that were well-protected by BCG.

473   Interestingly, within lung homogenates IL-17 correlates with disease; conversely, IL-17-

474   producing or RoRγt-expressing CD4 T cells correlate with protection. We hypothesize that a

475   cellular feature other than IL-17 production, such the capacity for self-renewal, migration, or

476   plasticity, is associated with Mtb control. The CC strains will be invaluable in further unraveling

477   the role of IL-17 and IL-17-producing T cells in immunity to TB.

478   While extrinsic factors, many that disrupt cell mediated immunity, increase the risk of

479   developing TB, what drives the emergence of disease in most people is largely unknown. Host

480   genetics affects susceptibility to TB both in mice and people, and genetic loci have been

481   identified that increase TB risk (i.e., MSMD, polygenic) (17, 31-33) The factors that influence

482   BCG efficacy remain controversial. We hypothesize that genetic factors contribute to variation in

483   the effectiveness of BCG against pulmonary Mtb. Our analysis shows that while certain mouse

484   genotypes cannot be protected by BCG, the inherent susceptibility or resistance of the mouse

19

485   strain is not a predictor of vaccine efficacy. This suggests that a priori, vaccines are appropriate

486   interventions even in individuals that might have a genetic susceptibility to TB.

487   One caveat is that we have tested the efficacy of only a single vaccine in CC mice. The

488   variation in host responses of CC mice could be leveraged as an ideal platform for testing the

489   efficacy of preclinical vaccines in the context of a genetically diverse background. It is possible

490   that vaccines that work through alternate mechanisms may be able to protect CC mice that are

491   refractory to BCG vaccination. Instead of optimizing vaccine concepts in a single mouse

492   genotype (e.g., C57BL/6 mice), identifying vaccines that are effective in numerous different

493   genotypes (i.e., genetically distinct individuals), as modeled by CC or DO mice, represents an

494   alternative paradigm for preclinical vaccine development. Our data establish CC strains as a

495   novel animal resource that can facilitate the discovery of correlates of vaccine-induced

496   protection and accelerate preclinical vaccine testing.

Materials and Methods

Mice

Mice were initially tested in four batches, with each batch including six different CC strains and C57BL/6J mice as a reference strain. Female C57BL/6/J (#0664) were purchased from The Jackson Laboratory. Female mice from 24 CC strains were purchased from the UNC Systems Genetics Core Facility (SGCF) between July 2018 and August 2019. The 24 CC strains used in this study include: CC001/Unc, CC002/Unc, CC003/Unc, CC004/TauUnc, CC010/GeniUnc, CC011/Unc, CC012/GeniUnc, CC023/GeniUnc, CC024/GeniUnc, CC025/GeniUnc, CC026/GeniUnc, CC031/GeniUnc, CC032/GeniUnc, CC035/Unc, CC036/GemUnc, CC037/TauUnc, CC040/TauUnc, CC041/TauUnc, CC042/GeniUnc, CC044/Unc, CC045/GeniUnc, CC059/TauUnc, CC061/GeniUnc, CC072/TauUnc. More information regarding the CC strains can be found at http://csbio.unc.edu/CCstatus/index.py?run=AvailableLines.information.

All animal studies were conducted using the relevant guidelines and regulations and approved by the Institutional Animal Care and Use Committee at the University of Massachusetts Medical School (UMMS) (Animal Welfare A3306-01), using the recommendations from the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the Office of Laboratory Animal Welfare.

Bacterial strains

Bacteria expressing yellow fluorescent protein (sfYFP) were generated as described (17). Mice were vaccinated with 10^5 CFU BCG (BCG-SSI strain, resuspended in 0.04% Tween/PBS) subcutaneously (100 uL/injection).

Aerosolized Mtb infection of mice

21

523 Mice were infected by the aerosol route as previously described (18). Frozen bacterial stocks

524 were thawed and sonicated for 1 minute and then diluted into 5 ml of 0.01% Tween-80 in PBS.

525 The diluted bacterial suspension was aerosolized to infect mice using a Glas-Col chamber

526 (Terre Haute). The average number of bacteria delivered into the lung was determined for each

527 experiment by plating lung homogenate from 4-5 mice 24 hours after infection and ranged

528 between 40-150 CFU/mouse.

529

530 Bacterial burden in lung and spleen

531 Infected mice were euthanized at pre-determined timepoints, and the left lung or whole spleen

532 were homogenized using 2 mm zirconium oxide beads (Next Advance) in a FastPrep

533 homogenizer (MP Biomedicals). Tissue homogenates were serially diluted and plated on 7H11

534 agar plates (Hardy Diagnosis). CFU was enumerated after 19-21 days of incubation at 37°C

535 and 5% CO_2.

536

537 Lung cell preparation

538 Single cell suspensions were prepared by homogenizing lungs using a GentleMACS tissue

539 dissociator (Miltenyi), digesting with 300 U/ml collagenase (Sigma) in complete RPMI (10%

540 FBS, 2 mM L-Glutamine, 100 units/ml Penicillin/Streptomycin, 1 mM Na-Pyruvate, 1X Non-

541 essential amino acids, 0.5X Minimal essential amino acids, 25 mM of HEPES, and 7.5 mM of

542 NaOH) at 37°C for 30 minutes, and followed by a second run of dissociation using the

543 GentleMACS. Suspensions were then sequentially filtered through 70 μm, treated with ACK

544 lysis buffer (Gibco), then filtered through 40 μm strainers.

545

546 Flow cytometric analysis

22

547 Cells were first stained with Zombie Fixable Viability dye (Biolegend) for 10 minutes at room

548 temperature (RT), after which cells were stained with antibodies in autoMACS running buffer

549 (Miltenyi) containing 5 ug/ml of anti-mouse CD16/32 (BioXcell) for 10 minutes at 4°C. In some

550 experiments, MAIT cells were identified using the MR-1-PE tetramer (NIH tetramer core facility)

551 at 37°C for 1hr. Cell were then stained with surface antibody cocktail for 20 minutes at 4°C. To

552 measure transcription factor (TF) expression, cells were fixed and permeabilized for 30 minutes

553 at RT, followed by staining for 30 minutes with antibodies to the different TF at RT using the

554 Foxp3/TF staining buffer set (ThermoFisher). Samples were acquired on the Aurora (Cytek).

555

556 For intracellular staining, cells were first stimulated for 5 hours with either anti-CD3 antibodies (1

557 ug/mL, Biolegend) or MTB300 peptide megapool (21) which contains 300 peptides representing

558 90 antigens, and contains epitopes recognized by both murine CD4 and CD8 T cells. GolgiPlug

559 (BD Biosciences) was also given at this time (0.5ug/well). Viability and surface staining were

560 performed as described, after which cells were fixed instead with Fixation/Permeabilization kit

561 for 20 minutes at 4°C. Cells were then stained with an intracellular cytokine antibody cocktail for

562 20 minutes at 4°C.

563

564 To inactivate the bacteria, samples were fixed with 1% paraformaldehyde/PBS for 1 hour at

565 room temperature and then washed with MACS buffer. For both cell surface and intracellular

566 markers, new markers were introduced in the repeat experiments for a deeper dive into the

567 immunological characterization. Flow data were analyzed using FlowJo v10.7.1.

568 Polyfunctionality was examined using the SPICE platform (34).

569

570 LEGENDplex cytokine analysis

571 Lung cells were stimulated with MTB300 peptide megapool (1ug/mL) or anti-CD3 (1ug/mL) at a

572 concentration of 10^6 cells/mL for 24hrs. Culture supernatants were filtered through 0.2um filter,

573 after which cytokine levels were measured with LEGENDplex kits from Biolegend. Specifically,

574 we used their mouse T helper cytokine panel (IFNγ, TNF, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-13,

575 IL-17a, IL-17f, IL-21, IL-22) as well as their mouse proinflammatory chemokine panel (RANTES,

576 MIP-3α, Eotaxin, TARC, KC, MCP-1, MIG, IP-10, MIP-1α, MIP-1β, BLC, LIX, MDC) for a total of

577 26 cytokines and chemokines.

578

579 <u>Lung homogenate</u>

580 Cytokines and chemokines in lung homogenates obtained four weeks after infection were

581 measured using the 31-plex discovery Assay (MD31) from Eve technologies (Calgary, Canada).

582

583 <u>Histology analysis</u>

584 The right middle lung lobe from each mouse was immersion fixed in 10% formalin immediately

585 following euthanasia. Lung tissue was processed, embedded in paraffin, sectioned at 5um, and

586 stained with hematoxylin and eosin at the Morphology Core Facility at UMass Chan Medical

587 School for histological examination by a board-certified veterinary pathologist (GB). Blind

588 classification (BCG vs unvaccinated) within each CC strain was determined based on the

589 assumption that BCG vaccination would have three possible effects on lungs: (i) reduce

590 necrosis/neutrophilic infiltrates; (ii) increase lymphocytes and/or plasma cell abundance; and (iii)

591 reduce granuloma size. Five of the 24 CC strains were not evaluated because of timing of the

592 study. Accuracy of the blind classification was assessed by an independent investigator (AFO).

593 All glass slides were subsequently re-evaluated to more closely examine and describe

594 histopathology features attributable to BCG vaccination and attempt to find unique responses.

595 Glass slides were then digitized by Vanderbilt University Medical Center (Nashville, TN) and

596 whole images generated using Aperio ScanScope (Buffalo Grove, IL, USA) at 400 times normal

597 magnification with jpeg compression. Representative images were captured using Aperio

598 ImageScope magnified 40 and 400 times and imported into GraphPad Prism 9 to generate the

599 panel Figure. Digital images were not adjusted. Whole slide imaging and quantification of

600 immunostaining were performed in the Digital Histology Shared Resource (DHSR) at Vanderbilt

601 University Medical Center.

602

603 Univariate correlations

604 To evaluate how C57BL/6 mice, protected CC strains, and non-protected CC strains compared

605 in their relationships between lung cytokine production and Mtb burden, a series of univariate

606 Pearson correlations were conducted. Lung homogenate cytokines and chemokines

607 measurements in the form of concentrations were natural logged and each of their Pearson

608 correlations with the log10 lung CFU or spleen CFU were evaluated, with both the cytokine and

609 CFU measurements separated into groups of C57BL/6, protected CC strains, and non-protected

610 CC strains. A Benjamini-Hochberg multiple hypothesis correction was applied to account for the

611 number of correlations calculated. A further comparison of the Pearson correlation coefficients

612 between the BCG-vaccinated and unvaccinated mice within the C57BL/6, protected CC, and

613 non-protected CC categories was conducted. To do so, the correlation coefficients underwent a

614 Fisher's z transformation and then were compared by a z-test, with the reported p values

615 ("pdiff") being Benjamini-Hochberg corrected p values of this comparison.

616

617 Multivariate analyses

618 Multivariate analyses involved combining all four immunological datasets in this study: the lung

619 homogenate cytokine/chemokine measurements, the ICS, the LEGENDplex cytokine analysis,

620 and the T cell phenotypic flow cytometry. The analyses also included all CC mice used in the

621 study in addition to the control C57/BL6 mice.

622

623 For the ICS, the data were analyzed as combinatorics of the seven markers measured in order

624 to best capture polyfunctionality, and further divided by CD4 T cell, CD8 T cell, CD4$^-$CD8$^-$ T

625 cell, and CD3$^-$ categories. The features were filtered by a minimum mean and variance

626 percentage of 1e-6. Only MTB300 stimulated samples were used from the ICS dataset. The

627 flow cytometry data was in the form of percentages of CD3$^+$ or CD3$^-$ cells, while the

628 immunoassay data was in the form of concentrations. The combined dataset was natural log

629 transformed and subsequently z-scored by measurement across all samples. For models that

630 involved BCG-induced protection, these z-scored values were further transformed by

631 subtracting the features from each vaccinated mouse and the mean of the features of the

632 unvaccinated mice from the same mouse strain.

