UNIVERSITY OF CALIFORNIA,
IRVINE

Peromyscus leucopus – a reservoir of zoonotic agents
and a model for understanding infection tolerance

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in Biomedical Sciences

by

Ana Milovic

Dissertation Committee:
Professor Alan G. Barbour, Chair
Professor Anthony D. Long
Professor Bert L. Semler
Professor Michael McClelland

2024

Dedication

For Mia

Table of contents

List of Figures

List of Tables

Acknowledgments

Part of Chapter 4 is used as it appears in the material published by Elife in 2023 under the title The white-footed deermouse, an infection-tolerant reservoir for several zoonotic agents, tempers interferon responses to endotoxin in comparison to the mouse and rat. Co-authors are Ana Milovic*, Jonathan V. Duong, Alan G. Barbour. Article is accessible at https://doi.org/10.7554/eLife.90135.2. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Authors of the papers with asterisks denote the first authors.

Vita

<u>Education</u>

2024   Doctor of Philosophy, Biomedical Sciences, University of California, Irvine
2011   Masters of Science in Medicine in Medical Microbiology, University of Cape Town
2006   Bachelors of Science, Biological Sciences, Belgrade University

<u>Experience</u>

2019-2024   Graduate Research Associate, with Dr. Alan Barbour
             Department of Microbiology and Molecular Genetics
             University of California, Irvine
2017-2019   Assistant Specialist
             Department of Microbiology and Molecular Genetics
             University of California, Irvine, USA
2009-2010   Scientific officer, study coordinator,
             Foundation for Innovative New Diagnostics
             Institute of Infections Disease and Molecular Medicine, Cape Town, South Africa
2007-2009   Research Assistant with Dr. Mark Nicol
             Institute of Infections Disease and Molecular Medicine, Cape Town, South Africa
2005-2006   Research Assistant with Dr. Miroslav Vrvic
             Institute of Chemistry, Technology and Metallurgy, Belgrade, Serbia
             Belgrade University

<u>Academic honors</u>

2021   Graduate Dean's Excellence Fellowship
        Department of Microbiology & Molecular Genetics, University of California, Irvine
2009   Scholarship Award
        Department of Medical Microbiology and South African Tuberculosis Vaccine Initiative
2008   Academic Award based on academic merit during the first year of MSc studies
        University of Cape Town
2007   Academic Award granted by the Dean of Faculty of Health Sciences
        University of Cape Town

<u>Publications</u>

1.   **Milovic A***, Gach J, Forthal D, Chatzistamou I, Olivarria G, Lane TE, Barbour AG. The Peromyscus leucopus model of SARS-CoV-2 infection of lungs and brain. In preparation. (2024)
2.   **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)
3.   Balderrama-Gutierrez G*, **Milovic A***, Cook VJ, Islam MN, Zhang Y, Kiaris H, Belisle JT, Mortazavi A, Barbour AG. An Infection-Tolerant Mammalian Reservoir for Several Zoonotic Agents Broadly Counters the Inflammatory Effects of Endotoxin. **mBio**. (2021)

4.   **Milovic A***, Bassam K, Shao H, Chatzistamou I, Tufts DM, Diuk-Wasser M, Barbour AG. Lactobacilli and other gastrointestinal microbiota of Peromyscus leucopus, reservoir host for agents of Lyme disease and other zoonoses in North America. **PLoS One**. (2020)

5.   Bassam K*, **Milovic A**, Barbour AG. Genome Sequences of Three Lactobacillus Species Strains of the Stomach of the White-Footed Deermouse (Peromyscus leucopus). **Microbiol Resour Announc**. (2019)

6.   Boehme CC*, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, **Milovic A**, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins MD. Rapid molecular detection of tuberculosis and rifampin resistance. **N Engl J Med**. (2010)

       * First author

Selected meeting presentations

       Milovic A*, Gach J, Forthal D, Barbour AG, Novel outbred small animal model in SARS-CoV-2 infection. Poster presented at ASM Microbe, 2023, Houston, TX

       Milovic A*, Duong JV, Barbour AG, Inflammation profile of Peromyscus leucopus to TLR2 and TLR4 agonists. Poster presented at Biology of Spirochetes, Gordon Research Conference, 2022, Ventura, CA

Career development training

2020   Biosafety Level III training
      University of California, Irvine
2013   Regulatory requirements for medical devices
      UC Irvine Extension, University of California, Irvine
2009   Infectious Disease Clinical Research Training Program
      George Washington University, School of Medicine, and Health Sciences
2007   Biosafety Level III training
      University of Cape Town, South Africa

Professional affiliations

2017   American Society for Microbiology

Abstract of the Dissertation

Peromyscus leucopus – a reservoir of zoonotic agents and a model for understanding infection

tolerance

by

Ana Milovic

Doctor of Philosophy in Biomedical Sciences

University of California, Irvine 2024

Professor Alan G. Barbour, Chair

Infection tolerance is the ability to minimize damage caused by pathogens or the host's response to them. A prime example of such resilience is the white-footed deermouse, or *P. leucopus*, a reservoir of agents of zoonoses like Lyme disease, anaplasmosis, or Powassan virus encephalitis. Despite being persistently infected with pathogens, deermice can strike a delicate balance, keeping the pathogens in check while avoiding maladaptive host response that could lead to inflammation-induced damage. This unique resilience makes deermice an excellent animal model for understanding infection tolerance: they don't get sick and remain fit for their populations' proliferation. As a result, they live longer than most other rodents of the same size. A deeper understanding of how deermice moderate inflammation and other damaging host responses, including sepsis, could explain why some patients with certain infections experience more prolonged or severe disease courses. These individuals may lack the capacity that deermice possess to avoid sickness. Identifying and characterizing factors of infection tolerance in deermice that reduce inflammation elicited by microbes or their toxins may thus lead to more effective therapies.

In this dissertation, I describe the microbiome of *P. leucopus* from a closed colony using metagenomics and compare it to that of *M. musculus* and a natural population of P. leucopus. Our findings reveal that deermice have a higher number of lactobacilli in their gastrointestinal microbiota. We also discovered new species of lactobacilli that are specific to deermice. As the microbiome influences the immune system, I employ sequencing and immunology techniques to examine tolerance in immune-related tissues of *P. leucopus*.

Immunity evolved to protect the host from pathogens, but inflammatory molecules can cause pathology, chronic conditions, and accelerated aging. The second part of the study explains some of the mechanisms for striking tolerance without developing pathology through inflammation in *P. leucopus* while infected with *Borrelia hermsii* or SARS-CoV-2. This study suggests that *P. leucopus* responds to pathogens by recruiting and activating leucocytes during the initial stage of infection while regulating the activity of these cells with anti-inflammatory cytokines. This model of infection tolerance was evaluated and characterized compared to *M. musculus* by the early response to endotoxins and bacterial infection. Each rodent species responded by activating a different immune response route. While *P. leucopus* displayed an alternatively activated macrophage profile and reduced transcription of endogenous retroviruses, as well as consistently moderate pathogenesis, *M. musculus* displayed classically activated macrophages and higher expression of endogenous retroviral proteins. We are the first to report infection of *P. leucopus* with SARS-CoV-2 and describe the host response in the brain.

CHAPTER ONE: Introduction and Literature Review

Overview of the Dissertation

The overall objective of the dissertation is to accurately assess the potential for infection tolerance in *Peromyscus leucopus*. I studied *Peromyscus leucopus'* transcriptome during infection and its microbiome, completing this species' "omics" mosaic to attain this goal. Experiments in this work are based mainly on sequencing techniques, either in describing microbiomes or analyzing differential gene expression in immune-relevant tissues. The central question is whether *P. leucopus* shows significant characteristics of immune tolerance in comparison to two additional rodent animal models. Results strongly point to a phenomenon of tolerance, the primary focus of this work.

In the Introduction Chapter, I begin by describing *P. leucopus* as a zoonosis reservoir of Lyme disease pathogen *Borrelia burgdorferi*. Then, I examine its potential as a valuable animal model compared to *M. musculus* due to its tolerance against infection. Since later chapters focus on inflammation to lipopolysaccharide (LPS), the introduction discusses inflammation in sepsis and its effect on the organism at both the tissue and the cellular level, with examples of neutrophils and macrophages in the early stages of the immune response. Besides the response to endotoxins, the study focuses on zoonoses, including borreliosis and COVID-19, in terms of animal reservoirs and vectors. Thus, the introduction also discusses zoonoses caused by these pathogens. I begin my dissertation considering *P. leucopus'* microbiota in Chapter 2.

In Chapter 2, I describe the microbiota of *P. leucopus* across sexes and between two rodent species. This chapter focuses on *Lactobacillus,* which originates in the stomach. After discussing our analysis of natural populations of *P. leucopus* and colony animals in terms of beta diversity, I describe a new species of *Lactobacillus* found in *P. leucopus* stomach biofilm.

Chapter 3 describes the innate immunity of *P. leucopus* four hours post-challenge with LPS compared to *M. musculus*.

Chapter 4 examines genome-wide transcription in the blood of *P. leucopus*, the outbred line of *M. musculus*, and *Rattus norvegicus* adjusted for white cell concentrations.

Chapter 5 describes the immunity of *P. leucopus* to infection with SARS-CoV-2 in the lungs and brain and plasma markers.

The last chapter discusses our findings and future directions, combining all the results, starting with microbiota, considering *P. leucopus'* response to external stimuli and pathogens, and proposing translational solutions based on our study.

