Normalizing the environment recapitulates adult human immune traits in laboratory mice

Lalit K. Beura1, Sara E. Hamilton2, Kevin Bi3, Jason M. Schenkel1, Oludare A. Odumade2†, Kerry A. Casey4†, Emily A. Thompson1, Kathryn A. Fraser1, Pamela C. Rosato1, Ali Filali-Mouhim2, Rafick P. Sekaly4, Marc K. Jenkins1, Vaiva Vezys1, W. Nicholas Haining3, Stephen C. Jameson2 & David Masopust1

Our current understanding of immunology was largely defined in laboratory mice, partly because they are inbred and genetically homogeneous, can be genetically manipulated, allow kinetic tissue analyses to be carried out from the onset of disease, and permit the use of tractable disease models. Comparably reductionist experiments are neither technically nor ethically possible in humans. However, there is growing concern that laboratory mice do not reflect relevant aspects of the human immune system, which may account for failures to translate disease treatments from bench to bedside1–8. Laboratory mice live in abnormally hygienic specific pathogen free (SPF) barrier facilities. Here we show that standard laboratory mouse husbandry has profound effects on the immune system and that environmental changes produce mice with immune systems closer to those of adult humans. Laboratory mice—like newborn, but not adult, humans—lack effector-differentiated and mucosally distributed memory T cells. These cell populations were present in free-living barn populations of feral mice and pet store mice with diverse microbial experience, and were induced in laboratory mice after co-housing with pet store mice, suggesting that the environment is involved in the induction of these cells. Altering the living conditions of mice profoundly affected the cellular composition of the innate and adaptive immune systems, resulted in global changes in blood cell gene expression to patterns that more closely reflected the immune signatures of adult humans rather than neonates, altered resistance to infection, and influenced T-cell differentiation in response to a de novo viral infection. These data highlight the effects of environment on the basal immune state and response to infection and suggest that restoring physiological microbial exposure in laboratory mice could provide a relevant tool for modelling immunological events in free-living organisms, including humans.

Given reported species-specific differences in immune responses1–8, we compared the distribution and differentiation of memory CD8+ T cells between mice and humans. CD8+ T cells are crucial for adaptive immune control of intracellular infections and cancer, and their distribution and differentiation relate directly to their function. We assessed nonlymphoid distribution by examining available specimens of normal cervical tissue from premenopausal adult women. We found that CD8+ T cells were integrated within the mucosa, exhibited a tissue-resident memory T-cell (T_{RM}) phenotype, and comprised ~1.5 million cells (Fig. 1a, b and data not shown). These findings are consistent with previous reports that adult (unlike neonatal) human nonlymphoid tissues are abundantly populated with T_{RM} cells10–12. In contrast, CD8+ T cells were almost completely absent from cervical sections from adult inbred laboratory mice (C57BL/6 strain) housed under SPF conditions (Fig. 1a, b). We then compared major CD8+ T-cell lineages in the blood of adult humans and laboratory mice, focusing on species-specific markers that define functionally homologous populations of naive, central memory (T_{CM}), and terminally differentiated effector memory CD8+ T (T_{EM}) or T_{EMRA} cells (Fig. 1c). Memory CD8+ T cells were much scarcer in laboratory mice than in adult humans and were almost entirely comprised of T_{CM} rather than T_{EM} or T_{EMRA} cells. Also unlike humans, laboratory mice lacked CD27lo/granzyme B+ effector differentiated memory CD8+ T cells, which are thought to respond most immediately to infection13,14 (Fig. 1c). Thus, memory CD8+ T cells in laboratory mice were scarcer and strikingly different from those in adult humans and,

