Title: Genetic and Environmental interactions contribute to immune variation in rewilded mice

Authors: Oyebola Oyesola[1#*], Alexander E. Downie[2#], Nina Howard[1], Ramya S. Barre[2], Kasalina Kiwanuka[1], Kimberly Zaldana[1], Ying-Han Chen[3], Arthur Menezes[2], Soo Ching Lee[1], Joseph Devlin[3], Octavio Mondragón-Palomino[1], Camila Oliveira Silva Souza[1], Christin Herrmann[3], Sergei B. Koralov[3,4], Ken Cadwell[3,5], Andrea L. Graham[2,6]*, P'ng Loke[1]*

Affiliations:

[1] Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institute of Health; Bethesda, MD, USA.

[2] Department of Ecology and Evolutionary Biology, Princeton University; Princeton, NJ, USA.

[3] Kimmel Center for Biology and Medicine at the Skirball Institute, New York University, Grossman School of Medicine; New York, NY, USA.

[4] Department of Pathology, New York University Grossman School of Medicine; New York, NY, USA.

[5]University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.

[6] Santa Fe Institute, Santa Fe, NM, USA

# These authors contributed equally

*Corresponding Author(s). Email: oyebola.oyesola@nih.gov, algraham@princeton.edu, png.loke@nih.gov.

One-Sentence Summary: Rewilding mice reveals how genotype and environment combine to determine inter-individual immune variation.

Abstract: The relative and potentially synergistic contributions of genetics and environment to inter-individual immune response variation remain unclear, despite critical implications of such variation in both medicine and evolutionary biology. Here, we quantify interactive effects of genotype and environment on immune traits by investigating different inbred mouse strains rewilded in outdoor enclosures and infected with the parasite, *Trichuris muris*. Whereas cytokine response heterogeneity was primarily driven by genotype, cellular composition heterogeneity was shaped by interactions between genotype and environment with genetic mediated differences decreasing following rewilding, but less dramatically for T cells than for B cells. Importantly, immune variation was associated with altered parasite burdens. These results indicate that nonheritable influences interact with genetic factors to shape immune variation, with synergistic impacts on the deployment and potential evolution of defense mechanisms.

Main Text:

An individual's immune phenotype is shaped by some combination of genetic factors and nonheritable influences such as environmental exposure (including infection history and the microbiome) *(1-9)*. However, the relative contributions of heritable and nonheritable factors to inter-individual immune variation in mammals remains controversial despite the importance of such variation for both medicine and evolutionary biology. For example, variation in immune responses can determine whether an individual will experience severe or asymptomatic infection *(10, 11)*, and whether severity arises due to failure to control pathogens or excessive collateral tissue damage following defective immune regulation *(3)*.

Recent studies on the human immune system have aimed at identifying the relative contributions of genetic and environmental factors to variation in immune phenotypes among healthy individuals *(1, 6, 7)*, as well as during infection *(7)* or inflammatory conditions *(12)*. Such studies tend to draw upon analysis of immunological divergence between identical twins *(7)* or genetic heritability estimation for immune traits through functional genomics *(1)*. These methods are powerful but do not readily enable quantification of interactive effects of genetics and environment. In most cases, variation not attributable to genetics is attributed to environment alone, rather than the possibility that effects of environment are differentially amplified in different genotypes, or vice-versa. Important context-dependency in immune function *(13)* is thus missing from these calculations. For example, what if the impact of environment upon memory T cell frequencies depends upon host genotype, or if, conversely, the impact of genotype upon memory T cell frequencies depends upon environment? Evolutionary biology is explicitly interested in such context dependencies because they provide the raw materials for adaptive evolution, indeed, genotype-by-environment interactions are common and substantial in effect for a variety of traits *(14, 15)* including disease outcomes *(16, 17)*. Here, we deploy statistical frameworks that quantify interactions in addition to the independent (or main) effects of genetics and environment to elucidate causes of variation in experimental immunology.

Controlled experiments with mice could help decipher interactions between genetic and environmental effects on the immune system, but most studies in mice instead aim to reduce environmental variation to discover genetic factors regulating cellular and molecular components of immunity *(3, 18-20)*. This approach not only ignores interactions but also gives us only partial insight into direct genetic effects by ignoring the extent of the measured genetic effect that is mediated by the particular environment. We have taken a decidedly different approach of using an outdoor enclosure system to introduce laboratory mice of different genotypes into a natural environment that we term "rewilding" *(21)*. We track behavior outdoors *(22)*, recover the mice for analysis, and then investigate genetic and environmental contributions to immune phenotypes *(20, 21, 23)*. For example, using mice with mutant alleles in inflammatory bowel disease susceptibility genes (*Nod2* and *Atg16l1*), we found that the genetic mutations affected the production of cytokines in response to microbial stimulation, whereas immune cell composition was more influenced by environment *(20, 23)*. We have also found that rewilded C57BL/6 mice become more susceptible to infection with the intestinal nematode parasite *Trichuris muris (21)*. However, those experiments explored limited genetic variation and thus did not examine whether interactions between genetics and environment would influence immune phenotype and helminth susceptibility *(4, 24)*.

To quantify relative contributions and interactions between genetic and nonheritable influences on heterogeneity in immune profiles and helminth susceptibility, here we compared

C57BL/6, 129S1, and PWK/PhJ mice kept in a conventional vivarium versus those that were rewilded (Fig1A). These host strains differ by up to 50 million single nucleotide polymorphisms (SNPs) and short insertions/deletions (indels) *(25, 26)*; for perspective on this range of variation, the entire human population is estimated to contain approximately 90 million SNPs and indels *(27)*. We rewilded mice (n = 72) or kept them in laboratory housing (n = 63) for 2 weeks and then infected them with approximately 200 eggs of the intestinal helminth *Trichuris muris* (n = 61) or left them uninfected (n = 74), returning them to the outdoor or vivarium environment for a further 3 weeks. We used *Trichuris muris* intestinal nematode, since, we have previously shown that rewilded C57BL/6 mice had bigger and higher *Trichuris muri*s worm burden compared their laboratory controls (*21*). We conducted two replicate experiments (Block 1, n = 61; Block 2, n = 74). We collected blood and mesenteric lymph nodes (MLN) and completed differential Complete Blood Counts (CBC). To assess immune cell composition, we analyzed peripheral blood mononuclear cells (PBMCs) by flow cytometry with a lymphocyte panel (table S1) and MLN cells with both a lymphocyte and myeloid cell panel (table S1 and S2). To assess cytokine responses, we measured plasma cytokine concentrations and stimulated MLN cells with microbial antigens. Single cell RNA sequencing (scRNAseq) of MLN cells enabled phenotyping of both immune cell composition and function. We also assessed worm burden and worm prevalence in all mice exposed to *T. muris*. Serology and PCR screening panels tested for over 30 pathogens indicated that the mice had no other detectable infections (table S3).