Peromyscus

Peromyscus leucopus as a zoonotic reservoir

Peromyscus, commonly known as deermice, are ecologically important mammals with over 50 species distributed across North America, spanning from Alaska to Central America and from the Atlantic to the Pacific coasts. Their habitats range from wetlands and beaches to forests, prairies, deserts, and high altitudes (1), (2). Given their broad habitat range extending from rural to suburban environments, they often come into contact with human populations, particularly where urban and natural landscapes intersect. This interaction has led to extensive research on *Peromyscus leucopus*, the white-footed mouse, recognized as a natural host and a key reservoir for several tick-borne zoonotic diseases, including Lyme disease, babesiosis, anaplasmosis, relapsing fever, ehrlichiosis, and viral encephalitis (3). Additionally, a close relative of *Peromyscus leucopus*, *Peromyscus maniculatus*, is known as a reservoir for Hantavirus, with both species displaying a notable tolerance to infections that can be severe and sometimes fatal in humans (3).

Despite often being referred to as a "mouse," the deermouse belongs to the *Cricetidae* family, which also includes hamsters and voles, rather than the *Muridae* family, which includes *Mus musculus*, a house mouse and *Rattus* species. The genetic divergence between *Peromyscus* and *Mus* is estimated to be around 27 million years (figure 1 (4). Due to the significant divergence between

genera, molecular tools used for studying *M. musculus* are incompatible with *P. leucopus*. The considerable role of *P. leucopus* as a disease reservoir, coupled with its resilience to infections, has prompted the genetic exploration of this species, including the sequencing and annotation of its nuclear genome (5) mitochondrial genomes (6) which established the necessary tools to use it as a research model.

Figure 1Peromyscus leucopus and distance phylogram
The right panel depicts the simplified phylogeny depicting Cricetidae family and Muridae family, diverged roughly 25 million years ago {Bedford, 2015 #195 and the left shows an image of *Peromyscus leucopus* (John White CC BY 4.0)

Peromyscus as an animal model

Peromyscus is an ecologically important animal in North America due to its abundance and its status as a zoonoses reservoir. It has heretofore been used for studying natural variation, oxidative stress, wound healing, aging, metabolism of sugar and lipids, hematology, pathogenesis, and behavioral studies (7), (1), (8), (9), (10), (11). However, the conventional model for biomedical research has been *M. musculus*. Although *M. musculus* shares many genetic homologies with humans, many established strains lack genetic diversity and do not reflect variations of the human population (12), (8). Although inbred strains increase study reproducibility, the downside is that they are not a good model for studying the diversity of responses to stimuli. Despite both *Peromyscus* and *Mus* being small rodents, they have different origins, physical characteristics, and behavioral traits. *M. musculus,* originally native to Asia, has spread worldwide and has adapted to living closely with

humans in environments such as homes and farms. They are less common in wild habitats and are considered a commensal species living off human resources (13). In contrast, *Peromyscus* are independent of humans, though their habitats can overlap (14). *Mus* are domesticated and docile in captivity, while *Peromyscus* are difficult to handle. Additionally, *Mus* has many strains but comprises only one species, while the genus *Peromyscus* comprises many species, one of which is *P. leucopus*. Studies using *P. leucopus* as an animal model have thus far focused mainly on infection burden with *B. burgdorferi* in a natural population, but also studies of aging and immunity to bacterial and viral pathogens (15), (16), (17), (18), (19), (20), (21), (22). Aging studies point out the longevity of *P. leucopus* - a species that can live in captivity for up to 8 years, while *M. musculus* is reported to live for 3.5 years (10). One explanation for this discrepancy is that *P. leucopus* makes less reactive oxidative species in mitochondria and has a better system of scavenging them while causing less damage to lipids and muscle (23), (24). Finally, the *P. leucopus* harbor and maintain one or more pathogens and facilitate their transmission via ticks or direct contact without suffering disease or fitness consequences, a phenomenon less commonly observed in *Mus* species. This immunology trait, known as tolerance, observable in both laboratory and field studies of *P. leucopus*, underscores its potential as a model for understanding host-pathogen interactions and disease resistance mechanisms (25), (3).

Tolerance in infection

Tolerance is defined as persistent infection with low morbidity (figure 2, left panel). Tolerance in infection refers to a state where the host maintains persistent infection with minimal morbidity, reflecting a finely tuned immune response that manages pathogen levels without inducing harmful inflammation.
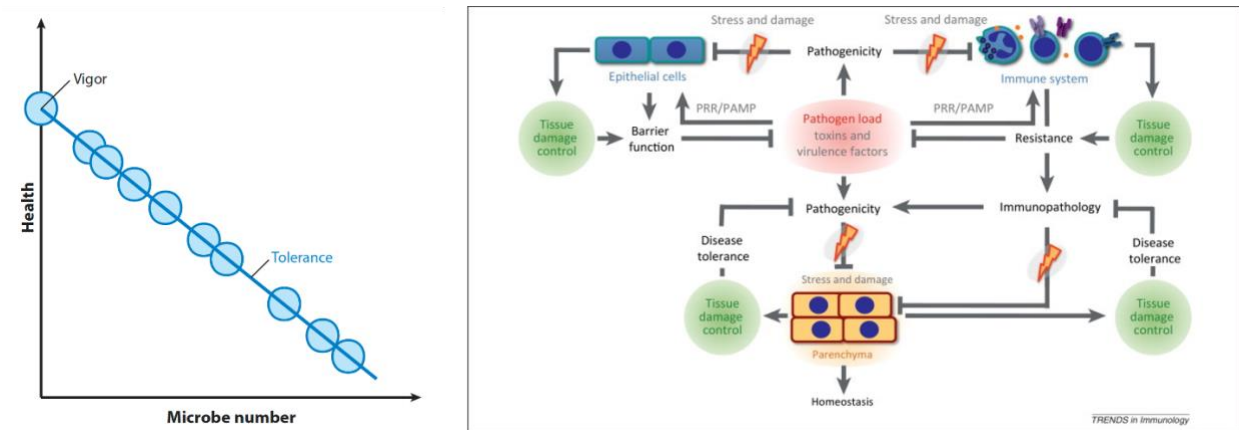
Figure 2 Graphical representation of tolerance
The left panel is a graphical representation of tolerance and vigor. In this context, resistance would be inverse pathogen load. Figure adapter from *(26)*. The right panel shows tissue damage control in host-microbe interactions and maintaining homeostasis *(27)*

In the context of infections, pathogens such as bacteria, viruses, and fungi pose threats that call for an immune response. However, the strength and nature of this response can produce critically different outcomes. The immune system has two significant roles:

- Ability to clear pathogens (resistance)

- Ability to reduce the impact of the level of pathogens (tolerance)

Tissue damage control is critical to the interaction between hosts and microbes. While the host balances eradicating pathogens and maintaining homeostasis, pathogens can disrupt this balance by releasing toxins and virulence factors that cause cellular stress and damage to these tissues, a key factor in disease pathogenicity (figure 2, right panel). When the body recognizes pathogens through pattern recognition receptors (PRRs) engaging with pathogen-associated molecular patterns (PAMPs), it triggers immune responses that aim to reduce pathogen numbers, thereby decreasing their harmful impact.

The concept of disease tolerance was first identified as a defense mechanism in plants against infections. It refers to the ability of a host to endure infection without experiencing any decrease in their overall fitness. This concept was later expanded to include the fruit fly Drosophila melanogaster. It was discovered that mutations in inflammatory cytokine genes affect the survival rate of the host

5

following an infection without changing the level of pathogens (27). But bats are probably the most notorious mammals for harboring pathogens asymptomatically. An example is the Marburg virus reservoir bat, which has been shown to have constitutively expressed and drastically diversified type I interferons. The same species show expansion of receptors on the natural killer cells that may result in overall inhibitory balance with this cell population (28), (29).

If homeostasis post-infection is not achieved, an overactive immune response can lead to tissue damage and disease, as seen in severe inflammation or septic shock cases in which the body's reaction to infection causes more harm than the pathogens themselves.

Out-of-control inflammation and sepsis

Sepsis contributes to at least 1.7 million adult hospitalizations and at least 350,000 deaths annually in the United States (30). Annual worldwide estimates predict 30 million cases of sepsis with 6 million deaths. Early diagnosis and antibiotics reduce fatality rates in high-income countries, while in most of the world, sepsis is considered an underlying condition in poor outcome cases (31).

The term sepsis was defined initially by Hippocrates as σῆψις, or rot, and connoted the foul-smelling rotting of a wound. Even in Ancient Greece and Egypt, well before Pasteur and any knowledge of microbes, people postulated that illness could result from invisible causes. Egyptians understood that sickness comes not only from outside but also from dangerous, invisible content of the intestines that becomes life-threatening when exposed to the blood (32). Another biological breakdown defined by the ancient Greeks is pepsis, translated as digestion or fermentation, a phenomenon I discuss later in the introduction. Both sepsis and pepsis are biological breakdowns facilitated by microbes, but one breaks down matter like tissues, which causes sickness, and the other helps digest food in the gut.

In many languages, sepsis translates as "blood poisoning." Since the discovery of antibiotics, it has become clear that blood poisoning continues even after bacteria are cleared, leading to the

conclusion that sepsis is a complex reaction triggered by bacteria in an unsequestered environment or necrotic or injured tissue (sterile sepsis in the case of bone fractures, ischemia or hemorrhagic shock), but is driven by exaggerated immune reaction causing dysregulated systemic inflammation (33), (34). Sepsis is traditionally diagnosed as the presence of microorganisms in the blood to which the host has a systemic response with a heterogeneous spectrum of pathologies (35), (36). Not only bacteria can cause sepsis, but also fungi and parasites. Sepsis can be a reaction to bacterial components like LPS from Gram-negative bacteria and lipoteichoic acid from Gram-positive bacteria. Pathology of sepsis is an extreme case of an uncontrolled immune response, underscoring the importance of striking a delicate balance between appropriately eradicating the infection and avoiding damage to the host.