---

**Figure 1** Laboratory mice, like neonatal but not adult humans, lack differentiated memory CD8+ T-cell subsets. a, b, CD8+ T-cell density in cervical tissue from adult laboratory mice (n = 5) and humans (n = 3). Representative immunofluorescence staining of frozen sections (scale bars, 50 μm) is shown. Red, CD8+; blue, 4',6-diamidino-2-phenylindole (DAPI)-stained nuclei. c, CD8+ T-cell phenotypes were compared among adult human blood (n = 13), adult laboratory mouse blood (n = 10), and human cord blood (n = 8) in two independent experiments by fluorescence flow cytometry (representative plots shown). Top panels are gated on CD8+ CD4− cells and highlight naive (blue), T_{CM} (green) and T_{EM} or T_{EMRA} (red) cells, as defined by conventional lineage markers in each species. Bottom panels are gated on antigen-experienced subsets (green and red quadrants defined above). d, Enumeration of granzyme B+CD8+ T-cell frequencies in antigen-experienced subsets. Significance was determined using unpaired two-sided Mann–Whitney U-test (b) or Kruskal–Wallis (analysis of variance, ANOVA) test (d). *P < 0.05, **P < 0.01, ****P < 0.0001; error bars indicate mean ± s.e.m.

1Center for Immunology, Department of Microbiology and Immunology, University of Minnesota, Minneapolis, Minnesota 55414, USA. 2Department of Immunology, Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota 55414, USA 3Department of Pediatric Oncology, Dana-Farber Cancer Institute, and Pediatric Hematology and Oncology, Children’s Hospital, Boston, Massachusetts 02115, USA. Present addresses: Department of Pediatrics, University of California San Diego, Rady Children’s Hospital, San Diego, California 92123, USA (O.A.O.); Department of Respiratory, Inflammation and Autoimmunity, MedImmune LLC, Gaithersburg, Maryland 20878, USA (K.A.C.). © 2016 Macmillan Publishers Limited. All rights reserved
in non-lymphoid tissues from laboratory (Lab.) and pet store mice compared by quantitative immunofluorescence microscopy (QIM). Immunofluorescence staining of frozen sections of indicated tissues (c; n = 8 animals per group; scale bars, 50 μm; red, CD8; blue, DAPI (nuclei)). Significance was determined using unpaired two-sided Mann–Whitney U-test. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; error bars indicate mean ± s.e.m.