The immune cell composition of PBMCs was examined by unsupervised clustering of individual mice (Fig S1) and visualized by principal component analyses (PCA). The PCA suggested strong effects of environment on variation on the PC1 axis (Fig. 1B. and Fig. 1C) and of genetics on variation along the PC2 axis (Fig. 1B. and Fig. 1C), while infection displayed minimal effect (Fig. S1). Indeed, when we used multivariate distance matrix regression analysis (MDMR) *(28)* to quantify effects of Genotype (i.e., strain), Environment (i.e., rewilded vs control) and Infection and their interactions as predictor variables (Fig. 1C), we found that Genotype and Environment had a significant effect on variation in immune composition, not only as independent variables but also through interactions between Genotype and the Environment (Gen*Env) explaining significant variance in immune composition (Fig. 1D). For example, Genotype-associated variance on the PC2 axis between mouse strains was greater in the laboratory setting than outdoors (Fig. 1B and 1C), with substantial differences in CD44 expression, a marker for T cell activation and memory, on CD4+ T cells among all the strains for lab-housed mice. Although, rewilding produced a larger increase in expression of CD44 on CD4+ T cells in C57BL/6 than in the other strains, CD44 expression in C57BL/6 and PWK/PhJ mice exhibited similar levels of expression when rewilded but different levels when lab housed. Rewilded C57BL/6 mice also had more CD4+Tbet+ cells after infection, indicating a stronger T_H1 response to *T. muris* in this mouse strain (Fig. 1F). Overall, these results. together with those from the CBC analysis (Fig S1), suggest that cellular composition in the blood is driven by interactions of Genotype with Environment and Infection, with genetic differences among inbred strains being reduced by the rewilded environment.

In contrast with the blood, PCA analysis of immune composition in the draining mesenteric lymph nodes (MLNs) with the lymphoid panel (Fig. 2A and Fig. S2) and the myeloid panel (Fig. S2), showed prominent effects of Genotype on variation along the PC1 axes (Fig. 2A and Fig. 2B) with effects of Environment and *T. muris* Infection on the PC2 (Fig. 2A and Fig. 2B) and PC3 axes respectively (Fig. S2). As seen in the PCA plot, MDMR analysis showed a

significant effect of Genotype, Environment and Exposure to *T. muris* (Infection) in determining immune variation (Fig. 2C). In addition, critically, an interactive effect of Genotype, Environment and Infection (Gen*Env*Inf) also contributed to the variation in immune composition in the MLN (Fig. 2C). For example, *T. muris* infection had a significant effect on cellular composition of the draining MLNs (Fig. 2D) with increased proportion and abundance of B cells, especially in the 129S1 and the C57BL/6 strains, and especially outdoors (Fig. S2). Additionally, we observed that the morphology of the MLNs was quite different among mouse strains after rewilding, and this is reflected in the total cellular counts from the MLNs (Fig. 2D). The PWK/PhJ mice had smaller lymph nodes overall that were not expanded in size compared to C57BL/6 and 129S1 mice after rewilding and *T. muris* infection, illustrating a Gen*Env interaction that could be statistically quantified (Fig. 2E). Interestingly, we also found that expression of CD44 on CD4 T cells was influenced by Genotype (Fig. 2F and Fig. 2G) with highest expression of CD44 on the PWK/PhJ mice while expression of CD44 on B cells, which usually depicts antigen experienced B cells (*29*) was predominantly influenced by Environment and exposure to *T. muris* parasites (Fig. 2F and G). For example, PWK/PhJ mice expressed higher levels of CD44 on CD4+ T cells across environments (Fig. 2F), while rewilded mice of all genotypes had a higher percentage of CD44 expressing B cells than their counterparts in the vivarium (Fig. 2H, 2I and Fig. S2).

Based on our previous study *(20)*, we hypothesized that Genotype would have a larger effect on cytokine responses than on immune cell composition. For plasma cytokine levels, we found no statistically significant interactions among Genotype, Environment, and Infection; instead, circulating IFN-γ levels were especially high in infected C57BL/6 mice in both lab and rewilded (Fig. S3A and S3B) and the main effect of Genotype indeed explained more variance than Environment (Fig. S3C) despite the strong effects of the different experimental replicates. Measurements of cytokine responses to microbial stimulation for MLN cells *in vitro* showed a similar pattern, with Genotype having the biggest effect size on variation (Fig. S3). However, our analysis also showed that the effect of Genotype on cytokine responses following stimulation of MLN cells with microbial antigens can be modulated by Environment and Infection (Gen*Env and Gen* Inf interactions). These results confirm our previous observations that genetics influence cytokine responses more strongly than environmental influences (*20)*. Here, however, we add evidence that the environment neither amplified nor eroded genetic effects on cytokine levels in the plasma but modulated responses during stimulation of MLN cells with microbial antigens.

Single cell RNA sequencing (scRNA-seq) can estimate both immune cell composition and function. Mesenteric lymph node cells (n=49,727) from individual mice (n=122) identified 23 major immune cell subsets visualized by UMAP (Fig. 3A). PCA analysis of the different immune cellular subsets from the single cell sequencing analysis (Fig. S4) reveals contributions of Genotype and Environment to the variation among individual mice along the PC1 and PC2 axes (Fig. 3B). Similarly, and in accordance with the flow cytometric data, MDMR analysis of the scRNAseq compositional dataset also showed a significant effect of Genotype, Environment and Exposure to *T. muris* (Infection) in determining immune variation as fixed predictor variables in addition to a huge block effect (Fig. 3C). Furthermore, an interactive effect of Genotype and Environment (Gen*Env) also explained the variation in immune composition as assessed by scRNAseq (Fig. 3C). For example, the increase in Follicular B cells with rewilding was especially heightened in C57BL/6 mice, while decreases in CD4 T cell abundance from

rewilded naïve mice occurred for both 129S1 mice and C57BL/6 mice, but not PWK/PhJ mice. (Fig. 3D and 3E).