Treatment of sepsis includes antibiotics followed by source control; that is, determining the presence of any abscesses, necrotic tissue, or infected device followed with fluid resuscitation to restore blood pressure and organ perfusion, sometimes including vasopressor agents and inotropic therapy combined to increase cardiac output (37), (38). Additional therapies besides antibiotics and fluid resuscitation include managing the host immune system with steroids and immune suppressive agents. One such approach uses recombinant human activated protein C (rhAPC), the first systemic sepsis agent used as a therapeutic for modulating coagulation and inflammation. This approach saw modest success before it was withdrawn since the trial data did not hold in practice (39), (40), (41), (42). Besides antibiotics, there is no treatment for sepsis, and in the case of sterile sepsis and systemic infection caused by bacterial toxins like LPS, antibiotics are useless. So why is there no effective treatment against a deadly disease we have known for millennia? Is the problem our insufficient understanding of the mechanisms of sepsis? Are we using appropriate research models?

Animal models in sepsis

M. musculus is the most likely choice of animal model in medical research, but there are certain caveats to consider in cases of systemic infection. The mice model of sepsis is usually an animal challenged with LPS, an endotoxin, or an animal afflicted with a puncture of a protective barrier like caecum. The problem is that the mouse might succumb to the inflammation; thus, the organ failure we see in humans will not occur. Additionally, it is technically challenging to provide supportive care for a mouse model (36), (43). Mice used in research are usually inbred, but the severity of sepsis varies individually among humans, and the mouse model does not account for that (44), (45). Sepsis in humans and mice further differs on the cellular level.

Immunity of sepsis at the cellular level and innate immunity

On a cellular level, sepsis in humans is driven by uncontrolled innate immunity. Cells of innate immunity are mainly phagocytic cells: macrophages, monocytes, granulocytes, dendritic cells, and, to a lesser extent, lymphocytes. When these cells are not causing a systemic response, they recognize, ingest, and kill microbes. In humans and the most abundant animal model in research, *M. musculus,* the number of circulating leucocytes is 3 –10 × 10^6 cells per milliliter of blood (3 –10 × 10^9 cells per L) (46). Even though *M. musculus* is the most commonly used experimental animal for understanding the immune system, it is an imperfect model since there are significant differences between mice and humans. In mice, lymphocytes are the predominant cell type at 75% to 90% of the total leucocytes in healthy wild-type animals, while neutrophils comprise 20% to 30 % of the white blood cells (WBC). In humans, lymphocytes comprise 30% to 50% of WBC in the peripheral blood. The most abundant cell population of peripheral blood leucocytes in humans are neutrophils at 50% to 70%, while in mice, that percentage is 10% to 25% (47), (48), (46), (49), (50). Due to these divergences between mice and humans, an animal model whose blood composition is closer to

humans could prove more effective in sepsis research, especially emphasizing cells involved in innate immunity like neutrophils and macrophages and effector molecules – interferons.

<u>Neutrophils</u>

Neutrophils, the predominant cells in human peripheral blood, initiate the immune response to PAMPs, leading to microbial phagocytosis (51). PAMPs are recognized by a large family of PRRs on innate immune cells, or inside the cytosol, namely in neutrophils, macrophages, and dendritic cells. The scope of the response is further determined by the downstream effects of expression of PRR response genes (34), (52). Microbial surfaces are the primary example of PAMPs but also LPS, flagellin, unmethylated CpG motifs characteristic of bacteria but not vertebrate genomic DNAs, and poly IC resembling double-stranded RNA in some viruses (53), (34). Self-derived host molecules can be causative agents of inflammation in cases of cellular damage or stress named danger-associated molecular patterns (DAMPs), and those include granulolysins, defensins, lactoferrin, cathepsin G, HMGB1, urate crystals, ATP, also enzymes (ATPases) or peptides (HMGB1) (34).

One of the receptors on mainly macrophages and neutrophils are lectins. Lectins bind sugar polymers on bacterial and fungal surfaces, initiating phagosome formation and fusion with a lysosome and breaking down microbes in a phagolysosome. Besides sugar-binding receptors, macrophages and neutrophils have G-protein-coupled receptors (GPCRs) which, coupled with intracellular GTP binding heterotrimeric protein, signal Rac/Rho cascade and increase the antimicrobial activity of phagocytes (54). These subsequently influence actin polymerization towards cell migration, motility, and neutrophil adhesion. Besides cytoskeletal rearrangements, GPCRs activate c-Jun kinase and p38 MAP kinase cascades, leading to transcriptional activation (55). Examples of G-protein coupled receptors are fMet-Leu-Phe (fMLP) receptors binding to fMet residue of bacteria – a unique protein initiating amino acid in prokaryotes (56), (57). Signaling with fMet coupled with GPCR induces the production of reactive oxygen species in the phagolysosome by

directly interacting with NADPH oxidase complex (58). Another example of GPCR is C5a binding seven membrane-spanning receptor (C5aR1) binding to complement fragment C5a. Complement fragment C5a is like generated by a convertase in the activated complement cascade. C5a either binds to complement receptor 1 (CR1), activating phagocytosis of opsonized pathogens, or GPCR C5a receptor initiating association of NADPH complex and like fMet promotes the production of reactive oxidative species (ROS) (59) (54). Signaling with C5a and fMet activates NADPH oxidase or phagocyte oxidase components. Superoxide and hydrogen peroxide produced here are electron donors for myeloperoxidase (MPO), one of the most abundant proteins in neutrophils, making up 5% of all proteins in the cell. MPO uses hydrogen peroxide to oxidize available substrates in the environment and can produce hypochlorous and hypothiocyanous acid and various organic and inorganic radicals (58), (60). Antioxidant scavenging enzymes like superoxide dismutase (SOD) or catalase dismutate reactive superoxide anion to hydrogen peroxide to prevent local tissue damage. Besides superoxide anion, reactive nitrogen species are another toxic product aimed at eradicating pathogens. Inducible nitric oxide synthase (NOS2) is another protein contributing to oxidative stress expressed among other stimuli by activating fMet via kinase cascade. Nitric oxide (NO) synthesized by constitutively expressed endothelial eNOS is a smooth muscle local mediator with paracrine function causing vasodilation, discovered as a first gaseous second messenger and named molecule of the year in 1992 (61). However, in the presence of oxidative species abundant in the inflammation, NO will react with superoxide and form highly reactive peroxynitrite (ONOO-) (62). Peroxynitrite will oxidize fatty acids and break the lipid chain disrupting membranes, interact with nucleic acids and proteins, and inhibit cell respiration by inhibiting the electron transport chain in mitochondria and slowing the rate of glycolysis by reducing the substrates and cofactors (63). NO synthesis requires nitrogen atoms from substrate arginine and cofactors NADPH and O_2, producing citrulline and NO on a ferrous heme. Another enzyme producing NO is NOS2 - induced by LPS binding

to surface cell receptors. This enzyme is regulated on a transcriptional level since a dimerized unit is the only post-translational requirement for NOS2 to produce NO while the substrate allows it (64).

Besides their role in oxidative bursts and recognizing and phagocytosis of pathogens, neutrophils contain granules with peptides such as defensins and cathelicidins, further priming them for antimicrobial roles. When neutrophil granules fuse with the phagosome, these peptides disrupt the microbial membrane, forming a pore (65). Cathelicidins are activated when cleaved by neutrophil elastase (NE), forming an active antimicrobial peptide (66). Neutrophil elastase inside the granules produces microbicidal activity against Gram-negative bacteria; it can cleave membrane protein A on the outer membrane of *Escherichia coli*. Chemically, this is serine endopeptidase, the most abundant of four neutrophil serine proteases at a concentration of 67,000 molecules (~5 mM) per granule. NE can cleave many substrates, accounting for 80% of the total protease hydrolysis activity in the human body since it cleaves neutral non-aromatic dipeptides like elastin, collagen, and fibronectin. Consistent with the phenomenon of innate immunity, NE can sometimes lead to disease worsening and promote inflammation. Higher numbers of NE in the tissue point to disease severity (67), (68). There are ways to naturally modulate the abundance and localization of NE by antiproteases: α1-antitrypsin, α2-macroglobulin, or secretory leukocyte proteinase inhibitor (SLPI). Recently, a new therapeutic NE inhibitor, Sivelestat, was introduced to treat acute lung injury or acute respiratory distress syndrome with mixed results (figure 3) (69), (70), (67), (71). There is potential therapeutic value in blocking NE in inflammation caused by pathogens or autoimmune disorders. Figure 3 shows additional drugs blocking NE function in various clinical trial stages.