633

634 For the PLSR and PLSDA models, elastic net feature selection was performed in 100 trials of 5-

635 fold cross validation, where features that were present in at least 80% of trials were included in

636 the final model. Each model was compared to two null models: one with permuted lung CFUs

637 and one with randomly chosen features. The mean squared error of the prediction of lung CFU

638 (for the PLSR) or the classification accuracy (for the PLSDAs) was compared to that of the null

639 models in 50 trials of 5-fold cross validation.

640

641 Quantification and Statistical Analysis

642 Statistical analysis was performed using Graphpad Prism 9. P-values were calculated using

643 unpaired t test, one-way ANOVA, or two-way ANOVA as indicated in the figure legends.

644 Multivariate analyses were performed in R (v3.6.0 and 4.0.4) using the ropls Bioconductor

645 package.

646

653 **Author contribution**

654 Conceptualization: R.L., S.M.B.; methodology, R.L. S.M.B.; investigation, R.L., D.G., T.W.,

655 A.F.O., K.C.; formal analysis, R.L., D.G., G.L.B., writing- Original draft, R.L., D.G., S.M.B;

656 writing- review and editing, R.L., D.G., C.S.L.A., G.L.B., M.T.F., C.M.S., S.M.B.; supervision,

657 R.L., D.A.L, S.M.B.; funding acquisition, S.M.B.

658

659 **Declaration of interests**

660 The authors declare no competing interests.

661

667   **Figure legends**

668   Figure 1. BCG-induced protection in 24 CC strains and C57BL/6 mice.

669   (A) Lung and (D) spleen CFU of 24 CC mice (n=120/group; top) or all C57BL/6 mice

670   (n=20/group; bottom) 4 weeks post infection. Each point represents an individual mouse. Line,

671   median. Student's t-test; ****, $p < 0.0001$. (B) Protection in the lung and (E) spleen was

672   assessed by comparing the CFU in BCG vs. unvaccinated mice for each strain (n=5/group)

673   using the Benjamini-Hochberg procedure to determine the false discovery rate (FDR) (top). The

674    $\Delta\log_{10}$CFU was calculated for each strain (bottom). The open bar represents C57BL/6 mice

675   (n=20/group). *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$. (C) Correlation in the lung

676   and (F) spleen between susceptibility (y-axis, defined by CFU in unvaccinated group) and the

677   protection conferred by BCG (x-axis, defined by $\Delta\log_{10}$CFU), The data is relative to the C57BL/6

678   reference strain so that the graph is centered on C57BL/6 mice, and the different CC strains are

679   divided into four quadrants based on their relative susceptibility and protection.

680

681   Figure 2. Lung histology of CC mice after BCG vaccination and Mtb challenge

682   Representative lung histology of three BCG-protected CC strains (CC025, CC037, CC072)

683   (left), and three unprotected CC strain (CC003, CC004, CC024) (right). Original magnification,

684   40X (inset, 400X). Scale bar, 500 μm (inset, 50 μm).

685

686   Figure 3. BCG-protected CC mice have an altered lung microenvironment after Mtb infection.

687   (A) Lung homogenate IFNγ levels from BCG vaccinated or unvaccinated CC mice at 4 weeks

688   post infection. A two-way ANOVA was performed using the original FDR method of Benjamini

689   and Hochberg. ***, $p < 0.001$; ****, $p < 0.0001$. (B) – (D) Pearson correlation between lung CFU

690   and (B) IL-17, LIF, KC, IL-6, MIP-1α; (C) IFNγ and TNF; and (D) LIX, RANTES, and IP-10

691   comparing unvaccinated and BCG vaccinated C57BL/6 mice, and unprotected or protected CC

692   mice. R values indicate the Pearson correlation coefficients for unvaccinated (blue) and BCG

28

693   (red) groups, with the accompanying p values indicating the Benjamini-Hochberg corrected

694   significances of those correlations. Within each category of C57BL/6, not protected CC, and

695   protected CC mice, the correlations between BCG vaccinated and unvaccinated mice were

696   compared using a Fisher's z transformation of the correlation coefficients followed by a z test,

697   and the significance of the difference in the correlations is reported as a Benjamini-Hochberg

698   corrected pdiff. Significant p values are bolded.

699

700   Figure 4. Frequency of T cell populations in protected and unprotected CC mice following Mtb

701   infection.

702   (A) Proportion of $CD3^+$, $CD3^+CD4^+$ and $CD3^+CD8^+$ T cells in the lungs of unvaccinated and BCG

703   vaccinated CC mice at 4 weeks post infection. (B) Proportion of memory CD4 and CD8 T cells

704   in the lungs of either unvaccinated or BCG vaccinated CC mice at 4 weeks post infection, as

705   defined by $CD44^+$ and various combinations of CD62L and CD127 expression; $CD62L^+CD127^+$

706   (central memory), $CD62L^-CD127^-$ (effector), $CD62L^-CD127^+$ (effector memory). (C) Proportion

707   of resident memory CD4 and CD8 T cells in the lungs of either unvaccinated or BCG vaccinated

708   CC mice at 4 weeks post infection, as defined by $CD44^+CD103^+CD69^+$. (D) Proportion of Th1,

709   Treg and Th17 cells in the lungs of either unvaccinated or BCG vaccinated CC mice at 4 weeks

710   post infection, as defined by $CD44^+$ and $Tbet^+$ (Th1), $Foxp3^+$ (Treg) or RoRγt (Th17) expression.

711   (E) Proportion of CD4 and CD8 T cells expressing either CXCR3 or CX_3CR1 in the lungs of

712   either unvaccinated or BCG vaccinated CC mice at 4 weeks post infection. (F) PCA of the T cell

713   phenotyping data. Each point represents the average of 5 mice within a given mouse strain that

714   are either BCG vaccinated or unvaccinated. Open symbols, unvaccinated; closed symbols,

715   BCG vaccinated. Each color represents a different mouse strain. (A–E) The data represent the

716   mean ± SEM from one experiment. Two-way analysis of variance with the Benjamini and

717   Hochberg multiple comparison method. The FDR was set to 0.05 and the numbers in the figures

718   are the q value.

719

720   Figure 5. Cytokine responses in CC strains following Mtb infection.

721   (A) Lung cells 4 weeks post infection were stimulated for 24 hours with the MTB300 peptide

722   pool. IFNγ, IL-2, TNF and IL-17A were measured in the supernatants. (B) and (C) Lung cells

723   were stimulated for 5 hours with the MTB300 peptide pool. Frequencies of CD4 or CD8 T cells

724   that produced IFNγ, IL-2, TNF or IL-17 was determined using intracellular cytokine staining. (A–

725   C) The data represent the mean ± SEM from one experiment. Two-way analysis of variance

726   with the Benjamini and Hochberg multiple comparison method was used to determine

727   significance.

728

729   Figure 6. Multivariate correlates of protection across all 24 CC strains

730   (A) A partial least squares regression (PLSR) model was performed to associate BCG-induced

731   changes in lung immune features with BCG-induced changes in lung CFU. Each data point

732   represents a vaccinated mouse that has the mean of the unvaccinated mice from the same

733   strain subtracted from it. The scores of the top two latent variables of the model are shown, with

734   the color of each data point representing the changed in lung CFU induced by BCG vaccination.

735   The percentages of the variance in the lung features captured by each latent variable are shown

736   as percentages in the axes. (B) The bar graph depicts the loadings of the first latent variable of

737   the PLSR model in (A). Positive loadings indicate immune features that increase from BCG

738   vaccination in samples that are on the positive x-axis in the scores plot, which generally

739   indicates they have worse protection. Negative loadings conversely indicate immune features

740   that generally increase from BCG vaccination in samples that are protected by vaccination. (C)

741   The heatmap illustrates all features selected to be included in the model, indicating they were

742   important for associating strains with the lung protection continuum. Colors are representations

743   of z-scored features, with a different color representing each 10% quantile of the data. Columns

744   are averages of the mice in the CC strains (n=5 for most cases). Feature names are colored

745   based on the dataset from which they come. In (B) and (C), features from the ICS dataset

746   indicate production of the cytokines listed and no production of any measured cytokines not

747   listed.

748

749   Figure 7. Identifying signatures of protection in categories of CC strains

750   (A), (C), (E) Scores plot (top) and heatmaps (bottom) of selected features for PLSDA models.

751   Each dot represents a single mouse. The two colors in the scores plots signify the two classes

752   being compared. An ellipse is drawn around the 95% confidence interval for each class. Percent

753   of variance of the immune feature data is shown on the axes. Heatmaps show features that

754   were most important for distinguishing the two classes. Colors are representations of z-scored

755   features. Columns are averages of the mice in the CC strains, with vaccinated and

756   unvaccinated mice being separated (n=5 for most cases). Feature names are colored based on

757   the dataset from which they come (blue=T cell phenotyping, brown=lung homogenate cytokines,

758   green=ICS, magenta=LEGENDplex). (B), (D), (F) PLSDA models performed as in (A), (C), and

759   (E), except that features are the changes induced by BCG vaccination as in Figure 6. (A) and

760   (B) examine only resistant CC strains, (C) and (D) examine only susceptible CC strains, and (E)

761   and (F) are comparisons of protected resistant and non-protected susceptible mice, which we

762   define as "good" and "poor" immunity, respectively. The color legend at the bottom of the figure

763   indicates the protection/susceptibility quadrant for each CC strain as defined in Figure 1C.

764

765   Figure 8. Th1/17 cells correlate with BCG-mediated protection in a subset of protected CC mice.

766   (A) Four weeks after infection, lung cells were stimulated for 5 hours with anti-CD3 mAb,

767   MTB300 peptides or nothing (media). Flow plots of IFNγ and IL-17 producing CD4 T cells from

768   CC037 and C57BL/6 mice are shown, which is representative of two independent experiments

769   with n=5 mice/group. (B) Frequencies of CD4 T cells that expressed both IFNγ and IL-17 was

770   determined using intracellular cytokine staining. (C) Representative flow plots of CD4 T cells

771   expressing either T-bet or RORγt in CC037 or C57BL/6 mice. (D) Frequencies of CD4 T cells

772   that expressed both T-bet or RORγt in CC037 or C57BL/6 mice. (E) Proportion of CD4 T cells

773   that express various combinations of IFNγ, TNF, IL-2 or IL-17 in either unvaccinated or BCG

774   vaccinated mice at 4 weeks post Mtb infection. Plots are generated using the SPICE platform

775   (34). (F) Volcano plots. Correlation between 60 immune features and the lung CFU determined

776   for CC037 and C57BL/6 mice (from Batch 1). Dotted line, $p = 0.05$. *, $q < 0.05$, corrected with

777   the Benjamini and Hochberg multiple comparison method (FDR = 0.05). (G) Correlation of

778   IFNγ$^+$IL-17$^+$ cells and bacterial burden in the lungs of either unvaccinated or BCG vaccinated

779   CC037 mice at 4 weeks post Mtb infection. Colors represent BCG vaccinated mice (red) or

780   unvaccinated mice (blue). Spearman r and p value are shown. Data is pooled from two

781   experiments. (B, D) Data is representative of two independent experiments with n=5

782   mice/group. One-way analysis of variance corrected with the Benjamini and Hochberg multiple

783   comparison method (FDR = 0.05).

