

eScholarship@UMassChan

Host-pathogen genetic interactions underlie tuberculosis susceptibility in genetically diverse mice [preprint]

Item Type	Preprint
Authors	Smith, Clare M.;Baker, Richard E.;Proulx, Megan K.;Mishra, Bibhuti B.;Long, Jarukit E.;Kiritsy, Michael C.;Bellerose, Michelle;Olive, Andrew J.;Murphy, Kenan C;Papavinasasundaram, Kadamba;Boehm, Frederick;Reames, Charlotte;Sassetti, Christopher M
Citation	<p>bioRxiv 2020.12.01.405514; doi: https://doi.org/10.1101/2020.12.01.405514 . Link to preprint on bioRxiv.</p>
DOI	10.1101/2020.12.01.405514
Rights	The copyright holder for this preprint is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .
Download date	2025-04-23 23:37:46
Item License	http://creativecommons.org/licenses/by-nc-nd/4.0/
Link to Item	https://hdl.handle.net/20.500.14038/29620

1 **Host-pathogen genetic interactions underlie tuberculosis susceptibility in**
2 **genetically diverse mice**

3

4 **Short title:** Collaborative Cross and Tuberculosis

5

6 **Authors:**

7 Clare M. Smith^{1,2*}, Richard E. Baker¹, Megan K. Proulx¹, Bibhuti B. Mishra^{1,3}, Jarukit E.
8 Long¹, Sae Woong Park⁴, Ha-Na Lee⁴, Michael C. Kiritsy¹, Michelle M. Bellerose¹,
9 Andrew J. Olive¹, Kenan C. Murphy¹, Kadamba Papavinasasundaram¹, Frederick J.
10 Boehm¹, Charlotte J. Reames¹, Rachel K. Meade², Brea K. Hampton^{5,6}, Colton L.
11 Linnertz⁵, Ginger D. Shaw^{5,7}, Pablo Hock⁵, Timothy A. Bell⁵, Sabine Ehrh⁴, Dirk
12 Schnappinger⁴, Fernando Pardo-Manuel de Villena^{5,7}, Martin T. Ferris⁵, Thomas R.
13 Ioerger⁸, and Christopher M. Sassetti^{1,9*}

14

15 **Affiliations:**

16 ¹Department of Microbiology and Physiological Systems, University of Massachusetts
17 Medical School, Worcester, MA

18 ²Department of Molecular Genetics and Microbiology, Duke University, Durham, NC

19 ³Department of Immunology and Microbial Disease, Albany Medical College, Albany,
20 NY

21 ⁴Department of Microbiology and Immunology, Weill Cornell Medical College, New York,
22 NY

23 ⁵Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC

24 ⁶Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel
25 Hill, Chapel Hill, NC

26 ⁷Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill,
27 Chapel Hill, NC

28 ⁸Department of Computer Science and Engineering, Texas A&M University, College
29 Station, TX

30 ⁹Lead contact

31

32 *Correspondence: C.M. Sassetti (Christopher.Sassetti@umassmed.edu) and C.M.
33 Smith (clare.m.smith@duke.edu)

34

35

36 **Abstract:**

37 The outcome of an encounter with *Mycobacterium tuberculosis* (*Mtb*) depends on the
38 pathogen's ability to adapt to the heterogeneous immune response of the host.

39 Understanding this interplay has proven difficult, largely because experimentally
40 tractable small animal models do not recapitulate the heterogenous disease observed in
41 natural infections. We leveraged the genetically diverse Collaborative Cross (CC)
42 mouse panel in conjunction with a library of *Mtb* mutants to associate bacterial genetic
43 requirements with host genetics and immunity. We report that CC strains vary
44 dramatically in their susceptibility to infection and represent reproducible models of
45 qualitatively distinct immune states. Global analysis of *Mtb* mutant fitness across the CC
46 panel revealed that a large fraction of the pathogen's genome is necessary for

47 adaptation to specific host microenvironments. Both immunological and bacterial traits
48 were associated with genetic variants distributed across the mouse genome, elucidating
49 the complex genetic landscape that underlies host-pathogen interactions in a diverse
50 population.

51

52

53 **Keywords:** Host-pathogen interactions, tuberculosis, *Mycobacterium*, systems
54 genetics, genomics, genetic diversity, natural variation, bacterial genetics, TnSeq,
55 mouse models, Collaborative Cross, QTL mapping, complex traits.

56

57

58 **Introduction**

59 Infection with *Mycobacterium tuberculosis* (*Mtb*) produces heterogeneous outcomes
60 that are influenced by genetic and phenotypic variation in both the host and the
61 pathogen. Classic genetic studies show that host variation influences immunity to TB
62 (Abel et al., 2018; Comstock, 1978). Likewise, the co-evolution of *Mtb* with different
63 populations across the globe has produced genetically distinct lineages that
64 demonstrate variable virulence traits (Gagneux et al., 2006). The role of genetic
65 variation on each side of this interaction is established, yet the intimate evolutionary
66 history of both genomes suggests that interactions between host and pathogen variants
67 may represent an additional determinant of outcome (McHenry et al., 2020). Evidence
68 for genetic interactions between host and pathogen genomes have been identified in
69 several infections (Ansari et al., 2017; Berthenet et al., 2018), including TB (Caws et al.,

70 2008; Holt et al., 2018; Thuong et al., 2016), however the combinatorial complexity
71 involved in identifying these relationships in natural populations have left the
72 mechanisms largely unclear.

73

74 Mouse models have proven to be a powerful tool to understand mechanisms of
75 susceptibility to TB. Host requirements for protective immunity were discovered by
76 engineering mutations in the genome of standard laboratory strains of mice, such as
77 C57BL/6 (B6), revealing a critical role of Th1 immunity. Mice lacking factors necessary
78 for the production of Th1 cells or the protective cytokine interferon gamma (IFN γ) are
79 profoundly susceptible to *Mtb* infection (Caruso et al., 1999; Cooper et al., 1993; Cooper
80 et al., 1997; Flynn et al., 1993; Saunders et al., 2002). Defects in this same immune
81 axis cause the human syndrome Mendelian Susceptibility to Mycobacterial Disease
82 (MSMD) (Altare et al., 1998; Bogunovic et al., 2012; Bustamante et al., 2014; Filipe-
83 Santos et al., 2006), demonstrating the value of knockout (KO) mice to characterize
84 genetic variants of large effect. Similarly, the standard mouse model has been used to
85 define *Mtb* genes that are specifically required for optimal bacterial fitness during
86 infection (Bellerose et al., 2020; Sasseti and Rubin, 2003; Zhang et al., 2013).

87

88 Despite the utility of standard mouse models, it has become increasingly clear that the
89 immune response to *Mtb* in genetically diverse populations is more heterogeneous than
90 any single small animal model (Smith and Sasseti, 2018). For example, while IFN γ -
91 producing T cells are critical for protective immunity in standard inbred lines of mice, a
92 significant fraction of humans exposed to *Mtb* control the infection without producing a

93 durable IFN γ response (Lu et al., 2019). Similarly, IL-17 producing T cells have been
94 implicated in both protective responses and inflammatory tissue damage in TB, but IL-
95 17 has little effect on disease progression in B6 mice, except in the context of
96 vaccination or infection with particularly virulent *Mtb* (Gopal et al., 2012; Khader et al.,
97 2007). The immunological homogeneity of standard mouse models may also explain
98 why only a small minority of the >4000 genes that have been retained in the genome of
99 *Mtb* during its natural history promote fitness in the mouse (Bellerose et al., 2020).
100 Thus, homogenous mouse models of TB fail to capture the distinct disease states,
101 mechanisms of protective immunity, and selective pressures on the bacterium that are
102 observed in natural populations.

103

104 The Collaborative Cross (CC) and Diversity Outbred (DO) mouse populations are new
105 resources that more accurately represent the genetic and phenotypic heterogeneity
106 observed in outbred populations (Churchill et al., 2004; Churchill et al., 2012). These
107 mouse panels are both derived from the same eight diverse founder strains but have
108 distinct population structures (Saul et al., 2019). DO mice are maintained as an outbred
109 population and each animal represents a unique and largely heterozygous genome
110 (Keller et al., 2018; Svenson et al., 2012). In contrast, each inbred CC strain's genome
111 is almost entirely homozygous, producing a genetically stable and reproducible
112 population in which the phenotypic effect of recessive mutations is maximized (Shorter
113 et al., 2019; Srivastava et al., 2017). Together, these resources have been leveraged to
114 identify host loci underlying the immune response to infectious diseases (Noll et al.,
115 2019). In the context of TB, DO mice have been used as individual, unique hosts to

116 identify correlates of disease, which resemble those observed in non-human primates
117 and humans (Ahmed et al., 2020; Gopal et al., 2013; Niazi et al., 2015). Small panels of
118 the reproducible CC strains have been leveraged to identify host background as a
119 determinant of the protective efficacy of BCG vaccination (Smith et al., 2016) and a
120 specific variant underlying IFN γ production and protective immunity to tuberculosis
121 (Smith et al., 2019). While these studies demonstrate the tractability of the DO and CC
122 populations to model the influence of host diversity on infection, dissecting host-
123 pathogen interactions requires the integration of pathogen genetic diversity.

124

125 We combined the natural but reproducible host variation of the CC panel with a
126 comprehensive library of defined *Mtb* mutants to characterize the interactions between
127 host and pathogen. Using over 60 diverse mouse strains, we report that the CC panel
128 encompasses a broad spectrum of TB susceptibility and immune phenotypes, including
129 outlier lines that model non-canonical immune states. Through “Transposon
130 Sequencing” (TnSeq), we quantified the relative fitness of *Mtb* mutants across the CC
131 panel and specific immunological knockout strains, allowing us to infer the
132 microenvironments encountered by the bacterium in each animal and define a large
133 fraction of the bacterial genome that is necessary for adapting to these diverse immune
134 states. Association of these immunological and bacterial fitness traits with distinct
135 Quantitative Trait Loci (QTL) highlighted the polygenic nature of TB susceptibility and
136 identified discrete Host-Interacting-with Pathogen QTL (*Hip*QTL) that represent a new
137 strategy to understand these epistatic interactions between genomes.

138

139 **Results:**

140 **The spectrum of TB disease traits in the CC exceeds that observed in standard**
141 **inbred mice.**

142 We infected a panel of 52 CC lines and the 8 founder strains with *Mtb*. To enable TnSeq
143 studies, the animals were infected via the intravenous route with a saturated library of
144 *Mtb* transposon mutants, which in sum produce an infection that is similar to the wild
145 type parental strain (Bellerose et al., 2020; Sasseti and Rubin, 2003). Groups of 2-6
146 (average of n=3) male mice per genotype were infected, and the bacterial burden after
147 four weeks of infection was assessed by plating (colony forming units, CFU) and
148 quantifying the number of bacterial chromosomes in the tissue (chromosome
149 equivalents, CEQ). These two metrics were highly correlated ($r=0.88$) and revealed a
150 wide variation in bacterial burden across the panel (**Figure 1A** and **S1**). The
151 susceptibility of the inbred founder strains was largely consistent with previous studies
152 employing an aerosol infection (Smith et al., 2016). As expected, B6 mice were
153 relatively resistant to infection, along with 129S1/SvlmJ (129) and NOD/ShiLtJ (NOD)
154 strains. Lung CFU varied by less than 10-fold between these genetically similar animals,
155 while the wild derived WSB/EiJ (WSB) founder was highly susceptible. In contrast to the
156 standard inbred lines, lung bacterial burden varied by more than 1000-fold across the
157 more diverse CC panel, ranging from animals that are significantly more resistant than
158 B6, to mice that harbored more than 10^9 bacteria in their lungs (**Figure 1A**). Bacterial
159 burden in the spleen also varied several thousand-fold across the panel and was
160 moderately correlated with lung burden ($r=0.43$) (**Table S1** and **Figure S1**). Thus, the

161 CC panel encompasses a much greater quantitative range of susceptibility than
162 standard inbred lines.

163

164 As an initial assessment of the disease processes in these animals, we correlated
165 bacterial burden and lung cytokine abundance with measures of systemic disease such
166 as weight loss and sufficient morbidity to require euthanasia (“earliness of death”). In
167 general, correlations between these metrics indicated that systemic disease was
168 associated with bacterial replication and inflammation (**Figure 1B** and **Figure S1**). Lung
169 CFU was correlated with weight loss, mediators that enhance neutrophil differentiation
170 or migration (CXCL2 (MIP-2; $r=0.79$), CCL3 (MIP-1a; $r=0.77$), G-CSF ($r=0.78$) and
171 CXCL1 (KC; $r=0.76$)), and more general proinflammatory cytokines (IL-6 ($r=0.80$) and
172 IL-1 α ($r=0.76$)) (**Figure S1**). These findings are consistent with previous work in the DO
173 panel, that found both proinflammatory chemokines and neutrophil accumulation to be
174 predictors of disease (Ahmed et al., 2020; Gopal et al., 2013; Niazi et al., 2015).

175

176 The reproducibility of CC genotypes allowed us to quantitatively assess the heritability
177 (h^2) of these immunological and disease traits. The percent of the variation attributed to
178 genotype ranged from 56-87% (mean=73.4%; **Table S2**). The dominant role of genetic
179 background in determining the observed phenotypic range allowed a rigorous
180 assessment of outlier phenotypes. For example, despite the correlation between lung
181 CFU and weight loss ($r=0.57$), several strains failed to conform to this relationship
182 (**Figure 1C**). In particular, CC030/GeniUnc ($p=0.003$), CC040/TauUnc ($p=0.027$) and
183 A/J ($p=0.03$) lost significantly more weight than their bacterial burdens would predict

184 **(Figure 1C;** outlier genotypes determined by studentized residuals; noted by #).

