Dermal fibroblast cultures recapitulate differences between deermice and mice in responses to a Toll-like receptor agonist

Jonathan V. Duong[1], Aqsa Motiwala[1], William J. Hotz[1], Landen Gozashti[2,3], Anthony D. Long[4], and Alan G. Barbour[1,5,*]

[1] Department of Microbiology and Molecular Genetics, School of Medicine, University of California Irvine, Irvine, CA, United States
[2] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, United States
[3] Museum of Comparative Zoology, Harvard University, Cambridge, MA, United States
[4] Department of Ecology and Evolutionary Biology, School of Biological Sciences, University of California Irvine, Irvine, CA, United States
[5] Department of Medicine, School of Medicine, University of California Irvine, Irvine, CA, United States

*abarbour@uci.edu

Key words: Peromyscus, Mus musculus, innate immunity, endogenous retrovirus, cell-autonomous immunity, interleukin-11, Nrf2, arginine

Abstract
The white-footed deermouse *Peromyscus leucopus* is a primary reservoir for the agents of Lyme disease and other zoonoses in North America and manifests infection tolerance for the bacteria, protozoa, and viruses it hosts. In previous in vivo studies *P. leucopus* and *M. musculus* differed in the degree of sickness and profiles of biomarkers after exposure to bacterial lipopolysaccharide, a TLR4 agonist. As an approach for assessing immunity of mammals in nature and for longitudinal studies of colony animals in the laboratory, we evaluated using bulk and single cell RNA-seq primary dermal fibroblast cultures of *P. leucopus* and *M. musculus* in their short-term responses to a TLR2 agonist lipopeptide. By single cell RNA-seq cultures of both species comprised at least two types of fibroblasts, which were further differentiated in their responses to TLR agonists. With continued passage the mouse cell population lost viability, while the deermouse cell population spontaneously transformed into a cell line stably maintained under standard conditions. Bulk RNA-seq revealed distinctive profiles for deermouse and mouse cells in arginine metabolism gene expression, high baseline transcription of the antioxidant transcription factor Nfe2l2 (Nrf2) in deermouse fibroblasts, and the transcription of the aging-associated cytokine interleukin-11 in agonist-treated mouse fibroblasts but not deermouse fibroblasts. In both species' cultures there was increased transcription of several types of endogenous retrovirus (ERV) and transposable elements (TE) after exposure to the agonist. The transcribed ERV/TE sequences in *M. musculus* cells were generally longer in length and with greater potential for translation than sequences in treated *P. leucopus* cells. The results indicate feasibility of this in vitro model for both laboratory- and field-based studies and that inherent differences between deermice and mice in cell-autonomous innate immune responses and ERV/TE activation can be demonstrated in dermal fibroblasts as well as the animals themselves.

1 Introduction

As Higai et al. noted, the immune system's characteristics of both rapid divergence and high cell-to-cell variability "seem to be at odds with strong regulatory constraints imposed on the host immune response: the need to execute a well-coordinated and carefully balanced program to avoid tissue damage and pathological immune conditions" (1). To resolve this apparent paradox, established animal models, principally the house mouse *Mus musculus,* are commonly enlisted. But what if the question could be more profitably addressed using another species (2), such as one that in its natural environment thrives while persistently infected with microbes of trans-species transmission potential? Ideally, wild populations of the species would be accessible for field-based studies, and there would be vivarium-bred stock colonies available for laboratory-based studies.

Our nominee as an alternative to the house mouse for this question is the North American species *Peromyscus leucopus*, the white-footed deermouse (3, 4). Mouse-like in appearance and size, this and other deermice, such as *P. maniculatus*, belong to the family Cricetidae, along with hamsters, voles, and woodrats, and not the family Muridae, the taxonomic clade of laboratory mice and rats (5). *P. leucopus* is abundant on this continent and can be found in a variety of environments. Closed, heterogenous stock colonies of long-standing exist at different institutions. As the basis for

forward and reverse genetics, as well as for bulk and single cell RNA-seq, there are high quality, chromosome-scale genome assemblies of both *P. leucopus* and *P. maniculatus* with full annotation and millions of single nucleotide polymorphisms and other variants that have been catalogued to date (6, 7).

Another rationale for choosing *P. leucopus* is its public health importance as a reservoir for several agents of human zoonoses. Besides the Lyme disease agent *B. burgdorferi*, *P. leucopus* is also a natural host for the obligate intracellular bacterium of anaplasmosis, the apicomplexan protozoan of babesiosis, and the flavivirus of Powassan viral encephalitis. Deermice may be persistently infected, but these deermice remain largely free of disability or discernible effects on fitness (8). This phenomenon is known as infection tolerance (9, 10), and it is an example of the balance between a sufficiently effective defense and minimization of immune-mediated damage that Higai et al. and others have highlighted (11). Infected with the spirochete *Borreliella burgdorferi*, the principal agent of Lyme disease in North America, *P. leucopus* display few of the pathologic changes or inflammation observed in infected *M. musculus* (8). Another way that *P. leucopus* differs from *M. musculus* is a lifespan that is up to three-fold longer (12, 13). Infection tolerance and greater longevity for their body sizes are characteristics *P. leucopus* has in common with some bats (14, 15).

These distinctions between deermice and mice in their respective responses to infection inspired our comparative studies of the two species, using as an experimental model bacterial lipopolysaccharide (LPS) to elicit acute systemic inflammation (16, 17). *P. leucopus* displayed genome-wide RNA-seq profiles in the blood, spleen, and liver that distinguished it from *M. musculu*s. In general, the LPS-treated deermice were characterized by alternatively-activated macrophage polarization rather than the expected classically-activated, mitigation of the effects of neutrophil activation, and restrained type 1 and type 2 interferon responses. What more specifically distinguished *P. leucopus* from *M. musculus* were an inverted ratio of nitric oxide synthase 2 (Nos2) to arginase 1 (Arg1) transcription, remarkably high transcription of the Slpi gene for secretory leukocyte peptidase inhibitor (Slpi), and moderated interferon expression, either directly measured or by evidence of interferon stimulated genes (ISGs).

For the present study of *P. leucopus* and *M. musculus*, we reduced the subject of the experiment to one that retains characteristics of live animals yet would also serve for studies of wild animals that could be captured, sampled, and then released (18). Blood sampling from trapped then released animals is an established procedure for the study of the immunology or physiology of animals in nature (19). However, assays of whole blood or serum are limited to what a single volume of blood can provide and nothing more. If the spleen is the tissue (20), the capture is a terminal event for the animal. For field and laboratory work on Lyme disease a specimen commonly obtained has been punch biopsies of ear skin tissue, which are then subjected to culture or quantitative PCR for *B. burgdorferi* bacteria (21, 22). It occurred to us that this biopsied tissue could also be propagated ex vivo in the laboratory, thus allowing for a much expanded set of experiments with further rounds of multiplication.

As a proof-of-principle for *Peromyscus*, we cultivated primary dermal fibroblasts obtained from the ears of heterogenous stock *P. leucopus* or outbred *M. musculus* in a laboratory setting. Low-passage cultures of skin cells were then exposed to either a Toll-like receptor 2 (TLR2) agonist or buffer alone. TLR2, as a heterodimer with TLR1, is the pattern recognition receptor (PRR) for pathogen-associated molecular pattern (PAMP) represented by the bacterial lipoproteins of *B. burgdorferi* (23). The resultant specimens were subjected to bulk and single cell RNA-seq, for which the reference sets were genome-wide protein coding sequences of *P. leucopus* and *M. musculus*, as well as comprehensive sets of endogenous retrovirus-derived elements of each species (24). We also characterized a spontaneously transformed line of *P. leucopus* dermal fibroblasts that continued in its capacity for proliferation long after an analogous *M. musculus* fibroblast line had ceased to replicate.

2 Results

Primary cultures and serial passages. For these experiments full-thickness samples of freshly-excised ears of euthanized *P. leucopus* or *M. musculus* animals were cultivated in a type of medium and under conditions long-used for primary fibroblast culture (25). There was no attempt to preserve the spatial characteristics of skin tissue by addition of growth factors or other supplements. The oxygen concentration was that of an incubator with 5% CO_2 and at sea level without any adjustments to limit oxidative stress (26). Cells of the original cultures and then of subsequent passages were aliquoted and frozen, thereby preserving the history of the two lineages. Examples of early passage cells from *P. leucopus* and *M. musculus* in culture are shown in Figure 1.

In the initial cultures *M. musculus* cells were confluent on the bottom of the dishes by day 5 to 7, while the *P. leucopus* cells at the same starting inoculum achieved that state only by day 10 to 12. Doubling times for early passage populations of *P. leucopus* cells and *M. musculus* cells were ~56 h and~28 h, respectively. By passage 16 (or ~70 doublings), the time for the *M. musculus* cultures to reach confluency had lengthened to 10-14 d, with a doubling time the same as the *P. leucopus* population at that point. Within a few more passages the *M. musculus* cells failed to adhere

and growth further slowed. The emergence of an established cell line, as can occur with cultivation of mouse embryo cells (27), was not observed.

Under the same cultivation conditions, by passage 20 (~86 doublings) the time to confluency for *P. leucopus* cells was only 6-8 days, which corresponded to a doubling time of ~14 h. Thereafter the *P. leucopus* cells grew at this higher rate through more than 47 serial passages with no signs of the abnormalities in adherence or growth rate noted for the mouse cells. This apparent adaptation of the *P. leucopus* cells to in vitro life was accompanied by a change in cell morphology; they became shorter and more rounded with fewer extensions, while retaining their adherence capacity (Figure 1). Thereafter the *P. leucopus* dermal fibroblast culture was considered "spontaneously transformed" (28). For subsequent experiments passage 4 (P4) or passage 47 (P47) were used as representative of low and high passage

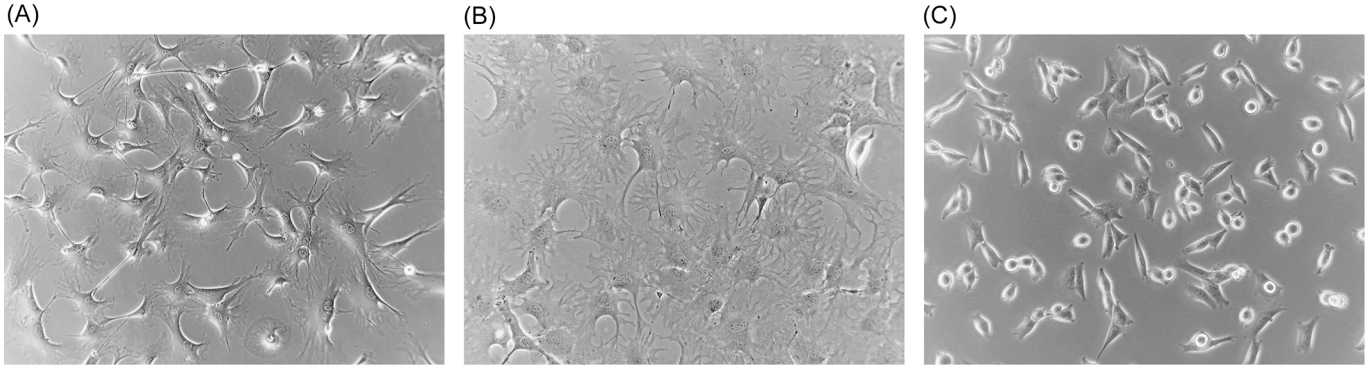

Figure 1. Phase microscopy photographs of cultures of dermal fibroblasts of *Peromyscus leucopus* (panels A and C) and *Mus musculus* (panel B). Panels (A) and (B) are of low passage (P4) cells, and panel C is of high passage (P47) transformed cells. Magnification, 150x.

populations.

Bulk RNA-seq. For the experiments with the lipopeptide TLR agonist Pam3CSK4, the sources of tissue for in vitro cultivation were 5 adult outbred *P. leucopus* (3 females and 2 males) and 5 adult outbred *M. musculus* (3 females and 2 males) (Table S1). Cultures at the same phase of growth were split in three and then subjected to the following exposures for 4 h: no treatment control or the lipopeptide TLR2 agonist Pam3CSK4 at 1 μg/ml or 10μg/ml. The cDNA libraries were mRNA stranded and yielded 1.03-1.42 x 10^8 paired-end 150 nt (PE150) reads for the 15 *P. leucopus* samples and 1.25-1.63 x 10^8 reads for *M. musculus*. Reference sets were the protein coding sequences (CDS) of the genome of a female *P. leucopus* of LL stock and the reference C57BL/6 genome of a female *M. musculus*. As a consequence of the more comprehensive isoform annotation for the mouse genome to that point, there were four times as many CDS sequences listed for *M. musculus* as for *P. leucopus*. Accordingly, for comparability, the first listed isoform for *M. musculus* was used for the reference set, resulting in nonredundant CDS sets of 22,760 for *M. musculus* and 22,654 for *P. leucopus* (Dryad Tables D1 and D2).

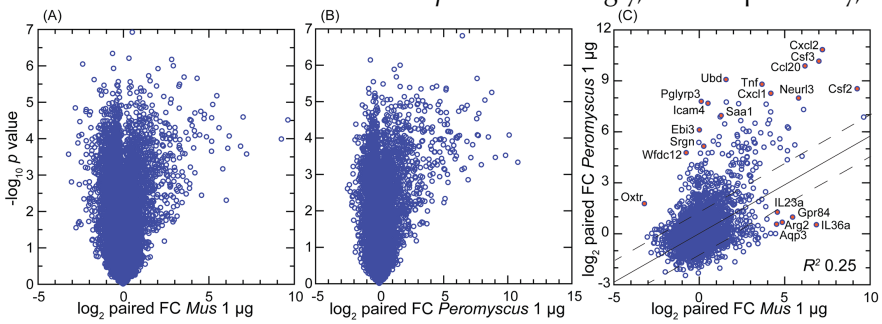

Figure 2. Volcano plot (panels A and B) and scatter plot (panel C) of treatment-to-control fold-change (FC) values from genome-wide bulk RNA-seq of primary dermal fibroblasts of *Mus musculus* (panel A) and *Peromyscus leucopus* (panel B). The treatment was 1 μg/ml of Pam3CSK4 lipopeptide for 4 h. In panels A and B false discovery rate p values were plotted against FC for respective species for 14,979 CDS in common and with TPM ≥10 for ≥1 sample (Dryad Table D1). In panel C the log-transformed FC values for each species are plotted against each other. The linear regression line with 95% confidence interval and coefficient of determination (R2) are shown. Selective CDS that up-regulated and differentially expressed for each are indicated by name and red fill. Data for analysis are in Dryad Table D3.

Distributions of log-transformed TPM values by cumulative count of genes were similar between species (Figure S1). There were 14,979 genes in common to both species and that met the criterion of TPM ≥10 in at least one of the individual 30 samples (Dryad Table D3). There was also similarity between species in the log-log plots of *p* values and fold changes of paired samples; for both species the numbers of DEGs with higher transcription after agonist exposure were greater than the number of DEGs with reduced transcription post-exposure (Figure 2 panels A and B). The mean coefficients of determination (R^2) values in log-transformed reads by CDS between paired values for the 1 μg and 10 μg per ml concentrations were 0.994 (0.992-0.996) for *M. musculus* and 0.964 (0.927-1.00)

for *P. leucopus* (p = 0.15), indicating low to negligible dose effect within the 10-fold range of the study (Dryad Tables D1 and D2). Unless otherwise noted, comparisons with paired controls were done with cells exposed to the 1 µg/ml concentration, as likely more representative of the tissue environment in the *B. burgdorferi*-infected skin of *P. leucopus* reservoir hosts.

By the gauge of fold-changes over paired control values, the deermouse fibroblasts had a higher number of upregulated DEGs than what was recorded for the mouse cells (Figure 2 panel C). Nearly identical results were found at the 10 µg/ml exposure for *M. musculus* (Figure S2), evidence that the lower number of upregulated DEGS in *M. musculus* was not likely attributable to greater resistance of mouse cells to the agonist. By this analysis some genes of relevance to innate immunity and that were upregulated in deermice cells but little if at all in mouse cells were Oxtr1 (oxytocin receptor) (29), Ebi3 (Epstein-Barr virus induced gene 3) (30), and the protease inhibitor Wfdc12 (WAP four-disulfide core domain 12), which is homologous to Slpi (31). For *M. musculus* two of the genes specifically upregulated in that species' cells were Gpr84 (G protein-coupled receptor 84), which is a loss-of-function allele in some inbred mouse strains (32), and the gene for the cytokine interleukin-36A, which is known to be active in skin (33).

By GO term analysis of the DEGs for each species (Figure S3), the responses of the deermouse and mouse cells were largely coherent, with the terms "innate immune response" (GO:0045087), "regulation of inflammatory response" (GO:0050727), and "inflammatory response" (GO:0006954) among the top 5 for each species and with p values $<10^{-20}$.