SUPPLEMENTAL INFORMATION

Supplementary tables

Table S1. A description of the CC strains analyzed in this study.

The 24 CC strains were initially analyzed in 4 different batches as indicated in A, after which select CC strains were repeated to confirm the initial findings and also for more in-depth immune analysis. MHC haplotype that each CC strain inherited is indicated in B. The reduction in Mtb burden following BCG vaccination, as indicated by ΔLung and ΔSpleen CFU, are listed for both the initial analysis by columns C through F, with values listed for the initial cohort (C and E) as well as for the repeat experiments (D and F). Note that only 13 of 24 strains have been repeated for in-depth analysis. Red text indicates experiments where BCG enhanced Mtb burden. Highlight samples indicate discordance between the initial experiment and the repeat experiment. C57BL/6 mice were included in all experiments as a control.

Table S2. Pathological descriptions of changes lung histology after BCG vaccination in CC mice.

A summary of the pathological features in the CC strains analyzed. A) Nineteen of the 24 CC strains that were available for analysis are indicated here. The ability of BCG to protect each strain (as defined in Figure 1) is indicated by B (N = not protected, P = protected). Histological sections were also compared between unvaccinated and BCG vaccinated mice following Mtb infection (as indicated in C) to see if changes in pathology could be detect (N = no, Y = yes). A description of the histological changes between PBS and BCG vaccinated mice are indicated in D.

Supplementary figure legends

810   Figure S1. Gating strategy for the analysis of the cell surface T cell phenotype.

811   (A) Lymphocytes were identified using FSC and SSC properties, after which viability dyes were

812   used to exclude dead cells. Singlet gates were used to exclude doublets, after which T cells and

813   non-T cells were identified using the lineage marker CD3. (B) NKT and MAIT cells were

814   identified from the CD3$^+$ gate using antigen-loaded MR1 tetramers, respectively. CD4$^+$, CD8$^+$

815   and CD4$^-$CD8$^-$ T cells were identified in the T cell population. (C) Expression of markers were

816   measured within the CD3$^-$ population. An example is the flow plot showing T-bet expression

817   within the CD3$^-$ population. (D) CD44 was used to identify antigen-experienced T cells.

818   Resident memory T cells (T$_{RM}$) were identified using CD103 expression. From the CD103$^-$

819   fraction, CD127 and CD62L expression identified central memory (T$_{CM}$), effector (T$_E$), and

820   effector memory (T$_{EM}$). Subset identification of various T cell subsets (ex: Tbet$^+$CD4$^+$ T cells)

821   and activation markers (ex: PD1, CD11a) were assessed among CD4$^+$, CD8$^+$ and CD4$^-$CD8$^-$ T

822   cell subsets. (E) Naïve T cells were identified within the CD44$^-$ compartment using CD62L

823   expression.

824

825   Figure S2. Frequencies of non-conventional T cells in CC mice. Frequencies of MR-1-restricted

826   MAIT cells and CD1d-restricted NKT cells were measured in the lungs of unvaccinated or BCG

827   vaccinated CC mice at 4 weeks post Mtb infection. The data in the graphs represent the mean ±

828   SD of one experiment. Two-way analysis of variance with the Benjamini and Hochberg multiple

829   comparsion method. The FDR was set to 0.05 and the numbers in the figures are the q value.

830

831   Figure S3. Gating strategy to analyze intracellular cytokine production by T cells.

832   (A) Lymphocytes were identified by their FSC and SSC properties, after which viability dyes

833   were used to exclude dead cells. Singlet gates were used to exclude doublets, after which T

834   cells were identified using the lineage marker CD3. (B) CD4, CD8 and CD4$^-$CD8$^-$ T cells were

835   identified in the CD3$^+$ population. Cytokine expression was determined in all 3 T cell sub-

836   populations. (C) Cytokine expression was also examined within the CD3$^-$ population.

837

838   Figure S4. Multivariate model accuracies.

839   (A) The accuracy of our protection model (Figure 6) was compared to that of two different types

840   of null models, where our measure of accuracy was mean squared error of prediction of Δlog

841   (lung CFU). Our PLSR model performed significantly better than both a null model consisting of

842   random features instead of our elastic-net chosen features and a null model where the

843   outcomes of Δlog (lung CFU) were scrambled. Our PLSR model and each of the null models

844   were run 50 times, and each dot represents one of these trials. Labeled p-values are from one-

845   way ANOVAs.

846   (B)   Our   models   highlighting   the   immune   features   that   distinguish   pairs   of

847   protection/susceptibility categories (Figure 7) were compared to two null models as in (A). Each

848   of our models, both the ones involving all mice and the protection models were significantly

849   better than both null models. Here, the accuracy metric is classification accuracy. Labeled p-

850   values are from one-way ANOVAs.

851

852   Figure S5. CD4 T cell polyfunctionality in select CC strains. Boolean gating was used to identify

853   sub-populations among CD4 T cells following MTB300 stimulation. An increase in IFNγ$^+$IL-17$^+$

854   CD4 T cells (that also produced TNF or IL-2) were detected in CC037, CC072, and CC001 but

855   not in C57BL/6. In contrast, a noticeable decrease in the frequency of IFNγ "single positive"

856   CD4 T cells was observed in all strains including C57BL/6. The data in the graphs represent the

857   mean ± SD of two experiment. One-way analysis of variance with corrected with the Benjamini

858   and Hochberg multiple comparison method (FDR = 0.05). q values are shown.

859

860   References
861

862   1.   WHO. Global Tuberculosis Report 2021. 2021.

863   2.   Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV, et al. Efficacy
864   of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published
865   literature. *JAMA.* 1994;271(9):698-702.

866   3.   Roy A, Eisenhut M, Harris RJ, Rodrigues LC, Sridhar S, Habermann S, et al. Effect of
867   BCG vaccination against Mycobacterium tuberculosis infection in children: systematic
868   review and meta-analysis. *BMJ.* 2014;349:g4643.

869   4.   Mangtani P, Abubakar I, Ariti C, Beynon R, Pimpin L, Fine PE, et al. Protection by BCG
870   vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clin*
871   *Infect Dis.* 2014;58(4):470-80.

872   5.   Abubakar I, Pimpin L, Ariti C, Beynon R, Mangtani P, Sterne JA, et al. Systematic review
873   and meta-analysis of the current evidence on the duration of protection by bacillus
874   Calmette-Guerin vaccination against tuberculosis. *Health Technol Assess.*
875   2013;17(37):1-372, v-vi.

876   6.   Fine PE. Variation in protection by BCG: implications of and for heterologous immunity.
877   *Lancet.* 1995;346(8986):1339-45.

878   7.   Flynn JL, Gideon HP, Mattila JT, and Lin PL. Immunology studies in non-human primate
879   models of tuberculosis. *Immunol Rev.* 2015;264(1):60-73.

880   8.   Ahmed M, Thirunavukkarasu S, Rosa BA, Thomas KA, Das S, Rangel-Moreno J, et al.
881   Immune correlates of tuberculosis disease and risk translate across species. *Sci Transl*
882   *Med.* 2020;12(528).

883   9.   Medina E, and North RJ. Resistance ranking of some common inbred mouse strains to
884   Mycobacterium tuberculosis and relationship to major histocompatibility complex
885   haplotype and Nramp1 genotype. *Immunology.* 1998;93(2):270-4.

886   10.   Aagaard C, Knudsen NPH, Sohn I, Izzo AA, Kim H, Kristiansen EH, et al. Immunization
887   with Mycobacterium tuberculosis-Specific Antigens Bypasses T Cell Differentiation from
888   Prior Bacillus Calmette-Guerin Vaccination and Improves Protection in Mice. *J Immunol.*
889   2020;205(8):2146-55.

890   11.   Woodworth JS, Clemmensen HS, Battey H, Dijkman K, Lindenstrom T, Laureano RS, et
891   al. A Mycobacterium tuberculosis-specific subunit vaccine that provides synergistic
892   immunity upon co-administration with Bacillus Calmette-Guerin. *Nat Commun.*
893   2021;12(1):6658.

894   12.   Kramnik I, and Beamer G. Mouse models of human TB pathology: roles in the analysis
895   of necrosis and the development of host-directed therapies. *Semin Immunopathol.*
896   2016;38(2):221-37.

897   13.   Nemeth J, Olson GS, Rothchild AC, Jahn AN, Mai D, Duffy FJ, et al. Contained
898   Mycobacterium tuberculosis infection induces concomitant and heterologous protection.
899   *PLoS Pathog.* 2020;16(7):e1008655.

900   14.   Noll KE, Ferris MT, and Heise MT. The Collaborative Cross: A Systems Genetics
901   Resource for Studying Host-Pathogen Interactions. *Cell Host Microbe.* 2019;25(4):484-
902   98.

903   15.   Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, et al. The Collaborative
904   Cross, a community resource for the genetic analysis of complex traits. *Nat Genet.*
905   2004;36(11):1133-7.

906   16.   Smith CM, Proulx MK, Olive AJ, Laddy D, Mishra BB, Moss C, et al. Tuberculosis
907   Susceptibility and Vaccine Protection Are Independently Controlled by Host Genotype.
908   *MBio.* 2016;7(5).

909   17.   Smith CM, Baker RE, Proulx MK, Mishra BB, Long JE, Park SW, et al. Host-pathogen
910   genetic interactions underlie tuberculosis susceptibility in genetically diverse mice. *Elife.*
911   2022;11.

912   18.   Lee J, Boyce S, Powers J, Baer C, Sassetti CM, and Behar SM. CD11cHi monocyte-
913   derived macrophages are a major cellular compartment infected by Mycobacterium
914   tuberculosis. *PLoS Pathog.* 2020;16(6):e1008621.

915   19.   Sakai S, Kauffman KD, Schenkel JM, McBerry CC, Mayer-Barber KD, Masopust D, et al.
916   Cutting edge: control of Mycobacterium tuberculosis infection by a subset of lung
917   parenchyma-homing CD4 T cells. *J Immunol.* 2014;192(7):2965-9.

918   20.   Smith CM, Proulx MK, Lai R, Kiritsy MC, Bell TA, Hock P, et al. Functionally Overlapping
919   Variants Control Tuberculosis Susceptibility in Collaborative Cross Mice. *mBio.*
920   2019;10(6).

921   21.   Lindestam Arlehamn CS, McKinney DM, Carpenter C, Paul S, Rozot V, Makgotlho E, et
922   al. A Quantitative Analysis of Complexity of Human Pathogen-Specific CD4 T Cell
923   Responses in Healthy M. tuberculosis Infected South Africans. *PLoS Pathog.*
924   2016;12(7):e1005760.

925   22.   Patankar YR, Sutiwisesak R, Boyce S, Lai R, Lindestam Arlehamn CS, Sette A, et al.
926   Limited recognition of Mycobacterium tuberculosis-infected macrophages by polyclonal
927   CD4 and CD8 T cells from the lungs of infected mice. *Mucosal Immunol.*
928   2020;13(1):140-8.

929   23.   Kipnis A, Irwin S, Izzo AA, Basaraba RJ, and Orme IM. Memory T lymphocytes
930   generated by Mycobacterium bovis BCG vaccination reside within a CD4 CD44lo CD62
931   ligand(hi) population. *Infect Immun.* 2005;73(11):7759-64.

932   24.   Lefford MJ. Transfer of adoptive immunity to tuberculosis in mice. *Infection and Immunity.*
933   1975;11(6):1174-81.