To investigate function in terms of cytokine production on a per cell basis, expression of genes (n=148) associated with cytokine activity (see Methods) was used to subset and re-cluster the cells, and they were visualized based on expression of cytokine genes and their original cellular identity (Fig. 3F). These analyses reveals that the more abundant CD4 T cells and Follicular B cells had the smallest percentage of cells that were expressing cytokine related genes (Fig. 3G), whereas less abundant CD8 effector cells, plasmablasts, and dark zone germinal center B cells had higher proportions of cells with cytokine activity. MDMR analysis of variation in cytokine profile shows that variation was explained by Genotype as suggested by our previous work using mutant mice on the C57BL/6 background *(20)* while other predictor variables like Environment, Exposure to *T. muris* (Infection) and interactions had no effect on cytokine gene expression profile variation as assessed by scRNA-seq (Fig. 3H). For example, strain specific differences are noticeable in the percentage of MLN cells with cytokine activity with the 129S1 and PWK/PhJ mice having more cytokine activity than the C57BL/6 mice (Fig. 3I). PCA visualization of cellular composition based on cells with cytokine gene expression alone also showed distinct Genotype differences along the PC1 axis (Fig S4). Hence, scRNA-seq analysis also supports the finding that cytokine response heterogeneity is driven primarily by Genotype, whereas cellular composition is more driven by interactions between Genotype and the Environment.

Finally, we investigated predictors of worm burden (Fig. 4). Despite all 74 *T. muris*-exposed mice receiving approximately the same infectious dose (200 eggs), worm burden was negative-binomially distributed among hosts, consistent with observations in humans (*30*) and wild animals (*31*) (Fig 4A). We therefore analyzed worm burdens using Generalized Linear Models with a negative binomial error distribution. We found a significant Gen*Env for worm burden (Fig 4B, $p = 0.04015$), whereby C57BL/6 mice harbored more worms than the other genotypes in the vivarium, but rewilding was associated with higher worm burdens in all genotypes. In other words, the relative susceptibility of the different host strains to *T. muris* depended upon environment (paralleling *(32))*. When we used logistic regression to analyze worm presence/absence at the experimental endpoint (reported as prevalence of infection among exposed mice in Fig. 4B), significant effects in the best model included only main effects of Genotype ($p=0.0001221$) and Environment ($p=0.0044835$), plus a significant effect of replicate experiment ($p=0.0329262$).

Interestingly, when we included PC1 and PC2 values from the MLN scRNA-seq analysis (Fig. 3B) as summary measures of immune variation among individual mice, significant effects in the best model of worm burden (Fig. 4C) included only main effects of Genotype ($p=0.0003322$), Environment ($p=0.0015615$), and PC2 scores ($p=0.0108213$), which had a significant negative association with worm burden. Loading factors on the PC2 axis (Fig. 4D) indicated that the dearth of T cells with an interferon signature (T.IFN) may be a driver of the relationship between high PC2 scores and decreased worm burden. Furthermore, the fact that PC2 statistically outcompeted Gen*Env suggests that environment-dependent differences among genotypes in worm burdens may hinge on immune factors captured on PC2 (Fig4D). These results are consistent with T_H1 responses being associated with increased susceptibility to helminth colonization *(33, 34)* and suggest that despite complexities in how immune phenotype is influenced by genetics and environment, once that immune phenotype emerges, established

"rules" of infection susceptibility apply (as in *(21)*). The results also demonstrate that genetic, environmental, and individual immune variation is associated with varied infection burden.

Our results support the hypothesis that the effect of even an extreme environmental shift on immune traits is modulated by genetics, and such modulations of phenotype through interactions between environment and genotype are an important source of variation in immune phenotypes. While we previously proposed that the immune cell composition for an individual is shaped by the environment *(20)*, we find here that immune composition in the peripheral blood is shaped by interactions between genetics and the environment and those in the mesenteric lymph node is shaped by interactions between genetics, environment and ongoing infection. We further find that phenotypic variation is influenced by social associations among pairs of rewilded mice *(22)*. This complexity has important ramifications for the course of natural selection on the immune system and diversity in immune genotypes. For example, because any given genotype may produce different immune responses in different ecosystems, environment can alter the ability of individuals to resist and tolerate infections; furthermore natural selection operating on such variation is likely to generate divergent alleles and allele frequencies in different environments *(13)*.

Quantification of such interactions is rare in immunological studies and is a valuable step forward in understanding the evolution and function of the immune system. We do also identify traits for which main (i.e., non-interaction) effects are dominant. For instance, heterogeneity in cytokine responses shows a stronger influence of genetics, consistent with human studies *(35)*. Furthermore, the Human Functional Genomics Project produced similar results to ours, that variation in T cell phenotypes is relatively more influenced by genetics, while B cell phenotypes are relatively more influenced by non-heritable environmental factors *(1)*. It is thus likely that genetics and environment likewise contribute to human immune phenotype in synergistic ways. However, it is also possible that differences in interpreting genetic vs environmental contribution to human immune variation arise because of a focus on different immunological readouts *(2, 3)*.

The reduction in genetically driven differences in laboratory mice under rewilding conditions may be relevant to the different prevalence of inflammatory conditions across geographical locations – e.g., immune phenotypes may be more extreme in the absence of intensive microbial exposures and therefore have a greater impact in genetically susceptible individuals. We also found here that the contributions of Genotype, Environment and helminth exposure could vary depending on tissue site analyzed, and since most human studies utilize peripheral blood, the effects of the infection status may appear less pronounced than if tissue samples were analyzed. Although, the results presented here using inbred mice, with homozygous alleles instead of the predominantly heterozygous state in humans, may present an unrealistic picture of genetic influences on immune variation. Nonetheless, we find surprising consistency between our data and data on human immune variation *(1, 35)*. Our results highlight how rewilding mice with controlled genetic backgrounds could be a bridge towards understanding immune variation between human individuals, while quantification of the interactions at this interface may help elucidate the evolution of the immune system.