Higher transcription in *Peromyscus* (P) than *Mus* (M)					Higher transcription in *Mus* than *Peromyscus*				
CDS	FC P/M cont.	P/M *t*-test *p* cont.	FC P/M Pam3CSK4	P/M *t*-test *p* PAM3CSK4	CDS	FC M/P cont.	M/P *t*-test *p* cont.	FC M/P PAM3CSK4	M/P *t*-test *p* PAM3CSK4
Arg1	14.6	6E-04	23.2	3E-04	Adamts9	2.20	2E-01	31.3	3E-04
Batf2	60.2	3E-06	113	4E-06	Aqp5	589	8E-09	73.9	2E-07
Bdkrb1	15.5	5E-07	14.8	4E-08	Arg2	5.65	2E-03	68.7	1E-05
Calhm5	4.56	2E-04	13.1	6E-07	Ass1	332	2E-06	1032	6E-07
Casp1	12.8	7E-06	26.6	1E-06	Ch25h	0.66	1E-01	10.1	8E-07
Cd40	2.02	8E-02	13.1	7E-06	Chi3l1	80.5	2E-06	6.24	9E-03
Cmpk2	8.21	1E-03	25.8	8E-05	Cxcl11	2862	2E-08	39.9	8E-04
Cxcl10	5.25	9E-04	12.6	7E-05	Cxcl5	15.7	5E-04	2.94	2E-04
Ddx60	13.3	5E-04	20.5	4E-06	Egln3	10.3	3E-03	11.8	4E-03
Dennd3	13.7	6E-06	48.4	3E-08	Ereg	28.8	3E-06	21.0	8E-05
Deptor	30.4	4E-06	27.2	6E-06	Fth1	43525	1E-09	74792	4E-10
Dpy19l2	14.6	4E-07	20.0	2E-07	Il11	1.92	6E-02	11.7	3E-05
Dtx3l	5.10	2E-06	16.2	2E-06	Il1rn	1530	4E-07	84.0	2E-06
Hpx	3.35	2E-03	14.4	5E-04	Il36a	1.74	2E-01	142	6E-06
Icam4	0.96	9E-01	187	3E-10	Il4r	12.1	2E-06	2.70	1E-03
Ifi205	15.8	8E-06	41.8	1E-06	Il6	17.2	1E-05	1.75	1E-01
Il15	2.66	2E-03	18.6	1E-07	Ism1	4.09	5E-02	12.7	6E-04
Irf7	4.29	4E-03	13.2	5E-05	Lcn2	1719	6E-07	12898	8E-09
Irgm2	5.55	5E-05	43.7	6E-06	Lrig1	4.41	2E-04	20.8	3E-07
Kctd12	15.9	7E-09	17.4	2E-08	Lrrc17	90.5	3E-06	45.6	3E-06
Mmp12	4372	9E-07	3114	7E-07	Mmp3	11.1	4E-03	23.4	4E-04
Mx2	24.3	6E-05	66.9	5E-06	Mt1	18.9	1E-06	46.4	9E-08
Neurl3	5.31	2E-02	21.7	2E-04	Mt2	34.6	2E-05	20.6	9E-06
Oas1	10.3	2E-04	24.7	8E-06	Nos2	14.9	2E-03	26.1	3E-05
Pglyrp3	4.95	2E-02	1243	2E-08	Rgs16	249	1E-06	640	6E-08
Pla2g2a	882	6E-08	222063	5E-12	Saa3	172	2E-04	9.81	7E-06
Pla2g5	58.3	9E-07	281	7E-08	Sema7a	5.87	5E-03	12.0	5E-05
Plek	2.60	6E-02	18.4	8E-06	Serpina3f	13.7	2E-03	3.83	1E-02
Rcan1	9.90	1E-07	15.7	6E-07	Serpina3g	14.1	1E-03	4.98	3E-03
Rgs5	375	1E-06	364	2E-06	Serpina3i	37.1	2E-04	38.7	2E-04
Rnf122	1.68	5E-02	39.2	3E-08	Slco2a1	54.7	3E-04	123	5E-07
Rsad2	1.50	4E-01	23.0	7E-04	Slfn2	199	7E-11	683	7E-08
S100a8	15.5	4E-03	29.0	4E-05	Tnfsf11	1465	3E-06	214	2E-07
Saa1	0.98	1E+00	57.7	7E-07	Tyk	7.50	5E-04	13.6	1E-03
Samd11	234	9E-11	97.8	3E-08	Wnt16	950	1E-07	164	6E-07
Selp	1.83	1E-01	78.3	1E-06					
Serpina3n	128	2E-05	1742	7E-06					
Slc2a6	2.99	3E-03	16.7	6E-08					
Slc7a2	14.7	3E-06	3.16	5E-03					
Sp110	104	3E-07	59.7	4E-08					
Srgn	0.45	4E-01	14.4	2E-04					
Stat4	323	2E-06	1359	5E-08					
Tm4sf19	17238	9E-12	6108	7E-09					
Tnf	0.34	5E-03	19.3	1E-04					
Tnfaip6	4.85	2E-05	20.5	6E-08					
Tnfsf18	50.3	4E-04	781	2E-06					
Tnfsf4	12.3	4E-04	79.7	1E-05					
Traf1	2.57	3E-02	25.4	2E-07					
Trim30d	1649	2E-07	2293	1E-07					
Ubd	0.87	8E-01	148	1E-05					

Table 1: Differentially-transcribed coding sequences (CDS) between *Peromyscus leucopus* and *Mus musculus* low passage dermal fibroblasts under conditions of no treatment (control) or treatment with the TLR2 agonist Pam3CSK4 at 1 µg/ml. The left side of table has CDSs that had higher transcription in *P. leucopus* (P) than *M. musculus* (M), and the table's right side has CDSs that had higher transcription in *M. musculus* than *P. leucopus*. The targeted RNA-seq data (Table S2) were normalized by total reads, adjusted for length as reads per kb, and then, for cross-species comparison, the ratio to Gapdh reads for same sample. The 5 different cell cultures for each species (Table S1) were analyzed as pairs within a species for both fold-change (FC) and *t*-test under either control (cont.) conditions or Pam3CSK4 treatment.

Further differences between the species' fibroblast cultures were apparent when the DEGs were categorized according to specificity for one species or another, defined here as a ≥10x difference between species in respective fold changes for a given gene (Table 1; Tables S2 and S3). In the most discriminating cases, such as Pla2g2a (phospholipase A2, group IIA) and the PRR Pglyrp3 (peptidoglycan recognition protein 3) (34) of *P. leucopus* or Fth1 (ferritin heavy polypeptide 1) and the siderophore-binding Lcn2 (lipocalin 2) of *M. musculus*, the orthologous gene was not detectably transcribed in the comparison species. For some genes (e.g. Mx2 for *P. leucopus*) the distinctions between species were present under both control and treatment conditions, while for other genes (e.g. Il11 for *M. musculus*) the species difference was only noted for cells exposed to the agonist.

Many of the genes or pathways that stand out in Figure 2, Table 1, Table S2, or Figure S3 merit further analysis. But to do justice for all is beyond the scope of this report. Accordingly, we limit the focus to the following: (1) genes that prior studies of experimental animals had shown had relevance for the phenomenon of infection tolerance demonstrated by *Peromyscus* species, and (2) DEGs in this study that were unexpected.

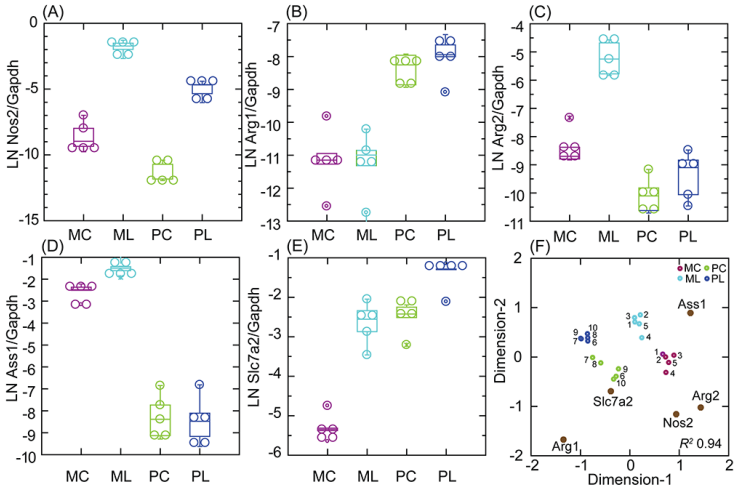

Figure 3. Transcription of nitric oxide and arginine metabolism CDS Nos2, Arg1, Arg2, Ass1, and Slc7a2 in targeted RNA-seq of primary dermal fibroblasts of *M. musculus* (M) or *P. leucopus* (P) without (C) or with (L) treatment with the lipopeptide Pam3CSK4 at 1 µg/ml. Panels (A)-(E) are box-whisker plots of individual CDS. Panel (F) is a Multi-Dimensional Scaling configuration plot of a analysis reduced to 2 dimensions (x- and y-axes) from the data for the box plots and with coordinates of CDS variables for the 10 samples indicated by gene names and symbols with dark brown fill. Pairs of individual lineages (1-5 for P and 6-10 for M; Table S1) are designated by numbers. The data for analyses are in Table S2.

Nitric oxide and arginine metabolism pathways. A notable finding in both studies of *P. leucopus* and *M. musculus* animals in response to LPS was a dichotomy between species in the blood and spleen in relationship between expression of nitric oxide synthase 2 (Nos2) and arginase 1 (Arg1) (16, 17). Deermice under treatment displayed a high Arg1-to-Nos2 ratio, which was consistent with the profile for alternatively activated macrophages. The mice instead had an inverted ratio of Nos2-to-Arg1, which is more typical for classically activated macrophage polarization (35, 36). The demonstration that LPS-treated, low passage fibroblasts of *P. leucopus* did manifest transcription of Nos2 above baseline values indicated that this species had the capacity for expressing Nos2.

In the present study we confirmed the transcription of Nos2 by RNA-seq with a different set of dermal fibroblast cultures and a different TLR agonist (Figure 3 and Table S3). The fold-change increase in Gapdh-normalized transcription of Nos2 between control and agonist-treated cells was similar in both species, but the expression of Nos2 was several fold lower at baseline in the deermouse cells. This was accompanied by a high baseline transcription of Arg1 with modest further elevation with agonist exposure for the *P. leucopus* cells. While the *M. musculus* fibroblasts, like the blood cells of animals, had low transcription of Arg1 under control and treatment conditions, transcription of Arg2, the gene for the mitochondrion-based type II arginase (37), was many fold higher than observed for deermouse cells for both conditions (Figures 2 and 3; Table 1).

Of note were other arginine metabolism genes that in their profiles of transcripton were distinguishing between the species (Figure 3 and Table 1). One of these was Ass1, the gene for arginosuccinate synthetase 1, which also is based in mitochondrion (38). Ass1 was little transcribed by *P. leucopus* cells but highly at baseline and further with treatment in *M. musculus* cells. A second was Slc7a2, which encodes a cationic amino acid transporter for L-arginine (39). Like Arg2 in the mouse cells, Slc7a2 transcription increased to high level in agonist-treated cells. Unlike Arg2, which was little transcribed in *P. leucopus*, Slc7a2, was at baseline at the same normalized level as for the treated *M. musculus*. Taken together, Nos2, Arg1, Arg2, Ass1, and Slc7a2, could in their multi-gene profiles differentiate not only between species but also between conditions in each species (Figure 3 panel F). The findings also point to differences between species in the metabolism of arginine and the extent to which this takes place in mitochondria of cultured fibroblasts.

Secretory leukocyte peptidase inhibitor (Slpi). Another feature distinguishing *P. leucopus* from *M. musculus* after exposure to LPS was a thousand-fold increase over baseline of transcription of Slpi in blood. At the same time there was only marginal elevation in the treated mouse samples over the low levels observed for controls (16). A heightened expression of Slpi over baseline was also seen in *P. leucopus* infected with the relapsing fever agent *Borrelia hermsii* as well as in a second experiment with LPS and with outbred instead of inbred mice (17). For LPS-treated cultures of ear skin fibroblasts, comparatively high transcription of Slpi was noted not just in the treated cells but in untreated cells as well (16).

In the present study with an independent set of cultures from *P. leucopus* and which were processed in parallel with cultures from outbred *M. musculus*, we confirmed high

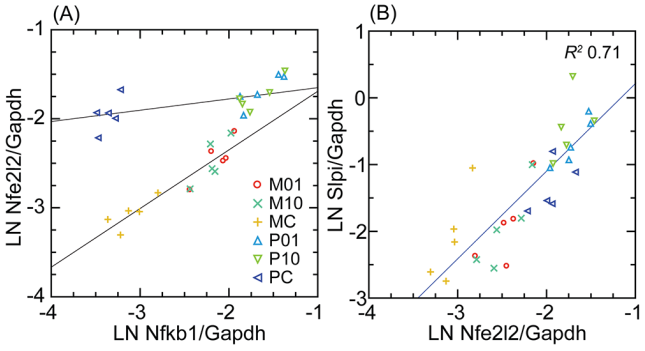

Figure 4. Scatter plots and linear regressions of normalized transcription of either Nfe2l2 (Nrf2) on Nfkb1 (panel A) or Slpi on Nfe2l2 (panel B) for *M. musculus* (M) and *P. leucopus* (P) dermal fibroblasts that either untreated (C) or treated with Pam3CSK4 lipopeptide at 1 µg/ml (01) or 10 µg/ml (10). For analysis of panel (A) the regression are separate for P and M. In panel (B) the regression was for both species and with the coefficient of determination (R2) for both. Data for analyses are in Table S2.

transcription Slpi in untreated cells and, as before, with a modest increase after TLR2 agonist exposure (Table S3). For this study we also documented the expression in the *P. leucopus* fibroblasts of Slpi protein by MALDI-TOF. The following three peptides, which cover 46% of the 106 amino acid processed protein (accession XP_028724460.1) were identified: KDSIKIGACPSISPAK, CTVPLPISRPVRRK, and KSGKCPTFQGRCMMLNPPNK.

The various pathways leading to Slpi expression are not well defined. To provide insight for this we identified among 365 CDS that were upregulated DEGs for either or both *P. leucopus* and *M. musculus* (Table S2) those that were highly correlated by R^2 calculation with Slpi transcription across species and conditions. The first transcription factor on the list of descending R^2 values was Nfe2l2, which encodes the protein nuclear factor, erythroid derived 2, like 2, also known as NRF2. Nfe2l2 is a transcriptional regulator of a number of genes involved in the adaptive response to oxidative and other cytotoxic stress (40).

Nfe2l2 transcription in the deermouse and mouse fibroblasts was compared with that of the more broadly active transcription factor Nfkb1 (Figure 4). For the control cells when Nfkb1 transcription was about the same in both species, Nfe2l2 expression was nearly as high for P. leucopus as it would be after exposure to the lipopeptide. While Nfe2l2 transcription in the treated *M. musculus* cells never reached the levels observed in *P. leucopus*, there was a still greater fold increase over baseline than was observed for deermouse cells. The relationship between Nfe2l2 and Slpi was direct and consistent across species and conditions, suggesting a regulatory role of Nfe2l2 in Slpi expression in *P. leucopus*.

PRRs and interferon-stimulated genes (ISG). The results with the fibroblast cultures replicated some of the previous findings with experimental animals. One observation for those PRRs and ISGs that were DEGs for either or both mouse and deermouse was that illustrated by the genes for the PRR RIG-I and Isg15 in panel (A) and panel (B) of Figure 5, namely transcription that was lower in *P. leucopus* than *M. musculus* at baseline but then exceeded that of the mouse with agonist exposure. Another observed pattern is shown in the right-most panel. As for *P. leucopus* animals the antiviral ISG Mx2 was in the fibroblast cells at an elevated level relative to mouse at baseline and displayed an even higher level with the treatment. Other genes demonstrating the second pattern in the fibroblasts in a replication of the findings for the experimental animals were the PRR Ifih1 (MDA5), the regulatory factor Irf7, and the antiviral ISG Oas1 (Figure S4).

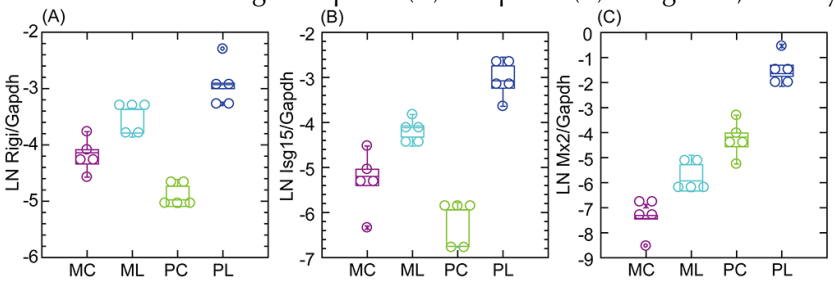

Figure 5. Box-whisker plots of log-transformed Gapdh-normalized transcription of the Pattern Recognition Receptor (PRR) CDS Rigi (panel A) or the Interferon Stimulated Genes (ISG) Isg15 (panel B) or Mx2 (panel C) in *M. musculus* (M) or *P. leucopus* (P) dermal fibroblasts without (C) or with treatment (L) with lipopeptide Pam3CSK4 at 1 µg/ml. Data for analyses are in Table S2.

Endogenous retrovirus and transposable elements. In a previous analysis of blood of deermice and mice injected with LPS, we observed that the two species differed in their transcription of an ERV envelope protein in response to the TLR4 agonist LPS (17). For the present study we used much expanded reference sets of ~1 million annotated ERV and derived transposable elements (TE) sequences for each species (Dryad Tables D4 and D5). The distribution of lengths of the ERVs in the complete sets was remarkably similar between species, with approximately equal proportions of sequences in different size classes, including those more the 5 kb in length (Figure S5). After exclusion of ~90% of the sequences that were 11-499 bp, there remained 104,932 sequences for the *M. musculus* reference set and 103,397 for *P. leucopus* (Dryad Table D6). For these sets the lengths ranged from 500 to 9329 bp for *M. musculus* and from 500 to 9543 bp for *P. leucopus*. The distributions in sequence lengths were skewed toward lower values and with a long tail to the right; the skewness statistic was 4.3 for mouse and 3.5 for deermouse. The median length (inter-quartile range) was 711 (572-950) bp for *M. musculus* and 742 (567-1164) bp for *P. leucopus*. For the control samples, mean values for the TPM-based transcription measure were 9.6 for *M. musculus* and 9.1 for *P. leucopus* sequences. TPM values ≥10 were observed for 5364 (5.1%) of the mouse sequences and 7336 (7.1%) for the deermice.

A closer look at the ERV/TE loci with higher transcription levels in cells under control conditions revealed a difference between cultures. If the sets of sequences were limited to the top 1000 in descending order of mean TPM (i.e. the top ~1% for each species), the species were comparable in terms of the distributions of the length-independent TPM values (Figure S6). What distinguished the mouse sequences were their overall longer absolute lengths. The mean (95% CI) lengths were 2622 (2525-2719) bp for mouse and 985 (942-1029) bp for deermouse sequences (*t*-test $p < 10^{-10}$; Kruskal-Wallis $p < 10^{-10}$); the respective medians were 2355 and 730 bp. The lengths difference was as marked for the top 100 (i.e. ~0.1%) of transcribed sequences by TPM values; mean lengths for the top 0.1% were 3019 (2742-3186) bp for mouse and 1146 (972-1319) for deermouse cultures.