25. Foster M, Hill PC, Setiabudiawan TP, Koeken V, Alisjahbana B, and van Crevel R. BCG-induced protection against Mycobacterium tuberculosis infection: Evidence, mechanisms, and implications for next-generation vaccines. *Immunol Rev.* 2021;301(1):122-44.

26. Steigler P, Verrall AJ, and Kirman JR. Beyond memory T cells: mechanisms of protective immunity to tuberculosis infection. *Immunol Cell Biol.* 2019;97(7):647-55.

27. Chen L, Wang J, Zganiacz A, and Xing Z. Single intranasal mucosal Mycobacterium bovis BCG vaccination confers improved protection compared to subcutaneous vaccination against pulmonary tuberculosis. *Infect Immun.* 2004;72(1):238-46.

28. Lindestam Arlehamn CS, Gerasimova A, Mele F, Henderson R, Swann J, Greenbaum JA, et al. Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. *PLoS Pathog.* 2013;9(1):e1003130.

29. Acosta-Rodriguez EV, Rivino L, Geginat J, Jarrossay D, Gattorno M, Lanzavecchia A, et al. Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. *Nat Immunol.* 2007;8(6):639-46.

30. Gideon HP, Hughes TK, Tzouanas CN, Wadsworth MH, 2nd, Tu AA, Gierahn TM, et al. Multimodal profiling of lung granulomas in macaques reveals cellular correlates of tuberculosis control. *Immunity.* 2022;55(5):827-46 e10.

31. Boisson-Dupuis S, Ramirez-Alejo N, Li Z, Patin E, Rao G, Kerner G, et al. Tuberculosis and impaired IL-23-dependent IFN-γ immunity in humans homozygous for a common TYK2 missense variant. *Sci Immunol.* 2018;3(30).

32. Boisson-Dupuis S, Bustamante J, El-Baghdadi J, Camcioglu Y, Parvaneh N, El Azbaoui S, et al. Inherited and acquired immunodeficiencies underlying tuberculosis in childhood. *Immunol Rev.* 2015;264(1):103-20.

33. Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H, Kobzik L, et al. Ipr1 gene mediates innate immunity to tuberculosis. *Nature.* 2005;434(7034):767-72.

34. Roederer M, Nozzi JL, and Nason MC. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytometry A.* 2011;79(2):167-74.
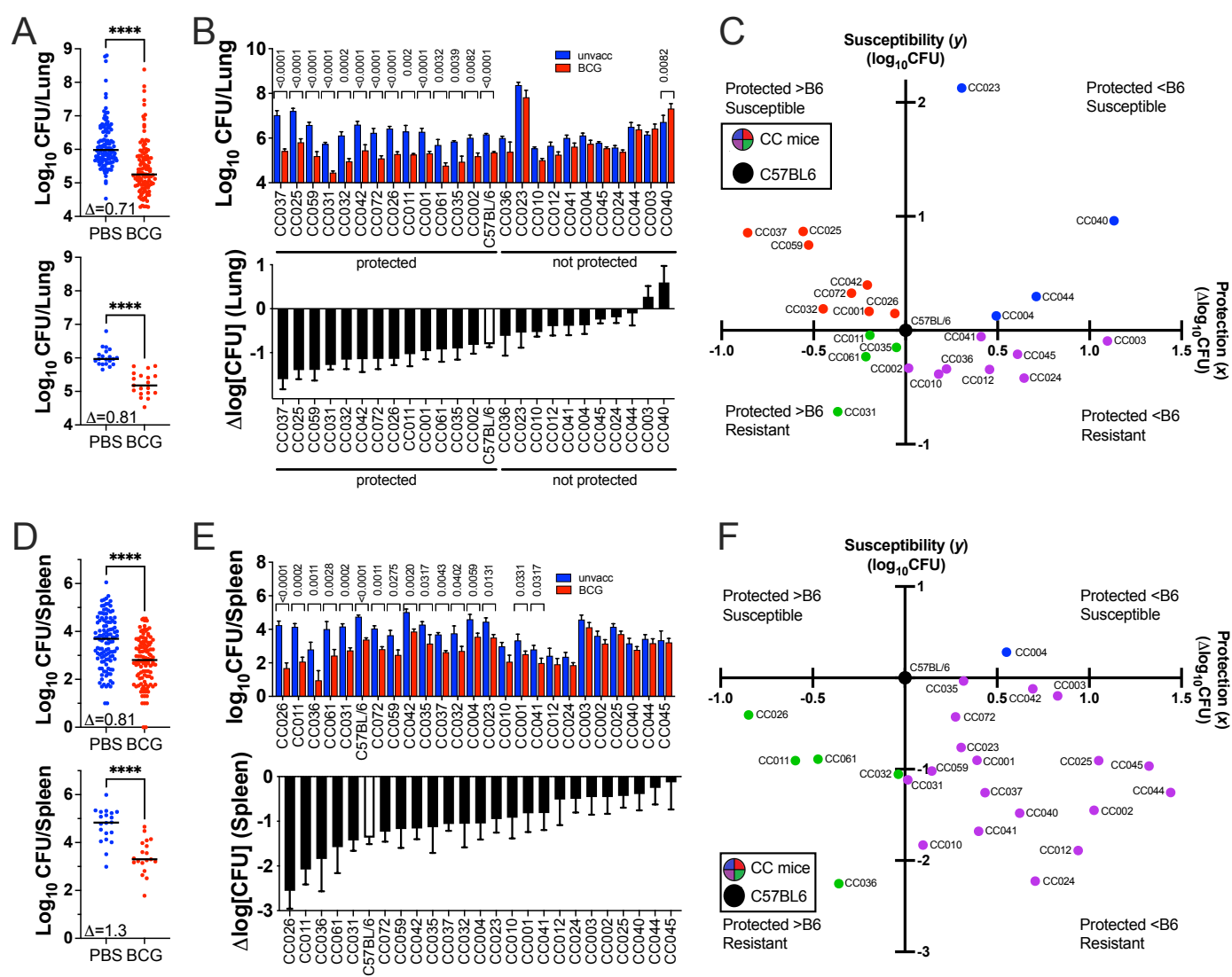
Mtb	BCG+Mtb	Mtb	BCG+Mtb
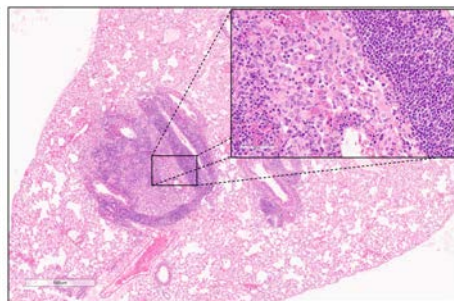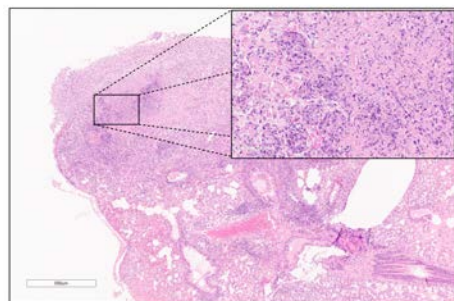
CC025
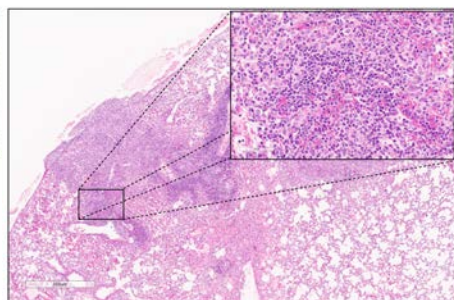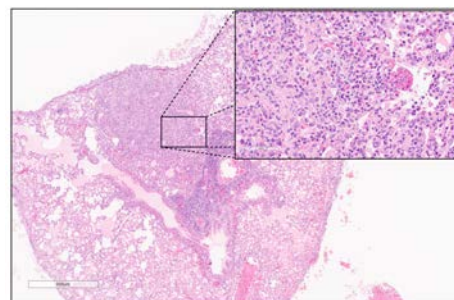
CC037
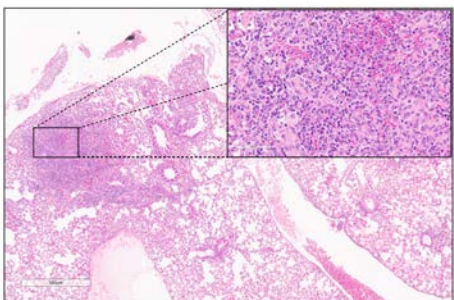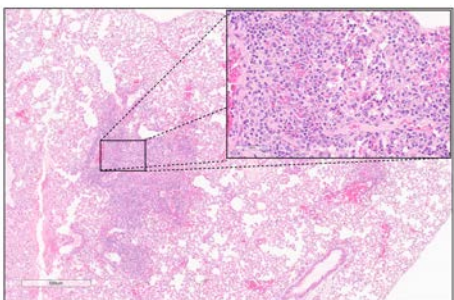
CC072