185 Similarly, CXCL1 abundance was significantly higher in CC030/GeniUnc ($p=0.001$) and
186 lower in CC056/GeniUnc ($p=0.040$), than the level predicted by their respective bacterial
187 burden **(Figure 1C;** outlier genotypes noted by †). Thus, these related disease traits can
188 be dissociated based on the genetic composition of the host.

189

190 The cluster of cytokines that was most notably unrelated to bacterial burden included
191 IFN γ and the interferon-inducible chemokines CXCL10 (IP10), CXCL9 (MIG), and CCL5
192 (RANTES) (Red cluster in **Figure 1B;** **Figure S1**) ($r<0.3$). Despite the clear protective
193 role for IFN γ (Cooper et al., 1993; Flynn et al., 1993), high levels have been observed in
194 susceptible mice, likely as a result of high antigen load (Barber et al., 2011; Lázár-
195 Molnár et al., 2010). While high IFN γ levels in susceptible animals was therefore
196 expected, it was more surprising to find a number of genotypes that were able to control
197 bacterial replication yet had very low levels of this critically important cytokine **(Figure**
198 **1D)**. This observation is likely due the inclusion of two founder lines, CAST/EiJ (CAST)
199 and PWK/PhJ (PWK) that display this unusual phenotype (Smith et al., 2016). To further
200 investigate, a separate cohort of B6 and CAST animals was infected by the aerosol
201 route, and the number of IFN γ producing T cells in lung and spleen was compared by
202 ELISPOT assays. At 4 weeks post-infection, B6 and CAST harbored comparable
203 burdens of *Mtb* in lung and spleen **(Figure 1E and 1F)**, and the infection elicited similar
204 numbers of *Mtb*-specific IFN γ producing cells in the spleen **(Figure 1H)**. In contrast,
205 while IFN γ producing cells were found in the lungs of B6 mice, none were detectable in
206 CAST **(Figure 1G)**. Thus, while CAST animals are capable of producing IFN γ -secreting

207 cells in response to *Mtb* infection, these cells do not appear to be involved in bacterial
208 control in the lung. As strains CC mice that share this immune profile also produced low
209 levels of IL-17 (**Figure 1D**), the mechanism(s) conferring protection in these animals
210 remain unclear. In sum, this survey of TB-related traits in the CC demonstrated a broad
211 range of susceptibility and the presence of qualitatively distinct and genetically
212 determined disease states.

213

214 **Genetic mapping of TB immunophenotypes**

215 Tuberculosis ImmunoPhenotype Quantitative Trait Loci (*TipQTL*), which were
216 associated with TB disease or cytokine traits, were identified using R/qtl2 (Broman et
217 al., 2019) and numbered in accordance with previously reported *TipQTL* (Smith et al.,
218 2019). Of the 32 TB-disease traits, we identified 9 individual metrics that were
219 associated with statistical confidence to a chromosomal locus. Of these, three were
220 associated with high confidence ($p \leq 0.053$), and six other QTL met a suggestive
221 threshold ($p < 0.2$; **Table 1**). Several individual trait QTL occupied the same
222 chromosomal locations. For example, spleen CFU and spleen CEQ, which are both
223 measures of bacterial burden and highly correlated, were associated with the same
224 interval on distal chromosome 2 (**Table 1**, *Tip5*; **Figure 2A** and **2C**). IL-10 abundance
225 was associated with two distinct QTL (**Table 1**). While IL-10 was relatively uncorrelated
226 with spleen CFU ($r=0.48$), one of its QTL fell within the *Tip5* bacterial burden interval on
227 chromosome 2 (**Figure 2A** and **2C**). At this QTL, the NOD haplotype was associated
228 with high values for all three traits (**Figure 2E**). Similarly, the correlated traits, CXCL1
229 abundance and lung CFU, were individually associated to the same region on

230 chromosome 7 (**Table 1**, *Tip8*; **Figure 2B** and **2D**). In this interval, the CAST haplotype
231 was associated with both low bacterial burden and CXCL1 (**Figure 2F**). At both *Tip5*
232 and *Tip8*, we found no statistical evidence that the positions of the associated QTL were
233 different (*Tip5* $p=0.55$; *Tip8* $p=0.27$; 400 bootstrap samples)(Boehm et al., 2019). These
234 observations support the role of a single causal variant at each locus that is responsible
235 for a pleiotropic trait. Coincident mapping can provide both additional statistical support
236 for QTL (p values by Fisher's combined probability test: Chr 7, $p=0.067$; Chr 2, $p=0.041$)
237 and suggests potential mechanisms of disease progression.

238

239 A number of factors can limit the statistical significance of QTL identified in the CC
240 population, including small effect sizes, the relatively low number of available
241 genotypes, and the genetic complexity of the trait. To determine if the statistically
242 suggestive QTL were likely to reflect genuine biological effects, we took an intercross
243 approach to validate the lung CFU QTL on chromosomes 7 and 15 (*Tip8* and *Tip10*,
244 **Table 1**). Given that the associations at both QTL were driven by the CAST haplotype
245 (**Figure 2F**), we generated an F_2 population based on two CC strains, CC029/Unc and
246 CC030/GeniUnc, that contained CAST sequence at *Tip8* and *Tip10*, respectively
247 (**Figure 3A**). 46 F_2 mice that were homozygous at either locus were infected with *Mtb*,
248 and lung CFU were enumerated as per the larger CC screen (**Table S3**). Compared to
249 F_2 mice that were not CAST at either locus, mice that contained CAST at *Tip8*, *Tip10* or
250 both loci had reduced CFU burden; however, the strongest predictor of significantly
251 lower bacterial burden was CAST at *Tip8* ($p=0.005$, 2-factor regression; **Figure 3B**).
252 This study provides strong independent support for a CAST-driven QTL on

253 chromosome 7 that controls lung CFU and verifies the validity of the genetic mapping in
254 this dataset.

255

256 ***Mtb* adapts to diverse hosts by utilizing distinct gene repertoires**

257 This survey of disease-associated traits indicated that the CC panel encompasses a
258 number of qualitatively distinct immune phenotypes. To determine if different bacterial
259 functions were necessary to adapt to these conditions, we used TnSeq to estimate the
260 relative abundance of each individual *Mtb* mutant after selection in each CC host
261 genotype. To serve as benchmarks of known immunological lesions, we also performed
262 TnSeq in B6 mice that were lacking the mediators of Th1 immunity, lymphocytes (*Rag2*^{-/-}
263 ^{-/-}) and IFN γ (*Ifn γ* ^{-/-}), or were lacking the immunoregulatory mediators that control
264 disease by inhibiting inflammation, nitric oxide synthase (*Nos2*^{-/-}) (Mishra et al., 2013) or
265 the NADPH phagocyte oxidase (*Cybb*^{-/-}) (Olive et al., 2018). The relative representation
266 of each *Mtb* mutant in the input pool versus the pools recovered from mouse spleens
267 was quantified (**Table S4**). Consistent with our previous work (Bellerose et al., 2020;
268 Sasseti and Rubin, 2003), we identified 234 *Mtb* genes that are required for growth or
269 survival in *Mtb* in B6 mice, based on significant underrepresentation of the
270 corresponding mutant after four weeks of *in vivo* selection. All but one of these genes
271 were found to be required in the larger panel, increasing confidence in this *Mtb* gene set
272 (**Figure 4A and 4B**). While the number of genes found to be necessary in each
273 genotype across the panel was largely similar, the composition of these gene sets
274 varied considerably. As more CC strains, and presumably more distinct immune states,
275 were included in the analysis, the cumulative number of genes necessary for growth in

276 these animals also increased. This cumulative gene set plateaued at ~750, after the
277 inclusion of approximately 20-25 genotypes (**Figure 4A**). This number of genes far
278 outnumbered those identified in the panel of immunodeficient KO strains (**Figure 4B**
279 and **Table S4**).

280

281 To verify that our TnSeq study accurately predicted the effect of the corresponding loss-
282 of-function alleles, we assessed the phenotypes of selected bacterial deletion mutants
283 in a small set of mouse genotypes that were predicted to result in differential selection.
284 Individual *Mtb* mutants lacking genes necessary for ESX-1 type VII secretion (*eccB1*),
285 siderophore-mediated iron acquisition (*mbtA*), phosphate transport (*pstC2*), glycerol
286 catabolism (*glpK*), and RNA processing (*maseJ*) were generated and tagged with a
287 unique molecular barcode. These mutants were combined with a barcoded wild type
288 parental strain, the resulting ‘mini-pool’ was subjected to *in vivo* selection in the same
289 manner as the TnSeq study, and the relative abundance of each mutant was
290 determined by sequencing the amplified barcodes. In each case, the difference in
291 relative abundance predicted by TnSeq was reproduced with deletion mutants (**Figure**
292 **4C**). In this simplified system, we were able to accurately quantify the expansion of the
293 bacterial population and calculate the “fitness” of each mutant relative to the wild type
294 strain. Fitness reflects the inferred doubling time of the mutant, where a fitness of 1 is
295 defined as wild type, and 0 represents a complete lack of growth. Even by this metric,
296 the deletion mutants displayed the differences in fitness between mouse strains that
297 was predicted by TnSeq (**Figure 4C**). The statistical significance of these differences in
298 abundance or fitness were similar for each mutant (between $p=0.009$ and $p=0.06$),

299 except for *mbtA* where the variation was higher, and confidence was modestly lower
300 ($p=0.07$ and $p=0.12$). This study also allowed us to estimate the sensitivity of the TnSeq
301 method, which could detect even the 30% fitness defect of the $\Delta glpK$ strain between the
302 B6 and CC018 animals (**Figure 4C**), a defect that was not observed in previous studies
303 in BALB/c mice (Bellerose et al., 2019; Pethe et al., 2010).

304

305 To also validate TnSeq predictions in a single-strain aerosol infection model, we used a
306 biotin biosynthetic mutant. *bioA* is necessary for biotin production and is essential for
307 growth in B6 mice (Woong Park et al., 2011). Our TnSeq study (**Table S4**) predicted
308 this mutant was less attenuated in the CAST background (ratio of input/selected = 12.1)
309 than in the B6 strain (ratio of input/selected = 42.2). Two weeks after aerosol infection,
310 we found that the $\Delta bioA$ mutant was cleared from the lungs and spleen of B6 mice but
311 displayed similar growth to WT in the lungs of CAST mice (**Figure 4D**). By 6 weeks post
312 infection the $\Delta bioA$ mutant had also been largely cleared from the lungs of CAST
313 (**Figure 4D**). Thus, while TnSeq was unable elucidate the complex kinetics of this
314 phenotype, it accurately predicted the relative levels of growth attenuation in these host
315 backgrounds.

316

317 **The immunological diversity of CC mice is reflected in the pathogen's genetic** 318 **requirements**

319 The distribution of *Mtb*'s requirements across the mouse panel suggested the presence
320 of two broad categories of genes. A set of 136 "core" virulence functions were required
321 in the majority of mouse genotypes, and a second larger set of 607 "adaptive" virulence

322 genes were required in only a subset of lines (**Table S4**). The core functions included a
323 number of genes previously found to be important in B6 mice, including those
324 necessary for the synthesis of essential cofactors, such as pyridoxine (*Pdx1*) (Dick et
325 al., 2010); for the acquisition of nutrients, such as siderophore-bound iron (*irtAB*)
326 (Ryndak et al., 2010), cholesterol (*mce4*) (Pandey and Sasseti, 2008), glutamine (*glnQ*
327 and *rv2563*) (Bellerose et al., 2020); and for type VII secretion (ESX1 genes) (Stanley et
328 al., 2003). Despite the importance of these core functions, a large range in the relative
329 abundance of these mutants was observed across the panel, and in some cases
330 specific immunological requirements could be discerned. Mutants lacking the major
331 structural components of the ESX1 system were attenuated for growth in B6 mice, as
332 expected. This requirement was consistently enhanced in mice lacking *Rag2*, *Ifny*, or
333 *Nos2* (**Figure 4E**), consistent with the preferential role of ESX1 during the initial stage of
334 infection before the initiation of adaptive immunity (Stanley et al., 2003). In contrast, the
335 attenuation of mutants lacking the *glnQ* encoded glutamine uptake system was relieved
336 in all four immunodeficient mouse lines (**Figure 4E**). In both cases, the differential
337 mutant abundance observed in these KO mice was reproduced, or exceeded, in the CC
338 panel.

339

340 The adaptive virulence functions included a number of genes previously thought to be
341 dispensable in the mouse model and were only necessary in CC strains. For example,
342 the alkyl hydroperoxide reductase, AhpC has been proposed to function with the
343 adjacently encoded peroxiredoxin, AhpD and is critical for detoxifying reactive nitrogen
344 intermediates *in vitro* (Chen et al., 1998; Hillas et al., 2000). However, deletion of *ahpC*

345 has no effect on *Mtb* replication in B6 or BALB/c mice (Springer et al., 2001), and we
346 confirmed that *ahpC* and *ahpD* mutations had no effect in any of the B6-derived lines. In
347 contrast, *ahpC*, but not *ahpD* mutants were highly attenuated in a small number of CC
348 strains (**Figure 4E**). Similarly, the four phospholipase C enzymes of *Mtb* (*plcA-D*) are
349 implicated in both fatty acid uptake and modifying host cell membranes but are
350 dispensable for replication in B6 mice (Le Chevalier et al., 2015). Again, while we found
351 that none of these genes were required in B6-derived KO lines, the *plcD* mutants were
352 specifically underrepresented in a number of CC mice (**Figure 4E**). These individual
353 bacterial functions are controlled by regulatory proteins, such as the extracytoplasmic
354 sigma factors. Despite the importance of these transcription factors in the response to
355 stress, only *sigF* has been consistently been shown to contribute to bacterial replication
356 in standard inbred lines of mice (Geiman et al., 2004; Rodrigue et al., 2006). Our study
357 assesses the importance of each sigma factor in parallel across diverse host genotypes
358 and identified a clear role for several of these regulators. *sigC*, *sigI*, *sigF*, *sigL*, and *sigM*
359 mutants were each significantly underrepresented in multiple lines of mice, and several
360 of these phenotypes were only apparent in the diverse CC animals (**Figure 4E**). As the
361 relative fitness of *Mtb* mutants is influenced by host immunity, we conclude that the CC
362 panel produces a wide variety of immune states, including those that are not present in
363 standard B6-derived strains.