References

1. Aguirre-Gamboa R, Joosten I, Urbano PCM, van der Molen RG, van Rijssen E, van Cranenbroek B, Oosting M, Smeekens S, Jaeger M, Zorro M, Withoff S, van Herwaarden AE, Sweep FCGJ, Netea RT, Swertz MA, Franke L, Xavier RJ, Joosten LAB, Netea MG, Wijmenga C, Kumar V, Li Y, Koenen HJPM. 2016. Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits. *Cell reports* 17: 2474-87
2. Ter Horst R, Jaeger M, Smeekens SP, Oosting M, Swertz MA, Li Y, Kumar V, Diavatopoulos DA, Jansen AFM, Lemmers H, Toenhake-Dijkstra H, van Herwaarden AE, Janssen M, van der Molen RG, Joosten I, Sweep FCGJ, Smit JW, Netea-Maier RT, Koenders MMJF, Xavier RJ, van der Meer JWM, Dinarello CA, Pavelka N, Wijmenga C, Notebaart RA, Joosten LAB, Netea MG. 2016. Host and Environmental Factors Influencing Individual Human Cytokine Responses. *Cell* 167: 1111-24.e13
3. Gause WC, Rothlin C, Loke Pn. 2020. Heterogeneity in the initiation, development and function of type 2 immunity. *Nature Reviews Immunology* 20: 603-14
4. Oyesola OO, Souza COS, Loke Pn. 2022. The Influence of Genetic and Environmental Factors and Their Interactions on Immune Response to Helminth Infections. *Frontiers in Immunology* 13
5. Poon MML, Byington E, Meng W, Kubota M, Matsumoto R, Grifoni A, Weiskopf D, Dogra P, Lam N, Szabo PA, Ural BB, Wells SB, Rosenfeld AM, Brusko MA, Brusko TM, Connors TJ, Sette A, Sims PA, Luning Prak ET, Shen Y, Farber DL. 2021. Heterogeneity of human anti-viral immunity shaped by virus, tissue, age, and sex. *Cell Rep* 37: 110071
6. Mangino M, Roederer M, Beddall MH, Nestle FO, Spector TD. 2017. Innate and adaptive immune traits are differentially affected by genetic and environmental factors. *Nature Communications* 8: 13850
7. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel Cesar JL, Furman D, Shen-Orr S, Dekker Cornelia L, Swan Gary E, Butte Atul J, Maecker Holden T, Davis Mark M. 2015. Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. *Cell* 160: 37-47
8. Abolins S, Lazarou L, Weldon L, Hughes L, King EC, Drescher P, Pocock MJO, Hafalla JCR, Riley EM, Viney M. 2018. The ecology of immune state in a wild mammal, Mus musculus domesticus. *PLOS Biology* 16: e2003538
9. Beura LK, Hamilton SE, Bi K, Schenkel JM, Odumade OA, Casey KA, Thompson EA, Fraser KA, Rosato PC, Filali-Mouhim A, Sekaly RP, Jenkins MK, Vezys V, Haining WN, Jameson SC, Masopust D. 2016. Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature* 532: 512-6
10. Mathew D, Giles JR, Baxter AE, Oldridge DA, Greenplate AR, Wu JE, Alanio C, Kuri-Cervantes L, Pampena MB, D'Andrea K, Manne S, Chen Z, Huang YJ, Reilly JP, Weisman AR, Ittner CAG, Kuthuru O, Dougherty J, Nzingha K, Han N, Kim J, Pattekar A, Goodwin EC, Anderson EM, Weirick ME, Gouma S, Arevalo CP, Bolton MJ, Chen F, Lacey SF, Ramage H, Cherry S, Hensley SE, Apostolidis SA, Huang AC, Vella LA, Betts MR, Meyer NJ, Wherry EJ. 2020. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* 369
11. Kuri-Cervantes L, Pampena MB, Meng W, Rosenfeld AM, Ittner CAG, Weisman AR, Agyekum RS, Mathew D, Baxter AE, Vella LA, Kuthuru O, Apostolidis SA, Bershaw L,

Dougherty J, Greenplate AR, Pattekar A, Kim J, Han N, Gouma S, Weirick ME, Arevalo CP, Bolton MJ, Goodwin EC, Anderson EM, Hensley SE, Jones TK, Mangalmurti NS, Luning Prak ET, Wherry EJ, Meyer NJ, Betts MR. 2020. Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci Immunol* 5

12. Ingelfinger F, Gerdes LA, Kavaka V, Krishnarajah S, Friebel E, Galli E, Zwicky P, Furrer R, Peukert C, Dutertre C-A, Eglseer KM, Ginhoux F, Flierl-Hecht A, Kümpfel T, De Feo D, Schreiner B, Mundt S, Kerschensteiner M, Hohlfeld R, Beltrán E, Becher B. 2022. Twin study reveals non-heritable immune perturbations in multiple sclerosis. *Nature* 603: 152-8
13. Lazzaro BP, Little TJ. 2009. Immunity in a variable world. *Philos Trans R Soc Lond B Biol Sci* 364: 15-26
14. Vieira C, Pasyukova EG, Zeng Z-B, Hackett JB, Lyman RF, Mackay TFC. 2000. Genotype-Environment Interaction for Quantitative Trait Loci Affecting Life Span in Drosophila melanogaster. *Genetics* 154: 213-27
15. Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* 10: 565-77
16. Hunter DJ. 2005. Gene–environment interactions in human diseases. *Nature Reviews Genetics* 6: 287-98
17. Mitchell SE, Rogers ES, Little TJ, Read AF. 2005. HOST-PARASITE AND GENOTYPE-BY-ENVIRONMENT INTERACTIONS: TEMPERATURE MODIFIES POTENTIAL FOR SELECTION BY A STERILIZING PATHOGEN. *Evolution* 59: 70-80
18. Graham AL. 2021. Naturalizing mouse models for immunology. *Nature Immunology* 22: 111-7
19. Hamilton SE, Badovinac VP, Beura LK, Pierson M, Jameson SC, Masopust D, Griffith TS. 2020. New Insights into the Immune System Using Dirty Mice. *The Journal of Immunology* 205: 3-11
20. Lin J-D, Devlin JC, Yeung F, McCauley C, Leung JM, Chen Y-H, Cronkite A, Hansen C, Drake-Dunn C, Ruggles KV, Cadwell K, Graham AL, Loke Pn. 2020. Rewilding Nod2 and Atg16l1 Mutant Mice Uncovers Genetic and Environmental Contributions to Microbial Responses and Immune Cell Composition. *Cell Host & Microbe* 27: 830-40.e4
21. Leung JM, Budischak SA, Chung The H, Hansen C, Bowcutt R, Neill R, Shellman M, Loke Png, Graham AL. 2018. Rapid environmental effects on gut nematode susceptibility in rewilded mice. *PLOS Biology* 16: e2004108
22. A. E. Downie OO, R. S. Barre, Q. Caudron, Y.-H. Chen, E. J. Dennis, R. Garnier, K. Kiwanuka, A. Menezes, D. J. Navarrete, O. Mondragon Palomino, J. B. Saunders, C. K. Tokita, K. Zaldana, K. Cadwell, P. Loke, & A. L. Graham. 2023. Social association predicts immunological similarity in rewilded mice.
23. Yeung F, Chen Y-H, Lin J-D, Leung JM, McCauley C, Devlin JC, Hansen C, Cronkite A, Stephens Z, Drake-Dunn C, Fulmer Y, Shopsin B, Ruggles KV, Round JL, Loke Pn, Graham AL, Cadwell K. 2020. Altered Immunity of Laboratory Mice in the Natural Environment Is Associated with Fungal Colonization. *Cell Host & Microbe* 27: 809-22.e6
24. Lai PS, Massoud AH, Xia M, Petty CR, Cunningham A, Chatila TA, Phipatanakul W. 2018. Gene-environment interaction between an IL4R variant and school