A further distinction between *P. leucopus* and *M. musculus* was for the sub-set of ERV/TE sequences that increased ≥5-fold in transcription in the presence of the lipopeptide over buffer alone (Figure 6; Table S4). For this comparison we

used Z scores, based on the distributions of lengths for the reference sets for each species. There were 178 sequences by the fold-change criterion in the *P. leucopus* set and 80 in the *M. musculus* set. Though the deermouse set outnumbered the mouse set, the total lengths of concatenated sequences were 174,055 bp for *P. leucopus* and 252,553 bp for *M. musculus*.

Of the 80 sequences in the mouse set of DEGs, 41 (51%) were ≥3000 bp, with a range of 3003 to 7995 and corresponding Z scores of 4.97-16.65 (Table S4). In contrast, only 4 (2.2%) of the 178 deermouse DEGs were ≥300o bp (Fisher exact $p < 10^{-6}$), with respective bp and Z score ranges of 3047-4018 and 3.11-4.58. The disparity in the lengths of DEG sequences between species led us to ask whether longer elements would more likely have open reading frames (ORF) with a potential for translation than shorter elements after adjusting for length. If this was the case for one species, did it hold for other? With the assumption this might be more apparent among the longer ERVs, we focused on the size range of 3 to 8 kb.

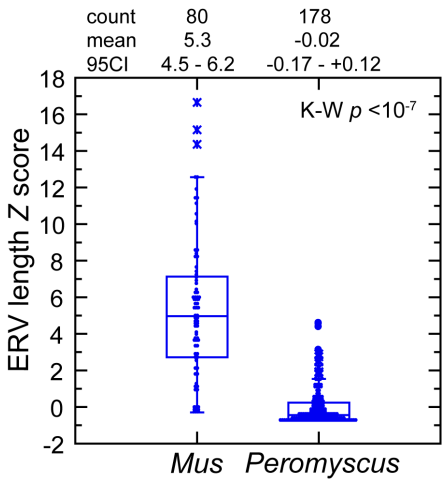

Figure 6. Box-whisker plots of a comparison *Mus musculus* and *Peromyscus leucopus* for the lengths of their upregulated ERV/TE sequences in low passage dermal fibroblasts exposed to the TLR2 agonist Pam3CSK4. Lengths are represented as Z scores from the distributions of lengths of all ERV/TE for each species. The non-parametric test is Kruskal-Wallis (K-W). Data for analysis are from Table S4.

We identified in these ERV/TE sequences for each species ORFs having an ATG as start codon (where A is position +1), were at least 30 codons, and had both a purine at position -3 and a G at position +4, which are features of a Kozak sequence for a protein translation initiation site in vertebrates (41). The identified ORFs were then used for searches of both gammaretrovirus protein sequences and mammalian proteins for those with similarity to Gag-pol polyproteins or retroviral ENV proteins. While there were no ERV sequences above 5 kb among the *P. leucopus* DEGs, 12 *M. musculus* DEGs ranged from 5151 to 7995 bp. (Table S4). These 12 sequences had in total 42 ORFs on either positive or minus strand encoding retroviral-like polypeptides. There was per ERV sequence a mean of 3.5 ORFs, ranging from of 30 to 723 amino acids and with a mean of 117 amino acids per ORF. Of the 42 ORFs, 30 (71%) were identified as part of a Gag-pol polyprotein and one (2%) was identified as a partial sequence for a retroviral envelope protein. Over the lengths of the 12 mouse ERV sequences, the mean proportion comprising these ORFs were 0.21 (0.14-0.27). An example was a 5596 bp sequence of chromosome 5 (positions 122331404-122337360) of the mouse reference genome; a 0.38 fraction of the sequence comprised two ORFs that met the specifying conditions. One would encode a 35 aa partial Gag-poly protein, and the second would encode a 728 aa polypeptide that, as one example of homology, was 95% identical to positions 647-1361 of the xenotropic MuLV-related virus VP62 (GenBank accession A1Z65).

The corresponding ORF proportion values for 12 ERV sequences that ranged from 5169 to 7778 bp and that were highly transcribed in the mouse fibroblasts under both control and treatment conditions were 0.19 (0.11-0.27) ($p > 0.2$). Thus, an ERV upregulated in transcription in presence of the agonist was neither more or less likely to have plausibly translated ORFs than those non-DEG ERVs of similar lengths.

For an analysis that included both *P. leucopus* and *M. musculus* ERVs, we applied the same filter for ORFs for DEGs that ranged in size from 3-4.5 kb. The mean proportions of the lengths of ERVs that comprised the ORFs were 0.04 (0.03-0.05) for 5 *P. leucopus* ERVs and 0.11 (0.08-0.13) for 28 *M. musculus* ERVs in that size range ($p = 0.05$).

To sum up, not only did the longer transcribed ERVs of mouse fibroblasts have, as expected, more qualifying ORFs than shorter ERVs, but the size-independent proportions of the sequences constituting those ORFs were disproportionately higher. In comparison, among the highly transcribed and DEG ERVs of *P. leucopus* fibroblasts, there were fewer long sequences. Moreover, among the ones that were DEGs, the cumulative ORF proportions were lesser for *P. leucopus* than was noted for *M. musculus*.

Interleukin-11. Unlike preceding examples, which were of correspondences between the findings of in vivo and in vitro systems, some distinguishing genes for the fibroblasts would not have been predicted by studies of specimens of whole blood, spleen, and liver. One we highlight here is interleukin-11 (IL-11), the gene for which was an upregulated DEG in mouse fibroblasts, but lowly transcribed in deermouse fibroblasts, regardless of condition (Table 1). IL-11 is a member of IL-6-type cytokine family; the specific receptor, IL-11Ralpha (Il11ra), is expressed by fibroblasts but not immune cells. An upstream transcription factor for IL-11 is NF-kappaB. IL-11 is considered a key factor in the inflammation of aging ("inflammaging") process, in part by promoting fibrosis (42, 43).

A targeted RNA-seq analysis of transcription of Nfkb1, Il6, Il11, and Il11ra, with Gapdh as the housekeeping gene for normalization across species, is shown in Figure 7. There were higher baseline levels of Il6 in mouse cells than

deermouse cells, a greater magnitude of Il6 elevation with agonist exposure in deermouse cells, and similar Nfkb1 levels between species. The finding of the genome-wide RNA-seq analysis for Il11 was confirmed. While Il11 transcription levels increased in mouse cells with both concentrations of Pam3CYSK4, in deermouse there was, if anything, a further decline from low levels at baseline. The scarce to absent Il11 transcripts in these low-passage *P. leucopus* cells was not accompanied by reduced expression of its receptor Il11ra, which was at comparatively high levels in deermouse control cells and, unlike mouse cells, did not decline with agonist exposure. These were the findings for the primary culture and early passage cells of *P. leucopus*. As described below for high passage cells, the capacity of *P. leucopus* for transcription of Il11 had not been lost.

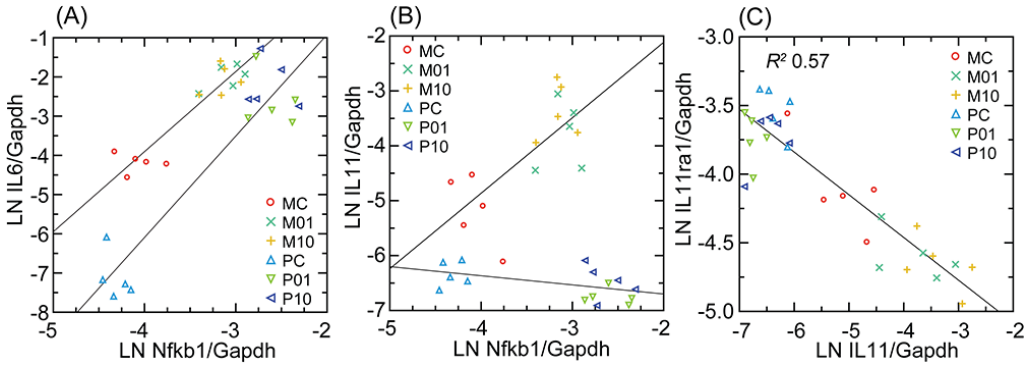

Figure 7. Interleukin-6 (Il6; panel A), interleukin-11 (Il11; panel B), and interleukin-11 receptor (Il11ra; panel C) Gapdh-normalized transcription compared in scatter plots and with linear regressions for *M. musculus* (M) and *P. leucopus* (P) dermal fibroblasts that were either untreated (C) or treated with Pam3CSK4 lipopeptide at 1 μg/ml (01) or 10 μg/ml (10). For analyses of panels (A) and (B) linear regressions are separate by species. For panel C the linear regression with coefficient of determination (R^2) was for both. Data for analyses are in Table S2.

Single-cell RNA-seq of low passage populations. At their outset the cultures of the full-thickness ear tissue included cartilage, capillaries and other small vessels, epidermis, hair, and hair follicles as well as the dermis that would be the presumptive source for the fibroblasts (44). What emerged from this initial mixture of cell types were adherent cells. These could be released from the polystyrene dish bottoms with trypsin, and when these were used to populate a fresh culture the cells soon became adherent again. They had the morphology of fibroblasts (Figure 1), and we had the bulk RNA-seq results, which for both species, confirmed the expression of fibroblast markers, such as the collagen gene Col1a1, fibroblast-specific protein 1 (S100a4), and vimentin (Vim). But heterogeneity was undefined.

As a first step to better characterize these primary cultures we carried out single cell RNA-seq of two sets of pooled cells, one for each species. Each pool contained equal parts of cells that were untreated, cells exposed to Pam3CYSK4 for 4 h, and cells exposed to LPS for 4 h. The reasoning was to represent in a single population cells after exposure to two different TLR agonists as well as cells at baseline and then processed together. Each species' sample yielded ~4 x 10^4 cells and, after downstream processing, indexing, library preparation, and sequencing, ~10^9 PE150 reads.

We approached the analysis without set ideas about markers to prioritize. Gene expression matrixes were normalized and dimensionally reduced by principal component analysis. UMAP projections and cell clustering were generated by the top principal components for each cell. There were 10 major clusters for each species (Figure 8 panel A), numbered 1-10 for *P. leucopus* and 11-20 for *M. musculus*. For both species' cells these organized in two super clusters, shown on the left and right in each graphic. The individual clusters for each species were in roughly the same spatial arrangement, with the exceptions of cluster 10 of *P. leucopus*, and cluster 20 of *M. musculus*, which did not have a clear counterparts in the comparison species. By analysis of single genes we found that transcription of two collagen genes, Col15a1 and Col8a1, specifically distinguished left and right super clusters, for both species, thus providing evidence of comparability of deermouse and mouse cultures in terms of cell type representation at this higher order.

Cross-species analogies among clusters were supported by predictions of what transcription factors would be predominant from GO term analysis of the genes contributing to cluster discrimination (Figure 8 panel B and Table S5). Transcription factors associated more generally with normal cell growth, such E2f1 and Rbl2, served to define the upper clusters (i.e. 5, 6, 7, and 9 of *P. leucopus* and 13-16 and 19 of *M. musculus*) of each of the left and right super clusters of deermouse and mouse. On the other hand, the lower poles of both super clusters, were identified with transcription factors, such as Nfkb1, Nfe2l2 (Nrf2), Stat1, and Stat3, more typically associated with host responses to a stressor, suggesting to us that these lower pole clusters were the cells exposed to one or the other TLR agonist. For half the clusters there were also between 2 and 8 genes that contributed to defining a cluster in the same relative location in both deermouse and mouse. These were clusters 1, 3, 4, 7, and 9 for *P. leucopus* and 11, 14, and 16-18 for *M. musculus*. These included the Epas1 (endothelial PAS domain protein 1) for clusters 3 and 11, and Csf1 (colony stimulating factor 1, macrophage) for clusters 1 and 11.

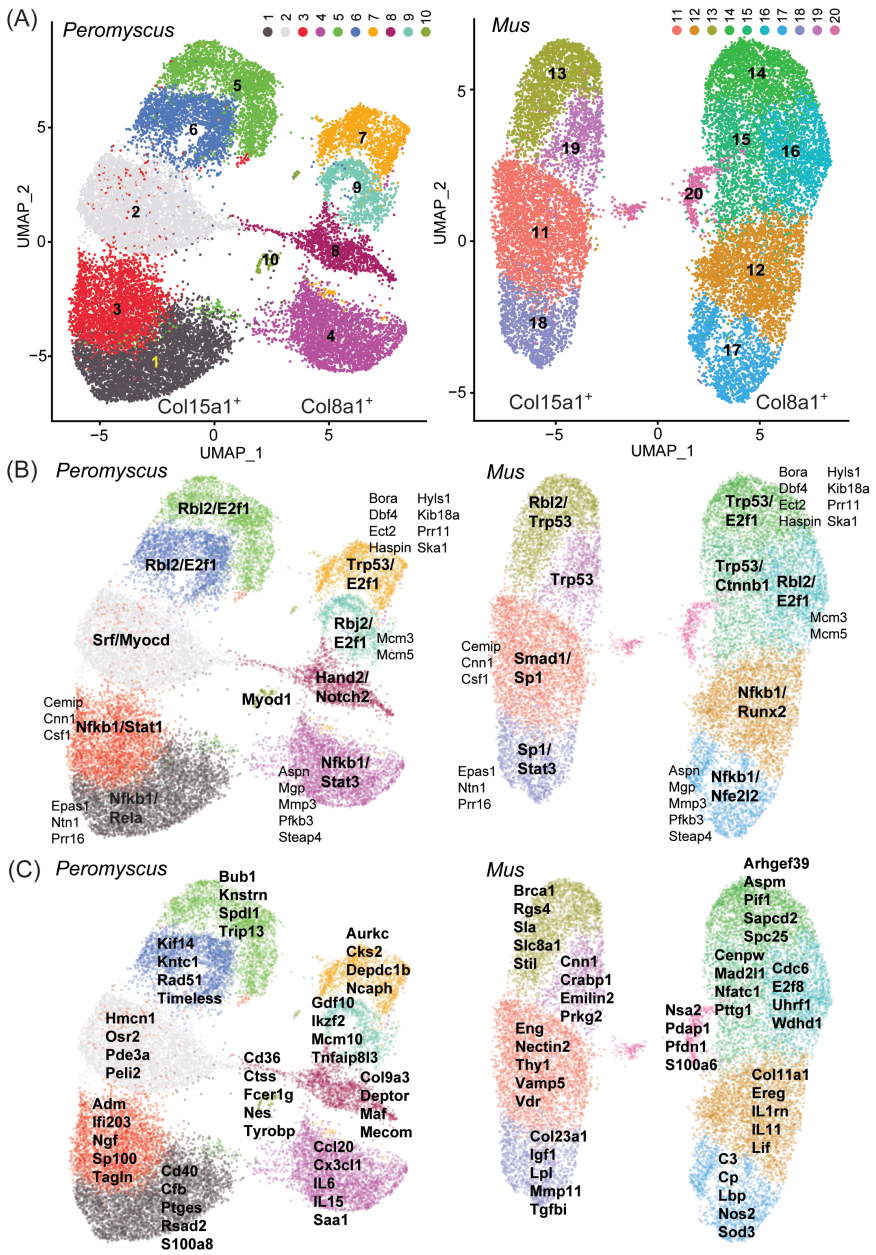

Figure 8. Single cell RNA-seq of low passage cultivated dermal fibroblasts of Peromyscus leucopus and Mus musculus with 2-dimensional UMAP projections of clusters of types of cells in pooled populations of cells that were untreated or exposed to either Pam3CSK4 or LPS. In panel (A) the clusters of each species are identified by number and color: 1-10 for P. leucopus and 11-20 for M. musculus. Two CDS that served as markers for the left and right super clusters were Col15a1 (left) and Col8a1 (right). In panel (B) transcription factors that were predicted by GO term analysis to be associated with signifying CDS for each cluster (Table S5) are indicated in bold font. CDS that defined the same clusters in both species are in smaller and regular font. In panel (C) CDS that were enriched in representation in the clusters and that discriminated to that degree between species are listed for each cluster.

If only the 4-5 top-ranked genes by degree of enrichment and that uniquely contributed to cluster identity were considered (Figure 8 panel C), we observed again in the lower halves of the two super clusters of each species a predominance of genes that are associated with innate immune and other host responses. Of note with regard to other aspects of this study, are the anti-viral effector Rsad2 and calgranulin component S100a8 in cluster 1, interleukin-6 and the chemokine Cxcl20 in cluster 4, interleukin-11 and the interleukin 1 receptor antagonist Il1rn in cluster 11, and Nos2 and complement component C3 in cluster 17. The small cluster 10 located between the two super clusters of *P. leucopus* stood out from the other 9 clusters of the deermouse fibroblasts and all 10 mouse clusters in the specifying contributions of dendritic cell-associated genes Cd36, Fcer1g (CD23), and Tyrobp (DAP12). Cluster 20 of *M. musculus* was defined in part by the genes listed in the panel C of figure but also by enrichment of lower magnitude for several ribosomal proteins (Table S5), an indication that this was a more rapidly dividing cell population, without a counterpart in the *P. leucopus* fibroblast sample.

Spontaneously transformed *P. leucopus* fibroblasts. Bulk RNA-seq was carried out on cultures of *P. leucopus* at passage P4 and passage P47 and which were controls or had been treated with Pam3CSK4 or LPS for 4 h. The number of PE150 reads ranged between 2.13-2.93 x 10^8 with a mean of 2.45 x 10^8. The primary objective was to identify genes that were transcribed in P4 cells but to little or no degree in P47 cells or vice versa. For this purpose, we compared the 3 combined samples of P4 against the 3 combined samples of P47 cells. By the criterion of a minimum fold-change of 4, maximum group mean TPM ≥10, and an FDR *p* value <0.01, 80 genes were differentially upregulated in P4 cells and 25 were upregulated (denoted by negative fold-change values) in P47 cells (Table 2 and Table S6). Genes that were little transcribed in P47 compared to P4 cells were Col8a1, Deptor, Eng, Il6, and Lbp, each of which were more closely associated with the right-side super clusters in the single cell experiment (Figure 8). On the other hand, upregulated genes in the transformed cells included Il1rn, Il11, and Timeless, which were associated with the left-side super cluster. This indicated that the transformed cells more likely trace their lineage to the left super cluster fibroblasts than the fibroblasts on the right.