364

365 **Differential genetic requirements define virulence pathways in *Mtb***

366 To more formally investigate the distinct stresses imposed on the bacterial population
367 across this host panel, we defined differentially required bacterial pathways. Upon

368 performing each possible pairwise comparison between the *in vivo* selected mutant
369 pools, we found 679 mutants whose representation varied significantly (FDR < 5%) in at
370 least two independent comparisons (**Table S4**). We then applied weighted gene
371 correlation network analysis (WGCNA) (Langfelder and Horvath, 2008) to divide the
372 mutants into 20 internally-correlated modules. Further enrichment of these modules for
373 the most representative genes (intramodular connectivity > 0.6) revealed that nearly all
374 modules contained genes that are encoded adjacently in the genome and many of
375 these modules consisted of genes dedicated to a single virulence-associated function
376 (**Figure 5A**). Module 3 contains two distally encoded loci both known to be necessary
377 for ESX1-mediated protein secretion, the primary ESX1 locus (*rv3868-rv3883*) and the
378 *espACD* operon (*rv3616c-rv3614c*). Similarly, other modules consisted of genes
379 responsible for ESX5 secretion (Module 7), mycobactin synthesis (Module 4), the Mce1
380 and Mce4 lipid importers (Modules 5 and 16), phthiocerol dimycocerosate synthesis
381 (PDIM, Module 8), PDIM transport (Module 16), and phosphate uptake (Module 14).
382 The 20 genes assigned to Module 6 included two components of an important oxidative
383 stress resistance complex (*sseA* and *rv3005c*) and were highly enriched for mutants
384 predicted to be involved in this same process via genetic interaction mapping (11/20
385 genes were identified in (Nambi et al., 2015), a statistically significant overlap (< 2.8e-10
386 by hypergeometric test). Thus, each module represented a distinct biological function.

387

388 Many pathway-specific modules contained genes that represented novel functional
389 associations. For example, the gene encoding the sigma factor, *sigC*, was found in
390 Module 1 along with a non-ribosomal peptide synthetic operon. Previous genome-wide

391 ChIP-seq and overexpression screens support a role for SigC in regulating this operon
392 (Minch et al., 2015; Turkarslan et al., 2015). Similarly, *rv3220c* and *rv1626* have been
393 proposed to comprise an unusual two component system that is encoded in different
394 regions of the genome (Morth et al., 2005). Both of these genes are found in Module 2,
395 along with the PPE50 and PPE51 genes that encode at least one outer membrane
396 channel (Wang et al., 2020)(**Figure 5A**). In both cases, these associations support both
397 regulatory and obligate functional relationships between these genes. 6 of the 20
398 modules were not obviously enriched for genes of a known pathway, indicating that
399 novel virulence pathways are important for adapting to changing host environments.

400

401 To explore the complexity of immune environments in the CC, we used the TnSeq
402 profiles of the 679 differentially fit *Mtb* mutants to cluster the mouse panel into 6 major
403 groups of host genotypes (**Figure 5B**). Two mouse clusters were significantly
404 associated with high CFU (yellow boxes, **Figure 5B**), one of which contained
405 susceptible *Nos2^{-/-}*, *Cybb^{-/-}*, *Ifny^{-/-}*, and *Rag2^{-/-}* animals. The high CFU clusters were
406 associated with alterations in the most diverse set of bacterial modules and
407 corresponded to an increased requirement for lipid uptake (Modules 5 and 16) and
408 ESX1, consistent with previous TnSeq studies in susceptible *Nos2^{-/-}* and C3HeB/FeJ
409 mice (Mishra et al., 2017). In addition, we identified a significant reduction in the
410 requirement for the oxidative stress resistance (Module 6) in the highest CFU cluster.
411 Despite these associations between bacterial genetic requirements and susceptibility,
412 the clustering of mouse genotypes was largely independent of overall susceptibility.
413 Similarly, while Module 1 was significantly associated with IFN γ levels, other bacterial

414 fitness traits were not highly correlated with cytokine abundance (**Figure S3**). Instead,
415 each major mouse cluster was associated with a distinct profile of *Mtb* genetic
416 requirements. This observation supported the presence of qualitatively distinct disease
417 states and relatively complex genetic control of immunity.

418

419 **Identification of host-pathogen genetic interactions**

420 To investigate the host genetic determinants of the bacterial microenvironment, we
421 associated *Mtb* mutant fitness profiles with variants in the mouse genome. When the
422 relative abundance of each *Mtb* mutant was considered individually, the corresponding
423 “Host Interacting with Pathogen QTL” (*Hip*QTL) were distributed across the mouse
424 genome (**Figure 6A**). 41 of these traits reached an unadjusted p value threshold of 0.05
425 and can be considered as robust for single hypothesis testing (*Hip1-41*, **Table S2** and
426 **S5**). These included *Hip*QTL associated with both *ahpC* and *eccD1*, that explain at
427 least a portion of the observed variable abundance of these mutants (**Figure 4E**). In
428 order to reduce complexity and increase the power of this analysis, we performed QTL
429 mapping based on the first principal component of each of the previously defined
430 modules of *Mtb* virulence pathways (**Figure 5A**). Three of these “eigentrains” were
431 associated with QTL at a similar position on chromosome 10 (**Figure 6B**),
432 corresponding to Module 3 (typeVII secretion, ESX1), Module 4 (mycobactin synthesis,
433 *mbt*), or Module 16 (cholesterol uptake, *mce4*). In all three cases, a single mutant from
434 the module was independently associated with a QTL at the same position as the
435 module eigentrait (Table S5, *Hip21*, *Hip22*, *Hip24*), and all genes in the corresponding
436 network cluster (**Figure 5A**) mapped to the same location (**Figure 6C-E**). While not all

437 individual traits mapped with high confidence, the coincidence of these multiple QTL
438 was statistically significant (**Figure 6B**).

439
440 Both the relative positions of the module-associated QTL and the associated founder
441 haplotypes indicated that a single genetic variant controlled the abundance of ESX1 and
442 *mbt* mutants (*Hip42*). Specifically, we found no statistical support for differentiating
443 these QTL based on position ($p=0.93$)(Boehm et al., 2019) and the same founder
444 haplotypes were associated with extreme trait values at both loci, though they had
445 opposite effects on the abundance of ESX1 mutants and *mbt* mutants (**Figure 6F**). We
446 conclude that a single haplotype has a pleiotropic effect on *Mtb*'s environment and has
447 opposing effects of the requirement for mycobactin synthesis and ESX1 secretion. The
448 relationship between this variant and the *mce4*-associated QTL (*Hip43*) was less clear,
449 as the statistical support for independent QTL was weak (ESX1 and *mce4* QTL $p=0.17$;
450 *mbt* and *mce4* QTL $p=0.08$) and the effects of founder haplotypes were similar but not
451 identical (**Figure 6F**). Some of this ambiguity may be related to the relatively small
452 range in trait values for *mce4*, compared to either ESX1 or *mbt* (**Figure 6G**). Based on
453 this data, we report two distinct *Hip*QTL in this region (*Hip42* and *43*, **Table S5**).

454
455 There was little overlap between disease related *Tip*QTL and the more specific *Hip*QTL
456 (**Figure S2**), indicating that the fitness of sensitized bacterial mutants can be used to
457 detect genetic variants that subtly influence the bacterial environment but not overtly
458 alter disease. We chose to further investigate whether *Hip*QTL might alter overall
459 bacterial disease using the most significant *Hip*QTL on chromosome 10 (*Hip42*). We

460 found that the founder haplotypes associated with extreme trait values at this QTL could
461 differentiate CC strains with significantly altered total bacterial burden, and the NOD and
462 WSB haplotypes were associated with higher bacterial numbers ($p=0.0085$ for spleen
463 CEQ; $p=0.027$ for spleen CFU; **Figure 6H**). Thus, not only could the *Hip*QTL strategy
464 identify specific interactions between host and bacterial genetic variants, but it also
465 appears to be a sensitive approach for identifying host loci that influence the trajectory
466 of disease.

467

468 **Identifying candidate genes underlying QTL.**

469 A pipeline was designed to prioritize genetic variants based on genomic and
470 tuberculosis disease criteria. We concentrated on three QTL: two that were highly
471 significant and with clear allele effects (*Tip5*, *Hip42*), and the *Tip8* locus which we
472 validated by intercross. For each QTL region, we identified genes that belonged to a
473 differentially expressed transcriptional module in mouse lungs following *Mtb* infection
474 (Moreira-Teixeira et al., 2020). Next, we identified genetic variants segregating between
475 the causal CC haplotypes in the gene bodies corresponding to these transcripts, and
476 prioritized missense or nonsense variants.

477

478 For the *Tip5* QTL underlying CEQ, CFU, and IL-10 levels, we identified nine candidate
479 genes with regulatory or splicing variants and two genes with missense variants specific
480 to the NOD haplotype. Of these high priority candidates, cathepsin Z (*Ctsz*) encodes a
481 lysosomal cysteine proteinase has previously been associated with TB disease risk in
482 humans (Adams et al., 2011; Cooke et al., 2008). The QTL underlying lung CFU and

483 CXCL1 abundance (*Tip8*), which was driven solely by the genetically divergent CAST
484 founder haplotype, contained over 50 genes (**Table S6**) and will need further
485 refinement. The QTL associated with the abundance of ESX1 and *mbt* mutants (*Hip42*)
486 had a complex causal haplotype pattern (AJ/B6/NZO high, 129/CAST/PWK
487 intermediate, NOD/WSB low) suggesting multiple variants might be impacting common
488 genes. Within this interval, we identified 13 genes expressed in response to *Mtb*
489 infection, three of which had SNPs fully or partially consistent with at least one of the
490 identified causal haplotype groups (**Table S6**). *Ank3* contains several SNPs in the 3'
491 UTR and other non-coding exons that differentiated NOD/WSB from the other
492 haplotypes. Similarly, *Fam13c* had two missense mutations following the same
493 haplotype pattern. For the AJ/B6 haplotype state, we identified a missense mutation and
494 several variants in the 3' UTR of *Rhobtb1*, which belongs to the Rho family of the Ras
495 superfamily of small GTPases (Goitre et al., 2014). Overall, the evidence supports a
496 role for *Rhobtb1* in a monogenic effect at the chromosome 10 locus. This evidence
497 includes both protein coding differences dividing AJ/B6 from the other haplotypes, a
498 potential expression/transcript regulatory difference that segregates the NOD/WSB
499 state from the remaining parental haplotypes, and a plausible role for this gene in
500 controlling intracellular trafficking (Long et al., 2020) and the opposing requirements for
501 ESX1 and mycobactin.

502

503 **Discussion**

504 Our immunological analysis of the CC panel identified correlates of TB disease
505 progression that were consistent with previous studies in both mice and human patients

506 (Ahmed et al., 2020; Niazi et al., 2015; Zak et al., 2016), as well as outlier strains that
507 produce distinct immunological states. For example, despite the generally strong
508 correlation between lung bacterial burden and disease, CC030/GeniUnc and
509 CC040/TauUnc mice suffered from more inflammation and wasting than would be
510 predicted from the number of bacteria in their lungs or spleens. This phenotype reflects
511 a failure of disease “tolerance”, which is proposed to be a critical determinant of
512 protective immunity both human patients and engineered mouse models (Ayres and
513 Schneider, 2012; Olive et al., 2018). Similarly, we identified a number of CC genotypes
514 that produce very low, or undetectable, levels of the protective cytokine IFN γ , but still
515 control lung bacterial replication. While a growing literature suggests that immune
516 responses distinct from the canonical Th1 response can control infection (Lu et al.,
517 2019; Sakai et al., 2016), these CC strains are the first example of an animal model in
518 which IFN γ appears to be dispensable. The reproducibility of the CC strains facilitated
519 the identification of these phenotypes and provides tractable models for their
520 characterization.

521

522 The ability to separate aspects of the immune response from disease progression
523 implied that these features are under distinct multigenic control. This conclusion is
524 supported by genetic mapping, which identified a number of variants that control distinct
525 aspects of the immune response to *Mtb*. The QTL identified in this study are generally
526 distinct from CC loci that control immunity to viruses (Ferris et al., 2013; Gralinski et al.,
527 2017; Noll et al., 2020) or another intracellular bacterial pathogen, *Salmonella* (Zhang et
528 al., 2019). However, *Tip8* and *Tip10* overlap with QTL previously defined via *Mtb*

529 infection of CC001xCC042 intercross population (Smith et al., 2019) suggesting that
530 common variants may have been identified in both studies. While the specific genetic
531 variants responsible for these QTL remain unknown, both coincident trait mapping and
532 bioinformatic analysis suggest mechanistic explanations for some QTL-phenotype
533 associations. In particular, a single interval on chromosome 2 controls CFU levels and
534 IL-10, and contains a variant in the *Ctsz* gene encoding Cathepsin Z. *Ctsz* is a strong
535 candidate considering its known roles in autophagy (Amaral et al., 2018), dendritic cell
536 differentiation and function (Obermajer et al., 2008), its upregulation in non-human
537 primates (Ahmed et al., 2020) and human patients with *Mtb* (Zak et al., 2016), and the
538 association of *CTSZ* variants with disease risk in human TB studies (Adams et al.,
539 2011; Cooke et al., 2008).