endotoxin exposure contributes to asthma symptoms in inner-city children. *Journal of Allergy and Clinical Immunology* 141: 794-6.e3
25. Lusis AJ, Seldin MM, Allayee H, Bennett BJ, Civelek M, Davis RC, Eskin E, Farber CR, Hui S, Mehrabian M, Norheim F, Pan C, Parks B, Rau CD, Smith DJ, Vallim T, Wang Y, Wang J. 2016. The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. *Journal of Lipid Research* 57: 925-42
26. Link VM, Duttke SH, Chun HB, Holtman IR, Westin E, Hoeksema MA, Abe Y, Skola D, Romanoski CE, Tao J, Fonseca GJ, Troutman TD, Spann NJ, Strid T, Sakai M, Yu M, Hu R, Fang R, Metzler D, Ren B, Glass CK. 2018. Analysis of Genetically Diverse Macrophages Reveals Local and Domain-wide Mechanisms that Control Transcription Factor Binding and Function. *Cell* 173: 1796-809.e17
27. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. 2015. A global reference for human genetic variation. *Nature* 526: 68-74
28. Zapala MA, Schork NJ. 2012. Statistical properties of multivariate distance matrix regression for high-dimensional data analysis. *Frontiers in genetics* 3: 190-
29. Camp RL, Kraus TA, Birkeland ML, Puré E. 1991. High levels of CD44 expression distinguish virgin from antigen-primed B cells. *Journal of Experimental Medicine* 173: 763-6
30. Maizels RM, Bundy DAP, Selkirk ME, Smith DF, Anderson RM. 1993. Immunological modulation and evasion by helminth parasites in human populations. *Nature* 365: 797-805
31. Babayan SA, Liu W, Hamilton G, Kilbride E, Rynkiewicz EC, Clerc M, Pedersen AB. 2018. The Immune and Non-Immune Pathways That Drive Chronic Gastrointestinal Helminth Burdens in the Wild. *Frontiers in Immunology* 9
32. Scott ME. 1991. Heligmosomoides polygyrus (Nematoda): susceptible and resistant strains of mice are indistinguishable following natural infection. *Parasitology* 103 Pt 3: 429-38
33. Klementowicz JE, Travis MA, Grencis RK. 2012. Trichuris muris: a model of gastrointestinal parasite infection. *Seminars in Immunopathology* 34: 815-28
34. Else KJ, Grencis RK. 1991. Cellular immune responses to the murine nematode parasite Trichuris muris. I. Differential cytokine production during acute or chronic infection. *Immunology* 72: 508-13
35. Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KTT, Deelen P, Ricaño-Ponce I, Schoffelen T, Jansen AFM, Swertz MA, Withoff S, van de Vosse E, van Deuren M, van de Veerdonk F, Zhernakova A, van der Meer JWM, Xavier RJ, Franke L, Joosten LAB, Wijmenga C, Kumar V, Netea MG. 2016. A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell* 167: 1099-110.e14
36. Bancroft AJ, McKenzie AN, Grencis RK. 1998. A critical role for IL-13 in resistance to intestinal nematode infection. *J Immunol* 160: 3453-61
37. Dixon H, Little MC, Else KJ. 2010. Characterisation of the protective immune response following subcutaneous vaccination of susceptible mice against Trichuris muris. *International Journal for Parasitology* 40: 683-93

38. Meskas J, Wang S, Brinkman R. 2020. flowCut — An R package for precise and accurate automated removal of outlier events and flagging of files based on time versus fluorescence analysis. *bioRxiv*: 2020.04.23.058545
39. Park LM, Lannigan J, Jaimes MC. 2020. OMIP-069: Forty-Color Full Spectrum Flow Cytometry Panel for Deep Immunophenotyping of Major Cell Subsets in Human Peripheral Blood. *Cytometry Part A* 97: 1044-51
40. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, Butte AJ, Bhattacharya M. 2019. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology* 20: 163-72