Figure 2 | CD8 T-cell subsets vary among feral, pet store, and laboratory mice. a. CD8 T-cell subsets in PBMCs from laboratory mice (n = 9), feral mice that were trapped in the wild (n = 10), and mice obtained from a pet store (n = 6); results from two independent experiments. b. Phenotypes of CD44hi/CD62Lhi (naive) and CD44hi (antigen-experienced) CD8 PBMCs from laboratory and pet store mice compared by fluorescence flow cytometry. c, d, CD8 T-cell density differences between mice could be solely due to genetics, or dependent on environment. Inbred mice are genetically homogeneous, which permits refined approaches and reductionist comparisons that are not possible in outbred populations, but they might have mutations that impair immune system development or function. Laboratory mice live in extremely hygienic filtered microisolator housing16. Conditions are so clean that mice with fatal immunodeficiency diseases often thrive in these environments owing to the absence of pathogen exposure. Thus, modern husbandry has evolved to the point in which laboratory mice, the standard model for biomedical research, acquire far less infectious experience than free-living (that is, ‘dirty’) mice or humans17–20. To test whether environmental conditions might affect the differentiation state and distribution of CD8 T cells, we co-housed inbred laboratory mice with pet store mice, which were not raised in ultra-hygienic barrier facilities. Adult pet store mice were introduced into cages containing adult laboratory (C57BL/6 strain) inbred mice. Within four weeks of co-housing, 22% of laboratory mice died, but no further mortality was observed after this point (Fig. 3c). Serological tests for common mouse pathogens revealed exposure to viral, bacterial, and helminth pathogens (but not murine cytomegalovirus; see Extended Data Table 1). Co-housing resulted in a constitutive increase in highly differentiated effector memory cells in laboratory mice that matched the pattern seen in outbred mice and humans, including the accrual of granzyme B+ and CD27lo cells (Fig. 3d). Moreover, nonlymphoid tissues in laboratory mice became populated by CD8 T cells expressing a Treg-cell phenotype (Fig. 3e and Extended Data Fig. 2). Expanding the cellular analysis on environment. Inbred mice are genetically homogeneous, which permits refined approaches and reductionist comparisons that are not possible in outbred populations, but they might have mutations that impair immune system development or function. Laboratory mice live in extremely hygienic filtered microisolator housing16. Conditions are so clean that mice with fatal immunodeficiency diseases often thrive in these environments owing to the absence of pathogen exposure. Thus, modern husbandry has evolved to the point in which laboratory mice, the standard model for biomedical research, acquire far less infectious experience than free-living (that is, ‘dirty’) mice or humans17–20. To test whether environmental conditions might affect the differentiation state and distribution of CD8 T cells, we co-housed inbred laboratory mice with pet store mice, which were not raised in ultra-hygienic barrier facilities. Adult pet store mice were introduced into cages containing adult laboratory (C57BL/6 strain) inbred mice. Within four weeks of co-housing, 22% of laboratory mice died, but no further mortality was observed after this point (Fig. 3c). Serological tests for common mouse pathogens revealed exposure to viral, bacterial, and helminth pathogens (but not murine cytomegalovirus; see Extended Data Table 1). Co-housing resulted in a constitutive increase in highly differentiated effector memory cells in laboratory mice that matched the pattern seen in outbred mice and humans, including the accrual of granzyme B+ and CD27lo cells (Fig. 3d). Moreover, nonlymphoid tissues in laboratory mice became populated by CD8 T cells expressing a Treg-cell phenotype (Fig. 3e and Extended Data Fig. 2). Expanding the cellular analysis on environment. Inbred mice are genetically homogeneous, which permits refined approaches and reductionist comparisons that are not possible in outbred populations, but they might have mutations that impair immune system development or function. Laboratory mice live in extremely hygienic filtered microisolator housing16. Conditions are so clean that mice with fatal immunodeficiency diseases often thrive in these environments owing to the absence of pathogen exposure. Thus, modern husbandry has evolved to the point in which laboratory mice, the standard model for biomedical research, acquire far less infectious experience than free-living (that is, ‘dirty’) mice or humans17–20. To test whether environmental conditions might affect the differentiation state and distribution of CD8 T cells, we co-housed inbred laboratory mice with pet store mice, which were not raised in ultra-hygienic barrier facilities. Adult pet store mice were introduced into cages containing adult laboratory (C57BL/6 strain) inbred mice. Within four weeks of co-housing, 22% of laboratory mice died, but no further mortality was observed after this point (Fig. 3c). Serological tests for common mouse pathogens revealed exposure to viral, bacterial, and helminth pathogens (but not murine cytomegalovirus; see Extended Data Table 1). Co-housing resulted in a constitutive increase in highly differentiated effector memory cells in laboratory mice that matched the pattern seen in outbred mice and humans, including the accrual of granzyme B+ and CD27lo cells (Fig. 3d). Moreover, nonlymphoid tissues in laboratory mice became populated by CD8 T cells expressing a Treg-cell phenotype (Fig. 3e and Extended Data Fig. 2). Expanding the cellular analysis
Co-housing with pet store mice changes the immune system of laboratory mice. a, Proportion of CD44hi (antigen-experienced) CD8+ T cells in laboratory mice housed in SPF conditions (light blue, n = 10 for each time point) or co-housed with pet store mice (green, n = 9 for each time point) or co-housed with pet store mice (green, n = 10 for each time point) from two independent experiments. b, CD8+ T-cell phenotypes in blood from SPF or co-housed laboratory mice. Representative flow plots (n = 15) are shown. c, Survival of SPF and co-housed laboratory mice (n = 26 SPF and 65 co-housed mice). d, Representative phenotypes of CD44hiCD8+ PBMCs isolated from laboratory mice co-housed with pet store mice for 100 days (n = 15) and non-co-housed age-matched controls (n = 10) were compared by fluorescence flow cytometry. e, Enumeration of CD8+ T cells in non-lymphoid tissues between laboratory and co-housed mice by QIM (n = 8 animals per group).