540
541 Using TnSeq as a multidimensional phenotyping method across this population
542 provided insight into how the diversity of host-derived microenvironments have shaped
543 the pathogen's genome. While *Mtb* is an obligate pathogen with no significant
544 environmental niche, only a minority of the genes in its genome have been found to
545 contribute to bacterial fitness in either laboratory media or individual inbred mouse
546 models, leaving the pressures that maintain the remaining genomic content unclear. We
547 find that approximately three times more genes contribute to bacterial growth or survival
548 in the CC population than in the standard B6 model. While some bacterial genetic
549 requirements could be associated with known immune pathways, most of the differential
550 pressures on bacterial mutants could not be attributed to these simple deficiencies in
551 known mechanisms of immune control. Instead, it appears that the CC population

552 produces a spectrum of novel environments, and that a relatively large fraction of the
553 pathogen's genome is needed to adapt to changing immune pressures. Differential
554 pressures on these adaptive virulence functions are similarly apparent in genomic
555 analyses of *Mtb* clinical isolates. Signatures of selection have been detected in ESX1-
556 related genes (Holt et al., 2018; Sousa et al., 2020), *phoPR* (Gonzalo-Asensio et al.,
557 2014), and the oxidative stress resistance gene *sseA* (de Keijzer et al., 2014),
558 suggesting that *Mtb* is exposed to similarly variable host pressures in genetically diverse
559 human and mouse populations. While the combinatorial complexity of associating host
560 and pathogen genetic variants in natural populations is daunting, the identification of
561 *Hip*QTL in the CC panel demonstrates that these inter-species genetic interactions can
562 be defined and characterized using experimentally tractable models of diversity.

563

564

565 **Acknowledgements:** We thank all members of Sasseti Lab, past and present for
566 technical help and discussions; Dr. Nathan Hicks and Dr. Sarah Fortune for kindly
567 providing the *rnaseJ* mutant, Dr. David Tobin for insightful manuscript comments; Dr.
568 Dennis Ko for QTL acronym creativity, Erin Curtis for mouse schema expertise and the
569 Systems Genetic Core at UNC for their help in procuring CC mice in timely fashion. This
570 work was supported by NIH grants AI132130 to C. M. Sasseti and FPMV;
571 U19AI100625 to FPMV and MTF; a fellowship from the Charles H. King Foundation to
572 C. M. Smith; and a HHMI Gilliam Fellowship A20-0146 to BKH. The genetic
573 characterization of the CC strains was supported in part by NIH grant U24HG010100 to
574 FPMV.

575

576 **Author contributions:** Conceptualization, C.M. Smith and C.M Sasseti; Methodology,
577 R.E.B, K.C.M, K.P, T.R.L; Investigation; C.M. Smith, M.K.P, B.B.M, J.E.L, M.C.K,
578 M.M.B, A.J.O, C.J.R, C.M. Sasseti; Validation, S.W.P, H.L, S.E, D.S; Formal Analysis,
579 R.E.B, F.J.B, R.K.M, B.K.H, C.L, M.T.F, T.R.I; Resources, G.D.S, P.H, T.A.B;
580 Visualization, C.M. Smith, R.E.B, M.K.P, R.K.M, C.M. Sasseti; Writing – original draft,
581 C.M Sasseti and C.M Smith; Writing – review and editing, all authors; Supervision,
582 C.M. Sasseti; Funding acquisition, C.M. Sasseti, M.T.F, F.P.M.V.

583

584 **Declarations of interests:** None.

585

586 **Figure titles and legends:**

587 **Figure 1. The spectrum of *M. tuberculosis* disease-related traits across the**

588 **Collaborative Cross.** (A) Average lung CFU (\log_{10}) across the CC panel at 4 weeks

589 post-infection. Bars show mean +/- SD for CFU per CC or parental strain; groups of 2-6

590 (average of n=3) mice per genotype. Bars noted with * indicate strains that were

591 statistically different from B6 ($P < 0.05$; 1-factor ANOVA with Dunnett's post-test). (B)

592 Heatmap of the 32 disease-related traits (\log_{10} transformed) measured including: lung

593 and spleen colony forming units (CFU); lung and spleen chromosomal equivalent

594 (CEQ); weight loss (% change); cytokines from lung, "earliness of death" (EoD)

595 reflecting the number of days prior to the end of experiment that moribund strains were

596 euthanized. Mouse genotypes are ordered by lung CFU. Scaled trait values were

597 clustered (hclust in R package heatmaply) and dendrogram nodes colored by 3 k-

598 means. Blue node reflects correlation coefficient $r > 0.7$; green $r = 0.3-0.6$ and red $r < 0.2$.
599 (C) Correlation of lung CFU and weight (% change) shaded by CXCL1 levels.
600 Genotypes identified as statistical outliers for weight are noted by #; CXCL1 by †. (D)
601 Correlation of lung CFU and IFN γ levels shaded by IL-17. Strains identified as outliers
602 for IFN γ noted by #. Each point in (C) and (D) is the average value per genotype. Outlier
603 genotypes were identified after linear regression using studentized residuals. (E-H) B6
604 and CAST traits measured at 4 weeks post aerosol infection (E) lung CFU; (F) spleen
605 CFU; (G) number of IFN γ producing lung T cells; (H) number of IFN γ producing spleen
606 T cells. Bar plots show the mean \pm SD. Significance between groups was determined
607 by Welch's t test. Groups consist of 4-10 mice per genotype from 2 independent
608 experiments. All mice in the initial CC screen were male; validation studies in panels E-
609 H included both sexes, where no statistical difference at the 4-week timepoint between
610 sexes was observed.

611
612 **Figure 2. Host loci underlying TB disease-related traits.** (A-B) Whole genome QTL
613 scans of (A) spleen CEQ, spleen CFU and IL-10 (B) lung CFU and CXCL1. (C) Zoom of
614 chromosome 2 loci. (D) Zoom of chromosome 7 loci. Thresholds were determined by
615 permutation analysis; solid line, middle dashed line, and lowest dotted lines represent P
616 = 0.05, $P = 0.1$, and $P = 0.2$. (E-F) Scaled phenotype value per haplotype at the QTL
617 peak marker. Each dot represents the mean value for a genotype.

618
619 **Figure 3. An F₂ intercross approach to validate QTL underlying lung CFU.** (A)
620 Cross schema between CC030 and CC029 CC strains that contain CAST informative

621 haplotypes at Chr 7 or 15. CC030 contains CAST at 30-45Mb on Chr 7 and CC029
622 contains CAST at 77-79Mb on Chr 15. The F₂ population based on these founders were
623 genotyped and informative segregants were infected with *Mtb*. (B) Lung CFU at 1-
624 month post-infection. F₂ mice are grouped by the four possible genotypes: CAST at
625 both 7 and 15, CAST at 7, CAST at 15, or non-CAST at both loci. To determine the
626 effect of CAST haplotype on lung CFU, data was analyzed by 2-factor ANOVA to
627 determine statistical significance.

628

629 **Figure 4. *Mtb* genetic requirements vary across diverse hosts.** (A) The number of
630 *Mtb* genes required for growth or survival in each diverse mouse strain across the panel
631 (Qval ≤ 0.05). Orange indicates the mutants required for each strain; turquoise shows
632 the cumulative requirement as each new host strain is added. (B) Venn diagram
633 showing the composition of *Mtb* gene sets required in each category of host (white,
634 largest circle), only required in the CC panel (grey), required in specific immunological
635 KO mice (blue) and genes required in B6 mice (red). In order to be called “essential” in
636 each mouse strain, *Mtb* genes had to be significantly over or unrepresented in at least 2
637 genotypes. (C) Each box shows Log₂ fold change (LFC) of individual mutants from the
638 TnSeq screen relative to the input pool in indicated mouse strains (top); Log₂ fold
639 change of the indicated deletion mutants relative to WT from a pooled mutant validation
640 infection (middle panel); Relative fitness calculated from (middle panel) to account for
641 generation differences in each host due to differential growth rate. Bars are the average
642 of 3-6 mice per mutant/genotype +/- SD. Statistical differences between mini-pool
643 validation groups was assessed by Welch’s t test. (D) Lung CFU and spleen CFU from

644 single strain low-dose aerosol infections of $\Delta bioA$ mutant or WT H37Rv strain in B6 and
645 CAST mice at 2- or 5-weeks post-infection. Dashed line indicates the limit of detection.
646 Each point indicates the average CFU +/- SD of 4-5 mice per group. (E) Log₂ fold
647 change of selected mutants from the TnSeq screen across the CC panel and
648 immunological KO mice. Each dot represents the average LFC per mouse genotype. All
649 mice in the complete mouse panel TnSeq screen were male; mice in the $\Delta bioA$ aerosol
650 validation were female; mice in the mini-pool validation studies were male and female
651 with no significant differences detected.

652

653 **Figure 5. *Mtb* virulence pathways associate with distinct host immune pressures.**

654 (A) Weighted gene correlation network analysis (WGCNA) of the 679 *Mtb* genes that
655 significantly vary across the diverse mouse panel. The most representative genes of
656 each module (intramodular connectivity > 0.6) are shown. (B) Mouse genotypes were
657 clustered based on the relative abundance of the 679 variable *Mtb* mutants. The six
658 major clusters (Cluster A-F) were associated with both CFU and the relative abundance
659 of mutants in each bacterial module (1-20; right hand-side with known functions).
660 Statistical analysis is described in Methods. Yellow shading indicates clusters
661 associated with lung CFU. * indicate modules significantly associated with specific
662 mouse clusters (P<0.05).

663

664 **Figure 6. Identification of ‘Host Interacting with Pathogen’ QTL mapping (*HipQTL*).**

665 (A) Manhattan plot of single *Mtb* mutant QTL mapping across the mouse genome. Each
666 dot represents an individual *Mtb* mutant plotted at the chromosomal location of its

667 maximum LOD score. Red dashed line indicates $P < 0.01$; Blue $P < 0.05$. (B)

668 Chromosome 10 QTL corresponding to *Mtb* eigentraits identified in network analysis in

669 **Figure 5.** Module 3 (type VII secretion, ESX1 operon; orange), Module 4 (Mycobactin

670 synthesis, *mbt*, green) and Module 16 (Cholesterol uptake, *mce4*; purple) are shown.

671 Solid and dotted lines indicated $P = 0.05$ and $P = 0.1$, respectively. (C-E) QTL mapping

672 of single *Mtb* mutants corresponding to the (C) ESX1 module, (D) *mbt* module and (E)

673 *mce4* modules. Coincidence of multiple QTL was assessed by the NL-method of Neto

674 et al (Neto et al., 2012). Thresholds shown are for $N=9$, $N=8$, and $N=6$ for panels C, D,

675 and E, respectively. (F) Parental founder effects underlying Module 3, 4 and 16 QTL.

676 Allele effects were calculated at the peak LOD score marker on chromosome 10. (G)

677 Distribution of log fold change (LFC) of representative single mutants from each

678 module; *eccCa1* (ESX1 module), *mbtE* (*mbt* module) and *mce4F* (*mce4* module)

679 relative to *in vitro*. Each dot is the LFC of the specified mutant in each CC mouse strain.

680 Box and whiskers plots (Tukey) of each trait indicate the median and interquartile range.

681 (H) Spleen CEQ and Spleen CFU for CC strains (box plots as in G). Mouse values are

682 grouped by the parental haplotype allele series underlying the chromosome 10 *Hip42*

683 locus (NOD/WSB vs AJ/B6/NZO). Each dot represents the average CFU/CEQ of each

684 CC genotype. Statistical differences in disease-associated traits and distinct haplotypes

685 groups was assessed by t-test. LOD, logarithm of the odds; LFC, log fold change; CEQ,

686 Chromosomal equivalents; CFU, colony forming units.

687

688 **Tables with title and legends:**

689

690 **Table 1. Disease-related Tuberculosis ImmunoPhenotype QTL (*Tip*QTL). Multiple**

691 QTL within the same interval and clear allele effects are designated with the same

692 *Tip*QTL number. Column headings: QTL, quantitative trait loci; Chr, chromosome; LOD,

693 logarithm of the odds; CEQ, chromosomal equivalents.