Figure and Figure Captions

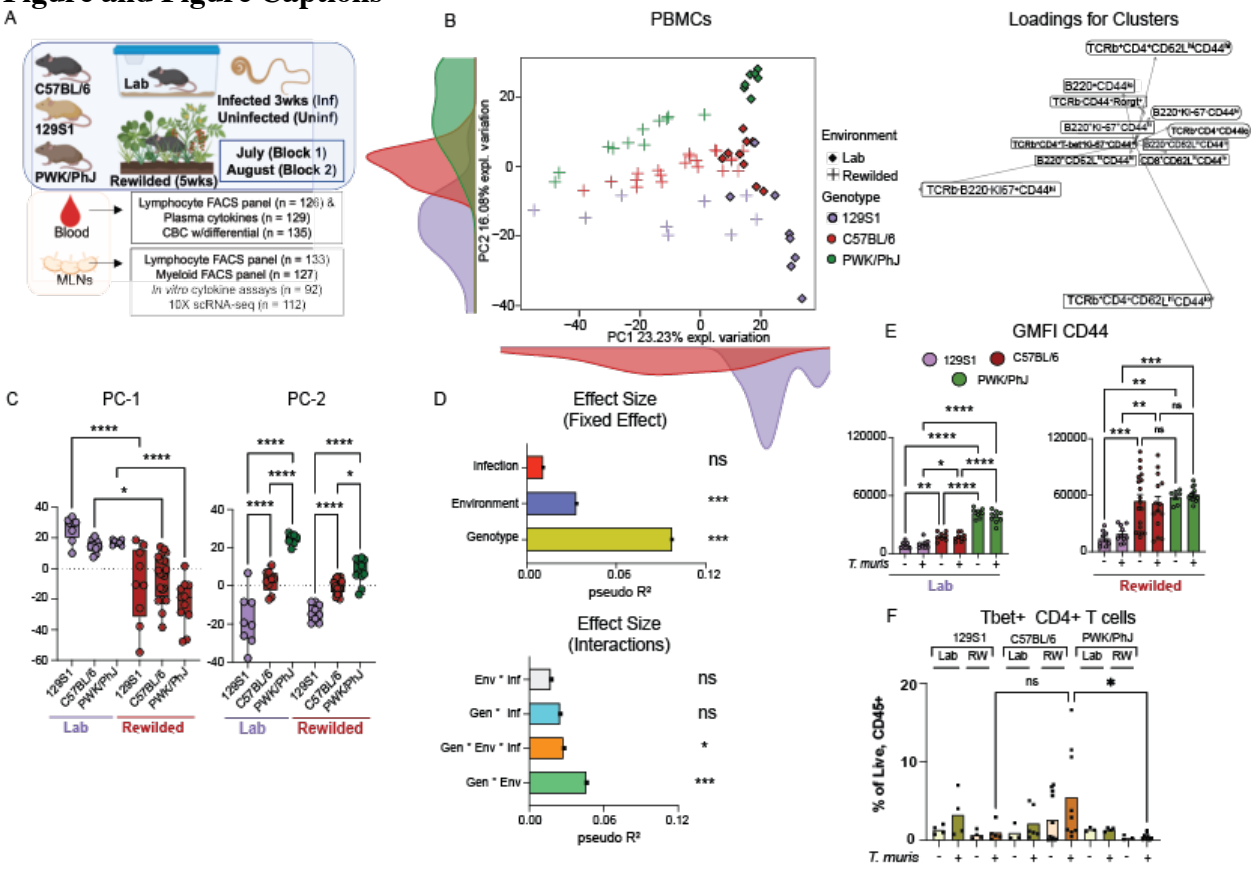

Fig. 1: Genotype and Environment interactions contribute to variation in immune cell composition in mouse peripheral blood mononuclear cells (PBMCs). (**A**) 51 C57BL/6, 41 129S1, 43 PWK/PhJ different mice strains (total=135) were used in these experiments. Some were kept in conventional vivarium (n = 63), and some were housed in the wild enclosure (Rewilded), (n = 72) for total of 5 weeks. In addition, some were exposed to 200 *T. muris* L3 eggs (n =61) while the others were left unexposed (n = 74). Experiments was repeated twice in July, Block 1 (n = 61) and August, Block 2 (n = 74). Blood was collected for Flow cytometry, Plasma cytokine assessment and CBC analysis. MLN cells were also collected for Single cell RNA sequencing (ScRNA-seq), Flow cytometry with two different panels - lymphocyte and myeloid panel as well as cytokine profiling of supernatants from MLN stimulated cells. Samples that fail quality control are not included in downstream analyses. (**B**) Principal Component Analysis (PCA) of cellular profiles for individual mice based on immune cell clusters identified by unsupervised clustering in the blood and the loading factors of key population along the PCA axis, shown for Block 2 experiment and (**C**) Bar plots showing variance on PC1 and PC2 axis of PCA plots in (B) for individual mice. (**D**) Bar plots showing the pseudo R^2 measure of effect size of predictor variables and interactions as calculated by multivariate distance matrix regression analysis (MDMR). (**E**) Bar plot for Geometric Mean Fluorescence Intensity (GMFI) of CD44 expression on blood CD4+ T cells. (**F**) % of Tbet+ CD4 T cells from Live, CD45+ T cells in PBMCs. Statistical significance was determined by One-way ANOVA test between groups. Data are displayed as mean ± SEM. *ns p>0.05; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001.*