Figure 3 | Co-housing with pet store mice changes the immune system of laboratory mice. a, Proportion of CD44hi (antigen-experienced) CD8+ PBMCs in laboratory mice housed in SPF conditions (blue, n = 10 for each time point) or co-housed with pet store mice (green, n = 10 for each time point) from two independent experiments. b, CD8+ T-cell phenotypes in blood from SPF or co-housed laboratory mice. Representative flow plots (n = 15) are shown. c, Survival of SPF and co-housed laboratory mice (n = 26 SPF and 65 co-housed mice). d, Representative phenotypes of CD44hiCD8+ PBMCs isolated from laboratory mice co-housed with pet store mice for 100 days (n = 15) and non-co-housed age-matched controls (n = 10) were compared by fluorescence flow cytometry. e, Enumeration of CD8+ T cells in non-lymphoid tissues between laboratory and co-housed mice by QIM (n = 8 animals per group).

Beyond CD8+ T cells revealed extensive and profound changes to many innate and adaptive immune cell lineages in diverse tissues from co-housed laboratory mice and increased levels of serum antibodies (Fig. 3f, g and Extended Data Fig. 2). Furthermore, principal component analysis of PBMC gene expression data in the space of all detected genes (~18,000) revealed a spatial shift of co-housed samples away from laboratory samples and towards pet store samples along the first principal component (PC1; Fig. 3h). Together, these data demonstrate that co-housing profoundly altered the status of the immune system.

Our data suggested that the immune systems of laboratory mice have features in common with those of neonatal humans, and that altering the environment reproduced phenotypic immune signatures of adult humans. To test this hypothesis more broadly, we queried expression-profiling data from maternal and neonatal cord PBMCs from an unaffiliated study21 with pet store versus laboratory and co-housed versus laboratory mouse signatures using gene set enrichment analysis (GSEA; Fig. 4a, b). The top 400 genes that were upregulated in pet store mice showed highly significant enrichment in adult human expression data, whereas the top 400 downregulated genes showed enrichment in neonatal humans. Laboratory mice acquired this gene expression program after co-housing.

To more deeply investigate similarities in transcriptional patterns among pet store mice, co-housed mice, and adult humans compared with laboratory mice and neonatal humans, we applied GSEA with the ImmuneSigDB database of immunological signatures22. We then used leading-edge metagene analysis of GSEA results to identify modules of co-regulated genes that were upregulated in human adult versus cord PBMCs, and those found in pairwise comparisons of laboratory, pet store, and co-housed mice. Overlap between the resulting metagenes was used to identify global similarities between each data set. We observed highly significant overlaps between metagenes that were upregulated in adult PBMCs compared with cord PBMCs and metagenes that were upregulated in PBMCs from pet store or co-housed mice compared with laboratory mice. These included numerous pathways related to innate and adaptive immune functions (Fig. 4c, Extended Data Fig. 3 and Supplementary Tables 1 and 2). Conversely, metagenes that were upregulated in human cord blood cells overlapped with those upregulated in laboratory mice. Thus, these functional modules represent a major axis of similarity in immune status between pet store mice and adult humans relative to laboratory mice and neonatal humans, and can be conferred on laboratory mice through co-housing with pet store mice.