Protected
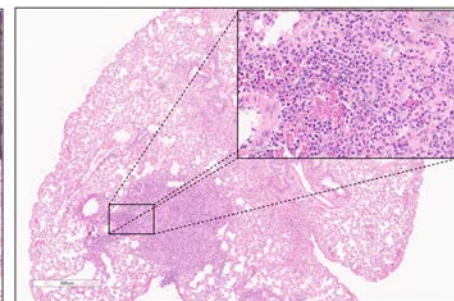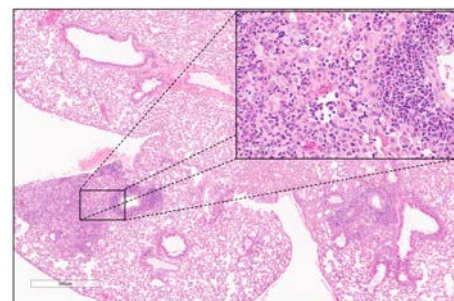
CC003
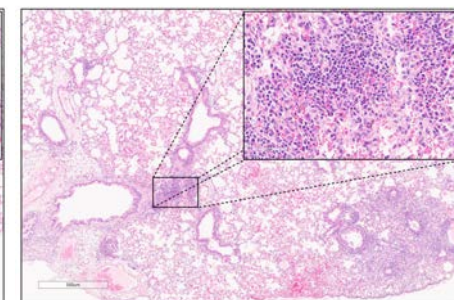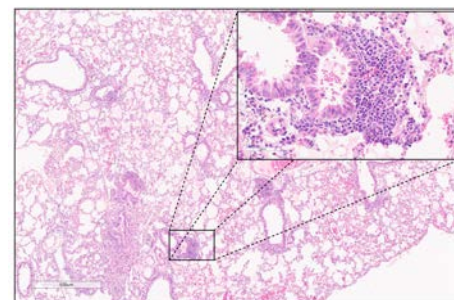
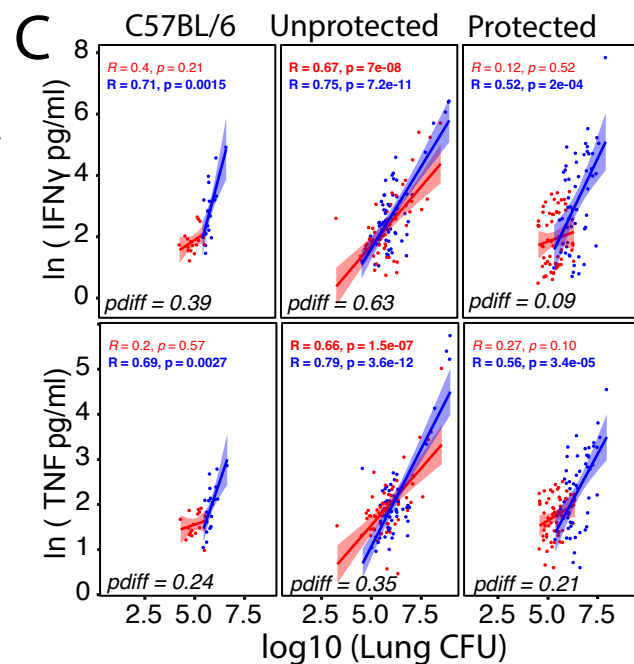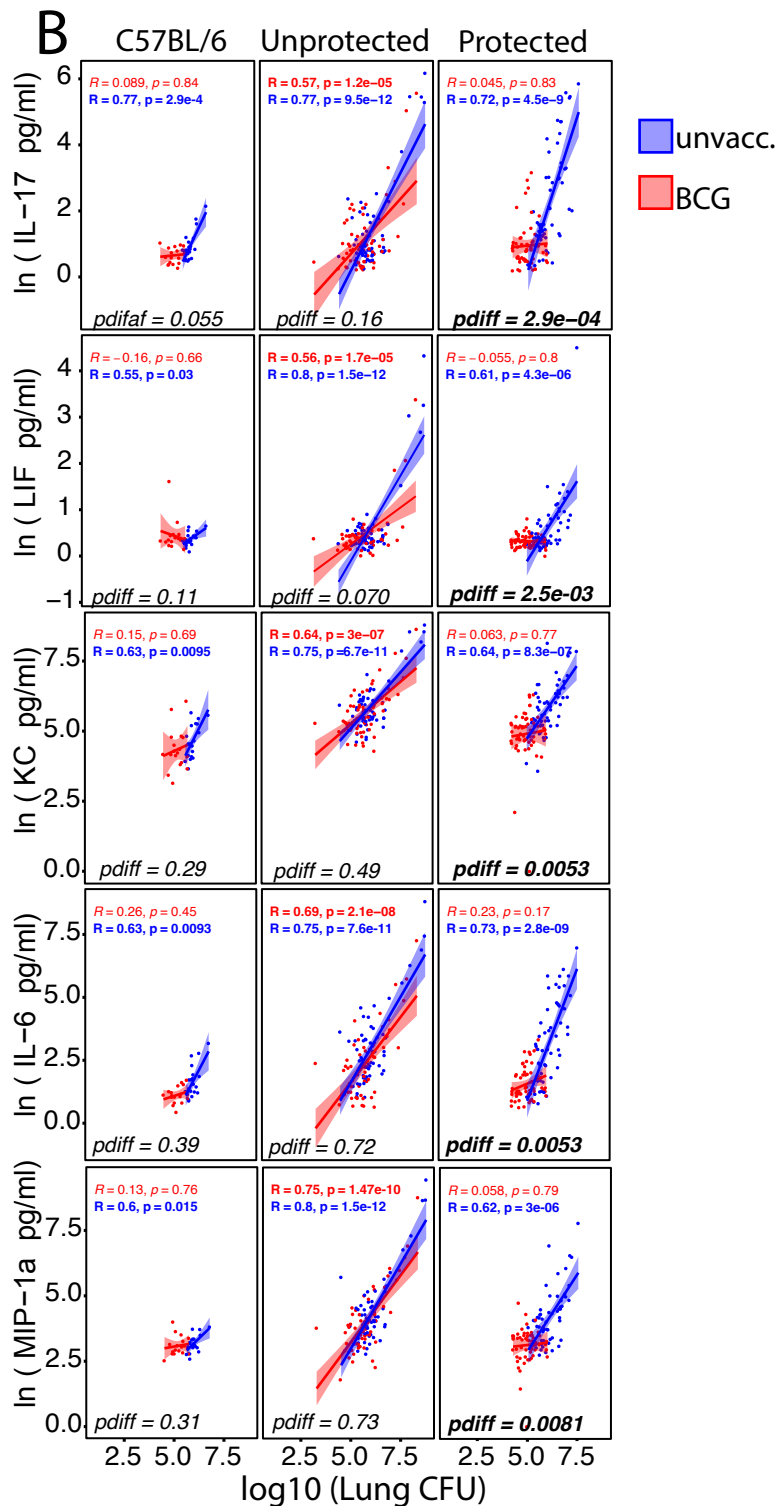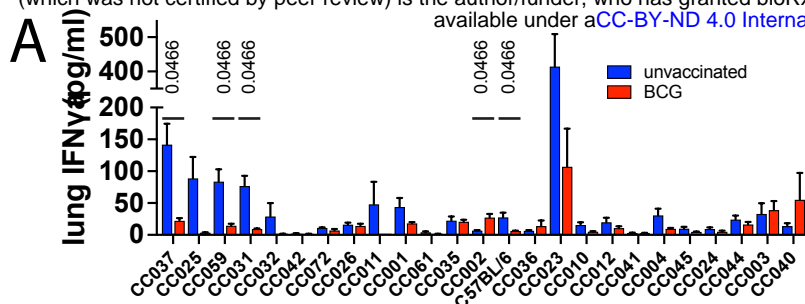CC024
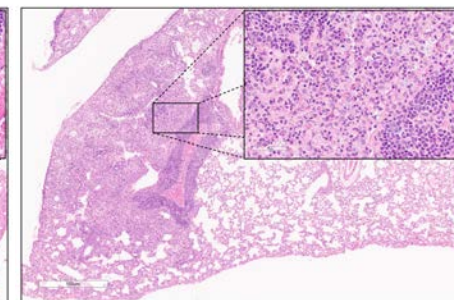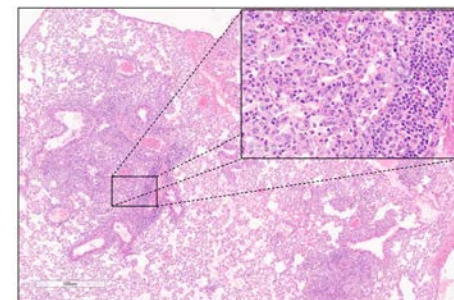
CC004

Not Protected
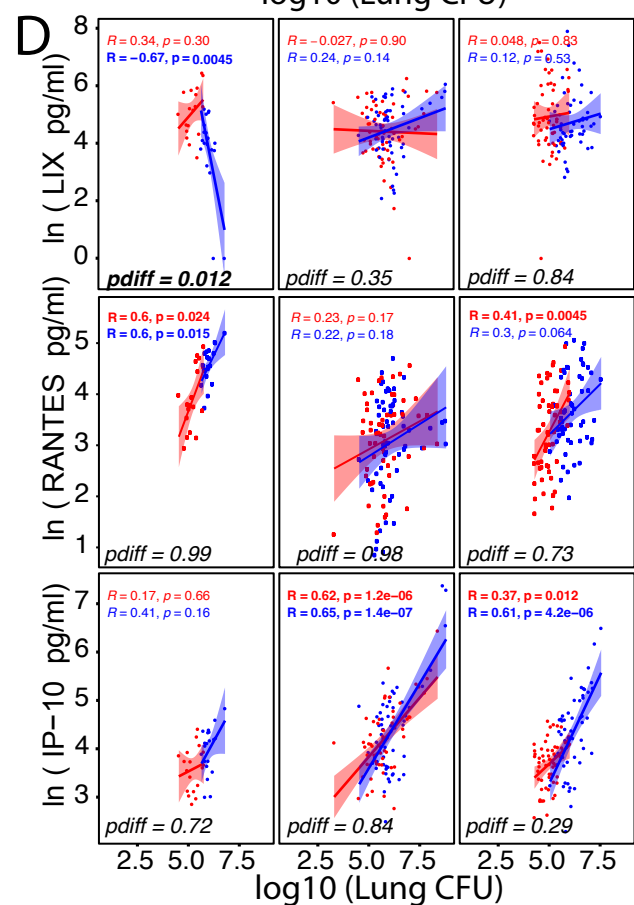
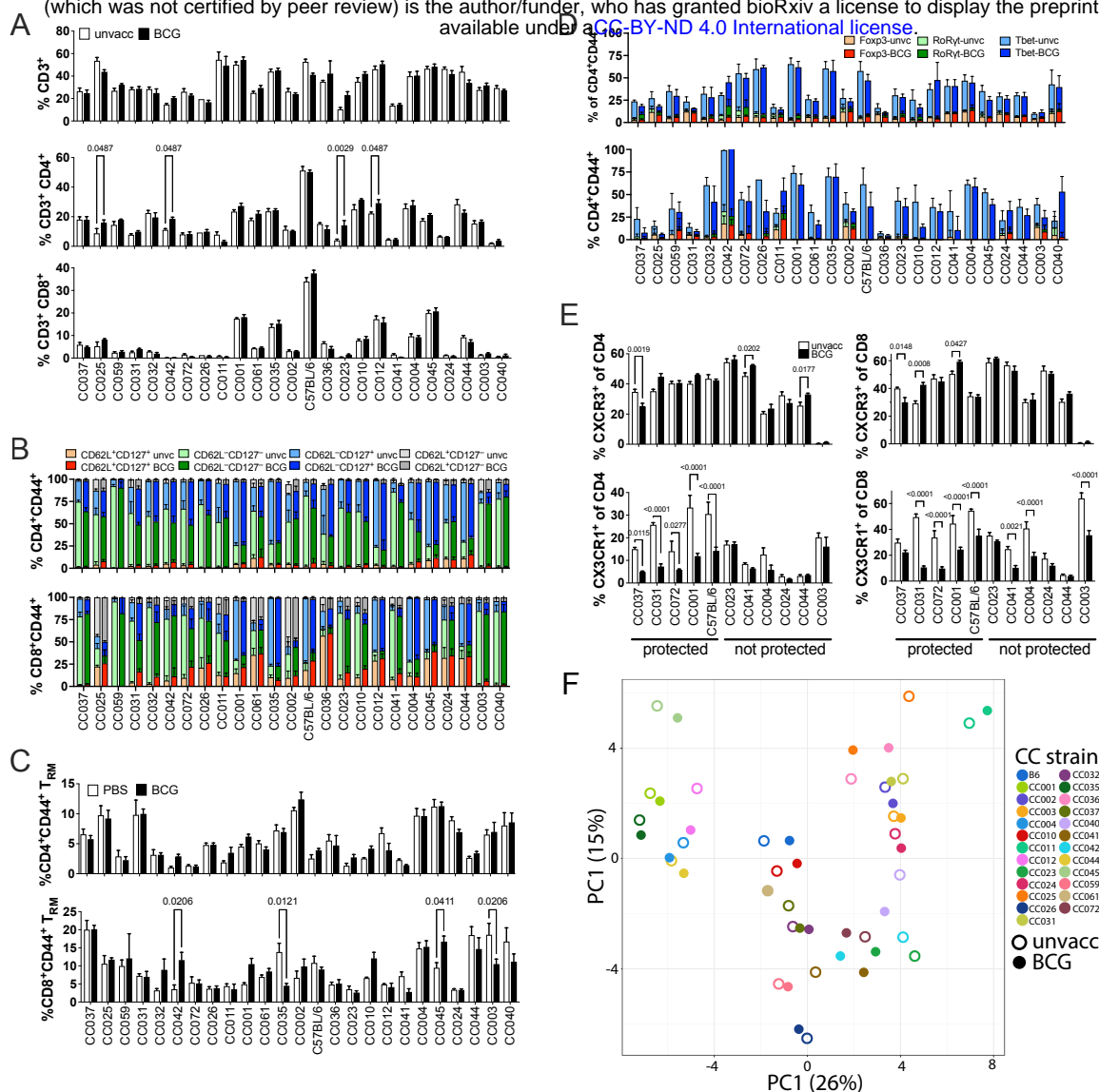Figure 4