In addition to microbicidal effects, activated neutrophils will undergo programmed cell death, releasing chromatin into extracellular space and forming neutrophil extracellular traps (NETs). This process, called netosis, was first described in neutrophils in 1996 (73). The proteins with antimicrobial properties released along with NETs can be directly harmful to tissues. Elastase and

myeloperoxidase are released with nets along with ROS, and although highly effective against pathogens, contribute to tissue injury (74), (75), (76).
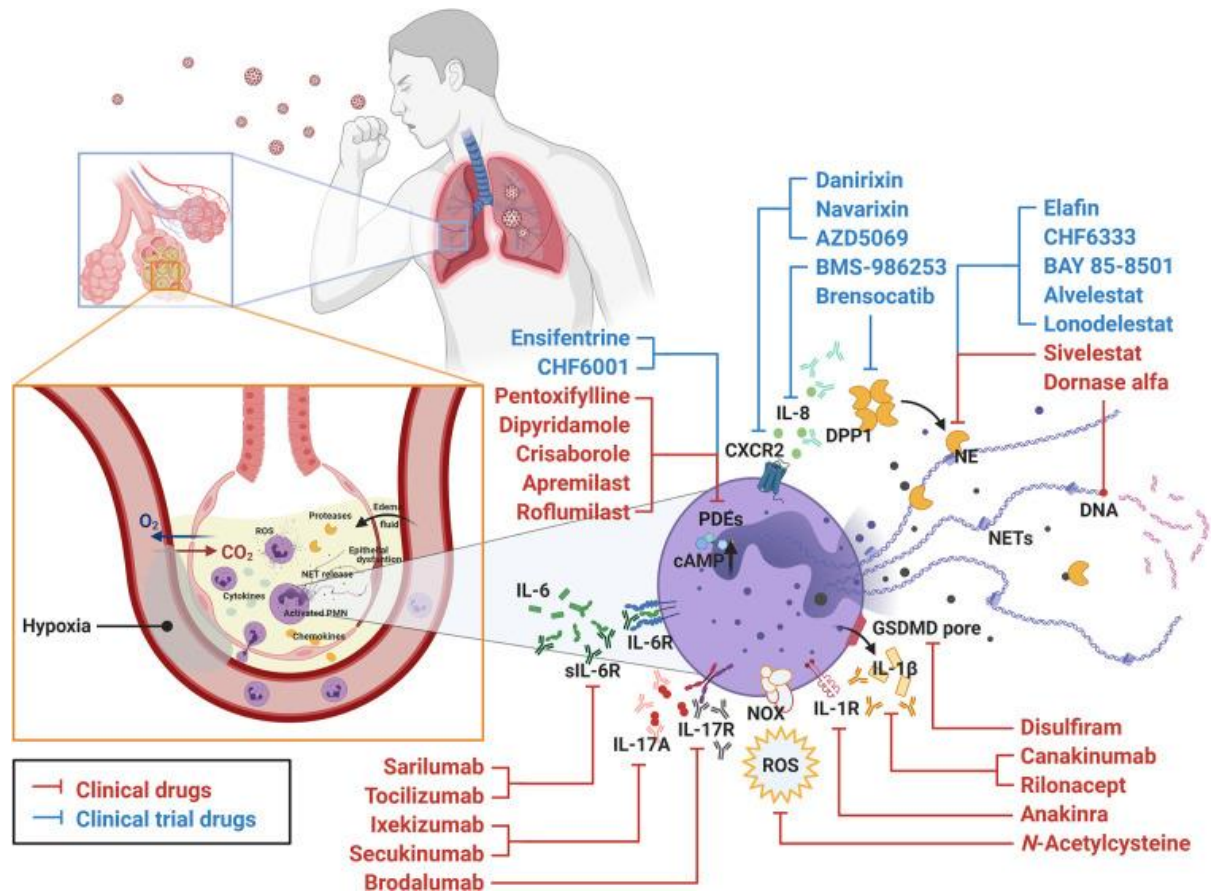
Figure 3 Drugs targeting neutrophils used in acute respiratory distress syndrome
Neutrophils recruited to the site of infection participate in the elimination of the virus, in this case - SARS-CoV-2, but they also contribute to the pathogenesis. The only approved drug for NE inhibition is Sivelestat. Clinical drugs (in red) and clinical trial drugs (in blue) are displayed with their targets in neutrophils. ARDS, acute respiratory distress syndrome; CXCR, C-X-C chemokine receptor; DPP1, dipeptidyl peptidase 1; GSDMD, gasdermin D; IL, interleukin; NE, neutrophil elastase; NET, neutrophil extracellular trap; NOX, NADPH oxidase; PDE, phosphodiesterase; PMN, polymorphonuclear leukocyte; ROS, reactive oxygen species. Figure adapted from *(72)*

The effect of neutrophils is that they can engulf pathogens and present their antigens to T cells, linking innate and adaptive immunity. Besides neutrophils, macrophages, monocytes, and dendritic cells are the body's phagocytes responsible for balancing innate immune response and inflammation.

Macrophages

Macrophages are a resident cell population in homeostasis. They are found in every tissue in the body regardless of the presence of blood vessels. They can undergo local proliferation independent of monocytes and are involved in organ maintenance. Macrophage phagocytosis allows for recycling cell debris; macrophages keep metabolite levels constant, transduce signals from their microenvironment, and initiate inflammation and curative processes. Macrophages have an assortment of receptors sensing their environment and can express some 4500 genes in response to nanomolar concentrations of LPS (77). This state of activation during inflammation is considered a classically activated macrophage when encountering pathogens and reacting to toll-like receptor (TLR) ligands (like LPS) and interferon-γ (IFNγ). In response to PRR, transcription factors like NF-κB, CREB, AP-1, and interferon-regulatory factors (IRFs) are activated to begin cytokine gene transcription. Effector transcription is a rapid process since RNA polymerase is already positioned on tumor necrosis factor (TNF) in homeostasis. TNF further stimulates transcription of inflammatory cytokines, including interleukin (IL)-6, IL-12, and the type I interferons, chemokines, lipid mediators, and antimicrobial peptides are similarly released from classically stimulated macrophages (78). Gene expression is accentuated by IFNγ and phosphorylation of transcription factor STAT1, responsible for the expression of 200 genes, including GM-CSF, IL-12p40, TNF-α, and IL-6. To prevent exaggerated inflammation, there are a variety of mechanisms in place: negative regulation of TLR signaling, anti-inflammatory cytokines (IL-4, IL-10, IL-13 IFNα, TGF-β), epigenetic modifications and miRNA, transcription factors (STAT3), and endogenous regulatory molecules (prostaglandin E2, resolvins, lipoxins) (78), (77). Which of these factors is more dominant can indicate the future action of macrophages.

Regulation of macrophages is a well-controlled and extensively researched process with IL-4 and IFNγ as crucial factors that substantially affect macrophage activity. That activation is described

based on the presence of these two cytokines (figure 3). But the reality is not binary. Populations of

macrophages from one tissue are characterized by an almost continuous spectrum of activation

phenotypes, and each state strikes a balance between inflammation and healing (figure 3). However,

in a disease state, this balance is disrupted, and one subset of macrophages could lead to disease

progression (79) (80), (81). Macrophage regulation is another example of tolerance and inflammation

trade-off.
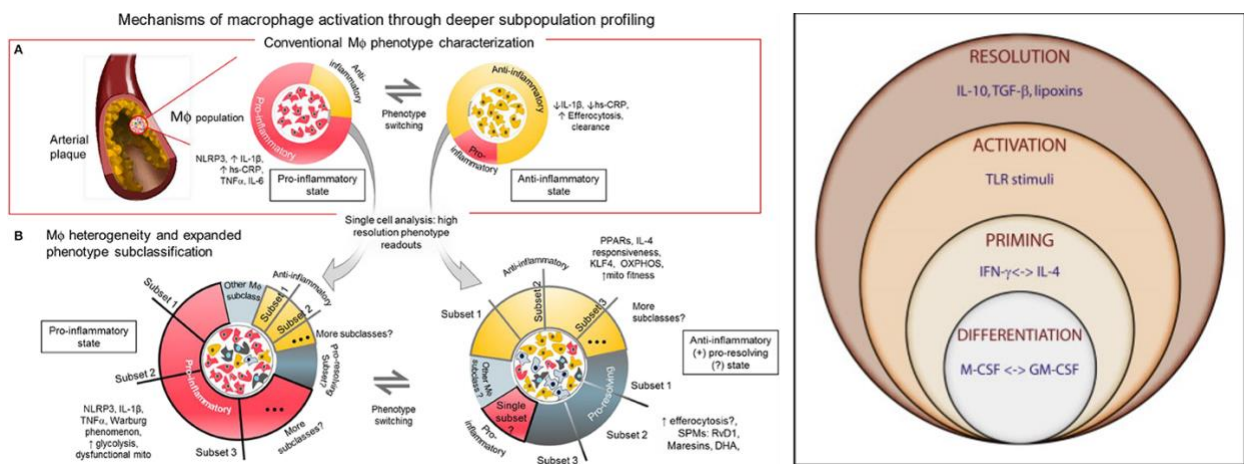
Figure 4 Macrophage activation pathways
The left panel shows subsets of macrophages as pro-inflammatory or anti-inflammatory. Figure
adapted from (79). The right panel shows the traditional way of understanding macrophage activation
as a classical, IFN-γ activated cell or IL-4 alternative polarization. Figure adapted from (80).
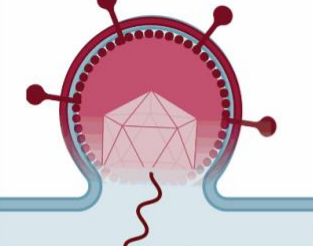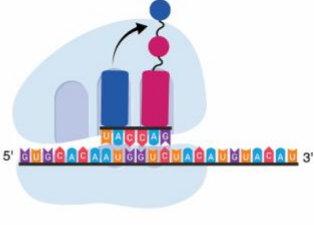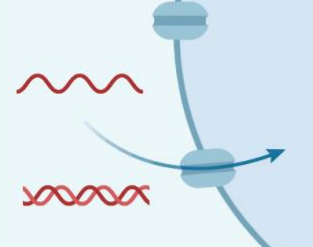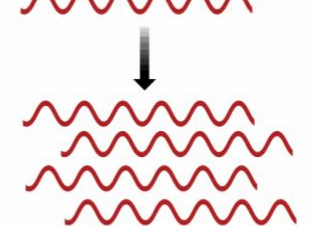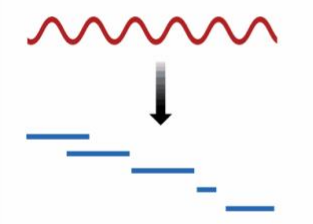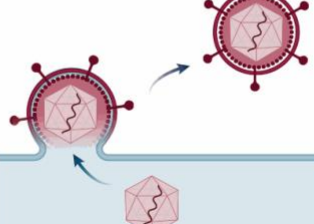
Interferons

Innate immunity's reactive components are myeloid cells and molecular products of cells affected

by infection, mainly interferons. The ability to produce interferons is not restricted to specific cell

types; instead, it can vary depending on the context of the infection and the overall state of the

immune response. Interferons have a wide range of effects, including direct antiviral activities,

modulating the immune response, affecting cell growth, and affirming an adaptive immune

response. Their production is tightly regulated, as dysregulation can lead to chronic inflammation or

autoimmunity (82).