We next tested whether mouse husbandry affected immune responses. We challenged mice with the intracellular pathogen *Listeria monocytogenes*, a bacterial infection that is often used to gauge immune function in laboratory mice. Compared to laboratory mice, both pet...
Figure 4 | Microbial experience matures mouse immune transcriptome from neonatal to adult human-like and affects immune system function. a, b, Enrichment of gene signatures among the indicated mouse group comparisons relative to human adult versus neonatal comparison. a, Experimental design. b, GSEA plots. Signatures consist of top 400 genes that were significantly differentially expressed in laboratory (n=8), co-housed (n=7), and pet store (n=8) mice. c, Pairwise overlaps of metagenes identified through leading-edge metagene analysis and corresponding Gene Ontology (GO) terms. d, Bacterial load in the liver 3 days after challenge with 8.5 × 10^4 colony-forming units (c.f.u.) of L. monocytogenes (n=9 for all except laboratory mice, n=8). e, Parasitic load in peripheral blood 5 days after P. berghei ANKA parasitized red blood cell challenge in laboratory (n=15), co-housed (n=19), and pet store (n=15) mice. f, Twenty-eight days after LCMV infection (Armstrong strain), the proportion of H-2D^b/gp33^+ CD8^+ T-cell MPECs (KLRG1^+ , CD127^+ ) and SLECs (KLRG1^+ , CD127^+ ) in PBMC from co-housed (n=8) and laboratory (n=9) mice. Cumulative data from two independent experiments. Data points in d–f represent individual mice. Significance was determined using Kruskal–Wallis (ANOVA) test (d), one-way ANOVA test (e) and unpaired two-sided t-test (f). *P < 0.05, **P < 0.01, ****P < 0.0001; error bars indicate mean ± s.e.m. NS, not significant. FDR, false discovery rate.
do not support an end to SPF studies. However, it is ironic that such an immunologically inexperienced organism has become de rigueur for studies of the immune system, as our data show that this compromises development of a human adult-like immune system. To maximize opportunities to translate novel treatments from preclinical studies to clinical therapies, it may be opportune to add ‘dirty’ mice to our repertoire of investigative tools. Much as the analysis of truly sterile 'germ-free' mice has revealed how influential commensal flora are on 'normal' physiology and immune system function, our study suggests that the mouse system in mice may not be fully ‘normalized’ without more complete microbial exposure16. Indeed, just as many autoimmune diseases do not manifest in genetically predisposed mice in the absence of commensal flora, certain infectious experiences have been shown to induce heterologous and innate immune memory, trigger autoimmune disease, and affect transplantation tolerance23–26. Forward genetic screens to reveal the function of immunological genes are ongoing in mice, and it might be beneficial to conduct these screens in a dirty mouse model.

More generally, dirty mice might be valuable for investigating aspects of the hygiene hypothesis, immune function and treatments for disease in the settings of transplantation, allergy, autoimmunity, and vaccination, and perhaps in disparate diseases that might involve the immune or inflammatory systems (such as cardiovascular disease and cancer)4. Such mice could supplement current models to either increase translational potential to human disease or to better inform the efficacy of preclinical prophylactic and therapeutic modalities, without sacrificing powerful experimental tools and approaches that cannot be used in human studies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 January; accepted 11 March 2016.

Published online 20 April 2016.


Supplementary Information is available in the online version of the paper.

Acknowledgements This study was supported by National Institutes of Health grants 1R01AI111671, R01AI084913 (to D.M.), R01AI116678, R01AI075168 (to S.C.J.) and a BSL-3 suite rental waiver grant from the University of Minnesota. We thank R. Ahmed for providing reagents for pilot studies, P. Southern and D. McKenna for tissue samples or cord blood, and all members of the BSL-3 mouse team (University of Minnesota).


Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.M. (masopust@umn.edu) or S.C.J. (james024@umn.edu).
Mice, co-housing and infections. Pet store mice were purchased from various Twin Cities area pet stores. Feral mice were trapped on a horse farm or rural outdoor petting zoo in Minnesota or Georgia, USA. Male or female pet store mice were introduced into the cages of 6–8-week-old C57BL/6 mice of the same sex purchased from the National Cancer Institute. Co-housing occurred within a BSL-3 facility. Age-matched C57BL/6 laboratory mice maintained in SPF facilities served as controls. The number of animals needed to reach statistical significance was determined on the basis of previous experience. All animals that survived the experimental treatment were included in the final analysis. No method of randomization was used to allocate animals to experimental groups. Investigators were not blinded to the group allocation during experiments. L. monocytogenes was grown in tryptic soy broth containing streptomycin to log phase growth. The indicated groups of mice were infected intravenously (i.v.) with 8.5 × 10^5 c.f.u. of wild-type L. monocytogenes (provided by J. Harty). Bacterial load in the spleen and liver was determined 3 days post-challenge as previously described. L. monocytogenes immune mice were generated by primary infection with recombinant L. monocytogenes expressing OVA (LM-OVA) (provided by H. Shen) 5 months before secondary challenge with wild-type L. monocytogenes. P. bergheri ANKA (provided by S. K. Pierce) was propagated by passage in mice and blood collection. One-million parasitized RBCs were injected intraperitoneally (i.p.) into the indicated mice. Parasitemia was measured by flow cytometry of peripheral blood. All mice were in accordance with the guidelines of the Institutional Animal Care and Use Committees at the University of Minnesota.