Figure 5
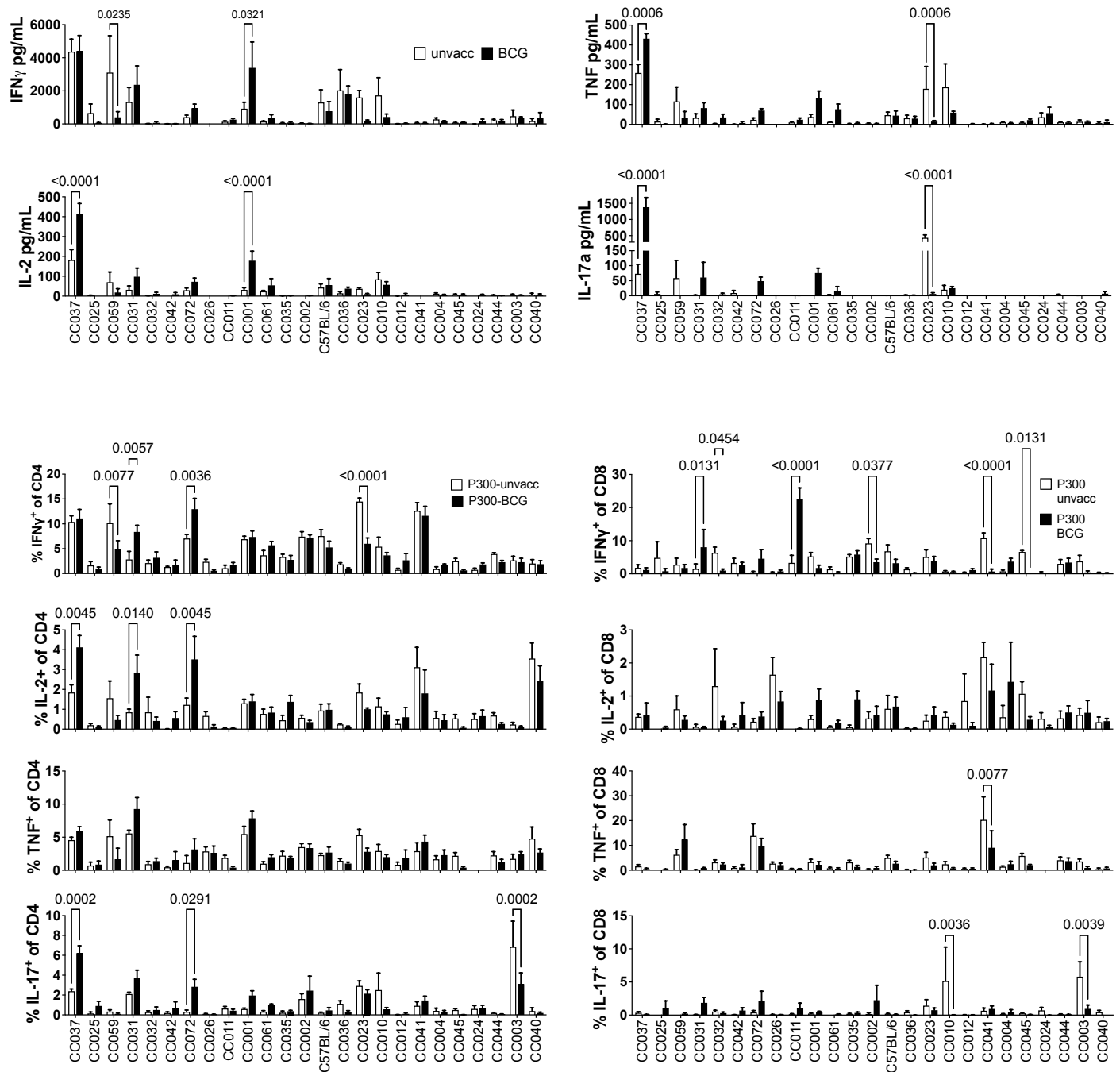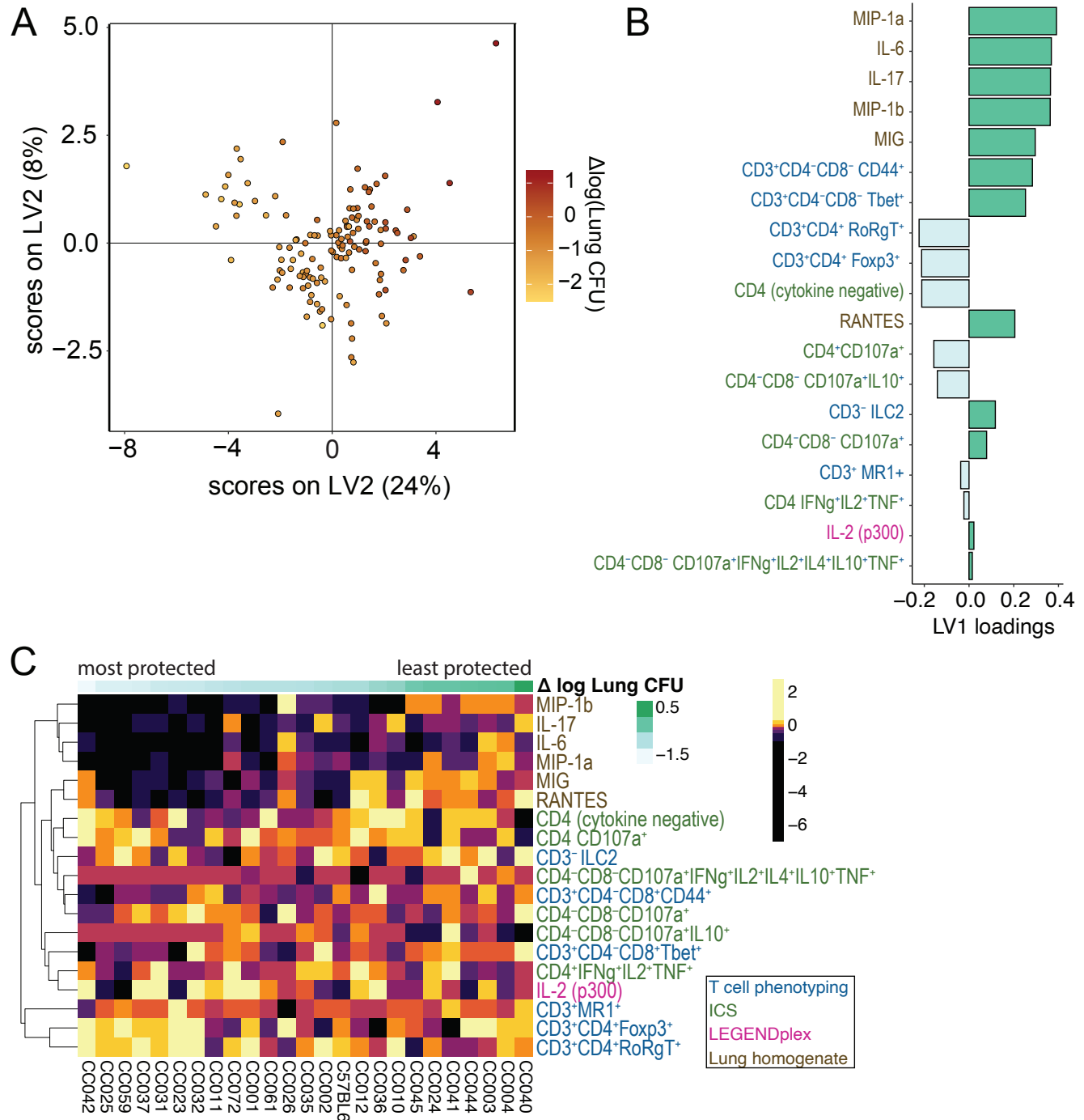
Figure 6
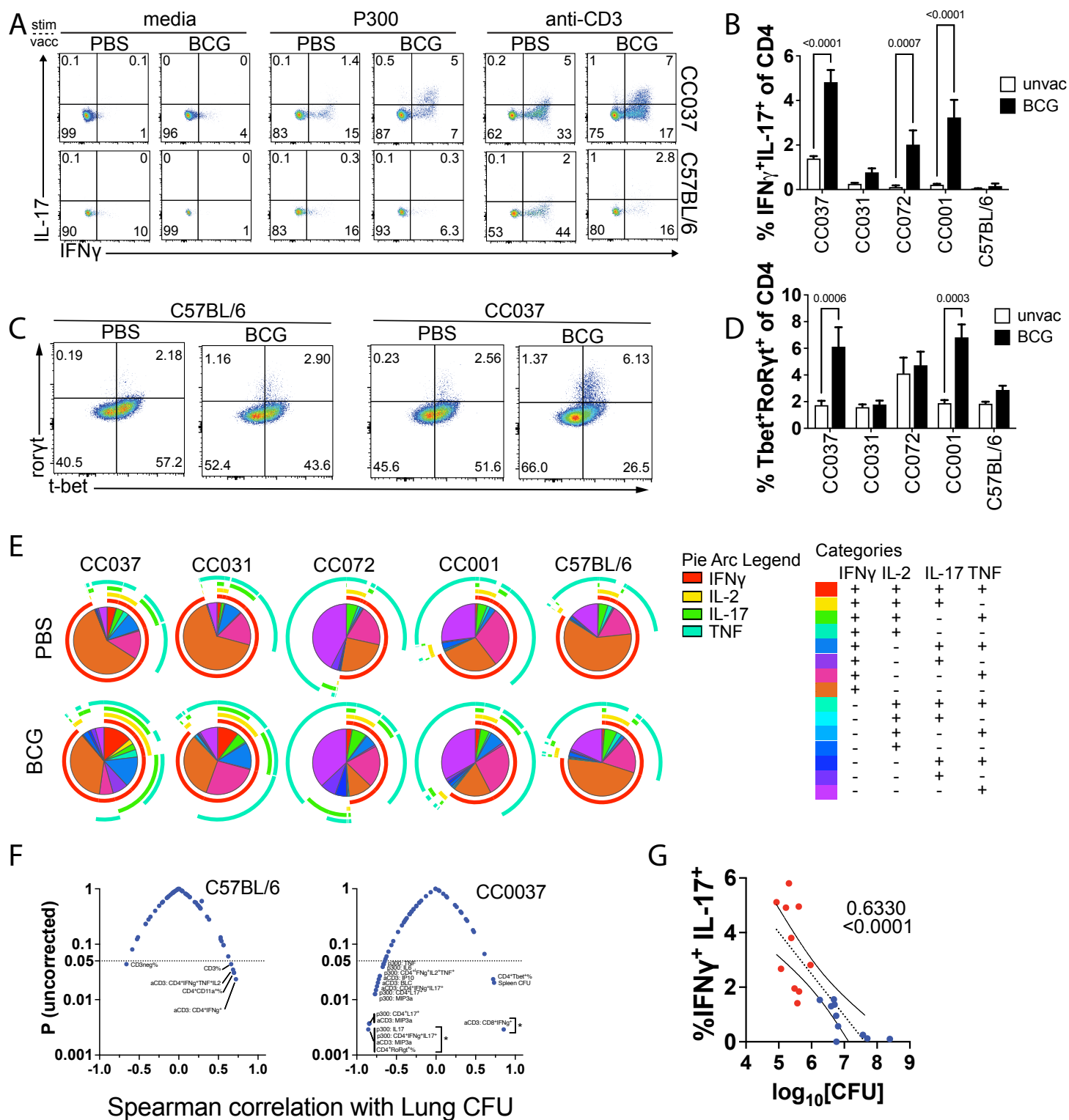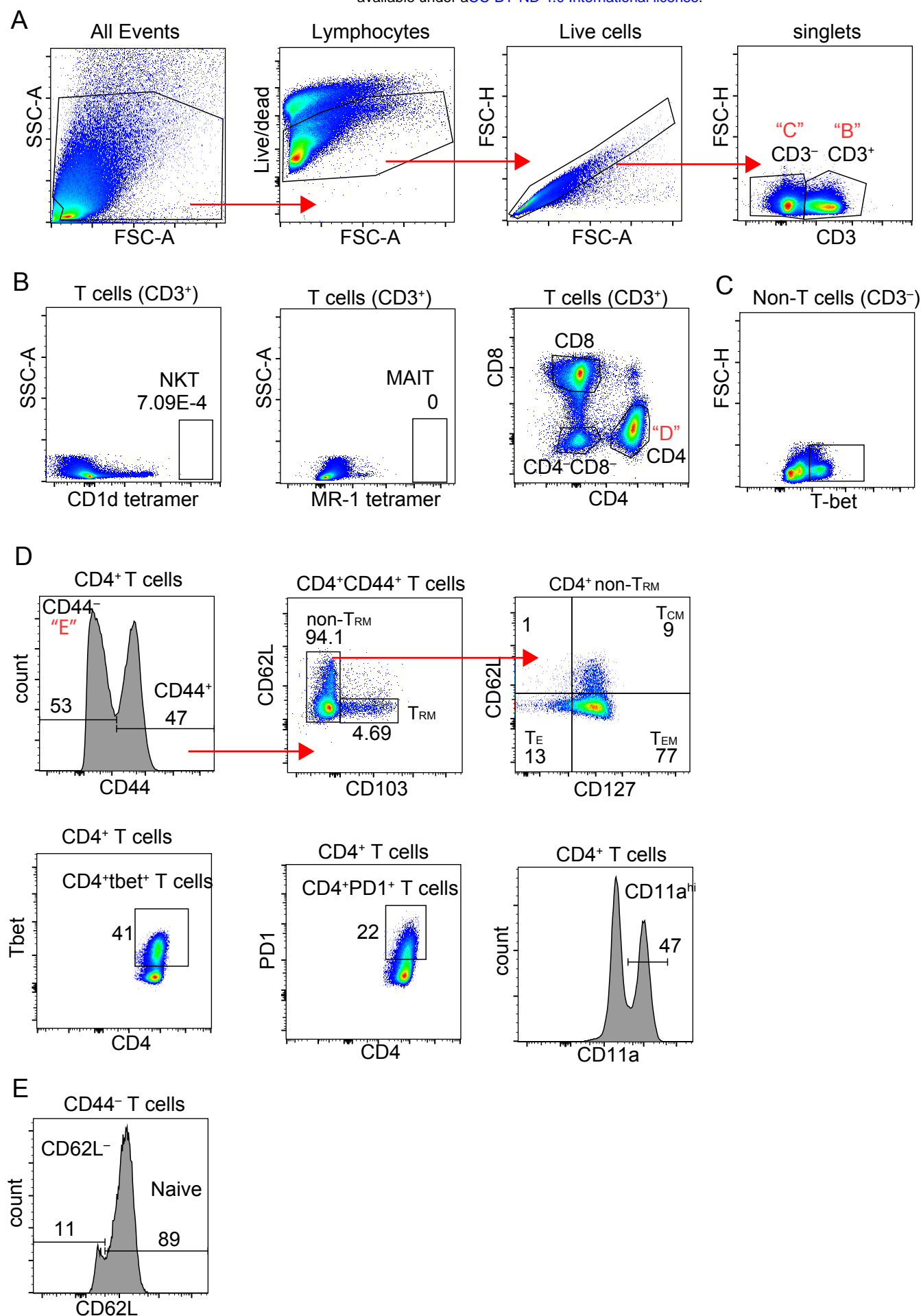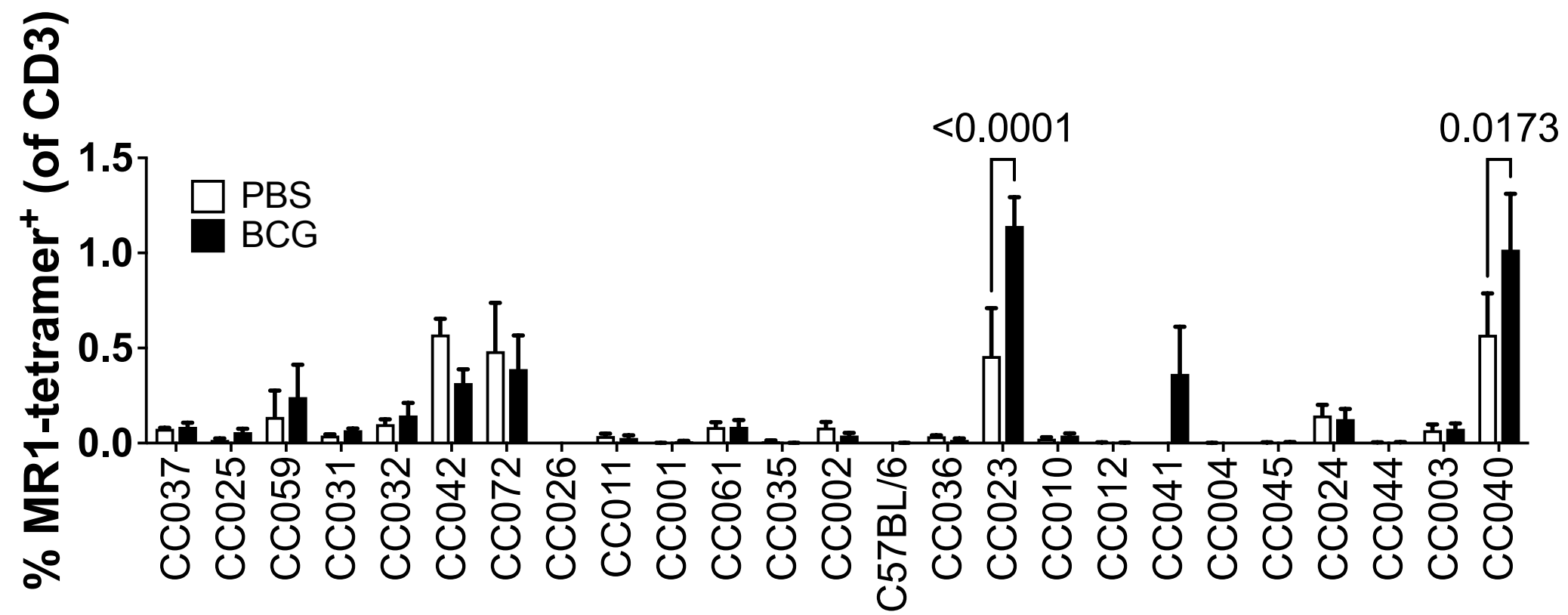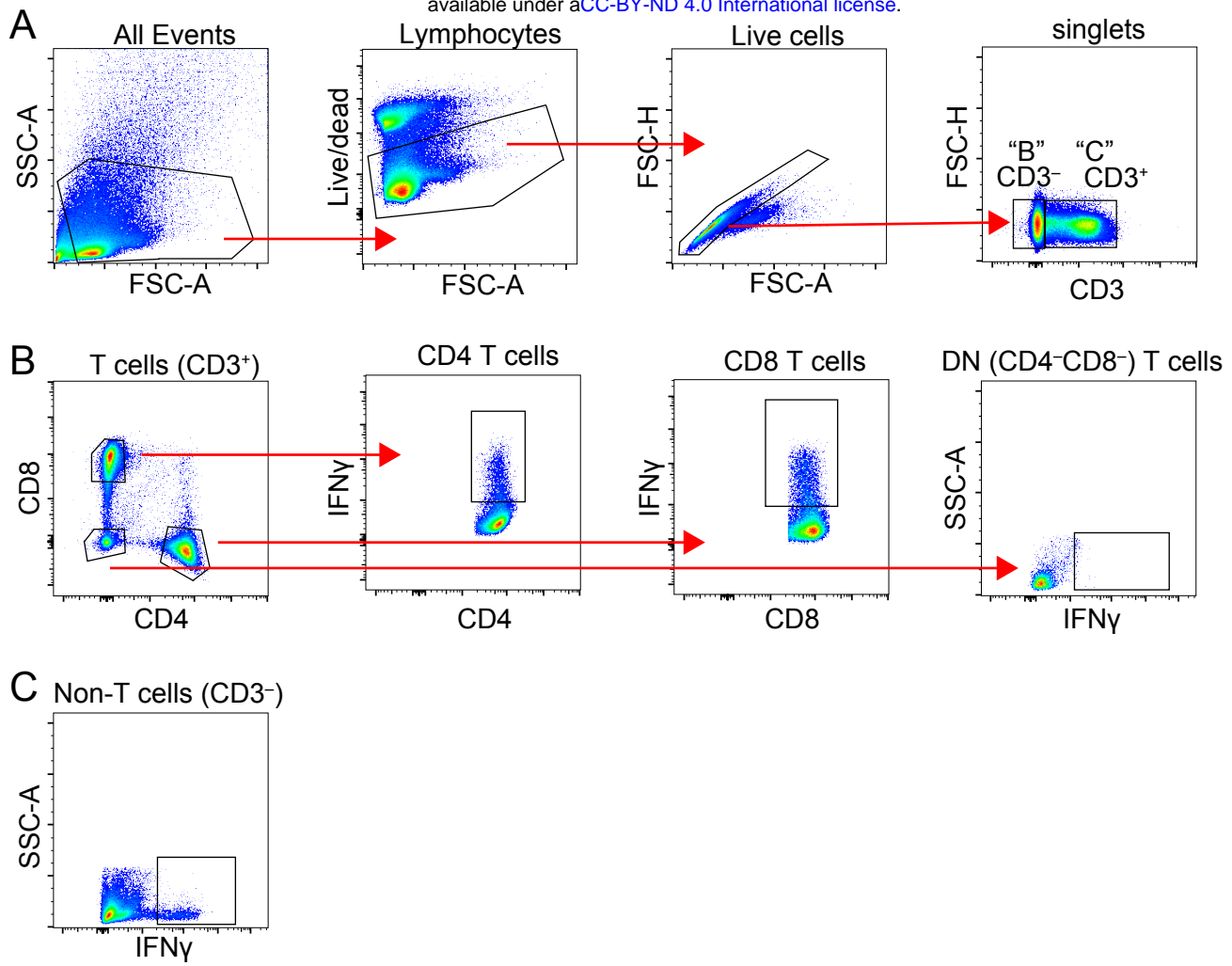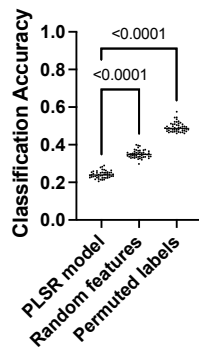
Spearman correlation with Lung CFU