694

QTL	Trait	Chr	LOD	P value	Interval start (Mb)	Peak (Mb)	Interval end (Mb)
<i>Tip5</i>	Spleen CEQ	2	9.14	2.38E-02	174.29	178.25	178.25
<i>Tip5</i>	Spleen CFU	2	7.04	2.19E-01	73.98	174.29	180.10
<i>Tip6</i>	IL-9	2	8.61	4.52E-02	33.43	41.48	41.4
<i>Tip6</i>	IL-9	2	7.85	1.26E-01	22.77	24.62	25.65
<i>Tip7</i>	IL-17	15	7.84	5.27E-02	67.98	74.14	82.11
<i>Tip8</i>	CXCL1	7	7.57	1.06E-01	30.43	45.22	46.72
<i>Tip8</i>	Lung CFU	7	7.47	1.17E-01	31.06	37.78	45.22
<i>Tip9</i>	IL-10	17	7.16	1.85E-01	80.98	82.47	83.55
<i>Tip10</i>	Lung CFU	15	7.13	1.86E-01	77.00	78.16	78.70

695

696

697 **Materials and Methods:**

698

699 *Ethics statement:*

700 Mouse studies were performed in strict accordance using the recommendations
701 from the Guide for the Care and Use of Laboratory Animals of the National
702 Institute of Health and the Office of Laboratory Animal Welfare. Mouse studies at
703 the University of Massachusetts Medical School (UMASS) were performed using
704 protocols approved by the UMASS Institutional Animal Care and Use Committee
705 (IACUC) (Animal Welfare Assurance Number A3306-01) in a manner designed to
706 minimize pain and suffering in *Mtb*-infected animals. Any animal that exhibited
707 severe disease signs was immediately euthanized in accordance with IACUC
708 approved endpoints. All mouse studies at UNC (Animal Welfare Assurance
709 #A3410-01) were performed using protocols approved by the UNC Institutional
710 Animal Care and Use Committee (IACUC).

711

712 *Mice:*

713 Male and female Collaborative Cross parental strains (A/J #0646; C57BL/6J #0664;
714 129S1/SvImJ #02448; NOD/ShiLtJ #01976; NZO/HiLtJ #02105; CAST/EiJ #0928,
715 PWK/PhJ #3715 and WSB/EiJ #01145) and single gene immunological knockout mice
716 were purchased from The Jackson Laboratory (*Nos2*^{-/-} #2609, *Cybb*^{-/-} #2365, *Ifny*^{-/-}
717 #2287) or Taconic (*RagN12*) and bred at UMASS. Male mice from 52 CC strains were
718 purchased from the UNC Systems Genetics Core Facility (SGCF) between July 2013
719 and August 2014. The 52 CC strains used in this study include: CC001/Unc,

720 CC002/Unc, CC003/Unc, CC004/TauUnc, CC005/TauUnc, CC006/TauUnc,
721 CC007/Unc, CC008/GeniUnc, CC009/Unc, CC010/GeniUnc, CC011/Unc,
722 CC012/GeniUnc, CC013/GeniUnc, CC015/Unc, CC016/GeniUnc, CC017/Unc,
723 CC018/Unc, CC019/TauUnc, CC021/Unc, CC022/GeniUnc, CC023/GeniUnc,
724 CC024/GeniUnc, CC025/GeniUnc, CC027/GeniUnc, CC028/GeniUnc, CC029/Unc,
725 CC030/GeniUnc, CC031/GeniUnc, CC032/GeniUnc, CC033/GeniUnc, CC034/Unc,
726 CC035/Unc, CC037/TauUnc, CC038/GeniUnc, CC039/Unc, CC040/TauUnc,
727 CC041/TauUnc, CC042/GeniUnc, CC043/GeniUnc, CC044/Unc, CC045/GeniUnc,
728 CC046/Unc, CC047/Unc, CC051/TauUnc, CC055/TauUnc, CC056/GeniUnc,
729 CC057/Unc, CC059/TauUnc, CC060/Unc, CC061/GeniUnc, CC065/Unc,
730 CC068/TauUnc. More information regarding the CC strains can be found at
731 <http://csbio.unc.edu/CCstatus/index.py?run=AvailableLines.information>.

732

733 CC030 x CC029 F₂ mice were generated at UNC by crossing CC030 and CC029
734 mice (purchased from the SGCF in 2016) to generate F₁s (both CC030 dam by
735 CC029 sires as well as CC029 dam by CC030 sires). The resulting F₁s were
736 subsequently intercrossed to generate F₂ mice with all possible grandparental
737 combinations. Male and female F₂ mice were shipped to UMASS for *Mtb*
738 infections.

739

740 All mice were housed in a specific pathogen-free facility under standard
741 conditions (12hr light/dark, food and water ad libitum). Mice were infected with

742 *Mtb* between 8-12 weeks of age. Male and Female mice were used, unless
743 otherwise noted.

744

745 *M.tuberculosis* strains:

746 All *M. tuberculosis* strains were grown in Middlebrook 7H9 medium containing
747 oleic acid-albumin-dextrose-catalase (OADC), 0.2% glycerol, and 0.05% Tween
748 80 to log-phase with shaking (200 rpm) at 37°C. Hygromycin (50 ug/ml) or
749 kanamycin (20 ug/ml) was added when necessary. The Δ *bioA* strain was made
750 by homologous recombination as previously described (Woong Park et al.,
751 2011). For pooled mutant infections, deletion strains were constructed using
752 ORBIT (Murphy et al., 2018), which included gene replacement by the vector
753 pKM464 carrying unique q-Tag sequences to identify each mutant for deep
754 sequencing. The *rnaseJ* mutant was also made by ORBIT and was kindly
755 provided by Dr. Nathan Hicks and Dr Sarah Fortune. Prior to all *in-vivo* infections,
756 cultures were washed, resuspended in phosphate-buffered saline (PBS)
757 containing 0.05% Tween 80, and sonicated before diluting to desired
758 concentration (see below).

759

760 *Mouse Infections:*

761 For TnSeq experiments, 1×10^6 CFU of saturated *himar1* transposon mutants
762 (Sasseti et al., 2003) was delivered via intravenous tail vein injection. Mice were
763 infected over 3 infection batches, as denoted in **Table S1**). At indicated time
764 points mice were euthanized, and organs were harvested then homogenized in a

765 FastPrep-24 (MP Biomedicals). CFU was determined by dilution plating on 7H10
766 agar with 20 ug/mL kanamycin. For library recovery, approximately 1×10^6 CFU
767 per mouse was plated on 7H10 agar with 20ug/mL kanamycin. After three weeks
768 of growth, colonies were harvested by scraping and genomic DNA was extracted.
769 The relative abundance of each transposon mutant was estimated as described
770 (Long et al., 2015).

771

772 Single strain aerosol infections were performed in a Glas-Col machine to deliver
773 50-150 CFU/mouse. At indicated time points, mice were euthanized, and organs
774 were harvested then homogenized in a FastPrep-24 (MP Biomedicals). CFU was
775 determined by dilution plating on 7H10 agar with 20 ug/mL kanamycin or
776 50ug/mL hygromycin as required.

777

778 Chromosomal equivalent (CEQ) was enumerated according to previously
779 published protocol (Lin et al., 2014; Muñoz-Elías et al., 2005). Cytokines and
780 chemokines were assayed from organ homogenates using the pro-inflammatory
781 focused 32-plex (Eve Technologies, Calgary, CA) or IFN γ DuoSet ELISA (R&D
782 Systems) according to manufacturer protocol. IFN γ release was assayed by
783 ELISPOT (BD Biosciences #551083) according to manufacturer protocol. Single
784 cell suspensions of lung or spleen tissue were stimulated for 18-24 hours with
785 4ug/ml purified protein derivative (PPD) and after development, spots were
786 quantified using a CTL Immunospot S5 Analyzer.

787 For pooled mutant infections, mice were infected with a pool of deletion mutants at
788 equal ratios via the intravenous route (1×10^6 CFU/mouse). At indicated time points,
789 approximately 10,000 CFU from the spleen homogenate of each mouse was plated on
790 7H10 agar. Genomic DNA was extracted for sequencing as described previously (Long
791 et al., 2015). Sequencing libraries spanning the variable region of each q-Tag were
792 generated using PCR primers binding to regions common among all q-Tags, similar to
793 previously described protocols (Bellerose et al., 2020; Martin et al., 2017). In each PCR,
794 a unique molecular counter was incorporated into the sequence to allow for the
795 accurate counting of input templates and account for PCR jackpotting. The libraries
796 were sequenced to 1,000-fold coverage on an Illumina NextSeq platform using a 150-
797 cycle Mid-Output kit with single-end reads. Total abundance of each mutant in the
798 library was determined by counting the number of reads for each q-Tag with a unique
799 molecular counter. Relative abundance of each mutant in the pool was then calculated
800 by dividing the total abundance of a mutant by the total abundance of reads for wild-
801 type H37Rv. The relative abundance was normalized to relative abundance at initial
802 infection (Day 0) and Log_2 transformed. Fitness was calculated as previously described
803 (Palace et al., 2014).

804

805 QUANTIFICATION AND STATISTICAL ANALYSIS

806 *TnSeq analysis:*

807 TnSeq libraries were prepared and counts of each transposon mutant were estimated
808 as described (Long et al., 2015). NCBI Reference Sequence NC_018143.1 was used
809 for H37Rv genome and annotations. In the majority of cases, two replicate mouse

810 libraries were used per genotype. Only a single TnSeq library was obtained for CC010,
811 CC031, CC037, CC059, CC016, and PWK/PhJ. Insertion mutant counts across all
812 libraries were normalized by beta-geometric correction (DeJesus et al., 2015), binned
813 by gene, and replicate values for each mouse genotype averaged. Mean values for
814 each gene were divided by the grand mean then \log_{10} transformed and quantile
815 normalized. The resulting phenotype values were used for both WGCNA and QTL
816 mapping.

817

818 To eliminate genes having no meaningful variation across the mouse panel, statistical
819 tests of \log_2 fold change (LFC) in counts between all possible pairs of mouse genotypes
820 were performed by resampling (DeJesus et al., 2015). 679 “significantly varying genes”
821 were identified whose representation varied significantly (FDR < 5%) in at least two
822 independent comparisons. For relative mutant abundance estimates, LFC in counts
823 between *in vitro*-grown H37Rv (6 replicate libraries) vs libraries from each mouse
824 genotype were determined by resampling as above.

825

826 *WGCNA analysis:*

827 Weighted gene correlation network analysis (WGCNA) was applied to categorize the
828 679 significantly varying genes into 20 internally-correlated modules (Langfelder and
829 Horvath, 2008). Modules were filtered (intramodular connectivity > 0.6) to obtain the
830 most representative genes. First principal component scores of module eigengenes
831 were used as phenotype values for QTL mapping after first winsorizing ($q=0.05$) using
832 the R package *broman* (<https://cran.r-project.org/web/packages/broman/index.html>).

833

834 In order to perform association analysis between modules of genes and clusters of mice
835 (Figure 5B), the mice were clustered based on the matrix of TnSeq LFCs for
836 significantly varying genes using *hclust* in R (with the “Ward.D2” distance metric). Then,
837 for each module of genes, the LFCs in each cluster of mice were pooled and compared
838 to all the other mice using a T-test, identifying modules with a mean LFC in a specific
839 mouse cluster that is significantly higher or lower than the average across all the other
840 mice. The resulting p-values over all combinations of gene modules and mouse
841 clusters were adjusted using Benjamini-Hochberg for an overall FDR < 0.05.

842

843 *Disease-related trait analysis and Heritability estimation:*

844 For the trait heatmap, trait values were clustered (*hclust* in R package *heatmaply*; traits
845 scaled as per default function) and dendrogram nodes colored by 3 k-means.

846 Correlation between disease-related TB traits was determined by Pearson’s correlation
847 and visualized using *corrplot* (ordered by *hclust*). Heritability (h^2) of the immunological
848 and TB disease-related traits was calculated by estimating the percent of variation
849 attributed to a genotype as previously described (Noll et al., 2020).

850

851 *Genotyping and QTL mapping:*

852 A subset of the inbred CC mice used in the analysis were genotyped on the GigaMUGA
853 array (Morgan et al., 2015) available from Neogen Inc. The inbred parents, F1s and F2
854 mice from the CC030xCC029 cross were genotyped on the MiniMUGA array (Sigmon et
855 al., 2020) at Neogen Inc.

856

857 For CC030 x CC029 F2 analysis, markers were filtered, restricting only to those
858 diagnostic between these strains (See **Table S3**). To assess the relative impact of the
859 Chr7 and 15 loci, a linear regression on CFU was run with CAST haplotypes on Chr7,
860 Chr15, and their interaction as factors.

861

862 For QTL mapping in the CC panel, the Most Recent Common Ancestor (Srivastava et
863 al., 2017) 36-state haplotypes were downloaded from the UNC Systems Genetics Core
864 Facility and simplified to 8-state haplotype probabilities (for the 8 CC founder strains),
865 which is appropriate for additive genetic mapping. We generated 36-state haplotype
866 probabilities from the individual CC mice genotyped on GigaMUGA and combined these
867 data with the MRCA data to obtain a common genome cache.

868

869 For CC QTL analysis, genotype and phenotype data were imported into R (version
870 3.6.1) and reformatted for R/qtl2 (version 0.20) (Broman et al., 2019). Individual TnSeq
871 and clinical trait phenotype values were winsorized ($q = 0.006$) as above. GigaMUGA
872 annotations were downloaded from the Jackson Laboratory, and markers were thinned
873 to a spacing of 0.1 cM using the `reduce_markers` function of R/qtl2. The final genetic
874 map contained 10,067 markers. QTL mapping was carried out using a linear mixed
875 model with LOCO (leave one chromosome out) kinship. For clinical trait scans, batch
876 (denoted by “block” in Table S1) was included as an additive covariate. Significance
877 thresholds for QTL were estimated using 10,000 permutations (`scan1_perm` function).
878 For each trait, the maximum LOD scores from the permutation scans were used to fit

879 generalized extreme value distributions, from which genome-wide permutation p-values
880 were calculated. LOD profiles and effect plots were generated using the plotting
881 functions of the R/qtl2 package. Multiple QTL at similar genetic locations were assessed
882 for independence using qtl2pleio with 400 bootstrap samples (Boehm et al., 2019). The
883 quantile-based permutation thresholding method of Neto et al. (Neto et al., 2012) was
884 used to assess the statistical significance of co-mapping traits. The NL-method, which
885 determines the LOD thresholds controlling genome wide error rate for a given P value
886 and “hotspot” size, was employed.