Interferons are part of a larger family of cytokines and, based on their function and receptor binding, are divided into three groups – type I, type II, and type III interferons. Type I interferons include various subtypes, with interferon-alpha (IFN-α) and interferon-beta (IFN-β) being the most well-known, binding to a common receptor complex interferon-alpha/beta receptor (IFNAR). There is only one type II interferon - interferon-gamma (IFN-γ). IFN-γ is distinguished by its receptor, the interferon-gamma receptor (IFNGR), and is produced mainly by adaptive immune response cells, whereas type I interferons are produced in response to viruses. The type III IFN is also a family of effector molecules with similar functions to cytokines of the type I family but restricted activity (83), (84), (82).

Since viruses can infect virtually any cell type, most cells can express type I interferons in response to PRRs (85). Similar to the effect of hormones on different cell types, where one hormone molecule activates a specific gene and promotes a particular function depending on the cell it bounds and the downstream signal cascade, so do the type I interferons activate signal transduction, leading to transcription of hundreds of interferon-stimulated genes (ISGs). In the human genome, ten percent of all genes can be regulated by interferons. Hence, ISGs have an encompassing effect on the cellular environment (figure 5) (83).

Figure 5 Type I interferon-stimulated genes and their targets
Graphical representation of type I interferon-stimulated genes innate immunity effects against viruses. Listed functions of ISGs include blocking pathogen entry in the cell, or nuclear import; inhibiting viral mRNA, and protein synthesis, viral replication and assembly. Abbreviations CH25H - Cholesterol 25-Hydroxylase, IFITM - Interferon Induced Transmembrane Proteins, NCOA7 - Nuclear Receptor Coactivator 7, TRIM5a - Tripartite Motif Containing 5, Alpha, MX - Myxovirus (influenza virus) resistance protein, APOBEC - Apolipoprotein B mRNA Editing Catalytic Polypeptide-like, IFI16 - Interferon Inducible protein 16, PKR - Protein Kinase R, also known as EIF2AK2, IFIT - Interferon-Induced Proteins with Tetratricopeptide Repeats, ZAP - Zinc Finger Antiviral Protein, is encoded by the ZC3HAV1 gene (Zinc Finger CCCH-Type, Antiviral 1), PARP12, or Poly(ADP-Ribose) Polymerase 12, SLFN11 - Schlafen Family Member 11, SAT1 - Spermidine/Spermine N1-Acetyltransferase 1, ISG20 - Interferon Stimulated Gene 20, The OAS - 2'-5'-Oligoadenylate Synthetase, CNP - Cyclic Nucleotide Phosphodiesterase, GBP5 - Guanylate Binding Protein 5. Adapted from (83).

Due to their vast repertoire of action, interferons are well-regulated and short-lived, but their effects

are lingering (86). There are over 450 ISGs with various functions, from blocking viral entry to

interrupting viral replication and assembly, and more effectors are constantly added (87), (88).

Peromyscus leucopus zoonoses

Borreliosis

P. leucopus is a crucial reservoir for the spirochete *Borrelia burgdorferi* in North America, and serves as a primary source of infection for ticks in the larval stage, particularly the black-legged tick (Ixodes scapularis), which is the primary vector for Lyme disease in humans (figure 6) (89), (90). Deer are adult tick hosts, but they are not competent reservoirs for Lyme disease since they are dead-end for bacteria, while humans are accidental hosts, and they develop disease if infected. (91). The first presenting sign of disease in humans is erythema migrans rash, followed by malaise and fever. Complications could occur months or years after initial infection and can present as prominent neurological symptoms, including deficits in concentration, cognition, and memory loss. (92)
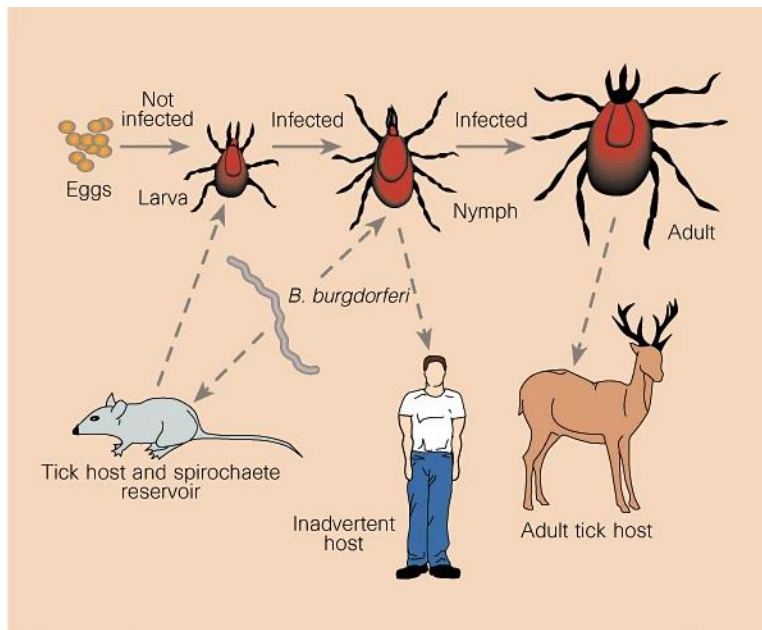
Figure 6 The life cycle of Ixodes tick vectors of Borrelia burgdorferi, the spirochaete agent of Lyme disease
The life cycle of Ixodes tick vectors of *Borrelia burgdorferi*, the spirochaete agent of Lyme disease. Ixode tick larva gets its first blood meal from a small rodent – *P. leucopus*; at this point, if *P. leucopus* is carrying *Borrelia burgdorferi*, the tick will get infected, and next time it takes a blood meal as a nymph, it will infect a host. A host for a nymph can be either *P. leucopus* again or inexpertly human, while deer is the final host *(91)*

On the other hand, the relationship between *P. leucopus* and *B. burgdorferi* is an example of tolerance. The white-footed deermouse is an efficient carrier of B. *burgdorferi* because it can harbor the bacterium without showing signs of illness, allowing the pathogen to persist in natural environments (3), (93). To that extent, the prevalence of Lyme disease in certain areas is closely tied to the population density of *P. leucopus*, as higher deermouse populations can lead to increased numbers of infected ticks. The prevalence of infection in *P. leucopus* with *B. burgdorferi* in the circulation is 10%, while antibody prevalence is seasonal and ranges from 57 - 95% by the end of summer (3).

Upon infection with *B. burgdorferi*, antigens bind to toll-like receptor (TLR) 2 and activate myeloid differentiation primary response 88 (MyD88). Additionally, *B. burgdorferi* antigens trigger the activation of the nuclear factor kappa B (NF-κB) pathway, which is pivotal in the immune response. This sequence of events produces pro-inflammatory cytokines and chemokines, attracting neutrophils, monocytes, and lymphocytes to the site of infection while producing type I interferons (IFNα and IFNβ) and type II interferon - IFNγ. Following innate immunity, *P. leucopus' Th*1 lymphocytes and mononuclear cells produce neutralizing antibodies, specifically IgG1 and IgG3, effectively fighting infection (94).

Another borreliosis-causing agent is *Borrelia hermisii*, a pathogen of relapsing fever. This disease is characterized by recurring episodes of fever, headache, muscle and joint aches, and nausea (95). The fever episodes typically last for several days, followed by periods of recovery, and then may recur. Relapsing fever can be transmitted through the bite of an infected soft tick, while the reservoirs in nature are small mammals (96). *P. leucopus* can be experimentally infected with *B. hermisii* (20).

COVID-19

Coronaviruses are positive-sense single-stranded RNA viruses (+ssRNA) that exhibit a distinctive coronal morphology when observed under an electron microscope. This characteristic crown-like

structure, from which their name is derived (with "corona" being the Latin word for "crown"), is attributed to the presence of spike glycoproteins on their viral envelope. They are classified into four genera; SARS-CoV-2 belongs to beta coronaviruses. Further genomic characterization puts SARS-CoV-2 in the same subgenus as SARS-CoV, and MERS-CoV-2 known to cause severe diseases with mortality up to 10% and 35% (97). The rest of the coronaviruses infecting humans - HKU1, NL63, OC43, and 229E - are associated with mild respiratory tract infections colloquially described as the common cold (98).