Besides the near- or complete loss expression of several of the innate immunity-associated transcription factors, cytokines, chemokines, and ISGs that featured in the comparison with the low-passage mouse fibroblasts, there was also diminished transcription of MHC class I protein genes in the P47 cells. In addition to the upregulated Il11 and Il1rn, which notably were lowly transcribed in low passage deermouse cells in comparison to mouse (Table 1), two cancer-associated genes, Brca1 and Myc, were more highly transcribed in P47 than in P4 cells.

A secondary objective was to identify in the transformed cell line those genes that not only were still transcribed but also increased in expression after exposure to either or both of the TLR agonists. If this was the case, it would indicate retention of some signaling pathways of cell autonomous immunity (Table S7). One of these was Slpi, which was a thousand-fold lower in transcription in P47 cells than the P4 cells across all conditions (Table 2), but still was ten-fold higher in transcription in the agonist-treated P47 cells than in the controls. Another was Cxcl1, which was a hundred-fold higher in transcription in Pam3CSK4-treated P47 cells than in the controls. The chemokine Cxcl10 was not one of those diminished in transcription in P47 versus P4 cells, but, like Cxcl1, it displayed greater than ten-fold up-regulation in transcription with the TLR2 agonist but not the TLR4 agonist.

The reads were also aligned with *P. leucopus* reference set of 103,397 ERVs of ≥500 bp. Across the 6 samples, 10,843 (10.5%) ERVs had a maximum group mean TPM ≥ 5. Of these, 832 (7.7%) of ERV sequences were more highly transcribed in low-passage than in high-passage cells across all 3 conditions by the criterion of FDR *p* <0.01 and absolute fold-change ≥2 for 3 samples for each passage history (Table S8). By the same criteria, 1136 (10.5%) were more highly-transcribed in high-passage cells (denoted in the table as negative fold-change values).

The distinguishing DEGs of P47 cells were of interest because of their association with an ostensibly immortalized state. The high passage cells' upregulated DEGs were in general longer (mean of 1189 vs. 1082 bp; *p* = 0.01) and more highly transcribed (mean of mean TPM of 328 vs. 81; *p* = 0.003) than low passage cell DEGs. Of the 14 high passage cell DEGs of ≥5 kb size, 3 were >9 kb, with lengths (and chromosome locations) of 9543 (NC_051064.1:6812577-6822120),

CDS name	CDS length (bp)	TPM group mean	log_2 fold change	FDR *p* value		CDS name	CDS length (bp)	TPM group mean	log_2 fold change	FDR p value
Rspo3	825	10	13.3	2E-08		Il6r	1377	51	3.6	2E-75
Ly6e	411	104	11.2	1E-83		Irf9	1278	38	3.6	3E-23
Vcam1	2208	696	11.0	6E-48		Saa1	369	13	3.5	1E-05
Icam1	1485	273	11.0	2E-32		Nfkbiz	2205	34	3.5	6E-06
Tfpi2	705	80	10.7	3E-99		Irf7	1401	77	3.0	3E-11
Cfb	2301	376	10.7	6E-35		Cxcl1	300	1396	2.8	8E-03
Ldoc1	423	243	10.6	3E-178		Eng	1968	50	2.8	1E-89
Saa3	369	75	10.6	2E-17		Col6a2	3081	363	2.8	9E-41
Pla2g2a	441	86	10.3	3E-19		Ccl7	462	1599	2.8	5E-04
Cxcl5	498	102	10.2	6E-48		Nod1	2814	76	2.7	5E-56
Csf3	627	110	10.2	7E-17		Ifih1	3069	28	2.7	3E-15
Ifi203	2709	38	10.1	1E-76		Stat2	2853	38	2.6	4E-15
Adamts5	2799	42	10.0	1E-89		Deptor	1236	17	2.5	7E-13
Ndn	978	62	10.0	2E-143		Sting1	1140	103	2.4	6E-22
Cxcl3	309	71	10.0	1E-20		Sod2	690	616	2.4	3E-06
Efemp1	1479	240	10.0	1E-31		Ikbke	2196	97	2.4	1E-14
Rsad2	1083	23	9.9	2E-23		Ifngr1	1404	163	2.3	1E-35
Prrx2	747	121	9.6	4E-64		H-2 Q10	1206	127	2.3	2E-38
Sema3d	2346	76	9.4	4E-27		Bst2	519	61	2.3	1E-06
Cxcl12	282	22	9.4	9E-34		Rigi	2781	35	2.2	1E-06
Dpt	606	785	9.3	3E-63		H-2 DDa	1086	105	2.2	4E-32
Col3a1	4398	723	9.2	3E-67		Nfkb1	2910	93	2.2	4E-10
Ago1	2565	17	9.2	3E-150		H-2_LDa	1215	216	2.2	5E-43
Col11a1	5568	16	9.1	1E-79		Gbp2b	2325	67	2.0	6E-03
Cd248	2292	27	9.1	3E-226						
Dcn	1092	688	9.1	7E-45		Tlr2	2355	60	0.1	9E-01
Ednrb	1329	14	9.0	2E-86		Tgfb1	1173	128	-0.2	3E-01
Slpi	396	329	9.0	2E-74		Actb	1128	6659	-0.3	2E-01
Lbp	1446	89	8.9	3E-96						
Lum	1017	50	8.8	2E-89		Insig1	774	127	-2.0	2E-32
Ccl27a	474	15	8.6	1E-46		Il11	1092	15	-2.0	9E-15
Icam4	930	11	8.3	2E-16		Lpl	1425	747	-2.1	2E-24
Plek	1053	24	7.8	5E-93		Timeless	3585	17	-2.1	2E-28
Col8a1	2238	48	7.8	0E+00		Mmp2	1986	22	-2.2	2E-28
Srgn	621	16	7.2	6E-17		Ung	1305	27	-2.5	1E-34
S100b	279	13	7.0	2E-28		Brca1	5469	21	-2.6	3E-55
Ctsk	990	174	6.8	0E+00		Myc	1317	64	-2.6	3E-56
Il1r2	1308	11	6.5	2E-95		Aqp1	810	698	-4.6	4E-88
Bdkrb1	996	93	6.3	1E-75		Col6a3	7350	328	-4.7	2E-151
Col5a3	5223	37	5.9	1E-58		Cxcl14	300	165	-5.1	3E-117
Irgm1	1239	38	5.9	8E-17		Mmp9	2157	13	-5.4	4E-16
Stat4	2244	43	5.9	2E-188		Ca3	783	120	-5.8	2E-290
Igf1	462	105	5.6	5E-145		Abcg2	1971	144	-6.7	0E+00
Il6	639	90	5.4	1E-14		Gzma	789	61	-6.8	1E-38
Sdc4	591	1303	5.2	2E-24		Col18a1	3972	65	-6.8	0E+00
Fscn1	1482	185	5.1	2E-274		Sema7a	1995	24	-7.3	4E-193
Cxcl11	306	133	4.8	2E-04		Il1rn	480	36	-7.4	2E-99
C4b	5226	22	4.7	4E-213		Krt14	1428	16	-8.0	4E-134
Rtp4	1224	24	4.6	7E-17		Adamtsl2	2775	89	-8.1	0E+00
H-2 DDa	738	71	4.2	4E-08		H19	351	3554	-8.9	0E+00
Mx2	1986	149	4.1	7E-16		Dkk2	780	133	-9.0	0E+00
C1r	2118	778	4.0	3E-58		Col4a5	5076	85	-9.8	0E+00
Col6a1	3108	91	3.9	5E-182		Kpna7	1503	133	-11.7	1E-239
Bcl3	1344	41	3.9	6E-19		Cck	348	315	-13.1	3E-74
Tagln	606	195	3.7	2E-123						

Table 2. Differentially-transcribed coding sequences (CDS) between low passage (P4) and high passage (P47) dermal fibroblasts of Peromyscus leucopus from bulk RNA-seq analysis of fibroblast cells without or with exposure to Pam3CSK4 or LPS. The comparison was of the samples from 3 conditions for the P4 cells with the samples from the 3 conditions of the P47 cells. The table lists by columns the CDS names, CDS lengths in base pairs (bp), the TPM group mean, the log2 of fold change (FC) between the P4 cells compared to the P47 cells, and the false discovery rate (FDR) p value. The CDS comparatively upregulated in P4 cells are indicated by yellow fill of the cells. The CDS comparatively upregulated in P47 cells are indicated by blue fill. Selected examples of CDS that were not differentially transcribed are indicated by gray fill. The product names of the CDSs are provided in Table S3. The data are drawn from the fuller roster of CDS and results of Table S6.

9432 (NC_051076.1:33708849-33718281), and 9208 (NC_051063.1:54636550-54645758) bp. Not only do the three have ORFs for components of Gag-pol polyproteins, including reverse transcriptases, but they also would encode retroviral envelope (Env) proteins. For the ERVs of 9543 and 9432 bp, the CDS for Env sequences aligned with feline leukemia virus (FLV) proteins along their full-lengths and without frame shifts or in-frame stop codons. The Env CDSs of the 9543 bp and 9432 bp ERV sequences have a purine at -3 position and ATGG for positions +1-4, thus plausibly translatable. The 1983 nt sequence of the predominant FLV-related Env polyprotein mRNA transcribed by the transformed *P. leucopus* fibroblasts encodes a polypeptide of 661 aa with a gp70 portion from positions 32 to 463 and the more conserved, transmembrane p15E portion from positions 464 to 638 (GenBank accession PV892740).

In contrast, the 6 low passage DEGs of ≥5 kb, fold-change ≥2x, and FDR p <0.01 were of a maximum length of 7503 bp (Table S8). The 6 ERV sequences each comprised ORFs for parts of Gag-pol polyproteins, but none would discernibly encode a retroviral envelope protein, in part or whole. This and other findings about the immortalized fibroblasts showed that amidst the silence of many genes of inflammation consequence, such as Il6, other genes, such as some ERV sequences and Il11, were effectively released from repression.

3 Discussion

Previous comparisons of *P. leucopus* with *M. musculus* used laboratory-bred animals that were euthanized at experiment's end (16, 17, 45, 46). Acknowledging that in vivo studies deserve their pre-eminence in research, we asked whether pre-mortem primary cultures of a tissue could stand-in for the animals themselves for some research questions. For field-based studies this would allow for the release of captured animals after measurements and specimens have been taken. This was one rationale for the application of wing tissue biopsies for studies of bats in nature (47, 48). For laboratory-based investigations of stock colonies longitudinal designs with non-lethal sampling of tissue at different times become feasible.

An indisputable advantage of tissue culture is the capability to freeze the cells indefinitely, permitting replication of an experiment or the application of procedures not available at the time of collections. With their tissues propagated and cultures split, animals can serve as their own controls in experiments. There can be interventions that could not be feasibly tested with live animals, because of toxicity at the organismal level or impeded delivery to the tissue or organ of interest. Experiments on cultured cells do not suffer from these restrictions. These manipulations include RNAi and other non-transgenic gene silencing methods or CRISPR gene editing applications before this technology is achieved for the whole animal.

These are justifications for use of cultured cells. But their validity as approximations of live animals rests on two assumptions. The first is that cells cultured from a non-model organism, e.g. *P. leucopus*, sufficiently resemble those of a more established model organism, i.e. the house mouse, that one can exploit extant databases and literature on the model organisms. If there are many incongruities or misalignments, the findings for the deermouse might be limited in impact to that one genus. The second assumption is that in vitro experiments' results generally line up with those in vivo experiments for the same organism. In other words, if a gene is upregulated in the animal under a certain condition, is it also upregulated in the tissue culture cells.

In considering the first assumption, we noted the report that established cultures of *P. leucopus* fibroblasts had similar doubling times to those of *M. musculus* (49). Unlike cultured fibroblasts of humans and some other large mammals, both deermouse and mouse cells had telomerase activity and did not undergo replicative senescence. The observed decline in the present study of the *M. musculus* culture within a few passages is provisionally attributed to greater oxygen sensitivity of the mouse cells (26) and not replicative senescence. Notwithstanding this possibly consequential distinction between species in response in oxidative stress, the primary or early passage populations of *P. leucopus* and *M. musculus* dermal fibroblasts had more in common than differences by genome-wide RNA-seq (Figure S1; Figures 2-3).

By scRNA-seq there were also analogies between species in the types of cells represented in the primary cultures (Figure 7). This was evident in the UMAP plots, as well as in the genes contributing to cluster identities (Table S5). Col15a1 and Col8a1 empirically distinguished the left and right super clusters of both deermouse and mouse in this experiment (Figure 7), but these genes are not generally recognized as markers for the predominant types of fibroblasts in mouse skin: papillary fibroblasts in the upper dermis and reticular fibroblasts in the lower dermis (50). Some markers known for their specificities for fibroblast types in intact skin would be expected to decline in expression in culture (51), thus reducing the numbers of informative characteristics available for cell origin typing. Nevertheless, there were some enrichments of expression that allow for tentative assignment of origins for some clusters. Examples were Crabp1 (cellular retinoic acid binding protein 1) and Col23a1 for left sided clusters and Aspn (asporin) and Col11a1 for right-sided clusters (Table S5). These suggest to us that the left and right super clusters had origins in papillary fibroblasts and reticular fibroblasts of the dermis, respectively, of mouse and deermouse.

Justifying the second assumption is a challenge, what with differences in complexity and time scale between the in vivo and in vitro. But, as noted by Tyshkovskiy et al. (52), who cited examples from research on aging, a rationale for this assumption exists. One of those examples was the report of Ma et al. (53). These investigators cultivated fibroblasts from the skin of captured *P. leucopus* and *P. maniculatus*, along with several other small and medium sized mammals, and found that the RNA-seq and biochemical assay results with in vitro-cultivated cells generally conformed with known longevity traits of the animals themselves. We similarly noted correspondences between prior findings for blood, spleen, and liver from deermice and mice and the findings with cultivated dermal fibroblasts of those species.

One of these correspondences was Slpi, which encodes an inhibitor of proteases of neutrophils and other phagocytic cells. Among its functions, the secreted protein has a role in wound repair (54) and has an inhibitory effect on the formation of neutrophil extracellular traps (55). Of relevance for a study of an infection-tolerant reservoir for the Lyme disease agent was the report that Slpi-deficient mice infected with *B. burgdorferi* had more severe arthritis and greater bacterial burdens in their joints than their wild-type counterparts (56). In prior studies we observed that expression of Slpi increased orders of magnitude in the blood of *P. leucopus* treated with LPS or infected with a relapsing fever *Borrelia* species (16). For LPS-treated mice there was only a modest increase in Slpi transcription from a low level. These observations notwithstanding, it was not a given that dermal fibroblasts of *P. leucopus* would transcribe Slpi and produce the protein, but that was the finding here. Transcription was high at baseline in the deermouse cultures and increased further with TLR2 agonist exposures (Figure 4). In contrast, with the mouse cells there was the same comparatively low expression of Slpi as observed in the in vivo experiment. While there remain questions about regulation of Slpi expression, the tight correlation of Slpi transcription across species with that of the stress responsive transcription factor Nfe2l2 (Figure 4), also known as NRF2, indicates a direct regulatory role, as has been reported for the mouse (57). Nfe2l2 regulates the transcription of a variety cytoprotective proteins, including antioxidants, such as heme oxygenase 1, and chaperones (58). Figure 4 reveals a high constitutive level of transcription of Nfe2l2 by the *P. leucopus* fibroblasts compared to the control *M. musculus* cells. This is also a characteristic of the long-lived naked mole rat, an established experimental model for aging research (59). A possible explanation of the successful serial propagation of the deermouse cells, but not the mouse cells, in presence atmospheric concentrations of oxygen was a greater capacity to withstand oxidative stress.

Another distinguishing feature of *P. leucopus* cells of the blood and spleen was the near absence of transcription of Nos2 (16, 17), which encodes the mitochondrion-based, inducible nitric oxide synthase. This at the same time that Arg1 is transcribed at high levels at baseline and even higher with LPS exposure. As noted previously for *P. leucopus* fibroblasts (16), Nos2 is transcribed, both constitutively and at higher levels after exposure to an agonist of TLR4 or TLR2 (Figure 3). But this was starting at a level many-fold lower than for the *M. musculus* fibroblasts at baseline. Arg1 expression was hardly detectable in the mouse fibroblasts and substantially higher in deermouse cells under both control and treated conditions, similarly to what was observed in blood and spleen. What the in vitro study added to the picture of arginine metabolism in the two species was the differential expression of two other arginine metabolism genes associated with the mitochondrion, Arg2 and Ass1, and an arginine transporter, Slc7a2 (39), which, unlike Arg2 and Ass1, was transcribed at comparatively high levels both in controls and agonist-treated deermouse cells (Figure 3). These findings draw attention to the mitochondria and the dichotomy between *P. leucopus* and *M. musculus* in the sites of activity for arginine metabolism genes of recognized importance in immunity (60, 61).

As in the study of whole blood of *P. leucopus* and *M. musculus*, there were differences between species in transcription of ERV and TE sequences in cultured fibroblasts. In the study of LPS-treated animals this was restricted to selected retroviral envelope protein sequences, which decreased in transcription in the deermice while increasing in the mice (17). In the present study much larger sets of the ERV sequences of small to large size were used as references. These datasets were comparable between species in numbers and in the overall distributions of lengths and over the wide ranges of levels of transcription. Given the high redundancy of ERVs of all sizes in the mammalian genomes, confidently ascribing a location for the source of a ~150 nt cDNA read is a challenge, and this limitation, which applies to all studies of this sort, is acknowledged. That is why the emphasis was on the lengths of the ERVs as a class and not their precise identification. With that disclaimer, we found that the more highly expressed ERV sequences at baseline were generally longer in *M. musculus* than in *P. leucopus* (Figure S6) and that this was also true for those sequences that were upregulated DEGs for one species or another (Figure 7).