887

888 *Candidate gene prioritization:*

889 To identify potential candidate genes, we focused on three QTL that were either
890 statistically significant (*Tip5*, *Hip42*) or were validated by intercross (*Tip8*). For each
891 QTL interval (determined by Bayesian interval in qtl2), we identified mouse genes that
892 were in differentially expressed modules between infected lungs of resistant and
893 susceptible mouse strains (Moreira-Teixeira et al., 2020). Of these genes, we next used
894 the Sanger sequence data (Keane et al., 2011) to filter on genetic variants segregating
895 between CC haplotypes. Where there were clear causal haplotypes, we further filtered
896 to genes with missense or nonsense variants.

897

898

899 **Data availability:**

900 All relevant data to support the findings of this study are located within the paper and
901 supplemental files or are in the process of being made publicly available at the time of

902 this submission; all mouse phenotype data are being deposited in Mouse Phenome
903 Database (<https://phenome.jax.org>).

904

905 **Supplemental information titles and legends:**

906

907 **Table S1 - CC TB phenotypes.** TB disease-related phenotypes measured in the CC
908 and parental strains. Recorded values are the average of measurements from 2-6 mice
909 per genotype (average n=3). Mice were infected over 3 batches (denoted by “block”).
910 “Blaze” denotes genotypes with white head-spotting coat color trait (WSB haplotype for
911 *Kitt*; used as a positive control/proof-of-concept for QTL mapping as per (Aylor et al.,
912 2011; Smith et al., 2019).

913

914 **Table S2 - Heritability (H^2) estimates for each measured TB-disease associated**
915 **phenotype (Tuberculosis ImmunoPhenotypes).** H^2 was calculated from the
916 percentage of variation attributed to strain differences in each trait across the CC
917 strains, as previously described (Noll et al., 2020). Weight change is the percentage of
918 weight (grams), CFU/CEQ is \log_{10} transformed, cytokines are measured in pg/mL lung
919 homogenate and \log_{10} transformed.

920

921 **Table S3 - F_2 Intercross phenotype data.** Lung CFU measured in 46 F_2 mice derived
922 from CC030xCC029 intercross strategy. F_2 mice were CAST at chromosome 7, CAST
923 at Chromosome 15, CAST at both or CAST at neither locus. The infected F_2 cohort
924 included both male and female mice, as indicated.

925

926 **Table S4 - TnSeq Summary Table.** LFC values represent the log₂ fold-change (LFC)
927 between input and mouse-selected pools. "NA" indicates genes with fewer than 3
928 occupied TA transposon insertion sites for the indicated comparison. Qvals represent
929 adjusted p values comparing mutant abundance in input and selected pools. "NA"
930 indicates genes with fewer than 3 occupied TA transposon insertion sites for the
931 indicated comparison. Required *in vivo*: "TRUE" indicates the mutant is significantly
932 underrepresented (Qval <0.05) after in mouse-selection in at least two mouse strains.
933 Required in B6: "TRUE" indicates the mutant is significantly underrepresented (Qval
934 <0.05) after in selection in B6 mice. Required in KO mice: "TRUE" indicates the mutant
935 is significantly underrepresented (Qval <0.05) after in selection in *Rag*^{-/-}, *Nos2*^{-/-}, *Cybb*^{-/-},
936 or *Ifng*^{-/-} mice. Core gene set: TRUE indicates the mutant is significantly
937 underrepresented (Qval <0.05) in 30 mouse strains. "Module" corresponds to WGCNA
938 module number as illustrated in Figure 5A. Mouse strains are listed in the same order
939 as Figure 5B, with the corresponding cluster designation.

940

941 **Table S5 - HipQTL for single *Mtb* mutant QTL and eigentrait/module QTL.** *Hip1-41*
942 each represent host loci associated with the relative abundance of a single mutant
943 (p<0.05). *Hip42-46* correspond to *Mtb* eigentraits identified in network analysis in **Figure**
944 **5** (including significant p<0.05 and suggestive p<0.25). Figure column headings: QTL,
945 quantitative trait loci; *Mtb*, *Mycobacterium tuberculosis*; Module #, number determined
946 from WGCNA modules, ORF, open reading frame; ID, identification number; LOD,
947 logarithmic of the odds; Chr, chromosome.

948

949 **Table S6 - Candidate genes within QTL regions.** Prioritized candidates shown for
950 selected QTL. Candidates were prioritized by filtering on 1) differential expression
951 during *Mtb* infection, and 2) variants within TB-expressed genes that segregated
952 between informative CC haplotypes. Genes listed below contain non-synonymous
953 variants (ie. amino acid changes, regulatory mutations or splicing mutations) consistent
954 with the identified singly causal haplotype (NOD for *Tip5*; CAST for *Tip8*). *Hip42*
955 displayed a more complex haplotype pattern (WSB/NOD vs AJ/B6/NZO), and candidate
956 selection is discussed in the main text. Genes with missense or nonsense variants
957 (denoted by *).

958

959 **Figure S1 - Phenotypic relationships between TB disease-related traits.** Correlation
960 between 32-measured TB traits was determined by Pearson's correlation and visualized
961 using corrplot version 0.84 (ordered by hclust method "complete") in R version 4.0.3.
962 Violet indicates a positive correlation, and yellow indicates negative correlations. The
963 correlation coefficient for each trait comparison (r value) is noted on each square. EoD
964 (earliness of Death); CEQ (Chromosomal equivalents).

965

966 **Figure S2 - Visual representation of QTL mapped in the CC TnSeq infection**
967 **screen.** Tuberculosis ImmunoPhenotypes (*Tip*) QTL (QTL mapped by disease-
968 associated traits in CC mice), are shown in green. *Tip*QTL mapped by separate traits
969 that share similar founder effects were considered to be the same QTL and were named
970 accordingly. Host Interacting with Pathogen (*Hip*) QTL, (QTL mapped by individual

971 TnSeq mutant relative abundance profiles), are shown in purple. After WGCNA mutant
972 clustering and mapping with representative eigengenes from each module, QTL
973 mapped by module eigengenes are shown in magenta.

974

975 **Figure S3 - Module-trait associations.** Rows correspond to modules, columns to
976 clinical traits. Numbers in each cell give the Pearson correlation between the module
977 eigengene and the trait values across the 60-mouse panel (P values in parentheses).
978 Cells are colored by correlation as shown in the color legend (right).

979

980 **References**

981 Abel, L., Fellay, J., Haas, D.W., Schurr, E., Srikrishna, G., Urbanowski, M., Chaturvedi,
982 N., Srinivasan, S., Johnson, D.H., and Bishai, W.R. (2018). Genetics of human
983 susceptibility to active and latent tuberculosis: present knowledge and future
984 perspectives. *Lancet Infect Dis* 18, e64-e75.

985 Adams, L.A., Möller, M., Nebel, A., Schreiber, S., van der Merwe, L., van Helden, P.D.,
986 and Hoal, E.G. (2011). Polymorphisms in MC3R promoter and CTSZ 3'UTR are
987 associated with tuberculosis susceptibility. *Eur J Hum Genet* 19, 676-681.

988 Ahmed, M., Thirunavukkarasu, S., Rosa, B.A., Thomas, K.A., Das, S., Rangel-Moreno,
989 J., Lu, L., Mehra, S., Mbandi, S.K., Thackray, L.B., *et al.* (2020). Immune correlates of
990 tuberculosis disease and risk translate across species. *Sci Transl Med* 12.

991 Altare, F., Durandy, A., Lammas, D., Emile, J.F., Lamhamedi, S., Le Deist, F., Drysdale,
992 P., Jouanguy, E., Döffinger, R., Bernaudin, F., *et al.* (1998). Impairment of
993 mycobacterial immunity in human interleukin-12 receptor deficiency. *Science* 280, 1432-
994 1435.

- 995 Amaral, E.P., Riteau, N., Moayeri, M., Maier, N., Mayer-Barber, K.D., Pereira, R.M.,
996 Lage, S.L., Kubler, A., Bishai, W.R., D'Império-Lima, M.R., *et al.* (2018). Lysosomal
997 Cathepsin Release Is Required for NLRP3-Inflammasome Activation by Mycobacterium
998 tuberculosis in Infected Macrophages. *Front Immunol* 9, 1427.
- 999 Ansari, M.A., Pedergrana, V., C, L.C.I., Magri, A., Von Delft, A., Bonsall, D., Chaturvedi,
1000 N., Bartha, I., Smith, D., Nicholson, G., *et al.* (2017). Genome-to-genome analysis
1001 highlights the effect of the human innate and adaptive immune systems on the hepatitis
1002 C virus. *Nat Genet* 49, 666-673.
- 1003 Aylor, D.L., Valdar, W., Foulds-Mathes, W., Buus, R.J., Verdugo, R.A., Baric, R.S.,
1004 Ferris, M.T., Frelinger, J.A., Heise, M., Frieman, M.B., *et al.* (2011). Genetic analysis of
1005 complex traits in the emerging Collaborative Cross. *Genome Res* 21, 1213-1222.
- 1006 Ayres, J.S., and Schneider, D.S. (2012). Tolerance of infections. *Annu Rev Immunol* 30,
1007 271-294.
- 1008 Barber, D.L., Mayer-Barber, K.D., Feng, C.G., Sharpe, A.H., and Sher, A. (2011). CD4
1009 T cells promote rather than control tuberculosis in the absence of PD-1-mediated
1010 inhibition. *J Immunol* 186, 1598-1607.
- 1011 Bellerose, M.M., Baek, S.-H., Huang, C.-C., Moss, C.E., Koh, E.-I., Proulx, M.K., Smith,
1012 C.M., Baker, R.E., Lee, J.S., Eum, S., *et al.* (2019). Common Variants in the Glycerol
1013 Kinase Gene Reduce Tuberculosis Drug Efficacy. *mBio* 10, e00663-00619.
- 1014 Bellerose, M.M., Proulx, M.K., Smith, C.M., Baker, R.E., Ioerger, T.R., and Sasseti,
1015 C.M. (2020). Distinct Bacterial Pathways Influence the Efficacy of Antibiotics against
1016 Mycobacterium tuberculosis. *mSystems* 5.
- 1017 Berthenet, E., Yahara, K., Thorell, K., Pascoe, B., MERIC, G., Mikhail, J.M., Engstrand,
1018 L., Enroth, H., Burette, A., Megraud, F., *et al.* (2018). A GWAS on Helicobacter pylori
1019 strains points to genetic variants associated with gastric cancer risk. *BMC Biol* 16, 84.
- 1020 Boehm, F.J., Chesler, E.J., Yandell, B.S., and Broman, K.W. (2019). Testing Pleiotropy
1021 vs. Separate QTL in Multiparental Populations. *G3 (Bethesda)* 9, 2317-2324.

- 1022 Bogunovic, D., Byun, M., Durfee, L.A., Abhyankar, A., Sanal, O., Mansouri, D., Salem,
1023 S., Radovanovic, I., Grant, A.V., Adimi, P., *et al.* (2012). Mycobacterial disease and
1024 impaired IFN- γ immunity in humans with inherited ISG15 deficiency. *Science* 337, 1684-
1025 1688.
- 1026 Broman, K.W., Gatti, D.M., Simecek, P., Furlotte, N.A., Prins, P., Sen, \acute{S} ., Yandell, B.S.,
1027 and Churchill, G.A. (2019). R/qtl2: Software for Mapping Quantitative Trait Loci with
1028 High-Dimensional Data and Multiparent Populations. *Genetics* 211, 495-502.
- 1029 Bustamante, J., Boisson-Dupuis, S., Abel, L., and Casanova, J.L. (2014). Mendelian
1030 susceptibility to mycobacterial disease: genetic, immunological, and clinical features of
1031 inborn errors of IFN- γ immunity. *Semin Immunol* 26, 454-470.
- 1032 Caruso, A.M., Serbina, N., Klein, E., Triebold, K., Bloom, B.R., and Flynn, J.L. (1999).
1033 Mice deficient in CD4 T cells have only transiently diminished levels of IFN-gamma, yet
1034 succumb to tuberculosis. *J Immunol* 162, 5407-5416.
- 1035 Caws, M., Thwaites, G., Dunstan, S., Hawn, T.R., Lan, N.T., Thuong, N.T.,
1036 Stepniowska, K., Huyen, M.N., Bang, N.D., Loc, T.H., *et al.* (2008). The influence of
1037 host and bacterial genotype on the development of disseminated disease with
1038 *Mycobacterium tuberculosis*. *PLoS Pathog* 4, e1000034.
- 1039 Chen, L., Xie, Q.W., and Nathan, C. (1998). Alkyl hydroperoxide reductase subunit C
1040 (AhpC) protects bacterial and human cells against reactive nitrogen intermediates. *Mol*
1041 *Cell* 1, 795-805.
- 1042 Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis,
1043 W.D., Belknap, J.K., Bennett, B., Berrettini, W., *et al.* (2004). The Collaborative Cross, a
1044 community resource for the genetic analysis of complex traits. *Nat Genet* 36, 1133-
1045 1137.
- 1046 Churchill, G.A., Gatti, D.M., Munger, S.C., and Svenson, K.L. (2012). The Diversity
1047 Outbred mouse population. *Mamm Genome* 23, 713-718.