In case of SARS-CoV-2 human infection, the virus binds to host cells through an interaction between the receptor-binding domain (RBD) of the viral spike protein (S) and angiotensin-converting enzyme 2 (ACE2). The ACE2 receptor on the respiratory epithelium is the main entry point for the virus. Still, it is also expressed in the digestive system, kidneys, and heart (99). Besides receptor binding, host transmembrane serine protease 2 (TMPRSS2) and probably proteases cathepsin B and L (CatB and CatL) enable cell entry by the proteolytic cleavage of coronavirus S proteins. In the initial stages of infection, viral replication causes tissue damage, especially in the upper respiratory tract. Besides receptors and proteases, the virus requires an endoplasmatic reticulum and Golgi apparatus as well as translation factors for viral mRNA (100). During this time, viral components are exposed to sensory molecules like cytoplasmic PRRs of the host immune system. Also, the interaction of S-protein with the receptor induces changes in the membrane. These events activate the innate immune system, interferon release, and upregulation of ISGs. Besides ISGs, one of the first lines of defense is the upregulation of lymphocyte antigen six family member E (Ly6E) - identified as a coronavirus fusion inhibitor (101). The adaptive immune system is lagging since S-protein is glycosylated, and its epitopes are hindered from neutralizing antibodies (100). In studies with hamsters, which have been used as a model to understand SARS-CoV-2 infection, neutralizing antibodies typically start to be detected within the first week after infection. By days 7-14 post-infection, neutralizing antibody titers

often reach levels that can confer protection, as seen in some studies where re-challenged animals exhibit milder symptoms or no disease, suggesting that these antibodies are functional in terms of viral neutralization (102). The findings in hamsters are similar to those in humans (103).

An unanswered question remains - how did the spillover into the human population happen? Previous studies proposed critical amino acid changes for a possible gain of function in SARS-CoV-2 for spillover in nature from primary hosts. The main viral variable determining host selection is S protein - trimeric glycoprotein, containing two functional subunits. The first subunit (S1) binds to the host ACE2 and harbors RBD. The second subunit (S2) is responsible for the fusion of the viral and cellular membranes. Table 1 shows amino acid residues across species enabling contact between ACE2-RBD in SARS-CoV-2 infection. Data is adapted from (104), (105), (106).

Matching ACE2 amino acid residues with between *P. leucopus* and species susceptible to SARS-CoV-2 infection a good indication that the virus can cause infection in deermice, but will the host get sick? How severe the illness becomes depends on the species, and more importantly, severity varies within species based on genetic factors. Golden hamsters lose weight and get moribund, but widely available hamsters are descendants of a single male and female with very poor genetic diversity.

Table 1 Interaction of RBD with human ACE2

Amino acid residues suggesting enhanced ACE2-RBD contacts in SARS-CoV-2					
ACE2 in select species	AA in position for recognition by SARS-CoV-2 RBD				
	31	35	38	82	353
Human	K	E	D	M	K
Hamster	K	E	D	N	K
Peromyscus leucopus	K	E	D	N	K
Civet	T	E	E	T	K
Bat	K	K	D	N	K
Mouse	N	E	D	S	H
Rat	K	E	D	N	H
Abbreviations ACE2 - angiotensin-converting enzyme 2 , RBD - receptor-binding domain, AA - amino acid					

Some species of *Peromyscus* are successfully infected with SARS-CoV-2 showing mild disease (107), (108), (109), and a potential for spreading the virus (110). So far there was no study of infection of *P. leucopus* with SARS-CoV-2. Confirmed infection of *P. leucopus* with SARS-CoV-2 could mean

that this species would have the capacity for a viral spillback from the human population and become viral reservoirs of SARS-CoV-2.

Microbiome

Initially, bacterial activity was discussed in terms of sepsis. However, microbial activity in the body can be commensal or symbiotic, but it is not necessarily pathogenic. The community of microorganisms and their environment is referred to as the microbiota (111). The microbiome refers to the genome of microorganisms in the environment, and in vertebrates, it represents half of the complete genome of an organism (112), (113), (114). The microbiome of the host has multiple roles, including immunomodulation, breakdown of complex carbohydrates required for complete nutrition and other complex compounds (115), (116). Humans and mice owe as much to their microbiota as nuclear and mitochondria genomes of somatic cells when it comes to tolerating and resolving infections (114). One way to resist the colonization of pathogens is by maintaining the indigenous microbiome of the host (112). Describing and understanding the microbiome is therefore necessary when analyzing a complete genome of somatic or bacterial cells, especially in inflammation. Finally, manipulating the microbiome can improve outcomes in infection and cause tolerance (117), (118). Given the development of bait-delivered oral vaccines targeting *P. leucopus* and plans to genetically modify and release this species, pushing ahead on these interventional fronts, a better understanding of *Peromyscus* microbiota and the gastrointestinal tract is necessary (119). For that same reason, bats as tolerant vectors and reservoirs of infection are investigated for the role of the microbiome in infection (120), (121), (122). However, there is limited information in the literature on the gastrointestinal microbiota of *P. leucopus*, and the studies were restricted to data on bacterial 16S ribosomal RNA sequences (123). The metagenomic approach can fully describe the microbiota, including parasites and DNA viruses. Microbial species can then be further characterized using microbiology techniques described in Chapter 2 (124), (125).

CHAPTER TWO: Lactobacilli and other gastrointestinal microbiota of Peromyscus leucopus, reservoir host for agents of Lyme disease and other zoonoses in North America

Abstract

The cricetine rodent *Peromyscus leucopus* is an important reservoir for several human zoonoses, including Lyme disease, in North America. Akin to hamsters, the white-footed deermouse has been unevenly characterized in comparison to the murid *Mus musculus*. To further understanding of *P. leucopus*' total genomic content, we investigated gut microbiomes of an outbred colony of *P. leucopus*, inbred *M. musculus*, and a natural population of *P. leucopus*. Metagenome and whole genome sequencing were combined with microbiology and microscopy approaches. A focus was the genus *Lactobacillus*, four diverse species of which were isolated from forestomach and feces of colony *P. leucopus*. Three of the species—*L. animalis*, *L. reuteri*, and provisionally-named species "L. peromysci"—were identified in fecal metagenomes of wild *P. leucopus* but not discernibly in samples from *M. musculus*. *L. johnsonii*, the fourth species, was common in *M. musculus* but absent or sparse in wild *P. leucopus*. Also identified in both colony and natural populations were a *Helicobacter* sp. in feces but not stomach, and a *Tritrichomonas* sp. protozoan in cecum or feces. The gut metagenomes of colony *P. leucopus* were similar to those of colony *M. musculus* at the family or higher level and for major subsystems. But there were multiple differences between species and sexes within each species in their gut metagenomes at orthologous gene level. These findings provide a foundation for hypothesis-testing of functions of individual microbial species and for interventions, such as bait vaccines based on an autochthonous bacterium and targeting *P. leucopus* for transmission-blocking.

Introduction

Epigraph: "I have always looked at problems from an ecological point of view, by placing most emphasis not on the living things themselves, but rather on their inter-relationships and on their interplay with surroundings and events." René Dubos, 1981 (126), (127)

Peromyscus leucopus, the white-footed deermouse, is one of the most abundant wild mammals in central and eastern United States and adjacent regions of Canada and Mexico (128), (2). The rodent is an omnivore, consuming a variety of seeds, such as oak acorns, as well as insects and other invertebrates. Its wide geographic range extends from rural areas to suburbs and even cities, and it is especially common in areas where humans and wildland areas interface (129). Conditions permitting, *P. leucopus* is procreatively proliferant, with 20 or more litters during a female's period of fecundity (130).

Although commonly called a "mouse", this species and other members of the genus *Peromyscus* belong to the family Cricetidae, which includes hamsters and voles, and not the family Muridae, which includes *Mus* and *Rattus*. The pairwise divergence time for the genera *Peromyscus* and *Mus* is estimated to be ~27 million years ago (131), (132), approximately the time since divergence of the family Hominidae, the great apes and hominids, from Cercopithecidae, the Old World monkeys (131),(133). While only a minority of a birth cohort of *P. leucopus* typically survive the predation and winter conditions of their first year in nature (23), (15), in captivity *Peromyscus* species can live twice as long as the laboratory mouse or rat (10). *P. leucopus* differs in its social behavior and reproductive physiology from rodents that are traditional experimental models (9), (134).

P. leucopus also merits attention as a natural host and keystone reservoir for several tickborne zoonoses of humans (reviewed in (3)). These include Lyme disease, babesiosis, anaplasmosis, a form of relapsing ever, an ehrlichiosis, and a viral encephalitis. For humans these infections are commonly disabling and sometimes fatal, but *P. leucopus* is remarkably resilient in the face of

persistent infections with these pathogens, singly or in combination. How this species tolerates infections to otherwise thrive as well as it does is poorly understood.

P. leucopus' importance as a pathogen reservoir, its resilience in the face of infection, and its appealing features as an animal model (130), (7), prompted our genetic characterization of this species, beginning with sequencing and annotating its nuclear and mitochondrial genomes (6), (5). The present study represents the third leg of this project, namely the microbial portion of the total animal "genome" for this species. Given the development of bait-delivered oral vaccines targeting *P. leucopus* (135) and plans to genetically modify and release this species (136), (137), pushing ahead on these interventional fronts without better understanding *Peromyscus* microbiota, the gastrointestinal (GI) tract's in particular, seemed shortsighted.