Deermice have recently experienced massive invasions of diverse ERVs, resulting in a large expansion of KRAB zinc finger proteins (KRAB-ZFPs) responsible for ERV suppression (24, 62). This expansion of suppression machinery could enable deer mice to control ERV expression more effectively, with important implications for inflammation, immune response, and longevity (63-65). Consistent with this, previous work has shown evidence that P. leucopus may control ERV expression more effectively compared to other rodents (17). Our data here also generally support this hypothesis. However, we also observe expression of full-length ERVs in P. leucopus encoding all or most of the proteins of a leukemia-type retrovirus. Interestingly, this particular ERV recently arose in the deer mouse germline via interspecies

horizontal gene transfer, exists at low copy numbers, and encodes a full-length env genes that is expressed, suggesting that it might still function as an infectious virus (24). Such ERVs often evade host suppression machinery so this ERV's expression might not be too surprising. Nonetheless, these observations lead to important questions about the dynamics of ERV suppression in deer mice, and how more effective ERV suppression might impact immune response to other pathogens.

A more manageable number of ERV sequences to consider were the upregulated DEGs unique to transformed *P. leucopus* cells. Among the P47 DEGs were three ERVs of over 9 kb and that would encode Env proteins as well as Gag-pol polyproteins. Activation of RNA tumor viruses in spontaneously transformed mouse cell lines with or without inciting factors has long been noted (28, 66). *P. leucopus* provides another example of this, but this phenomenon in this species remains incompletely characterized. A future step would be documentation of expression of viral proteins by these longer ERVs, such as with specific antibodies, mass spectrometry, or assay for reverse transcriptase activity. While the P47 fibroblasts behaved as if immortal, including increased expression of the oncogene Myc (Table 2), other distinguishing features of transformed cells, such as their karyotypes, have not been examined.

In distinction to these parallels between the vitro and in vivo results, an in vitro finding lacking a counterpart in the animal studies was a difference between species in transcription of the cytokine interleukin-11. In the blood, spleen, or liver of mice and deermice there was scant to no transcription of Il11 under either baseline or agonist exposure conditions (16, 17). Did the low transcription of Il11 in primary cultures of deermouse fibroblasts fairly represent what occurs in other tissues or other pathologic conditions? Re-examination for the present study of results of RNA-seq of the skin of *P. leucopus* with or without infection with *B. burgdorferi* found no detectable transcription of Il11 (6). Whether there is an association of IL-11 with aging in *Peromyscus*, as has been reported for other species (42, 43), remains to be determined.

What we can consider here is the possible source of Il11 transcription in the fibroblasts. This was evident for the mouse fibroblasts: the right-sided cluster 12 (Figure 7). Leukemia inhibitory factor (Lif), another member of the IL-6-type cytokine family, and interleukin 1 receptor antagonist (Il1rn) also were markers of cluster 12 (Table S5). Il11, Il1rn, and Lif were more highly transcribed in P47 than in P4 deermouse cells (Table 2 and Table S6) and, unlike some chemokines, like Ccl2 and Cxcl10, unchanged in transcription in P47 cells exposed to Pam3CSK4 or LPS (Table S7). This suggests that with the spontaneous transformation and possible de-differentiation there was release of repression of these genes, allowing for constitutive expression. The pro-inflammatory cytokine IL-6 also had right-side associations, specifically cluster 4 of *P. leucopus* (Figure 7). But, unlike the other IL-6-type family members, Il6 transcription was much decreased in the transformed cells (Table 2), an indication of a different form of regulation.

As the distinguishing features between *P. leucopus* and *M. musculus* mount and the list lengthens, a relationship between the infection-tolerant phenotype and the greater longevity of this deermouse begins to emerge. We knew before of the low susceptibility of *P. leucopus* to what would be lethal doses of LPS for a mouse, the moderated expression of interferons and downstream genes, and a polarization profile of alternatively activated macrophages tilting toward an anti-inflammatory outcome. The present study adds these distinctions: divergent natural histories of cultured fibroblasts, constitutive expression of the stressor-response transcription factor Nfe2l2/NRF2, diminished expression of the aging-associated cytokine IL-11, and an indisposition to transcribe long ERVs with their potential to elicit recognition by PRRs and the inflammatory consequences of that.

4 Materials and Methods

Animals. The study was carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, Guide for the Care and Use of Laboratory Animals: Eighth Edition of the National Academy of Sciences, and according to ARRIVE Guidelines (https://arriveguidelines.org). University of California Irvine (UCI) protocols AUP-23-127 and AUP-24–042 were approved by its Institutional Animal Care and Use Committee.

Peromyscus leucopus, here also referred to as "deermice", were of the outbred LL stock that originated with 38 animals captured near Linville, NC, and thereafter comprised a closed colony without sib-sib matings at the Peromyscus Genetic Stock Center at the University of South Carolina (67). Deermice used in this study were each from different mating pairs and bred at the UCI vivarium. Outbred *Mus musculus* CD-1 stock and inbred *Mus musculus* FVB/NJ strain, here also referred to as "mice", were obtained from Charles River Laboratories and Jackson Laboratory, respectively. Rodents were maintained in the AAALAC-accredited UCI vivarium, with 2 to 5 animals per cage, a 16-h light/8-h dark schedule, room temperature of 21-23°C, and humidity of 30-70%. Food and water were available ad libitum; the diet was 2020X Soy Protein-Free Extruded Rodent Diets (Teklad). The rodents were euthanized by carbon dioxide overdose followed by exsanguination by open thorax cardiac puncture.

Tissue culture medium. The medium was RPMI 1640 medium (Gibco) supplemented with 10% heat-activated fetal bovine serum (Gibco), 2 mM L-glutamine, 100 μM L-asparagine, 50 μM 2-mercaptoethanol, 500U/ml penicillin G, and 500 μg/ml streptomycin (68).

Processing of tissue. Immediately after euthanasia, ~1 cm diameter of full-thickness ear tissue was excised with sterile instruments, immersed in 70% ethanol for 5 min, and allowed to dry in a biosafety cabinet. Once dry, the tissue transferred to a 1.8 ml polystyrene screwcap tube containing 1.5 ml medium with 2 mg/ml collagenase type I (Millipore Sigma). Ear tissue was minced while in the tube with scissors down to ~3 mm size, and then the suspension was incubated at 37°C on a horizontal shaker at 200 rpm for 75 min. The suspension was then pushed through a 70 μm cell strainer with the rubber end of a 5 ml syringe plunger into 10 ml of medium. The pass-through liquid was centrifuged at 400 x g at room temperature for 5 min. The supernatant was removed, the pellet was suspended in 10 ml medium, the suspension centrifuged again, and the supernatant was removed.

Tissue cultivation. After the final centrifugation, the cell pellet was suspended in 10 ml of pre-warmed medium. This was transferred to a 10 cm diameter polystyrene culture dish (Fisher Scientific FB0875712) with lid and incubated at 37°C in a humidified atmosphere with 5% CO_2. After 24 hr the medium was removed by aspiration, adherent cells were washed with pH 7.4 phosphate-buffered saline (PBS) once, and then fresh medium was added at that point and then every 3-4 d thereafter. The cultures were monitored by phase microscopy at 100x magnification with Olympus CK2 inverted phase contrast microscope. For sub-cultures medium was aspirated, the cells were washed once with PBS, and then to release the cells 2 ml of a solution of 0.02 mM trypsin and 0.48 mM EDTA (trypsin-EDTA) was added before a 5 min incubation. The resultant suspension of cells were transferred to a 15 ml polystyrene conical tube containing 8 ml of medium and centrifuged at 400 x g for 5 minutes. The supernatant was removed, and the cell pellet was resuspended in 10 ml of supplemented medium. A 1:1 ratio of 0.4% w/v trypan blue in PBS and cell suspension was used for manual counting of cells with a Petroff-Hausser counting chamber under an inverted microscope at 200x magnification. Once counted, 100,000 cells were seeded into a new 10 cm petri dish with 10 ml of pre-warmed medium. The remaining cell suspension was centrifuged at 400 x g for 5 min. The cell pellet was resuspended in 1 ml of medium with 10% cell culture grade dimethyl sulfoxide and then transferred to 2 ml screwcap cryogenic vials (Greiner) and stored in liquid nitrogen. After adherent cells of subcultures became 80-90% confluent, they were treated with trypsin-EDTA and counted. Population doublings were calculated with the initial 100,000 seeding count and the final count of confluency.

Photomicrographs. Digital images were taken on an Olympus CK2 inverted phase contrast microscope equipped with a LabCam Ultra (iDu Optics) adapter for an iPhone 14 Pro (Apple Computer) and with the 15x magnification eyepiece, 10x magnification objective lens, and 2x digital zoom. The final magnification was 150x for the 4032 x 3024 resolution High Efficiency Image File Format (HEIF) format files.

Chemicals. The lipopeptide Pam3CysSerLys4 (Pam3CSK4; Invivogen) and ion-exchange chromatography-purified Escherichia coli O111:B4 lipopolysaccharide (LPS) (Sigma-Aldrich) were dissolved as stock solutions at 1 mg/ml concentration in endotoxin-free sterile 0.9% (w/v) NaCl (saline). Stock solutions were stored at -20°C until the time of use when dilutions were made in saline. For experiments a TLR agonist the Pam3CSK4 was used at either 1 μg/ml or 10 μg/ml final concentration and LPS was used at 1 μg/ml final concentration.

Cell culture treatments. Cell cultures at the point of 80-90% confluency were treated with trypsin-EDTA, counted using a hemocytometer, and split into 3 wells each in a 6-well polystyrene tissue culture plate (Fisher Scientific FB0875712) at 3×10^5 cells per well containing 2 ml of medium. For the study of single cells there were 5×10^5 cells per well. Cells were incubated at 37°C in 5% CO2 for 24 h prior to treatment before either saline alone or Pam3CSK4 or LPS diluted in saline was added for final concentrations of 1 μg/ml or 10 μg/ml for Pam3CSK4 or 1 μg/ml for LPS. The suspensions were incubated at 37°C in 5% CO2 for 4 h.

RNA extraction. After incubation with TLR agonist or medium alone, medium was aspirated, and 300 μl of DNA/RNA Shield (Zymo Research) was added to adherent cells followed by addition of 300 μl of Zymo RNA Lysis Buffer (Zymo Research). The Quick-RNA Miniprep Plus Kit (Zymo Research) was used for the isolation of skin fibroblast RNA following manufacturer's instructions, including DNase treatment. RNA was eluted from the spin-column in 50 μl of Zymo Nuclease Free Water. Quantification, purity, and RNA integrity were assessed using High Sensitivity Qubit 2.0 fluorometer, Nanodrop ND-1000 spectrophotometer, and Agilent Bioanalyzer 2100. RNA integrity number (RIN) values for samples were >9.0.

Bulk RNA-seq. cDNA libraries were produced with the Illumina TruSeq Stranded mRNA kit. Library size normalization was done by the TMM (trimmed mean of M values) method (69). Multiplexed libraries were sequenced at UCI's Genomics Research and Technology Hub on an Illumina NovaSeq 6000 with paired-end chemistry, 150 cycles, and ~120-250 million reads per sample. Fastq files of reads were trimmed of adapter linkers and homopolymeric 5′ or 3′ ends. Trimmed reads were aligned with a length fraction of 0.35, a similarity fraction of 0.9, and a cost of 3 for a mismatch, insertion, or deletion to the nonredundant coding sequences (CDS) based on the transcripts of the reference

genomes for *Mus musculus* strain C57/BL6 (GCF_000001635.27_GRCm39) and *Peromyscus leucopus* LL stock (GCF_004664715.2). For the *M. musculus* transcripts the first listed isoform was used for the CDS (70). For the *P. leucopus* CDS additional manual annotation of loci without assigned gene names was carried out (71).

For analyses within a species and by condition but without pairing by individual cell line, transcripts per million (TPM), which are normalized RNA-seq values accounting for gene length and total reads for the sample, were used. For fold-change (FC) and false discovery rate (FDR) analyses of results paired by cell line, unique counts were first normalized by total reads for the corresponding sample across the species and then adjusted as reads per kilobase. For cross-species comparisons the ratio of length-adjusted, total count-normalized reads for a given gene to the house keeping gene Gapdh was used. The mean $\log_{10}$ Gapdh unique counts (standard deviation) were 4.8 (0.05) across 15 *P. leucopus* samples (i.e. 5 cell lines x 3 conditions) and 5.1 (0.05) across 15 *M musculus* samples. There was little or no discernible effect of the treatments on transcription of Gapdh. For unique reads of Gapdh the mean paired fold-changes (paired *t*-test *p* value) of Pam3CSK4 treated cells to control were 1.08 (0.16), 1.12 (0.08), 1.00 (0.90), and 1.13 (0.07) for mouse 1 µg/ml, mouse 10 µg/ml, deermouse 1 µg/ml, and deermouse 10 µg/ml sets, respectively (Dryad Table D3). Following the recommendation of Hedges et al. we used the natural logarithm (LN) of ratios (72).

Gene ontology (GO) term analysis. The analysis was implemented with the tools of Metascape (https://metascape.org) (73). re used genome-wide GO-term analysis using default settings and the selection of *Mus musculus* as the closest taxon for comparison. Similarity matrices were hierarchically clustered, and a 0.3 similarity threshold was applied to trim resultant trees into separate clusters. The lowest p-value term represented each cluster shown in the heatmaps. Besides the terms beginning with 'GO' of the Gene Ontology resource (http://geneontology.org) (74), other terms refer to Kegg Pathway database (https://www.kegg.jp) for 'mmu.' designations, WikiPathways database (https://www.wikipathways.org) for 'WP…' designations, and Reactome database (https://reactome.org) for 'R-MMU…' designations.

Transposable element sequences. We annotated transposable elements in both the *Mus musculus* C57BL/6 and *Peromyscus leucopus* genomes using a combined nonredundant database of lineage-specific and ancestral TEs in each species. To do this, we first retrieved curated TE models for *P. maniculatus* and *Mus musculus* from DFAM (https://doi.org/10.1186/s13100-020-00230-y), as well as all ancestral TE models for each species. Then, we clustered redundant models using CD-HIT-EST (version 4.8.1) (https://doi.org/10.1093/bioinformatics/bts565) with parameters -n 10 -c .8 -r 1 -i. We employed RepeatMasker (version 4.1.2) (https://www.repeatmasker.org) with parameters -pa 12 -excln -s -no_is -u -noisy -html -xm -a -xsmall to annotate TEs in each genome using our combined TE library as input. The names, genome locations, and lengths for identified ERVs and TEs are provided in Dryad Tables D5 for *M. musculus* and D6 for *P. leucopus*.

Single-cell RNA-seq. After incubation of 5 x 10^5 low passage (P4) P. leucopus cells or low passage M. musculus cells with medium and saline alone, medium with 10 µg/ml Pam3CSK4, or medium with 1 µg/ml LPS, the medium was aspirated from the wells. Adherent cells were washed with 2 ml of 1x PBS. A 0.5 ml volume of trypsin-EDTA solution was added to each well, and the cells were incubated at 37°C in 5% CO_2 for 5 min Detached cells from each condition were pooled into 15 ml polystyrene conical tubes by species, and 10 ml of medium was added to neutralize trypsin activity. The pooled cell suspensions for *P. leucopus* and *M. musculus* were centrifuged at 400 x g for 5 minutes and supernatant was removed. Cells of pellets were resuspended in Illumina Cell Suspension Buffer of the Single Cell 3' RNA Prep Kit (Illumina). Cell viability and counts were determined in triplicate using trypan blue stain and a hemocytometer counting chamber as described. Cells were diluted in Cell Suspension Buffer to a concentration of 5000 per µl concentration and a final total count was determined. Approximately 40,000 cells for each species were further processed using an Illumina Single Cell 3' RNA Prep Kit for mRNA capture, cDNA synthesis, and library preparation, according to manufacturer's instructions. cDNA quality control and fragment analysis were performed using a DNA High Sensitivity chip on an Agilent BioAnalyzer 2100. A post-library preparation quality control, fragment analysis, and KAPA library quantification (Roche) by quantitative polymerase chain reaction were performed. cDNA libraries were sequenced on an Illumina NovaSeq X Plus system with paired-end reads, 150 cycles, and incorporation of a 2% PhiX spike-in. The yields were 1.6 x 10^9 and 1.8 x 10^9 reads for *Mus musculus* and *Peromyscus leucopus*, respectively.

Single-cell RNA-seq analysis. The raw FASTQ data were processed using the DRAGEN v 4.4.2 software on the BaseSpace Sequencing Hub (Illumina). The reference genomes used were default GCF_000001635.20 genome for *M. musculus* and an imported genome (GCF_004664715.2) for *P. leucopus*. At the data analysis stage for the *P. leucopus* reference set, updated annotations with formal gene names (71) substituted for loci with incomplete descriptions in the original annotation. The DRAGEN output included filtered matrix, features, and barcodes files, which were subsequently analyzed in *R* (75) using Seurat v5.2.1 (https://satijalab.org/seurat) (76). Seurat facilitated Uniform Manifold Approximation and Projection (UMAP) clustering via principal component analysis (PCA), visualization of gene expression on UMAPs through feature plots, cluster modifications, and identification of the enriched genes for each

cluster. For the UMAP, principal components were used with the Seurat FindNeighbors setting from 1:10 and the FindClusters resolution set to 0.5. For the single-cell RNA-seq data the criteria for GO term analysis of signifying genes of clusters were an enrichment of ≥1.5 and p value $<10^{-4}$ (Table S5).