IFNγ	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−
IL-2	+	+	+	+	−	−	+	+	+	+	+	+	−	−	−
IL-17	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−
TNF	+	−	+	−	+	+	+	−	+	−	+	−	+	−	+
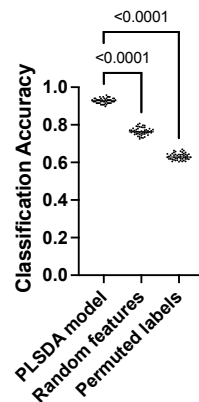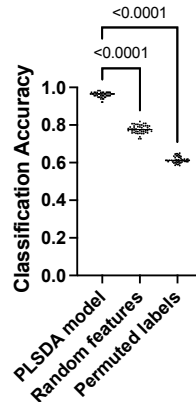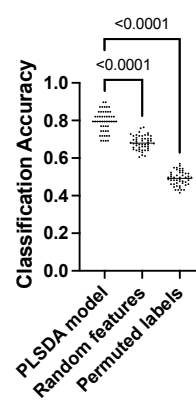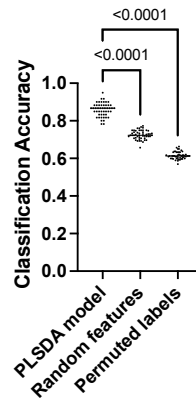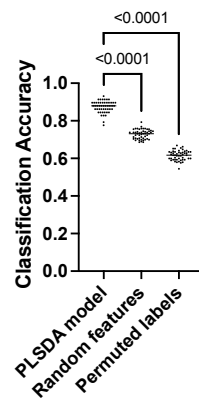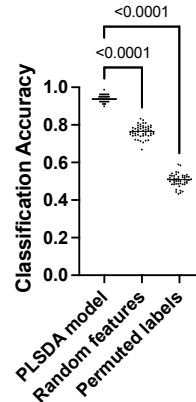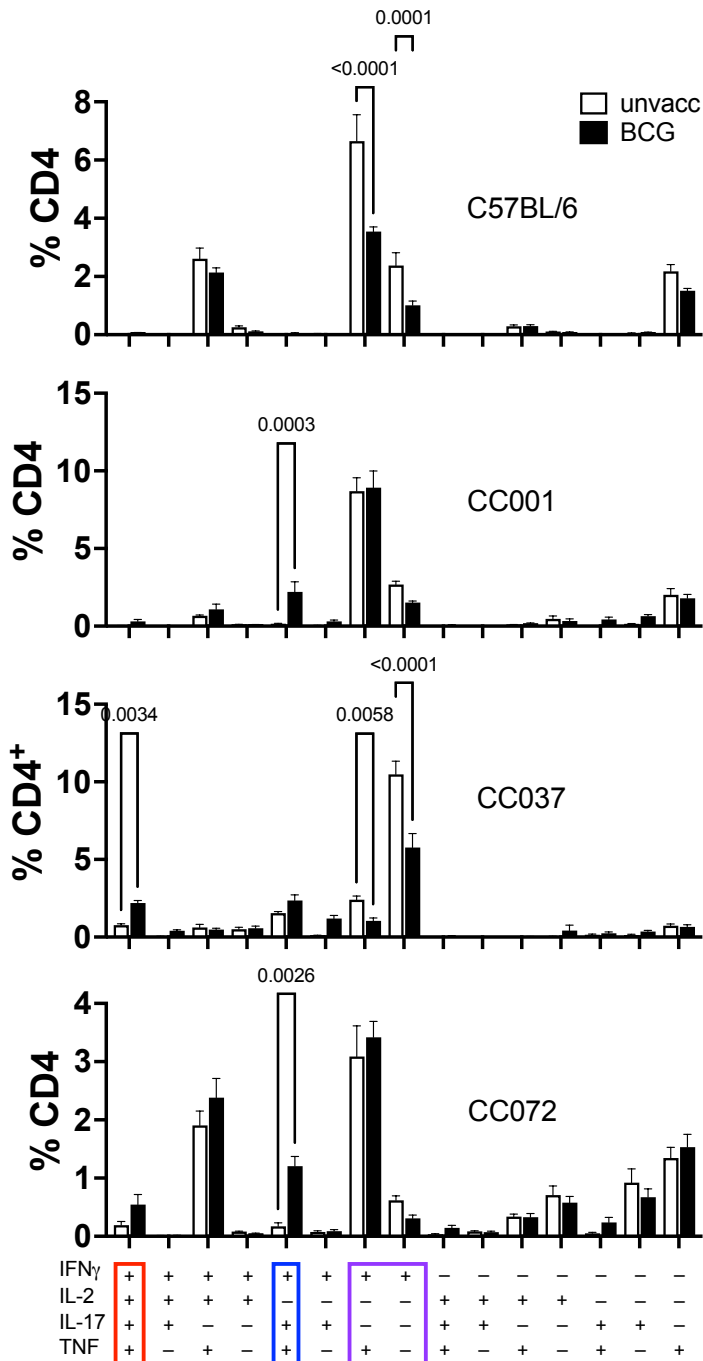
Table S1. A description of the CC strains analyzed in this study						
CC strain	Batch[a]	MHC Haplotype[b]	ΔLung CFU$_1$, [c] reduction (logCFU)	ΔLung CFU$_2$, [d] reduction (logCFU)	ΔSpleen CFU$_1$, [e] reduction (logCFU)	ΔSpleen CFU$_2$, [f] reduction (logCFU)
CC042/GeniUnc	1	b (129)	1.65	0.96	1.00	1.32
CC025/GeniUnc	3	pwk	1.40	n/a	0.43	n/a
CC059/TauUnc	2	b (129)/a	1.39	n/a	1.17	n/a
CC037/TauUnc	1	b (B6)	1.37	1.84	1.07	1.05
CC031/GeniUnc	3	b (B6)	1.22	1.35	1.26	1.60
CC023/GeniUnc	1	b (129)	1.17	-0.09	1.47	0.32
CC032/GeniUnc	1	b (B6)	1.16	n/a	1.06	n/a
CC011/Unc	3	g7	1.04	n/a	2.08	n/a
CC072/TauUnc	2	b (B6)	0.97	1.37	1.06	1.41
CC001/Unc	4	b (B6)	0.94	1.45	2.79	1.45
CC026/GeniUnc	2	cast	0.93	1.05	3.75	1.87
CC061/GeniUnc	1	b (129)	0.93	n/a	1.58	n/a
CC035/Unc	4	wsb	0.91	n/a	1.13	n/a
CC002/Unc	3	b (129)	0.83	n/a	0.46	n/a
CC012/GeniUnc	4	wsb	0.79	n/a	0.89	n/a
CC036/Unc	3	g7	0.62	n/a	1.84	n/a
CC010/GeniUnc	1	b (B6)	0.53	n/a	0.92	n/a
CC045/GeniUnc	4	z	0.24	n/a	0.13	n/a
CC024/GeniUnc	2	cast	0.15	0.24	0.38	0.62
CC041/TauUnc	2	b (129)	0.14	0.94	0.71	0.94
CC044/Unc	4	b (B6)/g7	0.08	0.13	0.12	0.39
CC003/Unc	3	g7	0.07	-0.61	0.77	0.15
CC004/TauUnc	4	b (129)	0.05	0.08	-0.16	-0.12
CC040/TauUnc	2	b (B6)	-0.64	-0.28	-0.35	1.14