Accordingly, we carried out a combined microbiologic and metagenomic study of the GI microbiome of *P. leucopus*. In our exploration of bacteria, protozoa, and DNA viruses in this species' gut, we included whole genome sequences from cultured organisms as well as large contigs assembled from metagenome reads of uncultured organisms in our in-depth exploration of this. Attention initially focused on animals of a stock colony that has for many years been the major source of animals for different laboratories and spin-off breeding programs, including our own, in North America. The study extended to samples of *P. leucopus* deermice in their natural environments and, for a comparative animal, vivarium-reared *M. musculus* under similar husbandry. While our investigations revealed similarities between the microbiota of the white-footed deermouse and the house mouse, there were also notable differences. These included a greater abundance and diversity of lactobacilli in *P. leucopus*. The investigation of four *Lactobacillus* species, particularly in their niches in the stomach of *P. leucopus*, was a special emphasis. A comparison of the GI microbiota of a natural population of *P. leucopus* and the stock colony animals revealed several species in common, albeit with larger variance among the wild animals.

Results

High coverage sequencing of fecal metagenome

At the start of our investigation there was only limited information in the literature on the GI microbiota of *P. leucopus*, and the studies were restricted to data on bacterial 16S ribosomal RNA genes sequences, as in the examples of references (138) and (123). For a fuller description of the microbiota, including parasites and DNA viruses, as well as bacteria, of representatives of this species, we took an untargeted, metagenomic approach. This was followed by bioinformatic characterization of the breadth of organisms represented in the metagenome and then assembly and annotation of large chromosomal contigs of selected microorganisms in the metagenome.

We began with deep sequencing of a library of DNA extracted from a pooled sample of fecal pellets collected from two adult males and two adult females of the LL stock colony. This yielded 332,279,332 paired-end reads of average length 247 nt, or ~8 x 10^{10} bp, after quality control and trimming of adapters. The mean % GC content was 47; 90% of the trimmed reads had PHRED scores of ≥ 30. The reads were characterized as to families of bacteria, parasites, and DNA viruses at the metagenomic server MG-RAST (http://mgrast.org). Annotated proteins accounted for 65% of the reads, followed by unknown proteins at 34%, and then ribosomal RNA (rRNA) genes at 0.8%. The rarefaction curve became asymptotic at ~200,000 reads and a species count of 9000. The alpha diversity for this pooled sample was 250 species. By phylum 94% of the matched reads were either Bacteroidetes (60%) or Firmicutes (34%). Higher level functional categories included carbohydrates (16.4% of reads), clustering-based subsystems (14.8%), protein metabolism (8.9%), amino acids and derivatives (7.8%), RNA metabolism (6.6%), and DNA metabolism (5.7%).

A portion of the DNA was also submitted to a reference laboratory for 16S rRNA determinations from metagenomic analysis of microbiota. As illustrated in figure 7, for the 20 most abundant taxa at the family or higher, there was concordance between the methods in the rankings. The most common

families by the metagenomic accounting were members of the gram-negative bacterial order Bacterioidales (*Bacterioidaceae*, *Porphyromonadaceae*, *Prevotellaceae,* and *Rikenellaceae*), the gram-positive phylum Actinobacteria (*Bifidobacteriaceae* and *Coriobacteriaceae*), or the gram-positive phylum Firmicutes (*Bacillaceae*, *Enterococcaceae*, *Lactobacillaceae*, unclassified Clostridiales, *Eubacteriaceae*, *Lachnospiraceae*, *Peptococcaceae*, *Ruminococcoceae*, Thermoanaerobacterales Family III, *Erysipelotrichaceae*, and *Veillonellaceae*). Two exceptions were organisms of the families *Spirochaetaceae* of the phylum Spirochaetes and *Helicobacteraceae* of the phylum Proteobacteria.
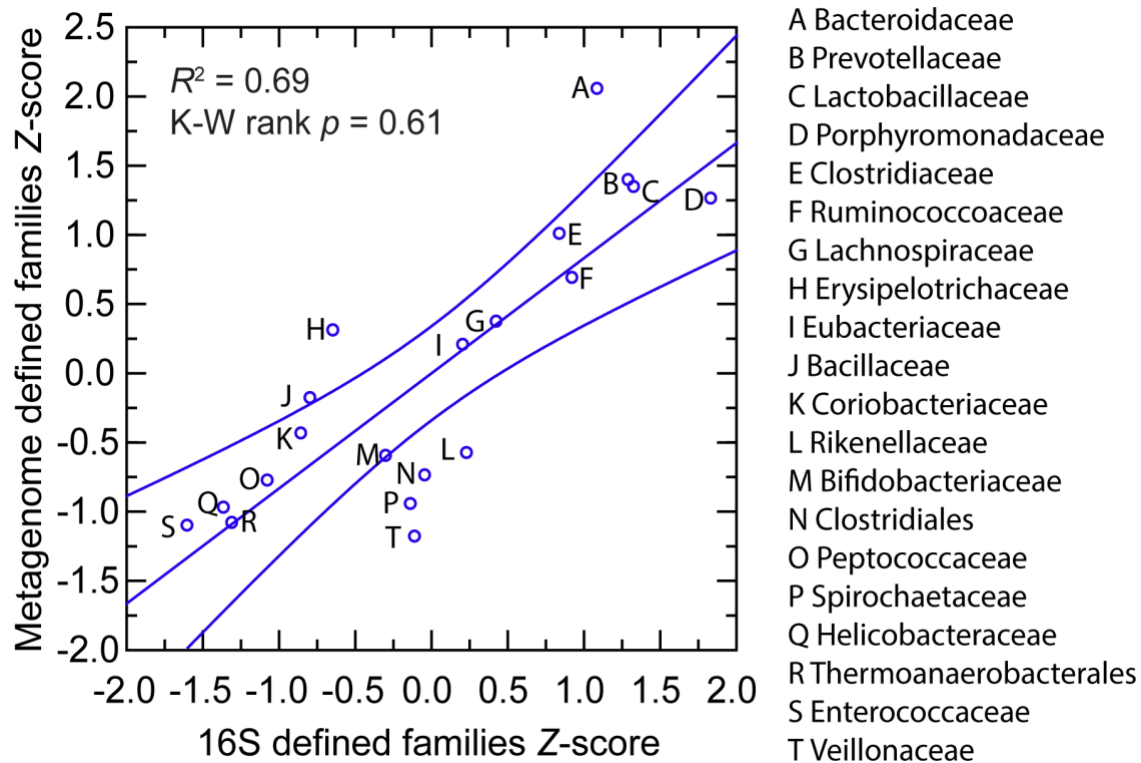
Figure 7 Scatter plot of relative metagenome abundance of P. leucopus
Scatter plot of relative abundances of commonly occurring bacterial families or orders in fecal metagenomes of *Peromyscus leucopus* LL stock by 16S ribosomal RNA gene criterion (*x*-axis) and by genome-wide metagenome criterion (*y*-axis).The values of the two methods were normalized by *Z*-score. The different taxa are indicated in the graph by capital letters, which are defined in the box to the right. The linear regression curve and its 95% confidence interval is shown. The coefficient of determination (R^2) value and the Kruskal-Wallis (K-W) test by ranks *p* value are given.

A de novo assembly yielded 16,945 ungapped contigs of ≥ 10 kb from 197,369,943 reads and totaling 385 Mb of sequence with an average coverage of 104X. Of the total, 219 contigs were ≥ 100 kb in length and with ≥ 30X coverage. These were used in searches of non-redundant nucleotide and protein databases for provisional classifications. The identified taxa included Bacteroidales, Clostridia, *Clostridiaceae*, Erysipelotrichales, *Lactobacillaceae*, *Muribaculaceae*, Firmicutes, and *Spirochaetaceae*. Three organisms represented among the high coverage contigs could be unambiguously classified as to species: *Lactobacillus animalis*, which is considered in detail below, and two *Parabacteroides* species: *distasonis* and *johnsonii*. Another *Lactobacillus* species represented among the highly ranked contigs could not be identified with a known species represented in the database (139).

Among the high coverage contigs were also representatives of Rhodospirillales of the class Alphaproteobacteria, *Mycoplasmataceae* of Mollicutes, and the little-characterized group of bacteria called Elusimicrobia (140). Nearly as prevalent were organisms closely related to the phylum-level designation *Candidatus* Melainabacteria (141). On the list of organisms identified by searches with metagenomic contigs of databases and cumulatively accounting for 95% of the matched reads were the unexpected finding of the protozoan taxon *Trichomonadidae* with 41,614 or 0.12% of the reads. *Enterobacteriaceae* at 0.3% accounted for a relatively small proportion of matched reads.

As is the case with human gut microbiome (142), Bacteroidaceae, Lachnospiraceae, Prevotellaceae, and Ruminococcaceae were abundant in the gut metagenome and cumulatively accounted for approximately half of the identified families in the *P. leucopus* sample. One difference between humans and this *P. leucopus* sample was the much higher prevalence in *P. leucopus* of the family Lactobacillaceae, which on average represented only ~0.2% of metagenomes in a European population (142) and by 16S sequencing ≤ 0.4% on the fecal microbiota in other studies (143). A

higher proportion of lactobacilli in the fecal microbiota was previously noted in other rodents (144), and specifically in *Peromyscus* species (138).

Selected taxa

<u>Escherichia coli</u>

Although *Enterobacteriaceae* were infrequently represented among the metagenomic sequences, their cultivability under routine laboratory conditions and the availability of a vast database for some species prompted our isolation of *Enterobacteriaceae* from LL stock *P. leucopus* fecal pellets on selective media. The predominant isolate on the plates was an *Escherichia coli*, which we designated LL2. The whole genome sequence of isolate LL2's chromosome and plasmids was sequenced and assembled using a hybrid of long reads and short reads (Table 2) for an overall coverage of 90X. The chromosome in two contigs of 3.4 Mb and 1.6 Mb totaled 5.0 Mb with a GC content of 50%.