MALDI-TOF mass spectrometry. After aspiration of medium, adherent P. leucopus dermal fibroblasts of low or high passage grown on 10 cm diameter polystyrene petri dishes were washed with 5 ml PBS. Adherent cells were lysed by the addition of 1 ml of Radioimmunoprecipitation Assay buffer (ThermoFisher) with protease inhibitor. The petri dish bottom was scraped with a cell scraper and homogenized. Cell lysate suspension was transferred to a sterile 1.5 ml polystyrene screwcap tube, and sonication was performed with alternating cycles of 15 s of sonication and 15 s on ice over 1 min. Lysates were then centrifuged at 21,000 x g for 10 min at room temperature, and the supernatant was decanted. The lysate was subjected in sodium dodecyl sulfate polyacrylamide gel electrophoresis on a 4-12% gradient Bis-Tris denaturing polyacrylamide gel (Invitrogen) on a NuPAGE apparatus. electrophoresis was performed, and Coomassie blue-stained bands of the desired size range were excised. In-gel digestion with trypsin was carried out as described by Shevchenko et al. (77). In brief, destaining was performed with a 1:1 acetonitrile and ammonium bicarbonate, followed by reduction with dithiothreitol, and alkalization with iodoacetamide. An overnight trypsin digestion at 37°C was carried out, and this was followed by extraction of the mixture of peptides, spotting them onto a plate reader, and air drying. Cell analytes were measured using matrix-assisted laser desorption/ionization (MALDI) on a Bruker UltrafleXtreme instrument at the UCI Department of Chemistry Mass Spectrometry Facility. The ribosomal protein S18 (Rps18; XP_028744221.1) amino acid sequence served as a control for the targeted analysis of secretory leukocyte peptidase inhibitor (Slpi) (XP_028724460.1). The reference for Slpi was the processed protein after cleavage of signal peptide. Spectra were acquired using a mass range from 800-6000 Daltons. FlexControl software version 4.0 was used for data collection and raw spectral data were imported into mMASS software version 5.5. Baseline correction and peak picking were performed using mMASS default settings.

Statistics. Means are presented with 95% confidence intervals (95CI) in parentheses. These were with asymmetrical confidence intervals for data that was not normally distributed. Parametric (t test) and non-parametric (Kruskal-Wallis) tests of significance were 2-tailed. Unless otherwise stated, it the t-test p-value that is given. Categorical variables were assessed by 2-tailed Fisher's exact test. False discovery rate (FDR) correction of p values for multiple testing was by the Benjamini-Hochberg method (78). Linear regression, coefficient of determination (R^2), Multidimensional Scaling (MDS), and General Linear Model (GLM) analyses were performed with SYSTAT v. 13.1 software (Systat Software, Inc). Box plots with whiskers display the minimum, first quartile, median, third quartile, and maximum.

Data resources. Table S1 includes links to BioProject and BioSample numbers for samples of the *P. leucopus* and *M. musculus* comparison, along with accession numbers for raw reads (SRA) and RNA-seq analysis (GEO). The low and high passage *P. leucopus* dermal fibroplast study is registered under BioProject PRJNA1169383 with BioSamples SAMN45236449-54, SRA accessions SRR31668268-73, and GEO accessions GSM9100082-87. The single cell RNA-seq study is registered under BioProject PRJNA1277016 with BioSamples SAMN49093172 for *P. leucopus* and SAMN49093173 for *M. musculus* dermal fibroblasts. The corresponding SRA accessions were SRR33982422 and SRR33982421, respectively. For this study the 5′ UTR and coding sequence for the interleukin-11 gene Il11 of *P. leucopus* LL stock was determined by sequencing of the mRNA using the RNA-seq reads. This was deposited at GenBank under accession PV639634; it corrects an error in the predicted protein coding sequence in the NCBI automated annotation for accession XM_037199381.1. Tables of large datasets deposited with the Dryad (https://datadryad.org) public repository and denoted in the text by the convention "Dryad Table Dx" are accessible at https://doi.org/10.5061/dryad.m905qfvdq.

Acknowledgements

The research was supported by National Institutes of Health grant AI-157513 (A.G.B. and A.D.L.). This work utilized resources of the UCI Genomics Research and Technology Hub (GRT Hub), parts of which are supported by NIH grants to the Comprehensive Cancer Center (P30CA-062203) and to the UCI Skin Biology Resource Based Center (P30AR075047), as well as to the GRT Hub for instrumentation (1S10OD010794-01 and 1S10OD021718-01). Some computations for this work were run on the Cannon cluster supported by the Faculty of Arts and Sciences Division of Science Research Computing Group at Harvard University.

References

1. Hagai T, Chen X, Miragaia RJ, Rostom R, Gomes T, Kunowska N, et al. Gene expression variability across cells and species shapes innate immunity. Nature. 2018;563(7730):197-202. doi: 10.1038/s41586-018-0657-2.

2. Zhao Y, Seluanov A, Gorbunova V. Revelations About Aging and Disease from Unconventional Vertebrate Model Organisms. Ann Rev Genet. 2021;55:135-59. doi: 10.1146/annurev-genet-071719-021009

3. Bedford NL, Hoekstra HE. Peromyscus mice as a model for studying natural variation. eLife. 2015;4:e06813. doi: 10.7554/eLife.06813

4. Havighorst A, Crossland J, Kiaris H. Peromyscus as a model of human disease. Semin Cell Devel Biol. 2017;61:150-5. doi: 10.1016/j.semcdb.2016.06.020

5. Dewey MJ, Dawson WD. Deer mice: "The Drosophila of North American mammalogy". Genesis. 2001;29(3):105-9. doi: 10.1002/gene.1011

6. Long A, Baldwin-Brown J, Tao Y, Cook V, Balderrama-Gutierrez G, Corbett-Detig R, et al. The genome of Peromyscus leucopus, natural host for Lyme disease and other emerging infections. Sci Advances. 2019;5(7):eaaw6441. doi: 10.1126/sciadv.aaw6441

7. Long PN, Cook VJ, Majumder A, Barbour AG, Long AD. The utility of a closed breeding colony of Peromyscus leucopus for dissecting complex traits. Genetics. 2022;221(1). doi: 10.1093/genetics/iyac026

8. Barbour AG. Infection resistance and tolerance in Peromyscus spp., natural reservoirs of microbes that are virulent for humans. Semin Cell Develop Biol. 2017;61(1):115-22. DOI: 10.1016/j.semcdb.2016.07.002

9. Medzhitov R, Schneider DS, Soares MP. Disease tolerance as a defense strategy. Science. 2012;335(6071):936-41. doi: 10.1126/science.1214935

10. Ayres JS, Schneider DS. Tolerance of infections. Ann Rev Immunol. 2012;30:271-94. doi: 10.1146/annurev-immunol-020711-075030

11. Mandl JN, Schneider C, Schneider DS, Baker ML. Going to bat(s) for studies of disease tolerance. Front Immunol. 2018;9:2112. doi: 10.3389/fimmu.2018.02112

12. Sacher GA, Hart RW. Longevity, aging and comparative cellular and molecular biology of the house mouse, Mus musculus, and the white-footed mouse, Peromyscus leucopus. Birth Defects Orig Artic Ser. 1978;14(1):71-96.

13. Cohen BJ, Cutler RG, Roth GS. Accelerated wound repair in old deer mice (Peromyscus maniculatus) and white-footed mice (Peromyscus leucopus). J Gerontol. 1987;42(3):302-7. doi: 10.1093/geronj/42.3.302

14. Pei G, Balkema-Buschmann A, Dorhoi A. Disease tolerance as immune defense strategy in bats: One size fits all? PLoS Pathogens. 2024;20(9):e1012471. doi: 10.1371/journal.ppat.1012471

15. Gorbunova V, Seluanov A, Kennedy BK. The World Goes Bats: Living Longer and Tolerating Viruses. Cell Metabol. 2020;32(1):31-43. doi: 10.1016/j.cmet.2020.06.013

16. Balderrama-Gutierrez G, Milovic A, Cook VJ, Islam MN, Zhang Y, Kiaris H, et al. An infection-tolerant mammalian reservoir for several zoonotic agents broadly counters the inflammatory effects of endotoxin. mBio. 2021;12(2). doi: 10.1128/mBio.00588-21. PMC8092257

17. Milovic A, Duong JV, Barbour AG. The infection-tolerant white-footed deermouse tempers interferon responses to endotoxin in comparison to the mouse and rat. eLife. 2024;12. doi: 10.7554/eLife.90135. PMC10945503

18. Machtinger ET, Williams SC. Practical guide to trapping Peromyscus leucopus (Rodentia: Cricetidae) and Peromyscus maniculatus for vector and vector-borne pathogen surveillance and ecology. J Insect Sci. 2020;20(6). doi: 10.1093/jisesa/ieaa028

19. Downie AA-OX, Barre RA-O, Robinson A, Yang J, Chen YA-O, Lin JA-O, et al. Assessing immune phenotypes using simple proxy measures: promise and limitations. Discov Immunol 2024 Jun 28;3(1):kyae010. doi: 10.1093/discim/kyae010

20. Zhong X, Lundberg M, Råberg L. Comparison of spleen transcriptomes of two wild rodent species reveals differences in the immune response against Borrelia afzelii. Ecol Evol. 2020;10(13):6421-34. doi: 10.1002/ece3.6377

21. Sinsky RJ, Piesman J. Ear punch biopsy method for detection and isolation of Borrelia burgdorferi from rodents. J Clin Microbiol. 1989;27(8):1723-7. doi: 10.1128/jcm.27.8.1723-1727.1989

22. Zawada SG, von Fricken ME, Weppelmann TA, Sikaroodi M, Gillevet PM. Optimization of tissue sampling for Borrelia burgdorferi in white-footed mice (Peromyscus leucopus). PLoS One. 2020;15(1):e0226798. doi: 10.1371/journal.pone.0226798

23. Oosting M, Ter Hofstede H, Sturm P, Adema GJ, Kullberg BJ, van der Meer JW, et al. TLR1/TLR2 heterodimers play an important role in the recognition of Borrelia spirochetes. PLoS One. 2011;6(10):e25998. doi: 10.1371/journal.pone.0025998

24. Gozashti L, Feschotte C, Hoekstra HE. Transposable element interactions shape the ecology of the deer mouse genome. Mol Biol Evol. 2023;40(4):msad069. doi: 10.1093/molbev/msad069

25. Hayflick L, Plotkin SA, Norton TW, Koprowski H. Preparation of poliovirus vaccines in a human fetal diploid cell strain. Am J Hyg. 1962;75:240-58. doi: 10.1093/oxfordjournals.aje.a120247

26. Parrinello S, Samper E, Krtolica A, Goldstein J, Melov S, Campisi J. Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. Nat Cell Biol. 2003;5(8):741-7. doi: 10.1038/ncb1024

27. Todaro GJ, Green H. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol. 1963;17(2):299-313. doi: 10.1083/jcb.17.2.299

28. Aaronson SA, Hartley JW, Todaro GJ. Mouse leukemia virus: "spontaneous" release by mouse embryo cells after long-term in vitro cultivation. Proc Natl Acad Sci U S A. 1969;64(1):87-94. doi: 10.1073/pnas.64.1.87

29. Szeto A, Sun-Suslow N, Mendez AJ, Hernandez RI, Wagner KV, McCabe PM. Regulation of the macrophage oxytocin receptor in response to inflammation. Am J Physiol Endocrinol Metab. 2017;312(3):E183-e9. doi: 10.1152/ajpendo.00346.2016

30. Wirtz S, Becker C, Fantini MC, Nieuwenhuis EE, Tubbe I, Galle PR, et al. EBV-induced gene 3 transcription is induced by TLR signaling in primary dendritic cells via NF-kappa B activation. J Immunol. 2005;174(5):2814-24. doi: 10.4049/jimmunol.174.5.2814

31. Hagiwara K, Kikuchi T, Endo Y, Huqun, Usui K, Takahashi M, et al. Mouse SWAM1 and SWAM2 are antibacterial proteins composed of a single whey acidic protein motif. J Immunol. 2003;170(4):1973-9. doi: 10.4049/jimmunol.170.4.1973

32. Perez CJ, Dumas A, Vallières L, Guénet JL, Benavides F. Several classical mouse inbred strains, including DBA/2, NOD/Lt, FVB/N, and SJL/J, carry a putative loss-of-function allele of Gpr84. J Hered 2013;104(4):565-71. doi: 10.1093/jhered/est023

33. Foster AM, Baliwag J, Chen CS, Guzman AM, Stoll SW, Gudjonsson JE, et al. IL-36 promotes myeloid cell infiltration, activation, and inflammatory activity in skin. J Immunol. 2014;192(12):6053-61. doi: 10.4049/jimmunol.1301481

34. Liu C, Xu Z, Gupta D, Dziarski R. Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules. J Biol Chem. 2001;276(37):34686-94. doi: 10.1074/jbc.M105566200

35. Viola A, Munari F, Sánchez-Rodríguez R, Scolaro T, Castegna A. The metabolic signature of macrophage responses. Front Immunol. 2019;10:1462. doi: 10.3389/fimmu.2019.01462

36. Murray PJ. Macrophage polarization. Ann Rev Physiol. 2017;79:541-66. doi: 10.1146/annurev-physiol-022516-034339

37. Dowling JK, Afzal R, Gearing LJ, Cervantes-Silva MP, Annett S, Davis GM, et al. Mitochondrial arginase-2 is essential for IL-10 metabolic reprogramming of inflammatory macrophages. Nat Commun. 2021;12(1):1460. doi: 10.1038/s41467-021-21617-2

38. Qualls JE, Subramanian C, Rafi W, Smith AM, Balouzian L, DeFreitas AA, et al. Sustained generation of nitric oxide and control of mycobacterial infection requires argininosuccinate synthase 1. Cell Host Microbe. 2012;12(3):313-23. doi: 10.1016/j.chom.2012.07.012

39. Thompson RW, Pesce JT, Ramalingam T, Wilson MS, White S, Cheever AW, et al. Cationic amino acid transporter-2 regulates immunity by modulating arginase activity. PLoS Pathogens. 2008;4(3):e1000023. doi: 10.1371/journal.ppat.1000023

40. O'Rourke SA, Shanley LC, Dunne A. The Nrf2-HO-1 system and inflammaging. Front Immunol. 2024;15:1457010. doi: 10.3389/fimmu.2024.1457010

41. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987;15(20):8125-48. doi: 10.1093/nar/15.20.8125

42. Cook SA, Schafer S. Hiding in plain sight: interleukin-11 emerges as a master regulator of fibrosis, tissue integrity, and stromal inflammation. Ann Rev Med. 2020;71(1):263-76. doi: 10.1146/annurev-med-041818-011649

43. Widjaja AA, Lim WW, Viswanathan S, Chothani S, Corden B, Dasan CM, et al. Inhibition of IL-11 signalling extends mammalian healthspan and lifespan. Nature. 2024;632(8023):157-65. doi: 10.1038/s41586-024-07701-9

44. Lendahl U, Muhl L, Betsholtz C. Identification, discrimination and heterogeneity of fibroblasts. Nat Commun. 2022;13(1):3409. doi: 10.1038/s41467-022-30633-9

45. Sheppard CH, Kazacos KR. Susceptibility of Peromyscus leucopus and Mus musculus to infection with Baylisascaris procyonis. J Parasitol. 1997;83(6):1104-11.

46. Bourgeois JS, McCarthy JE, Turk SP, Bernard Q, Clendenen LH, Wormser GP, et al. Peromyscus leucopus , Mus musculus , and humans have distinct transcriptomic responses to larval Ixodes scapularis bites. Infect Immun. 2025 Apr 8;93(4):e0006525. doi: 10.1128/iai.00065-25

47. Kacprzyk J, Locatelli AG, Hughes GM, Huang Z, Clarke M, Gorbunova V, et al. Evolution of mammalian longevity: age-related increase in autophagy in bats compared to other mammals. Aging (Albany NY). 2021;13(6):7998-8025. doi: 10.18632/aging.202852

48. Alcock D, Power S, Hogg B, Sacchi C, Kacprzyk J, McLoughlin S, et al. Generating bat primary and immortalised cell-lines from wing biopsies. Sci Rep. 2024;14(1):27633. doi: 10.1038/s41598-024-76790-3

49. Seluanov A, Hine C, Bozzella M, Hall A, Sasahara TH, Ribeiro AA, et al. Distinct tumor suppressor mechanisms evolve in rodent species that differ in size and lifespan. Aging Cell. 2008;7(6):813-23. doi: 10.1111/j.1474-9726.2008.00431.x

50. Driskell RR, Watt FM. Understanding fibroblast heterogeneity in the skin. Trends Cell Biol. 2015;25(2):92-9. doi: 10.1016/j.tcb.2014.10.001

51. Philippeos C, Telerman SB, Oulès B, Pisco AO, Shaw TJ, Elgueta R, et al. Spatial and Single-Cell Transcriptional Profiling Identifies Functionally Distinct Human Dermal Fibroblast Subpopulations. J Invest Dermatol. 2018;138(4):811-25. doi: 10.1016/j.jid.2018.01.016

52. Tyshkovskiy A, Ma S, Shindyapina AV, Tikhonov S, Lee SG, Bozaykut P, et al. Distinct longevity mechanisms across and within species and their association with aging. Cell. 2023;186(13):2929-49.e20. doi: 10.1016/j.cell.2023.05.002

53. Ma S, Upneja A, Galecki A, Tsai YM, Burant CF, Raskind S, et al. Cell culture-based profiling across mammals reveals DNA repair and metabolism as determinants of species longevity. eLife. 2016;5. doi: 10.7554/eLife.19130

54. Ashcroft GS, Lei K, Jin W, Longenecker G, Kulkarni AB, Greenwell-Wild T, et al. Secretory leukocyte protease inhibitor mediates non-redundant functions necessary for normal wound healing. Nat Med. 2000;6(10):1147-53. doi: 10.1038/80489

55. Zabieglo K, Majewski P, Majchrzak-Gorecka M, Wlodarczyk A, Grygier B, Zegar A, et al. The inhibitory effect of secretory leukocyte protease inhibitor (SLPI) on formation of neutrophil extracellular traps. J Leukocyte Biol. 2015;98(1):99-106. doi: 10.1189/jlb.4AB1114-543R

56. Yu Q, Tang X, Hart T, Homer R, Belperron AA, Bockenstedt LK, et al. Secretory leukocyte protease inhibitor influences periarticular joint inflammation in Borrelia burgdorferi-infected mice. eLife. 2025 May 20:14:RP104913. doi: 10.7554/eLife.104913

57. Schäfer M, Farwanah H, Willrodt AH, Huebner AJ, Sandhoff K, Roop D, et al. Nrf2 links epidermal barrier function with antioxidant defense. EMBO Mol Med. 2012;4(5):364-79. doi: 10.1002/emmm.201200219

58. Eisenstein A, Hilliard BK, Pope SD, Zhang C, Taskar P, Waizman DA, et al. Activation of the transcription factor NRF2 mediates the anti-inflammatory properties of a subset of over-the-counter and prescription NSAIDs. Immunity. 2022;55(6):1082-95.e5. doi: 10.1016/j.immuni.2022.04.015

59. Lewis KN, Wason E, Edrey YH, Kristan DM, Nevo E, Buffenstein R. Regulation of Nrf2 signaling and longevity in naturally long-lived rodents. Proc Natl Acad Sci U S A. 2015;112(12):3722-7. doi: 10.1073/pnas.1417566112

60. Banoth B, Cassel SL. Mitochondria in innate immune signaling. Transl Res. 2018;202:52-68. doi: 10.1016/j.trsl.2018.07.014

61. Canè S, Geiger R, Bronte V. The roles of arginases and arginine in immunity. Nature Rev Immunol. 2025;25(4):266-84. doi: 10.1038/s41577-024-01098-2

62. Bruno M, Maisha S, Mitra A, Costello K, Watkins-Chow D, Logsdon GA, et al. Young KRAB-zinc finger gene clusters are highly dynamic incubators of ERV-driven genetic heterogeneity in mice. bioRxiv. 2025:2025.02.26.640358.