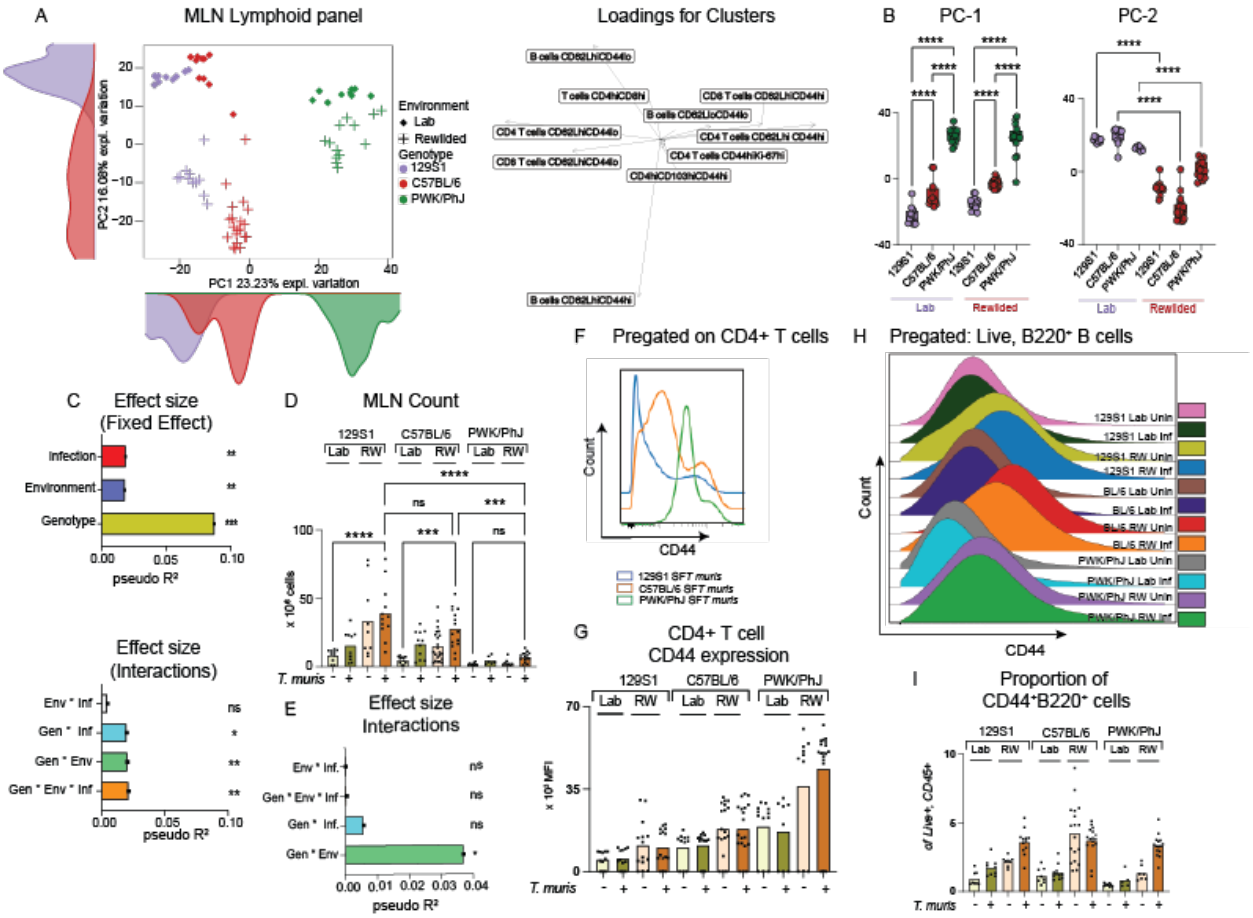

Fig. 2: Interactions between Genotype, Environment and Infection contribute to variation in immune composition in mouse mesenteric lymph node (MLN) cells. (**A**) PCA of profiles for individual mice from immune cell clusters identified by unsupervised clustering in the MLNs with the lymphoid panel and the loading factors of the main population along the PCA. (**B**) Bar plots showing variance on PC1 and PC2 axis of PCA plots in (A) for individual mice. (**C**) Bar plots showing the pseudo R^2 measure of effect size of predictor variables and interactions as calculated by MDMR. (**D**) MLN cell count of individual mice from each group and (**E**) pseudo R^2 measure of effect size of predictor variables and interactions as calculated by MDMR analysis of MLN cell counts. (**F**) Representative Histograms for CD44 expression on CD4+ T cells from *T. muris* infected and rewilded mice of each genotype with corresponding (**G**) Bar plots depicting GMFI of CD44 on MLN CD4+ T cells for individual mice. n = 5-15 mice per group, Block = 2. (**H**). Histograms of concatenated files for CD44 expression on B cells from different groups of mice in Block 2 with its corresponding (**I**) Bar plots depicting proportion of B cells expressing CD44 on MLN cells. n = 6-18 mice per group, Block = 2. Data are displayed as mean ± SEM. *ns p>0.05; * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001.*

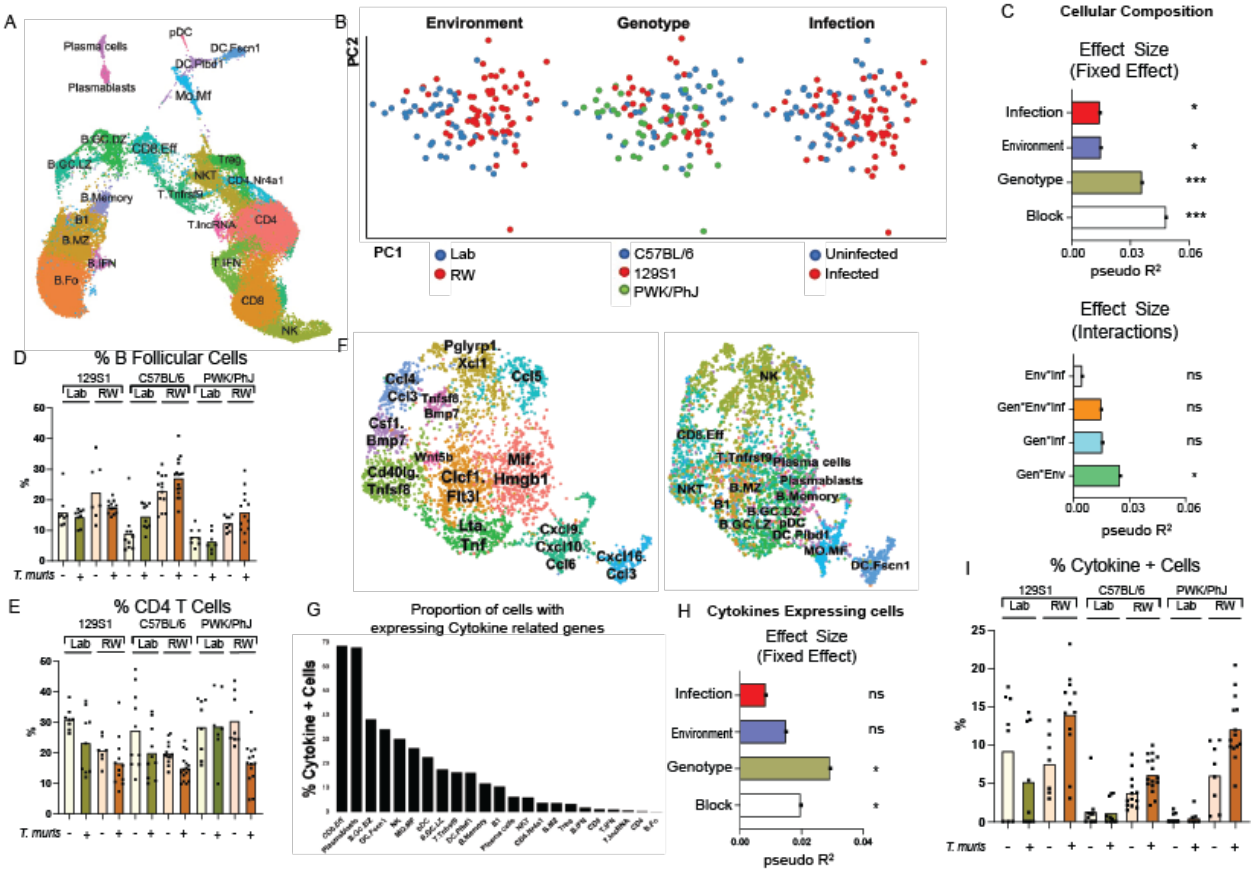

Fig 3: Single Cell RNA Sequencing Analysis of mesenteric lymph node (MLN) cells for assessing variation in cellular composition and cytokine profiles. (**A**) UMAP visualization of scRNASeq data identifying 23 major immune cell subsets in the MLN cells. (**B**) PCA of MLN cellular compositional (of the 23 subsets) data for individual mice as determined by scRNAseq analysis (**C**) Bar plot showing the pseudo R^2 measure of effect size of predictor variables and interactions as calculated by MDMR based on proportion of cells identified in (A). Bar plots showing percentage of B Follicular cells (**D**) and CD4 T cells (**E**) based on the Single cell sequencing analysis in (A) for individual mice. (**F**) Cells that express genes indicating cytokine activity were analyzed as a subset and re-clustered based on cytokine activity (left); or color coded by cell type (right). (**G**) Proportion of each cell type expressing cytokine genes of those identified in (F). (**H**) Bar plots showing the pseudo R^2 measure of effect size of predictor variables and interactions calculated by MDMR based on data from proportion of cells identified in (F). (**I**) Bar plots showing the % of cells expressing cytokine genes identified from individual mice in different groups from (H).