- 1048 Comstock, G.W. (1978). Tuberculosis in twins: a re-analysis of the Prophit survey. *Am*
1049 *Rev Respir Dis* *117*, 621-624.
- 1050 Cooke, G.S., Campbell, S.J., Bennett, S., Lienhardt, C., McAdam, K.P., Sirugo, G.,
1051 Sow, O., Gustafson, P., Mwangulu, F., van Helden, P., *et al.* (2008). Mapping of a novel
1052 susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. *Am J*
1053 *Respir Crit Care Med* *178*, 203-207.
- 1054 Cooper, A.M., Dalton, D.K., Stewart, T.A., Griffin, J.P., Russell, D.G., and Orme, I.M.
1055 (1993). Disseminated tuberculosis in interferon gamma gene-disrupted mice. *J Exp Med*
1056 *178*, 2243-2247.
- 1057 Cooper, A.M., Magram, J., Ferrante, J., and Orme, I.M. (1997). Interleukin 12 (IL-12) is
1058 crucial to the development of protective immunity in mice intravenously infected with
1059 mycobacterium tuberculosis. *J Exp Med* *186*, 39-45.
- 1060 de Keijzer, J., de Haas, P.E., de Ru, A.H., van Veelen, P.A., and van Soolingen, D.
1061 (2014). Disclosure of selective advantages in the "modern" sublineage of the
1062 Mycobacterium tuberculosis Beijing genotype family by quantitative proteomics. *Mol*
1063 *Cell Proteomics* *13*, 2632-2645.
- 1064 DeJesus, M.A., Ambadipudi, C., Baker, R., Sasseti, C., and Ioerger, T.R. (2015).
1065 TRANSIT--A Software Tool for Himar1 TnSeq Analysis. *PLoS Comput Biol* *11*,
1066 e1004401.
- 1067 Dick, T., Manjunatha, U., Kappes, B., and Gengenbacher, M. (2010). Vitamin B6
1068 biosynthesis is essential for survival and virulence of Mycobacterium tuberculosis. *Mol*
1069 *Microbiol* *78*, 980-988.
- 1070 Ferris, M.T., Aylor, D.L., Bottomly, D., Whitmore, A.C., Aicher, L.D., Bell, T.A., Bradel-
1071 Tretheway, B., Bryan, J.T., Buus, R.J., Gralinski, L.E., *et al.* (2013). Modeling host
1072 genetic regulation of influenza pathogenesis in the collaborative cross. *PLoS Pathog* *9*,
1073 e1003196.

- 1074 Filipe-Santos, O., Bustamante, J., Haverkamp, M.H., Vinolo, E., Ku, C.L., Puel, A.,
1075 Frucht, D.M., Christel, K., von Bernuth, H., Jouanguy, E., *et al.* (2006). X-linked
1076 susceptibility to mycobacteria is caused by mutations in NEMO impairing CD40-
1077 dependent IL-12 production. *J Exp Med* *203*, 1745-1759.
- 1078 Flynn, J.L., Chan, J., Triebold, K.J., Dalton, D.K., Stewart, T.A., and Bloom, B.R. (1993).
1079 An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis*
1080 infection. *J Exp Med* *178*, 2249-2254.
- 1081 Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S.,
1082 Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., *et al.* (2006). Variable host-
1083 pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* *103*,
1084 2869-2873.
- 1085 Geiman, D.E., Kaushal, D., Ko, C., Tyagi, S., Manabe, Y.C., Schroeder, B.G.,
1086 Fleischmann, R.D., Morrison, N.E., Converse, P.J., Chen, P., *et al.* (2004). Attenuation
1087 of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking
1088 the SigF alternate sigma factor and identification of SigF-dependent genes by
1089 microarray analysis. *Infect Immun* *72*, 1733-1745.
- 1090 Goitre, L., Trapani, E., Trabalzini, L., and Retta, S.F. (2014). The Ras superfamily of
1091 small GTPases: the unlocked secrets. *Methods Mol Biol* *1120*, 1-18.
- 1092 Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C.,
1093 Moreau, F., Laval, F., Daffé, M., Martin, C., Brosch, R., *et al.* (2014). Evolutionary
1094 history of tuberculosis shaped by conserved mutations in the PhoPR virulence
1095 regulator. *Proc Natl Acad Sci U S A* *111*, 11491-11496.
- 1096 Gopal, R., Lin, Y., Obermajer, N., Slight, S., Nuthalapati, N., Ahmed, M., Kalinski, P.,
1097 and Khader, S.A. (2012). IL-23-dependent IL-17 drives Th1-cell responses following
1098 *Mycobacterium bovis* BCG vaccination. *European journal of immunology* *42*, 364 - 373.
- 1099 Gopal, R., Monin, L., Torres, D., Slight, S., Mehra, S., McKenna, K.C., Fallert Junecko,
1100 B.A., Reinhart, T.A., Kolls, J., Báez-Saldaña, R., *et al.* (2013). S100A8/A9 proteins

- 1101 mediate neutrophilic inflammation and lung pathology during tuberculosis. *Am J Respir*
1102 *Crit Care Med* 188, 1137-1146.
- 1103 Gralinski, L.E., Menachery, V.D., Morgan, A.P., Totura, A.L., Beall, A., Kocher, J.,
1104 Plante, J., Harrison-Shostak, D.C., Schäfer, A., Pardo-Manuel de Villena, F., *et al.*
1105 (2017). Allelic Variation in the Toll-Like Receptor Adaptor Protein Ticam2 Contributes to
1106 SARS-Coronavirus Pathogenesis in Mice. *G3 (Bethesda)* 7, 1653-1663.
- 1107 Hillas, P.J., del Alba, F.S., Oyarzabal, J., Wilks, A., and Ortiz De Montellano, P.R.
1108 (2000). The AhpC and AhpD antioxidant defense system of Mycobacterium
1109 tuberculosis. *J Biol Chem* 275, 18801-18809.
- 1110 Holt, K.E., McAdam, P., Thai, P.V.K., Thuong, N.T.T., Ha, D.T.M., Lan, N.N., Lan, N.H.,
1111 Nhu, N.T.Q., Hai, H.T., Ha, V.T.N., *et al.* (2018). Frequent transmission of the
1112 Mycobacterium tuberculosis Beijing lineage and positive selection for the EsxW Beijing
1113 variant in Vietnam. *Nat Genet* 50, 849-856.
- 1114 Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A.,
1115 Agam, A., Slater, G., Goodson, M., *et al.* (2011). Mouse genomic variation and its effect
1116 on phenotypes and gene regulation. *Nature* 477, 289-294.
- 1117 Keller, M.P., Gatti, D.M., Schueler, K.L., Rabaglia, M.E., Stapleton, D.S., Simecek, P.,
1118 Vincent, M., Allen, S., Broman, A.T., Bacher, R., *et al.* (2018). Genetic Drivers of
1119 Pancreatic Islet Function. *Genetics* 209, 335-356.
- 1120 Khader, S.A., Bell, G.K., Pearl, J.E., Fountain, J.J., Rangel-Moreno, J., Cilley, G.E.,
1121 Shen, F., Eaton, S.M., Gaffen, S.L., Swain, S.L., *et al.* (2007). IL-23 and IL-17 in the
1122 establishment of protective pulmonary CD4+ T cell responses after vaccination and
1123 during Mycobacterium tuberculosis challenge. *Nat Immunol* 8, 369-377.
- 1124 Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation
1125 network analysis. *Bmc Bioinformatics* 9, 559.
- 1126 Lázár-Molnár, E., Chen, B., Sweeney, K.A., Wang, E.J., Liu, W., Lin, J., Porcelli, S.A.,
1127 Almo, S.C., Nathenson, S.G., and Jacobs, W.R., Jr. (2010). Programmed death-1 (PD-

- 1128 1)-deficient mice are extraordinarily sensitive to tuberculosis. *Proc Natl Acad Sci U S A*
1129 *107*, 13402-13407.
- 1130 Le Chevalier, F., Cascioferro, A., Frigui, W., Pawlik, A., Boritsch, E.C., Bottai, D.,
1131 Majlessi, L., Herrmann, J.L., and Brosch, R. (2015). Revisiting the role of
1132 phospholipases C in virulence and the lifecycle of *Mycobacterium tuberculosis*. *Sci Rep*
1133 *5*, 16918.
- 1134 Lin, P.L., Ford, C.B., Coleman, M.T., Myers, A.J., Gawande, R., Ioerger, T., Sacchettini,
1135 J., Fortune, S.M., and Flynn, J.L. (2014). Sterilization of granulomas is common in
1136 active and latent tuberculosis despite within-host variability in bacterial killing. *Nat Med*
1137 *20*, 75-79.
- 1138 Long, J.E., DeJesus, M., Ward, D., Baker, R.E., Ioerger, T., and Sasseti, C.M. (2015).
1139 Identifying essential genes in *Mycobacterium tuberculosis* by global phenotypic profiling.
1140 *Methods Mol Biol* *1279*, 79-95.
- 1141 Long, M., Kranjc, T., Mysior, M.M., and Simpson, J.C. (2020). RNA Interference
1142 Screening Identifies Novel Roles for RhoBTB1 and RhoBTB3 in Membrane Trafficking
1143 Events in Mammalian Cells. *Cells* *9*.
- 1144 Lu, L.L., Smith, M.T., Yu, K.K.Q., Luedemann, C., Suscovich, T.J., Grace, P.S., Cain,
1145 A., Yu, W.H., McKittrick, T.R., Lauffenburger, D., *et al.* (2019). IFN- γ -independent
1146 immune markers of *Mycobacterium tuberculosis* exposure. *Nat Med* *25*, 977-987.
- 1147 Martin, C.J., Cadena, A.M., Leung, V.W., Lin, P.L., Maiello, P., Hicks, N., Chase, M.R.,
1148 Flynn, J.L., and Fortune, S.M. (2017). Digitally Barcoding *Mycobacterium tuberculosis*
1149 Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *mBio* *8*.
- 1150 McHenry, M.L., Williams, S.M., and Stein, C.M. (2020). Genetics and evolution of
1151 tuberculosis pathogenesis: New perspectives and approaches. *Infect Genet Evol* *81*,
1152 104204.

- 1153 Minch, K.J., Rustad, T.R., Peterson, E.J., Winkler, J., Reiss, D.J., Ma, S., Hickey, M.,
1154 Brabant, W., Morrison, B., Turkarslan, S., *et al.* (2015). The DNA-binding network of
1155 *Mycobacterium tuberculosis*. *Nat Commun* 6, 5829.
- 1156 Mishra, B.B., Lovewell, R.R., Olive, A.J., Zhang, G.L., Wang, W.F., Eugenin, E., Smith,
1157 C.M., Phuah, J.Y., Long, J.E., Dubuke, M.L., *et al.* (2017). Nitric oxide prevents a
1158 pathogen-permissive granulocytic inflammation during tuberculosis. *Nature Microbiology*
1159 2.
- 1160 Mishra, B.B., Rathinam, V.A., Martens, G.W., Martinot, A.J., Kornfeld, H., Fitzgerald,
1161 K.A., and Sasseti, C.M. (2013). Nitric oxide controls the immunopathology of
1162 tuberculosis by inhibiting NLRP3 inflammasome-dependent processing of IL-1 β . *Nat*
1163 *Immunol* 14, 52-60.
- 1164 Moreira-Teixeira, L., Tabone, O., Graham, C.M., Singhanian, A., Stavropoulos, E.,
1165 Redford, P.S., Chakravarty, P., Priestnall, S.L., Suarez-Bonnet, A., Herbert, E., *et al.*
1166 (2020). Mouse transcriptome reveals potential signatures of protection and
1167 pathogenesis in human tuberculosis. *Nat Immunol* 21, 464-476.
- 1168 Morgan, A.P., Fu, C.P., Kao, C.Y., Welsh, C.E., Didion, J.P., Yadgary, L., Hyacinth, L.,
1169 Ferris, M.T., Bell, T.A., Miller, D.R., *et al.* (2015). The Mouse Universal Genotyping
1170 Array: From Substrains to Subspecies. *G3 (Bethesda)* 6, 263-279.
- 1171 Morth, J.P., Gosmann, S., Nowak, E., and Tucker, P.A. (2005). A novel two-component
1172 system found in *Mycobacterium tuberculosis*. *FEBS Lett* 579, 4145-4148.
- 1173 Muñoz-Elías, E.J., Timm, J., Botha, T., Chan, W.T., Gomez, J.E., and McKinney, J.D.
1174 (2005). Replication dynamics of *Mycobacterium tuberculosis* in chronically infected
1175 mice. *Infect Immun* 73, 546-551.
- 1176 Murphy, K.C., Nelson, S.J., Nambi, S., Papavinasasundaram, K., Baer, C.E., and
1177 Sasseti, C.M. (2018). ORBIT: a New Paradigm for Genetic Engineering of
1178 *Mycobacterial Chromosomes*. *mBio* 9.