63. Lima-Junior DS, Krishnamurthy SR, Bouladoux N, Collins N, Han SJ, Chen EY, et al. Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. Cell. 2021;184(14):3794-811.e19. doi: 10.1016/j.cell.2021.05.020

64. Liu X, Liu Z, Wu Z, Ren J, Fan Y, Sun L, et al. Resurrection of endogenous retroviruses during aging reinforces senescence. Cell. 2023;186(2):287-304.e26. doi: 10.1016/j.cell.2022.12.017

65. Kassiotis G. The immunological conundrum of endogenous retroelements. Ann Rev Immunol. 2023;41:99-125. doi: 10.1146/annurev-immunol-101721-033341

66. Todaro GJ. Evolution and modes of transmission of RNA tumor viruses. Parke-Davis Award lecture. Am J Pathol. 1975;81(3):590-606. doi: 10.1146/annurev-immunol-101721-033341

67. Joyner CP, Myrick LC, Crossland JP, Dawson WD. Deer mice as laboratory animals. ILAR. 1998;39(4):322-30. doi: 10.1093/ilar.39.4.322

68. Khan M, Gasser S. Generating primary fibroblast cultures from mouse ear and tail tissues. JoVE. 2016(107). doi: 10.3791/53565

69. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):R25. doi: 10.1186/gb-2010-11-3-r25

70. Milovic A, Barbour AG. Protein coding sequences (CDS) of the genome of a Mus musculus 2025. Dryad doi: 10.5061/dryad.hdr7sqvwr

71. Milovic A, Barbour AG. Protein coding sequences (CDS) of the genome of a Peromyscus leucopus. 2025. Dryad doi: 10.5061/dryad.fqz612k44

72. Hedges LV, Gurevitch J, Curtis PS. The meta-analysis of response ratios in experimental ecology. Ecology. 1999;80(4):1150-6.

73. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523. doi: 10.1038/s41467-019-09234-6

74. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9. doi: 10.1038/75556

75. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. v. 4.4.3 ed. 2025. https://www.r-project.org

76. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573-87.e29. doi: 10.1016/j.cell.2021.04.048

77. Shevchenko A, Tomas H, Havli J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols. 2006;1:2856-2860 doi: /10.1038/nprot.2006.468

78. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc Series B (Methodological). 1995:289-300.

Table S1. Line listing of samples under BioProject PRJNA1168383 and with their characteristics and accession numbers

					Control - saline				Pam3CSK4 1 µg/ml				Pam3CSK4 10 µg/ml			
number	animal ID	species	stock	sex	BioSample	PE150 reads	SRA	GEO	BioSample	PE150 reads	SRA	GEO	BioSample	PE150 reads	SRA	GEO
1	25417	*Peromyscus leucopus*	LL	female	SAMN45236433	62061469	SRR31668291	GSM9100090	SAMN45236432	57480172	SRR31668292	GSM9100089	SAMN45236431	57802220	SRR31668293	GSM9100088
2	25418	*Peromyscus leucopus*	LL	female	SAMN45236430	53885821	SRR31668294	GSM9100093	SAMN45236429	62753260	SRR31668295	GSM9100092	SAMN45236428	64192593	SRR31668262	GSM9100091
3	25459	*Peromyscus leucopus*	LL	female	SAMN45236424	66116931	SRR31668266	GSM9100096	SAMN45236423	55069478	SRR31668267	GSM9100095	SAMN45236422	71210765	SRR31668274	GSM9100094
4	25510	*Peromyscus leucopus*	LL	male	SAMN45236421	68185985	SRR31668285	GSM9100099	SAMN45236420	69777800	SRR31668296	GSM9100098	SAMN45236419	75123467	SRR31668297	GSM9100097
5	25558	*Peromyscus leucopus*	LL	male	SAMN45236427	68264338	SRR31668263	GSM9100102	SAMN45236426	67367713	SRR31668264	GSM9100101	SAMN45236425	70462671	SRR31668265	GSM9100100
6	MF1	*Mus musculus*	CD-1	female	SAMN45236442	70502563	SRR31668281	GSM9100069	SAMN45236441	68600982	SRR31668282	GSM9100068	SAMN45236440	72470228	SRR31668283	GSM9100067
7	MF2	*Mus musculus*	CD-1	female	SAMN45236445	62561601	SRR31668278	GSM9100072	SAMN45236444	58171570	SRR31668279	GSM9100071	SAMN45236443	59452728	SRR31668280	GSM9100070
8	MF3	*Mus musculus*	CD-1	female	SAMN45236447	69456772	SRR31668276	GSM9100075	SAMN45236448	62560040	SRR31668275	GSM9100074	SAMN45236446	68555317	SRR31668277	GSM9100073
9	MM1	*Mus musculus*	CD-1	male	SAMN45236436	63434915	SRR31668288	GSM9100078	SAMN45236435	82867487	SRR31668289	GSM9100077	SAMN45236434	77940881	SRR31668290	GSM9100076
10	MM2	*Mus musculus*	CD-1	male	SAMN45236439	67833497	SRR31668284	GSM9100081	SAMN45236438	77880118	SRR31668286	GSM9100080	SAMN45236437	60443235	SRR31668287	GSM9100079

Table S3. Product names of DEGs of CDS among DEGs

CDS	Product
Abcg2	ATP binding cassette subfamily G member 2
Actb	actin, beta
Adamts5	ADAM metallopeptidase with thrombospondin type 1 motif 5
Adamts9	ADAM metallopeptidase with thrombospondin type 1 motif 9
Adamtsl2	ADAMTS-like 2
Ago1	argonaute RISC catalytic subunit 1
Aqp5	aquaporin 5
Arg1	arginase 1
Arg2	arginase 2
Ass1	Argininosuccinate synthetase 1
Batf2	basic leucine zipper transcription factor, ATF-like 2
Bcl3	B cell leukemia/lymphoma 3
Bdkrb1	bradykinin receptor, beta 1
Brca1	breast cancer 1, early onset
Bst2	bone marrow stromal antigen 2
C1r	complement component 1, r subcomponent A
C4b	complement C4B
Ca3	carbonic anhydrase 3
Calhm5	calcium homeostasis modulator family member 5
Casp1	caspase 1
Cck	cholecystokinin
Ccl27a	C-C motif chemokine ligand 27A
Ccl7	C-C motif chemokine ligand 7
Cd248	CD248
Cd40	CD40
Cfb	complement factor B
Ch25h	cholesterol 25-hydroxylase
Chi3l1	chitinase 3 like 1 (Brp39)
Cmpk2	cytidine/uridine monophosphate kinase 2
Col11a1	collagen, type XI, alpha 1
Col18a1	collagen, type XVIII, alpha 1
Col3a1	collagen, type III, alpha 1
Col4a5	collagen, type IV, alpha 5
Col5a3	collagen, type V, alpha 3
Col6a1	collagen, type VI, alpha 1
Col6a2	collagen, type VI, alpha 2
Col6a3	collagen, type VI, alpha 3
Col8a1	collagen, type VIII, alpha 1
Csf3	colony stimulating factor 3 (granulocyte)
Ctsk	cathepsin K
Cxcl1	C-X-C motif chemokine ligand 1
Cxcl10	C-X-C motif chemokine ligand 10
Cxcl11	C-X-C motif chemokine ligand 11
Cxcl12	C-X-C motif chemokine ligand 12
Cxcl14	C-X-C motif chemokine ligand 14
Cxcl3	C-X-C motif chemokine ligand 3
Cxcl5	C-X-C motif chemokine ligand 5
Dcn	decorin
Ddx60	DExD/H box helicase 60
Dennd3	DENN domain containing 3
Deptor	DEP domain containing MTOR-interacting protein
Dkk2	dickkopf WNT signaling pathway inhibitor 2
Dpt	dermatopontin
Dpy19l2	Dpy-19 like 2
Dtx3l	deltex 3-like, E3 ubiquitin ligase
Ednrb	endothelin receptor type B
Efemp1	EGF-containing fibulin-like extracellular matrix protein 1
Egln3	Egl-9 family hypoxia-inducible factor 3
Eng	endoglin (CD105)
Ereg	epiregulin
Fscn1	fascin
Fth1	Ferritin heavy polypeptide 1
Gbp2b	guanylate binding protein 2b (Gbp1)
Gzma	granzyme A
H-2_DDa	MHC class I-like protein (XP_028718389.1)
H-2_DDa1	MHC class I-like protein (XP_037055648.1)
H-2_LDa-3	MHC class I-like protein (XP_028718386.2)
H-2_Q10a-1	MHC class I-like protein (XP_037055648.1)
H19	H19, imprinted maternally expressed transcript
Hpx	hemopexin
Icam1	intercellular adhesion molecule 1
Icam4	intercellular adhesion molecule 4
Ifi203	interferon activated gene 203
Ifi205	interferon activated gene 205
Ifih1	interferon induced with helicase C domain 1 (MDA5)
Ifngr1	interferon gamma receptor 1
Igf1	insulin-like growth factor 1
Ikbke	inhibitor of kappaB kinase epsilon
Il11	interleukin-11
Il15	interleuki-15
Il1r2	interleukin-1 receptor, type II
Il1rn	interleukin-1 receptor antagonist
Il36a	interleukin-36a
Il4r	interleukin-4 receptor
Il6	interleukin-6
Il6r	interleukin-6 receptor
Insig1	insulin induced gene 1

CDS	Product
Irf7	interferon regulatory factor 7
Irf9	interferon regulatory factor 9
Irgm1	immunity-related GTPase family M member 1
Irgm2	immunity-related GTPase family M member 2
Ism1	isthmin 1, angiogenesis inhibitor
Kctd12	potassium channel tetramerisation domain containing 12
Kpna7	karyopherin subunit alpha 7
Krt14	keratin 14
Lbp	lipopolysaccharide binding protein
Lcn2	lipocalin 2
Ldoc1	regulator of NFKB signaling
Lpl	lipoprotein lipase
Lrig1	leucine-rich repeats and immunoglobulin-like domains 1
Lrrc17	leucine rich repeat containing 17
Lum	lumican
Ly6e	lymphocyte antigen 6E
Mmp12	matrix metallopeptidase 12
Mmp2	matrix metallopeptidase 2
Mmp3	matrix metallopeptidase 3
Mmp9	matrix metallopeptidase 9
Mt1	metallothionein 1
Mt2	metallothionein 2
Mx2	MX dynamin-like GTPase 2
Myc	myelocytomatosis oncogene
Ndn	necdin, MAGE family member
Neurl3	neuralized E3 ubiquitin protein ligase 3
Nfkb1	nuclear factor of kappa light polypeptide gene enhancer in B cells 1, p105
Nfkbiz	nuclear factor of kappa light polypeptide gene enhancer in B cells inhibitor, zeta
Nod1	nucleotide-binding oligomerization domain containing 1
Nos2	nitric oxide synthase 2
Oas1	2'-5' oligoadenylate synthetase 1
Pglyrp3	peptidoglycan recognition protein 3
Pla2g2a	phospholipase A2, group IIA
Pla2g5	phospholipase A2, group V
Plek	pleckstrin
Prrx2	paired related homeobox 2
Rcan1	regulator of calcineurin 1
Rgs16	regulator of G-protein signaling 16
Rgs5	regulator of G-protein signaling 5
Rigi	RNA sensor RIG-I (Ddx58)
Rnf122	ring finger protein 122
Rsad2	radical S-adenosyl methionine domain containing 2 (Viperin)
Rspo3	R-spondin-3
Rtp4	receptor transporter protein 4
S100a8	S100 calcium binding protein A8 (calgranulin A)
S100b	S100 protein, beta polypeptide, neural
Saa1	serum amyloid A 1
Saa3	serum amyloid A 3
Samd11	sterile alpha motif domain containing 11
Sdc4	syndecan 4
Selp	Selectin, platelet
Sema3d	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, 3D
Sema7a	Sema domain, immunoglobulin domain (Ig), and GPI membrane anchor, 7A
Serpina3f	serine (or cysteine) peptidase inhibitor, clade A, member 3f
Serpina3g	serine (or cysteine) peptidase inhibitor, clade A, member 3g
Serpina3i	serine (or cysteine) peptidase inhibitor, clade A, member 3I
Serpina3n	serine (or cysteine) peptidase inhibitor, clade A, member 3N
Slc2a6	solute carrier family 2, member 6 (GLUT6)
Slc7a2	solute carrier family 7, member 2 (CAT2)
Slco2a1	solute carrier organic anion transporter family, member 2a1 (PGT)
Slfn2	schlafen 2
Slpi	secretory leukocyte peptidase inhibitor
Sod2	superoxide dismutase 2
Sp110	Sp110 nuclear body protein (Ipr1)
Srgn	serglycin (Prg)
Stat2	signal transducer and activator of transcription 2
Stat4	signal transducer and activator of transcription 4
Sting1	stimulator of interferon response cGAMP interactor 1
Tagln	transgelin
Tfpi2	tissue factor pathway inhibitor 2
Tgfb1	transforming growth factor, beta 1
Timeless	timeless circadian clock 1
Tlr2	Toll-like receptor 2
Tm4sf19	transmembrane 4 L six family member 19
Tnf	tumor necrosis factor
Tnfaip6	tumor necrosis factor alpha induced protein 6 (TSG-6)
Tnfsf11	tumor necrosis factor (ligand) superfamily, member 11 (RANKL)
Tnfsf18	tumor necrosis factor (ligand) superfamily, member 18
Tnfsf4	tumor necrosis factor (ligand) superfamily, member 4
Traf1	TNF receptor-associated factor 1
Trim30d	tripartite motif-containing 30D
Tyk	tyrosine kinase 2
Ubd	ubiquitin D (FAT10)
Ung	uracil DNA glycosylase
Vcam1	vascular cell adhesion molecule 1
Wnt16	Wingless-type MMTV integration site family, member 16

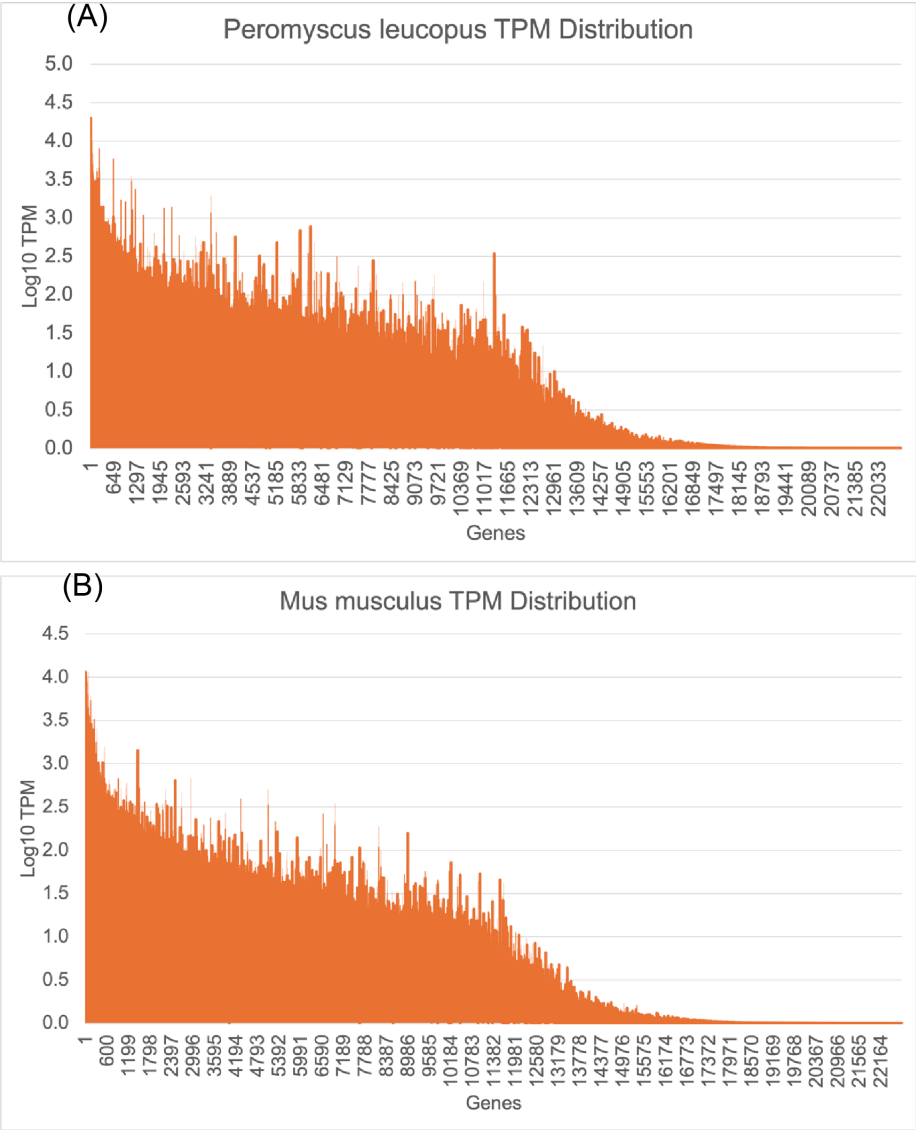

Figure S1. Histograms of the distributions of $\log_{10}$ of the mean (n=5) TPM values for 22,760 CDS of *Mus musculus* (panel A) and 22,654 CDS of *Peromyscus leucopus* (panel B) from genome-wide RNA-seq of primary dermal fibroblast cultures under control conditions. The *x*-axis is the cumulative CDS count. Data for analysis are in Dryad Tables D1 and D2.