Human tissue. Adult PBMC samples were collected from healthy volunteers at the University of Minnesota. Fresh cord blood samples were acquired from the Clinical Cell Therapy Laboratory at the University of Minnesota Medical Center. PBMC isolation has been described in detail elsewhere. After isolation, cells were frozen in 10^6 cells-per-ml aliquots in a cryopreservative solution (Sigma-Aldrich) for future phenotyping. Cervical tissue from premenopausal women was obtained from the Tissue Procurement Facility (BioNet, University of Minnesota). Cervical samples were frozen roughly 1–2 h after surgical resection. Informed consent was obtained from all subjects. The University of Minnesota Institutional Review Board approved all protocols used.

Intravascular staining, leukocyte isolation and phenotyping. An intravascular staining method was used to discriminate between cells present in the vasculature and cells in the tissue parenchyma. Briefly, animals were injected i.v. with biotin/fluorochrome-conjugated anti-CD45 through the tail vein. Three minutes after injection, animals were killed by cervical dislocation, and tissues were collected as described. Isolated mouse cells were surface-stained with antibodies against CD3 (145-2C11), CD4 (30-F11), CD11b (M1/70), CD11c (N418), NKP46 (29A.14), Ly6G (1A8), MHC II (la-e) (M5/114.15.2), CD8α (53-6.7), CD45.2 (104), CD4 (RM4-5), CD62L (MEL-14), CD44 (IM7), CD69 (H1.2F3), CD103 (M290), Ly6C (AL24), CD11c (B11), CD43 (57), CD27 (LG.3A10), PD-1 (RMP1-30 and J43), KLRC1 (2F1), CXCXR3 (CXCXR3-173), CD127 (SB199), e837 (DATK32), F4/80 (Cl A3-1), CXCR5 (2G8), CD38 (907), IgM (RM3-1), IgD (11-26c.2a), GL7 (GL7), CD19 (DSY), and B220 (RA3-6B2). Isolated human cells were surface-stained with antibodies against CD8α (R3B5), CD45RA (HL100), CCR7 (G043H7), CD27 (O323), and CD3 (SK7). All of the above antibodies were purchased from BD Biosciences, Biolegend or Affymetrix eBiosciences. Cell viability was determined using Ghost Dye 780 (Tonbo Biosciences). Intracelluar staining with phycoerythrin (PE)-conjugated granzyme B (Invitrogen), fluorescein isothiocyanate (FITC)-conjugated Ki67 (Invitrogen) and Af488-conjugated anti-mouse IgG (H+L) antibodies was performed using the CyTox/Cytoperm kit (BD Pharmingen) following the manufacturer’s instructions. Intracellular staining for transcription factors was performed using a transcription factor staining buffer set (Affymetrix eBiosciences) with antibodies against FosP3 (FKF-163), T-bet (4B10), Eomes (Dan11mag), Gat3 (L50-823) and Rorγt (Q31-378) following the manufacturer’s guidelines. Single positive staining for T-bet, Gat3, and Rorγt was used to identify Th1, Th2, and Th17 lineages, respectively. FITC-conjugated mouse lineage cocktail (Tonbo Biosciences) was used in combination with other recommended lineage markers to identify various innate lymphoid cell subsets. The stained samples were acquired using LSR1 or LSR Fortessa flow cytometers (BD) and analyzed with FlowJo software (Tree Star, Inc.).

Infectious agent screening. Laboratory mice, co-housed laboratory mice (after at least 30 days of co-housing) and pet store mice were screened using EZ-spot and Infectious agent screening. Mouse serum antibody titres were quantified using Ready-Set-Go! ELISA kits (Affymetrix eBiosciences) following the manufacturer’s instructions.