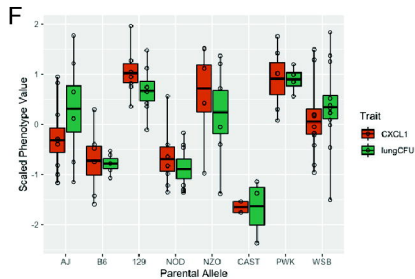
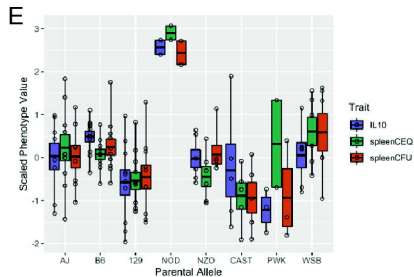
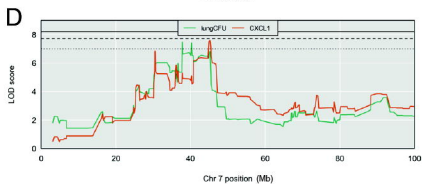
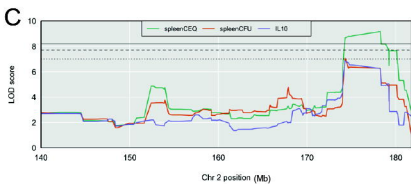
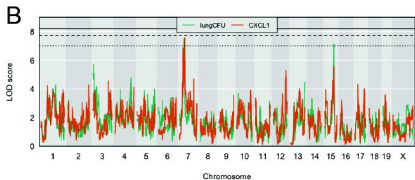
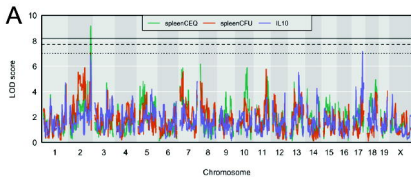
- 1179 Nambi, S., Long, J.E., Mishra, B.B., Baker, R., Murphy, K.C., Olive, A.J., Nguyen, H.P.,
1180 Shaffer, S.A., and Sasseti, C.M. (2015). The Oxidative Stress Network of
1181 *Mycobacterium tuberculosis* Reveals Coordination between Radical Detoxification
1182 Systems. *Cell Host Microbe* 17, 829-837.
- 1183 Neto, E.C., Keller, M.P., Broman, A.F., Attie, A.D., Jansen, R.C., Broman, K.W., and
1184 Yandell, B.S. (2012). Quantile-based permutation thresholds for quantitative trait loci
1185 hotspots. *Genetics* 191, 1355-1365.
- 1186 Niazi, M.K., Dhulekar, N., Schmidt, D., Major, S., Cooper, R., Abeijon, C., Gatti, D.M.,
1187 Kramnik, I., Yener, B., Gurcan, M., *et al.* (2015). Lung necrosis and neutrophils reflect
1188 common pathways of susceptibility to *Mycobacterium tuberculosis* in genetically
1189 diverse, immune-competent mice. *Dis Model Mech* 8, 1141-1153.
- 1190 Noll, K.E., Ferris, M.T., and Heise, M.T. (2019). The Collaborative Cross: A Systems
1191 Genetics Resource for Studying Host-Pathogen Interactions. *Cell Host Microbe* 25, 484-
1192 498.
- 1193 Noll, K.E., Whitmore, A.C., West, A., McCarthy, M.K., Morrison, C.R., Plante, K.S.,
1194 Hampton, B.K., Kollmus, H., Pilzner, C., Leist, S.R., *et al.* (2020). Complex Genetic
1195 Architecture Underlies Regulation of Influenza-A-Virus-Specific Antibody Responses in
1196 the Collaborative Cross. *Cell Rep* 31, 107587.
- 1197 Obermajer, N., Svajger, U., Bogyo, M., Jeras, M., and Kos, J. (2008). Maturation of
1198 dendritic cells depends on proteolytic cleavage by cathepsin X. *J Leukoc Biol* 84, 1306-
1199 1315.
- 1200 Olive, A.J., Smith, C.M., Kiritsy, M.C., and Sasseti, C.M. (2018). The Phagocyte
1201 Oxidase Controls Tolerance to *Mycobacterium tuberculosis* Infection. *J Immunol* 201,
1202 1705-1716.
- 1203 Palace, S.G., Proulx, M.K., Lu, S., Baker, R.E., and Goguen, J.D. (2014). Genome-wide
1204 mutant fitness profiling identifies nutritional requirements for optimal growth of *Yersinia*
1205 *pestis* in deep tissue. *mBio* 5.

- 1206 Pandey, A.K., and Sassetti, C.M. (2008). Mycobacterial persistence requires the
1207 utilization of host cholesterol. *Proc Natl Acad Sci U S A* *105*, 4376-4380.
- 1208 Pethe, K., Sequeira, P.C., Agarwalla, S., Rhee, K., Kuhen, K., Phong, W.Y., Patel, V.,
1209 Beer, D., Walker, J.R., Duraiswamy, J., *et al.* (2010). A chemical genetic screen in
1210 *Mycobacterium tuberculosis* identifies carbon-source-dependent growth inhibitors
1211 devoid of in vivo efficacy. *Nat Commun* *1*, 57.
- 1212 Rodrigue, S., Provvedi, R., Jacques, P.E., Gaudreau, L., and Manganelli, R. (2006).
1213 The sigma factors of *Mycobacterium tuberculosis*. *FEMS Microbiol Rev* *30*, 926-941.
- 1214 Ryndak, M.B., Wang, S., Smith, I., and Rodriguez, G.M. (2010). The *Mycobacterium*
1215 *tuberculosis* high-affinity iron importer, *IrtA*, contains an FAD-binding domain. *J*
1216 *Bacteriol* *192*, 861-869.
- 1217 Sakai, S., Kauffman, K.D., Sallin, M.A., Sharpe, A.H., Young, H.A., Ganusov, V.V., and
1218 Barber, D.L. (2016). CD4 T Cell-Derived IFN- γ Plays a Minimal Role in Control of
1219 Pulmonary *Mycobacterium tuberculosis* Infection and Must Be Actively Repressed by
1220 PD-1 to Prevent Lethal Disease. *PLoS Pathog* *12*, e1005667.
- 1221 Sassetti, C.M., Boyd, D.H., and Rubin, E.J. (2003). Genes required for mycobacterial
1222 growth defined by high density mutagenesis. *Mol Microbiol* *48*, 77-84.
- 1223 Sassetti, C.M., and Rubin, E.J. (2003). Genetic requirements for mycobacterial survival
1224 during infection. *Proc Natl Acad Sci U S A* *100*, 12989-12994.
- 1225 Saul, M.C., Philip, V.M., Reinholdt, L.G., and Chesler, E.J. (2019). High-Diversity Mouse
1226 Populations for Complex Traits. *Trends Genet* *35*, 501-514.
- 1227 Saunders, B.M., Frank, A.A., Orme, I.M., and Cooper, A.M. (2002). CD4 is required for
1228 the development of a protective granulomatous response to pulmonary tuberculosis.
1229 *Cell Immunol* *216*, 65-72.
- 1230 Shorter, J.R., Najarian, M.L., Bell, T.A., Blanchard, M., Ferris, M.T., Hock, P., Kashfeen,
1231 A., Kirchoff, K.E., Linnertz, C.L., Sigmon, J.S., *et al.* (2019). Whole Genome

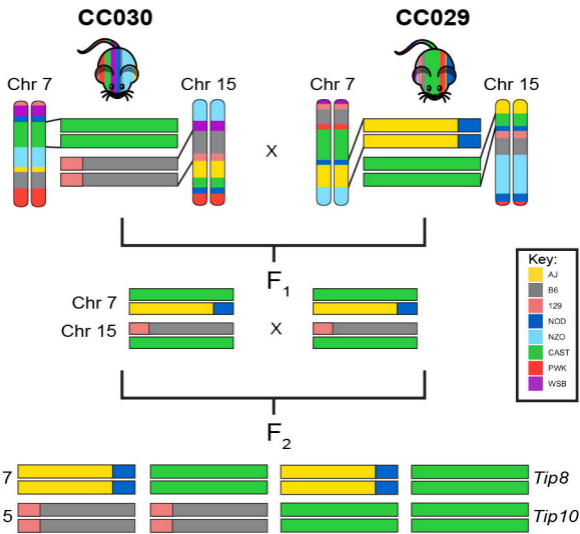
- 1232 Sequencing and Progress Toward Full Inbreeding of the Mouse Collaborative Cross
1233 Population. *G3 (Bethesda)* 9, 1303-1311.
- 1234 Sigmon, J.S., Blanchard, M.W., Baric, R.S., Bell, T.A., Brennan, J., Brockmann, G.A.,
1235 Burks, A.W., Calabrese, J.M., Caron, K.M., Cheney, R.E., *et al.* (2020). Content and
1236 performance of the MiniMUGA genotyping array, a new tool to improve rigor and
1237 reproducibility in mouse research. *bioRxiv*, 2020.2003.2012.989400.
- 1238 Smith, C.M., Proulx, M.K., Lai, R., Kiritsy, M.C., Bell, T.A., Hock, P., Pardo-Manuel de
1239 Villena, F., Ferris, M.T., Baker, R.E., Behar, S.M., *et al.* (2019). Functionally
1240 Overlapping Variants Control Tuberculosis Susceptibility in Collaborative Cross Mice.
1241 *mBio* 10.
- 1242 Smith, C.M., Proulx, M.K., Olive, A.J., Laddy, D., Mishra, B.B., Moss, C., Gutierrez,
1243 N.M., Bellerose, M.M., Barreira-Silva, P., Phuah, J.Y., *et al.* (2016). Tuberculosis
1244 Susceptibility and Vaccine Protection Are Independently Controlled by Host Genotype.
1245 *mBio* 7.
- 1246 Smith, C.M., and Sasseti, C.M. (2018). Modeling Diversity: Do Homogeneous
1247 Laboratory Strains Limit Discovery? *Trends Microbiol* 26, 892-895.
- 1248 Sousa, J., Cá, B., Maceiras, A.R., Simões-Costa, L., Fonseca, K.L., Fernandes, A.I.,
1249 Ramos, A., Carvalho, T., Barros, L., Magalhães, C., *et al.* (2020). Mycobacterium
1250 tuberculosis associated with severe tuberculosis evades cytosolic surveillance systems
1251 and modulates IL-1 β production. *Nat Commun* 11, 1949.
- 1252 Springer, B., Master, S., Sander, P., Zahrt, T., McFalone, M., Song, J.,
1253 Papavinasasundaram, K.G., Colston, M.J., Boettger, E., and Deretic, V. (2001).
1254 Silencing of oxidative stress response in Mycobacterium tuberculosis: expression
1255 patterns of *ahpC* in virulent and avirulent strains and effect of *ahpC* inactivation. *Infect*
1256 *Immun* 69, 5967-5973.

- 1257 Srivastava, A., Morgan, A.P., Najarian, M.L., Sarsani, V.K., Sigmon, J.S., Shorter, J.R.,
1258 Kashfeen, A., McMullan, R.C., Williams, L.H., Giusti-Rodríguez, P., *et al.* (2017).
1259 Genomes of the Mouse Collaborative Cross. *Genetics* 206, 537-556.
- 1260 Stanley, S.A., Raghavan, S., Hwang, W.W., and Cox, J.S. (2003). Acute infection and
1261 macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion
1262 system. *Proceedings of the National Academy of Sciences* 100, 13001-13006.
- 1263 Svenson, K.L., Gatti, D.M., Valdar, W., Welsh, C.E., Cheng, R., Chesler, E.J., Palmer,
1264 A.A., McMillan, L., and Churchill, G.A. (2012). High-resolution genetic mapping using
1265 the Mouse Diversity outbred population. *Genetics* 190, 437-447.
- 1266 Thuong, N.T., Tram, T.T., Dinh, T.D., Thai, P.V., Heemskerk, D., Bang, N.D., Chau,
1267 T.T., Russell, D.G., Thwaites, G.E., Hawn, T.R., *et al.* (2016). MARCO variants are
1268 associated with phagocytosis, pulmonary tuberculosis susceptibility and Beijing lineage.
1269 *Genes Immun* 17, 419-425.
- 1270 Turkarslan, S., Peterson, E.J., Rustad, T.R., Minch, K.J., Reiss, D.J., Morrison, R., Ma,
1271 S., Price, N.D., Sherman, D.R., and Baliga, N.S. (2015). A comprehensive map of
1272 genome-wide gene regulation in *Mycobacterium tuberculosis*. *Sci Data* 2, 150010.
- 1273 Wang, Q., Boshoff, H.I.M., Harrison, J.R., Ray, P.C., Green, S.R., Wyatt, P.G., and
1274 Barry, C.E., 3rd (2020). PE/PPE proteins mediate nutrient transport across the outer
1275 membrane of *Mycobacterium tuberculosis*. *Science* 367, 1147-1151.
- 1276 Woong Park, S., Klotzsche, M., Wilson, D.J., Boshoff, H.I., Eoh, H., Manjunatha, U.,
1277 Blumenthal, A., Rhee, K., Barry, C.E., 3rd, Aldrich, C.C., *et al.* (2011). Evaluating the
1278 sensitivity of *Mycobacterium tuberculosis* to biotin deprivation using regulated gene
1279 expression. *PLoS Pathog* 7, e1002264.
- 1280 Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Thompson, E., Suliman, S., Amon, L.M.,
1281 Mahomed, H., Erasmus, M., Whatney, W., Hussey, G.D., *et al.* (2016). A blood RNA
1282 signature for tuberculosis disease risk: a prospective cohort study. *Lancet* 387, 2312-
1283 2322.

- 1284 Zhang, J., Teh, M., Kim, J., Eva, M.M., Cayrol, R., Meade, R., Nijnik, A., Montagutelli,
1285 X., Malo, D., and Jaubert, J. (2019). A Loss-of-Function Mutation in the Integrin Alpha L
1286 (Itgal) Gene Contributes to Susceptibility to Salmonella enterica Serovar Typhimurium
1287 Infection in Collaborative Cross Strain CC042. *Infect Immun* 88.
- 1288 Zhang, Y.J., Reddy, M.C., Ioerger, T.R., Rothchild, A.C., Dartois, V., Schuster, B.M.,
1289 Trauner, A., Wallis, D., Galaviz, S., Huttenhower, C., *et al.* (2013). Tryptophan
1290 Biosynthesis Protects Mycobacteria from CD4 T-Cell-Mediated Killing. *Cell* 155, 1296 -
1291 1308.
- 1292



A



B