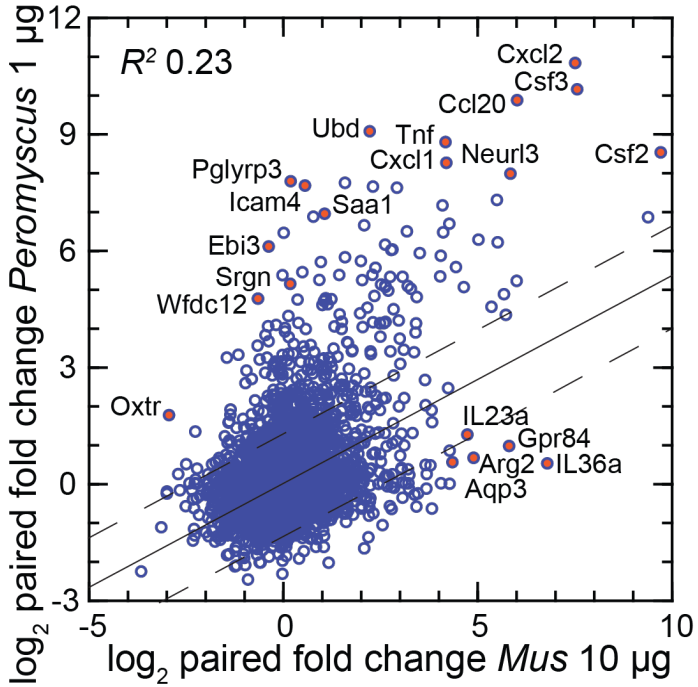

Figure S2.

Figure S2. Scatter plot of log-transformed FC values for each species against each other as described for Figure 2. The difference is that the comparison for *P. leucopus* is with *M. musculus* fibroblasts exposed to the 10 µg/ml concentration of Pam3CSK4. The linear regression line with 95% confidence interval and coefficient of determination (R^2) are shown. Selective genes that are up-regulated and differentially expressed (DEGs) for each species are indicated by name and red fill. Data for analysis are in Dryad Table D3.

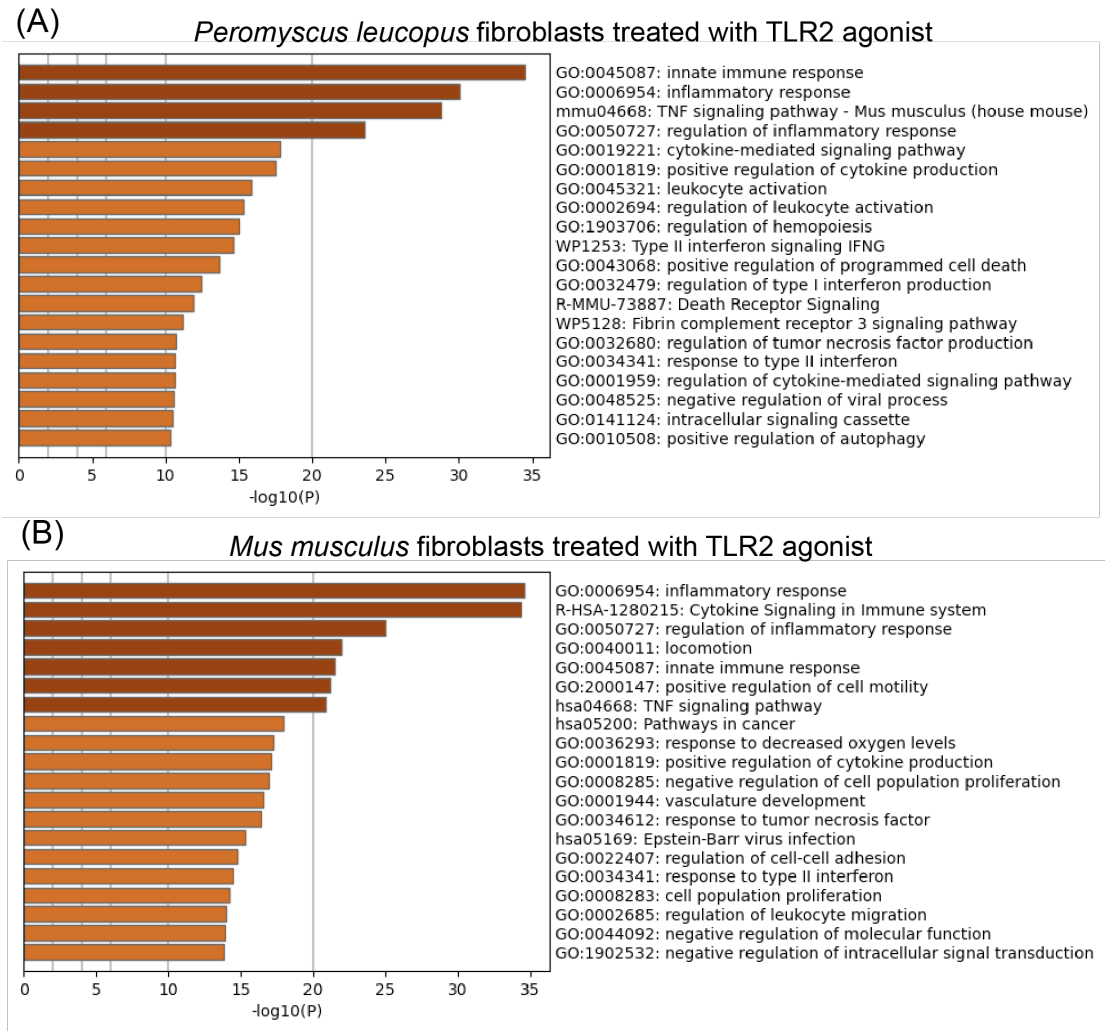

Figure S3. Horizontal bar graphs for the comparison of *Peromyscus leucopus* (panel A) and *Mus musculus (panel B)* in their responses to the TLR2 agonist Pam3CSK4 by RNA-seq and categorization of DEGs by Gene Ontology (GO) term enrichment. The analyses shown are of those sets of genes of each species that were higher in transcription in the treated cells than in the control cells. Data for analysis are in Table S2.

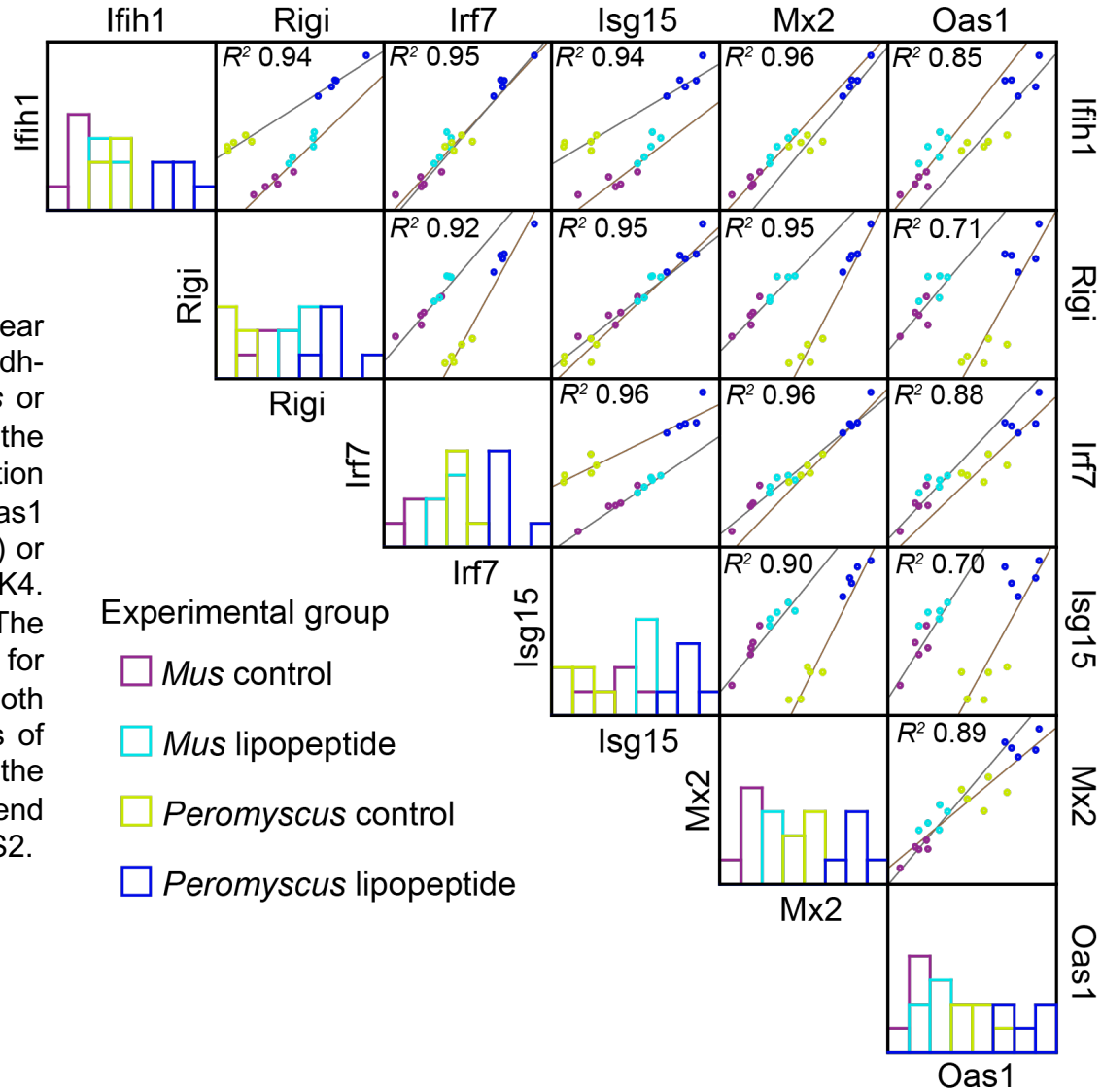

Figure S4. Matrix of scatter plots and linear regressions of log-transformed Gapdh-normalized transcriptions by *Mus musculus* or *Peromyscus leucopus* dermal fibroblasts of the PRRs Ifih1 (MDA5) and Rigi, transcription factor Irf7, and the ISGs Isg15, Mx2, and Oas1 for all pairs. The cells were without (control) or with treatment with the lipopeptide Pam3CSK4. The linear regressions are by species. The coefficients of determination (R^2) are for Geneal Linear Model calculations for both species. Histograms of density distributions of the values by experimental group for the variable of each *y*-axis are show at the left-end of each row. Data for analyses are in Table S2.

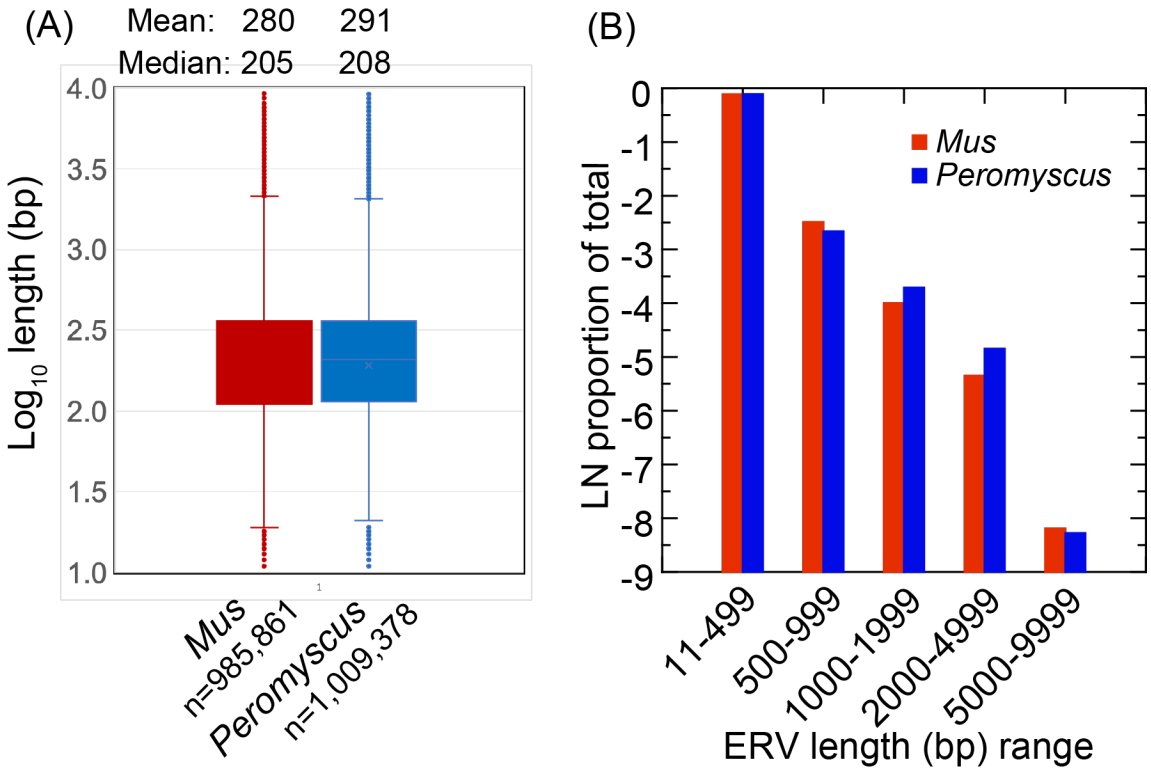

Figure S5. Comparison of the lengths of ERV/TE sequences in the 11 to 9999 bp range for *Mus musculus* and *Peromyscus leuccopus*. Panel A is a box-whisker plot. Panel B is a graph of natural logarithms (LN) of the proportion of sequences of the specified range (*x*-axis) out of the total number (n) of sequences for each species. Data for analysis in Dryad Tables D4 (*P. leucopus*) and D5 (*M. musculus*).

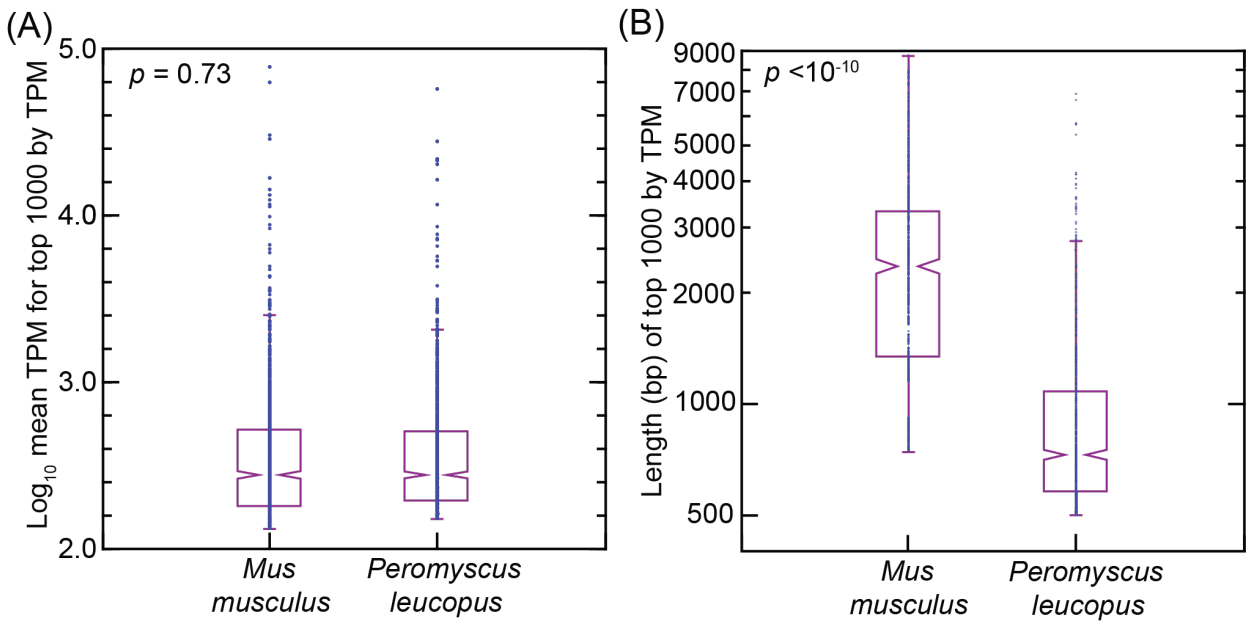

Figure S6. Box-whisker plots of TPM values (left panel) and lengths (right panel) for the top 1000 in descending order of mean for 5 samples of each species of length-adjusted TPM values for ERV/TE sequences of 500-9999 bp for *M. musculus* (n=104,932) and *P. leucopus* (n=103,397) for untreated dermal fibroblasts. The *p* values are for *t*-tests of log-transformed numbers. Data for analyses are from Dryad Table D6.