RNA isolation and microarray hybridization. For each sample, 1–3 × 10^6 PBMCs were used for RNA extraction. Cells were first homogenized using QIAshredder columns (Qiagen) and RNA was then extracted using an RNeasy kit (Qiagen) as per the manufacturer’s instructions. Following quality control, total RNA samples were processed using the Illumina TotalPrep-96 RNA Amplification Kit for High-Throughput RNA Amplification for Array Analysis. Samples were loaded onto the MouseRef-8 v2.0 Expression BeadChip (Illumina) and hybridized beadchips were scanned using the Illumina iScan Beadarray Reader. Basic quality metrics were checked using Illumina Genomestudio.

Bioinformatics analysis. Before analysis, mouse microarray data were quantile normalized using preproccessCore (Bioconductor) and batch correction was performed using the ComBat algorithm. Principal components analysis was performed in R. Raw human adult and neonatal cord PBMC microarray data were obtained from a previous unaffiliated study profiling the peripheral blood of 72 smoking or non-smoking women and the cord blood of their neonates (Gene Expression Omnibus, accession code GSE27727). Human microarray data were obtained from a previous unaffiliated study profiling the peripheral blood of 4,872 gene sets describing a wide range of cell states and experimental perturbations from immunology literature. The top 150 significantly enriched gene sets, as restricted by an FDR < 0.25 and ranked by P < 0.05, were subsetted for their leading edge genes. These genes were then clustered into metagenes using non-negative matrix factorization. The significance of overlap between pairs of metagenes was determined using a Fisher exact test (P < 1 × 10^-5). Metagenes were functionally annotated based on the significance of overlap between member genes and GO terms, as measured by hypergeometric test using the GOrilla enrichment analysis tool.

Statistics. Data were subjected to the D’Agostino and Pearson omnibus normality test to determine whether they were sampled from a Gaussian distribution. If a Gaussian model of sampling was satisfied, parametric tests (unpaired two-tailed Student’s t-test for two groups and one-way ANOVA with Bonferroni multiple comparison test for more than two groups) were used. If the samples deviated from a Gaussian distribution, non-parametric tests (Mann–Whitney U test for two groups, Kruskal–Wallis test for more than two groups) were used unless otherwise stated. Viances between groups were compared using an F test and found to be equal. All statistical analysis was done in GraphPad Prism (GraphPad Software Inc.). P < 0.05 was considered significant.

Extended Data Figure 1 | Frequency of CD8+ T-cell subsets in newborn versus adult humans. CD8+ T-cell subsets were defined in adult PBMCs (n = 13) and cord blood PBMCs (n = 8) by fluorescence flow cytometry based on the following markers: naive, CD45RAhiCCR7hi; TCM, CD45RAloCCR7hi; TEM, CD45RAloCCR7lo; TEMRA, CD45RAhiCCR7lo. Significance was determined using unpaired two-sided t-test. ***P < 0.001, ****P < 0.0001; error bars indicate mean ± s.e.m.
Extended Data Figure 2  | Co-housing laboratory mice with pet store mice induces accumulation of T<sub>RM</sub>-phenotype CD8<sup>+</sup> T cells and other innate cells in tissues of laboratory mice. a, CD8<sup>+</sup> T-cell density within the indicated tissues of adult laboratory mice (n = 5) and co-housed mice (n = 7). Representative immunofluorescence staining, CD8<sub>β</sub> (red), DAPI (nuclei, blue); scale bars, 50 μm. b, Phenotype of CD8<sup>+</sup> T cells was compared between laboratory mice (n = 9) and age-matched laboratory mice that were co-housed (n = 9, representative flow cytometry plots shown). Samples gated on CD44<sup>hi</sup> cells isolated from the indicated tissue (vasculature populations were excluded, see Methods). c, Enumeration of CD11b<sup>+</sup> granulocytes and Ly6C<sup>hi</sup> inflammatory monocytes in spleens of laboratory (n = 6) and co-housed (n = 6) mice. Significance was determined using unpaired two-sided Mann–Whitney U-test. **P < 0.01; error bars indicate mean ± s.e.m.
Extended Data Figure 3 | LEM metagene analysis. For each comparison, standard GSEA was performed using the ImmSigDB database of gene-sets. Genes in the top 150 enriched sets (FDR < 0.001, ranked by P value) were filtered to only leading edge genes and subsequently clustered into groups (metagenes) using an NMF algorithm. Hierarchical clustering of genes within individual metagenes was performed to obtain the final heatmap. Metagenes with qualitatively discernible 'blocks' of gene-set membership were annotated according to the identity of corresponding enriched gene-sets. Heatmaps for adult versus neonatal, pet store versus laboratory, co-housed versus laboratory, neonatal versus adult, laboratory versus pet store, and laboratory versus co-housed comparisons are shown. Individual genes within each metagene are listed in Supplementary Table 1. Pairwise overlaps between metagenes from different comparisons are visualized in Fig. 4c.
Extended Data Figure 4  | Environment altered antimicrobial resistance and CD8+ T-cell differentiation. Laboratory mice were co-housed with pet store mice as described in Figure 3. a, Bacterial load in the spleen 3 days after challenge with 8.5 × 10⁴ c.f.u. of *L. monocytogenes* (LM) in laboratory (*n* = 8), LM-immune (*n* = 9), co-housed (*n* = 9) and pet store mice (*n* = 9) in two independent experiments. b, Survival of laboratory (*n* = 15), co-housed mice (*n* = 19) and pet store mice (*n* = 15) after challenge with 10⁶ *P. berghei* ANKA parasitized RBCs in two independent experiments. c, Laboratory (*n* = 9) and co-housed (*n* = 8) mice were infected with LCMV. Four weeks later, LCMV-specific CD8+ T cells (identified with H-2Db/gp33 MHC I tetramers) were evaluated for expression of the indicated markers. Top row, gated on live CD8+ α+ T cells. Bottom three rows, gated on live CD8α+ H-2Dβ/gp33+ T cells. Significance was determined using Kruskal–Wallis (ANOVA) test (a) and log-rank (Mantel–Cox) test (b). *P < 0.05, ***P < 0.001, ****P < 0.0001; error bars indicate mean ± s.e.m.
Extended Data Table 1  |  Microbial exposure in laboratory, pet store and co-housed mice