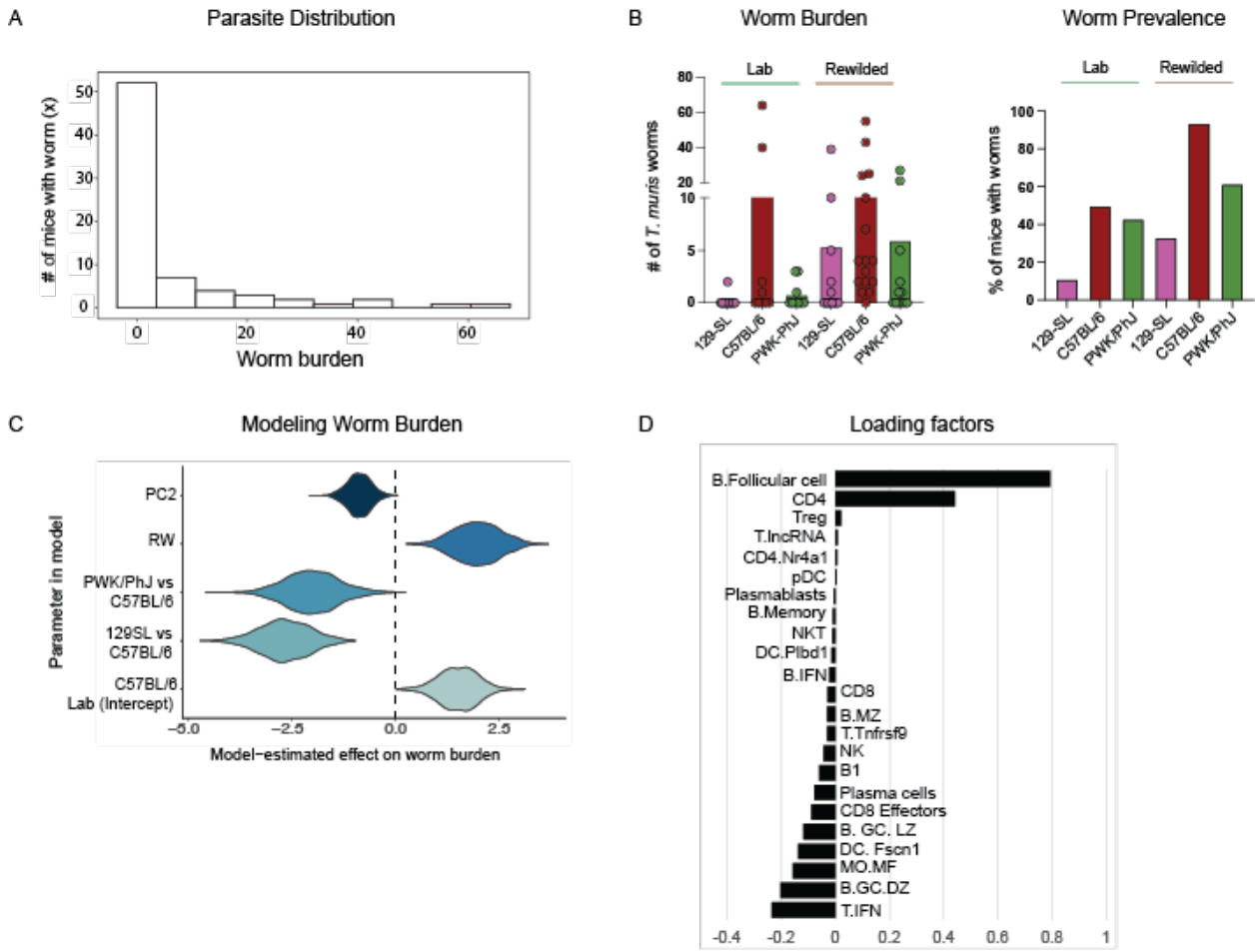

Fig 4: Variation in *Trichuris muris* worm burden is determined by Genetic, Environmental and immunological factors. (**A**) Histograms showing number of mice with different worm burdens, determined by endpoint cecal *Trichuris muris* worm count, followed a negative binomial distribution. (**B**) Bar plots showing worm burden as either (left) number of worms per mouse or (right) percentage of mice (Prevalence) infected with worms. (**C**) Model-estimated main effects on worm burden of host strain (C57BL/6 vs 129SL vs PWK/PhJ), environment (Lab vs RW), and PC2 of the scRNAseq data. The figure depicts 1000 simulated distributions around the mean, given the standard error, for each model-estimated effect on worm burden. (**D**) Loading Factors of for PC2 of the scRNAseq dataset.

Acknowledgements: We thank William Craigens, Christina Hansen and Felix Rozenberg for invaluable assistance in the field and at Stony Ford. We thank Mingming Zhao and Johnson Randall for help with setting up experiments in the lab and for help in maintaining the Joes' Flow App Software respectively. We thank Dr. Elia Tait Wojno and Dr. Lauren Webb for providing *Trichuris muris* parasite eggs. **Funding:** This research was supported by in part by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, NIH and K.C is supported through the NIH grant AI130945. A.E.D. acknowledges funding support from the National Science Foundation (Award # DGE-2039656). R.S.B. and A.L.G. acknowledge funding support from NJ ACTS (New Jersey Alliance for Clinical and Translational Science), which is supported in part by the New Jersey Health Foundation, Inc., and in part by a Clinical and Translational Science Award from the National Centre for Advancing Translational Science of the National Institutes of Health, under award number UL1TR003017. YC acknowledges funding from the Bernard Levine Postdoctoral Research Fellowship and the Charles H. Revson Senior Fellowship. SBK is supported in part by funding from NCI (CA269617). The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health. **Author contributions:** Conceptualization: OO, AED, KC, ALG, PL. Methodology: OO, AED, NH, RSB, KK, YC, ALG, SCL, JD, OMP. Investigation: OO, AED, RB, NH, KK, KZ, YC, AM, SCL OMP, COSS, CH, PL. Data curation and analysis: OO, AED, NH, AG, SC. Writing – original draft: OO, AED, NH, ALG, PL. Writing – review and editing: OO, AED, KC, ALG, PL. Visualization: OO, NH, JC, AED, ALG. Supervision: SBK, KC, ALG, PL. Funding acquisition: KC, ALG, PL. **Competing interests:** All authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplemental materials. Additional raw data, code and materials used in the analysis would be made publicly available at the point of publication