a)   Indicates batch that the CC strain were part of for the initial round of immune characterization

b)   Indicates MHC haplotype of each CC strain

c, e)   Lung and Spleen CFU of the initial round of CC infection

d, f)   Lung and Spleen CFU of the repeat cohort

[a] Strain	[b] BCG-induced protection	[c] Change in pathology	[d] How did the vaccinated mice differ?
CC003	N	N	No differences. Both groups had mild lesions with perivascular lymphocytic cuffs and a few macrophage-rich lesions
CC041	N	N	No differences. Both groups had mild lesions with perivascular lymphocytic cuffs and a few macrophage-rich lesions
CC024	N	N	No differences. Both groups had small hard-to-find lymphocytic lesions with perivascular and peribroncial lymphocytic cuffs
CC002	N	Y	Ctrl had mild lesions with perivascular lymphocytic cuffs. BCG mice had larger moderate sized macrophage-rich lesions. Lymphocytes were similarly abundant.
CC004	N	Y	Ctrl mice had mild to moderate lesions with lymphohistiocytic infiltrates. BCG mice have more lymphocytes with regions of dense perivascular/peribronchiolar lymphocytic cuffs.
CC036	N	Y	Ctrl mice have mild disease with lymphohistiocytic infiltrates. BCG mice also have lymphohistiocytic infiltrates with few very dense lymphocytic areas
CC040	N	Y	Ctrl mice have moderate necrotizing lesions with few lymphocytes. BCG mice have mild disease - small & hard to find lesions that are lymphocyte rich without necrosis.
CC044	N	Y	Unusual. Ctrl mice have lymphohistiocytic infiltrates with dense lymphocytic cuffs. BCG mice have lymphohistiocytic infiltrates with dense lymphocytic cuffs and lymphoid aggregates/follicules in parenchyma.
CC061	P	N	No differences. Unusual. Mild lesions. Both groups have dense lymphoid aggregates/follicles in parenchyma. Also, perivascular lymphocytic cuffs.
CC072	P	N	No differences. Mild disease. perivascular lymphocytic cuffs with a few macrophage-rich lesions. One mouse has a foreign-body granulomatous bronchitis/pneumonia with intralesional food/bedding/plant.
C57BL/6	P	Y	Ctrl mice had mild to moderate lymphohistiocytic lesions. BCG mice had with more, dense lymphocytic infiltrates
CC001	P	Y	Ctrl mice had moderate lesions with lymphohistiocytic infiltrates and equivocal necrosis. BCG mice had mild lesions with perivascular lymphocytic cuffs and a few macrophage-rich lesions without necrosis
CC023	P	Y	Ctrl mice had moderate to severe lesions with necrosis and minimal lymphocytes. BCG mice have moderate lesions with more lymphocytes and less necrosis.
CC025	P	Y	Ctrl mice have moderate disease with mild lymphocytic infiltrates and necrosis. BCG mice have abundant dense perivascular/peribronchiolar lymphocytic ifiltrates and lymphohistiocytic granulomas and less necrosis.
CC026	P	Y	Ctrl mice had moderate lesions with lymphohistiocytic infiltrates. BCG mice had mild lesions with lymphohistiocytic infiltrates and a few very dense lymphocytic areas
CC031	P	Y	Ctrl mice have moderate disease with lymphohistiocytic infiltrates. BCG mice have abundant dense perivascular/peribronchiolar lymphocytes and lymphohistiocytic granulomas
CC032	Protected	Y	Ctrl mice have mild disease with lymphohistiocytic infiltrates. BCG mice have more lymphocyte rich lesions that are small & hard to find
CC037	Protected	Y	Ctrl mice have moderate disease with lymphohistiocytic infiltrates. BCG mice have abundant dense perivascular/peribronchiolar lymphocytes and lymphohistiocytic granulomas
CC042	Protected	Y	Ctrl mice have mild to moderate disease with few lymphocytes and necrosis. BCG mice have mild disease with more lymphocytes, less necrosis
CC059	Protected	Y	Ctrl mice had mild disease with lymphohistiocytic infiltrates with small regions of necrosis. BCG mice have abundant dense perivascular/peribronchiolar lymphocytes and lymphohistiocytic granulomas without necrosis