<table>
<thead>
<tr>
<th>Viruses</th>
<th>Pet store</th>
<th>Laboratory</th>
<th>Co-housed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotavirus (EDIM)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mouse Hepatitis Virus</td>
<td>93.3</td>
<td>0</td>
<td>61.5</td>
</tr>
<tr>
<td>Murine norovirus</td>
<td>60</td>
<td>0</td>
<td>38.5</td>
</tr>
<tr>
<td>Mouse parvovirus NS1</td>
<td>53.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mouse parvovirus type 1</td>
<td>40</td>
<td>0</td>
<td>7.7</td>
</tr>
<tr>
<td>Mouse parvovirus type 2</td>
<td>46.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minute virus of mice</td>
<td>46.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Theiler's murine encephalomyelitis virus</td>
<td>60</td>
<td>0</td>
<td>38.5</td>
</tr>
<tr>
<td>Sendai virus</td>
<td>66.7</td>
<td>0</td>
<td>23.1</td>
</tr>
<tr>
<td>Ectromelia virus</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lymphocytic Choriomeningitis virus</td>
<td>6.7</td>
<td>0</td>
<td>7.7</td>
</tr>
<tr>
<td>Mouse adenovirus 1 and 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mouse cytomegalovirus</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Polyoma virus</td>
<td>6.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pneumonia virus of mouse</td>
<td>53.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reovirus</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Pet store</th>
<th>Laboratory</th>
<th>Co-housed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cilli-Associated Respiratory Bacillus</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mycoplasma pulmonis</td>
<td>73.3</td>
<td>0</td>
<td>30.8</td>
</tr>
<tr>
<td>Clostridium piliforme</td>
<td>26.7</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parasites/Protozoa/Fungi</th>
<th>Pet store</th>
<th>Laboratory</th>
<th>Co-housed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encephalitozoan cuniculi</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pinworm</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Mites</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Frequency of each indicated microbial exposure within the population was evaluated by serological analysis and/or PCR in laboratory (n = 4), pet store (n = 15) and co-housed (n = 13) mice. Each co-housed sample was collected from a different cage. Pinworm- and mite-specific PCR was performed on pooled pet store and co-housed samples.