E. coli LL2 had the following MLST schema types (http://pubmlst.org or http://enterobase.org): Achtman ST-278, Pasteur ST-357, and ribosomal protein ST-122394. The ribosomal protein profile was unique among thousands of isolates in the database. The 121 kb, 56.5 kb, and 91 kb plasmids of strain LL2 were similar to the following *E coli* plasmids, respectively: a 185 kb plasmid (NC_007675) found in an avian strain, a 58 kb plasmid (CP024858) of a multiply antibiotic-resistant human isolate, and an 89 kb plasmid (CM007643) in an organism isolated from sewage. *E. coli* LL2 was susceptible to ampicillin, ciprofloxacin, gentamicin, and sulfamethoxazole-trimethoprim by in vitro testing.

The chromosome was notable for the following: CRISPR-Cas1 and–Cas3 arrays; ISas1, ISNCY, IS3, IS110 and IS200 family transposases; restriction-modification systems; fimbria and curli biosynthesis and transport systems; type II toxin-antitoxin systems; and type II, type III and type VI secretion systems. The plasmids encoded fimbrial and pilin proteins, type I, type II, and type IV secretion systems, colicins, CdiA-type contact-dependent inhibition toxin, and three conjugative transfer systems, but no discernible coding sequences for antibiotic resistance.

Table 2 Resources for the microbiome study

Organism and strain	Description	BioProject	BioSample	SRA or MG-RAST [a]	Accession No.
Cultured *Lactobacillus johnsonii* LL8	2,045,501 bp WGS chromosome; 40 contigs	PRJNA589091	SAMN13266521	SRX7128459	WKKC00000000
Cultured *Lactobacillus johnsonii* LL8	75,746 bp plasmid	PRJNA589091	SAMN13266521	SRX7128459	CM019125
Cultured *Escherichia coli* LL2	Chromosome; 3,345,873 and 1,610,537 bp contigs	PRJNA533838	SAMN11468944	SRR9087223 SRR9087224	VBVB01000001 VBVB01000002
Cultured *Escherichia coli* LL2	121,192 bp plasmid	PRJNA533838	SAMN11468944	SRR9087223 SRR9087224	CM017030
Cultured *Escherichia coli* LL2	90,617 bp plasmid	PRJNA533838	SAMN11468944	SRR9087223 SRR9087224	CM017032
Cultured *Escherichia coli* LL2	56,474 bp plasmid	PRJNA533838	SAMN11468944	SRR9087223 SRR9087224	CM017031
Gut metagenome	50–450 kb contigs of fecal metagenome	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	JAAGKN000000000
Uncultured *Helicobacter* sp. LL4	290,716 bp 172,100 bp 203,294 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN577567 MN577568 MN577569
Uncultured *Helicobacter* sp. LL4	16S ribosomal RNA gene, partial	n.a. [b]	n.a.	n.a.	MT114577
Uncultured Candidatus Melainabacteria bacterium isolate LL20	270,170 bp chromosome fragment	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN577570
Uncultured Elusimicrobia bacterium LL30	232,820 bp 215,518 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN577571 MN577572
Uncultured Clostridiales bacterium LL40	438,773 bp chromosome fragment	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN577573
Uncultured *Spirochaetaceae* bacterium LL50	277,828 bp chromosome fragment	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN577574
Uncultured *Prevotella* sp. LL70	456,702 bp 379,405 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN990733 MN990734
Uncultured Rhodospirillales bacterium LL75	145,048 bp 150,471 bp 130,339 bp 104,625 bp 177,737 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN990728 MN990729 MN990730 MN990731 MN990732
Uncultured *Mycoplasmataceae* bacterium LL85	126,601 bp 100,243 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MN991199 MN991200
Uncultured *Muribaculaceae* bacterium LL71	125,326 bp 103,384 bp chromosome fragments	PRJNA540317	SAMN11533720	mgm4799371.3 mgm4799372.3	MT002444 MT002445
Uncultured *Tritrichomonas* sp. LL5	1,501 bp of small subunit ribosomal RNA	PRJNA540317	SAMN13920683	mgm4864879.3	MN120899
Uncultured *Tritrichomonas* sp. LL5	989 bp partial iron hydrogenase gene of hydrogenosome	PRJNA540317	SAMN13920683	mgm4864879.3	MN985504
Uncultured *Tritrichomonas* sp. LL5	5298 bp fragment with DNA polymerase type B, organellar and viral family protein	PRJNA540317	SAMN13920683	mgm4864879.3	MT002461
Uncultured *Lactobacillus* sp. ("peromysci") BI7442	*rpsA, ftsK, ftsZ, dnaA, dnaN, recD, ileS, recA, topA*	PRJNA593618	SAMN13482862	SRX7285441	MN792760 -MN792768
Uncultured *Lactobacillus animalis* 7442BI	51 large and small ribosomal proteins	PRJNA593618	SAMN13482862	SRX7285441	MN817867 -MN817918

Serial dilutions of feces of LL stock 20 animals (11 females and 9 males) in phosphate-buffered saline and plated on agar selective for gram-negative enteric bacteria yielded a mean (asymmetric 95% confidence interval) of 3,491 (677–18,010) colony-forming units (cfu) of *E. coli* per g of feces. This low density was consistent with the findings from metagenomic sequencing.

The origin of this *E. coli* strain in these animals is obscure. Its source may have been the vivarium for the breeding stock of the colony, which includes other *Peromyscus* species besides *P. leucopus*. But the strain appears to be stably maintained among the gut microbiota of this population

of *P. leucopus*. This adaptation and its susceptibilities to commonly-used antibiotics may make it a candidate as a vector for delivering oral vaccines to this species (145).

Lactobacillus

We isolated lactobacilli from fecal pellets of stock colony *P. leucopus* on plates of selective medium that were incubated under microaerophilic and hypercapnic conditions at 37 °C. Four different species were identified. The genomes of three of organisms, namely *L. animalis* strain LL1, *L. reuteri* strain LL7, and a new species, designated as *Lactobacillus* sp. LL6 and provisionally named as "L. peromysci", have been reported (139). The fourth genome, of the LL8 strain of *L. johnsonii*, is described first here (Table 2). *L. johnsonii*'s chromosome from cumulative contigs was 2,045,501 bp, about the same size as that of "L. peromysci" at 2,067,236 bp, but shorter than the 2,280,577 bp length for *L. animalis* and 2,205,740 bp for *L. reuteri*. The % GC content of "L. peromysci" at 33.5 was closer to *L. johnsonii* (34.4) than to either *L. animalis* (41.0) or *L. reuteri* (38.9).

Figure 8 is a distance phylogram of aligned sites of 16S ribosomal RNA genes for the four different lactobacilli. These were distributed across four major groups of the genus *Lactobacillus*. The phylogenetic relationships were examined in more depth by multilocus sequence typing of the 53 genes for ribosomal proteins. These were identified in the genomes, compared with other deposited sequences in the ribosomal MLST dataFbase (https://pubmlst.org) (146), concatenated, and then aligned with analogously concatenated DNA sequences from related species. Bacteria with identical sequences for the 53 ribosomal protein genes were not found in the rMLST database of 133,460 profiles. The % GC contents of the concatenated coding sequences were 39.5, 40.8, 42.2, and 42.3 for *L. johnsonii*, "L. peromysci", *L reuteri*, and *L. animalis*, respectively. Figure 9 shows the distance phylograms for ~20 kb of aligned positions for the four species, each grouped with other strains or species within their respective phylogenetic clusters.
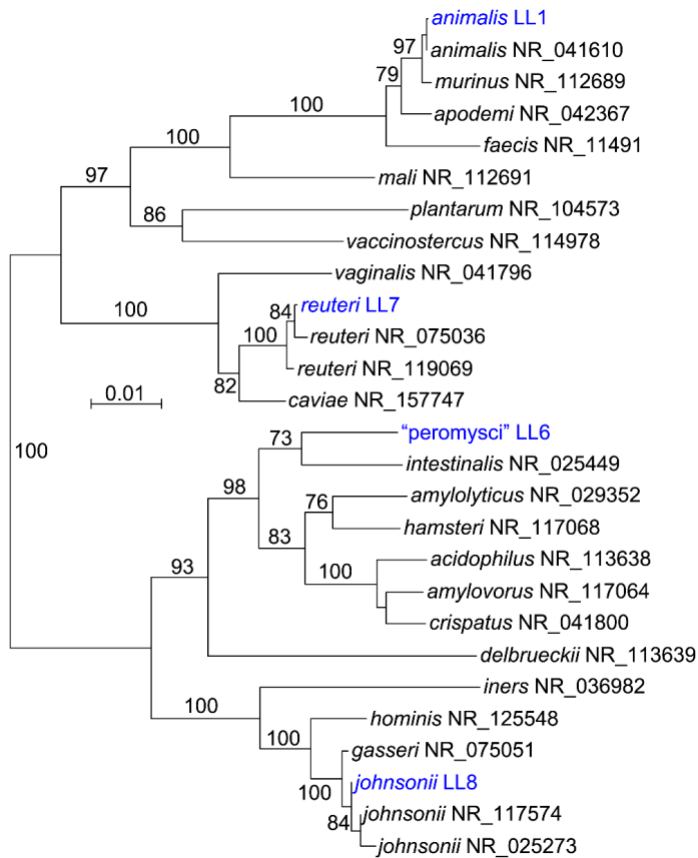
Figure 8 Neighbor-joining distance phylogram Lactobacillus *species*
Neighbor-joining distance phylogram of 1420 aligned positions of 16S ribosomal RNA genes of the culture isolates of four *Lactobacillus* species from *Peromyscus leucopus* and selected other *Lactobacillus* spp. The sources for the accession numbers for the strains are given in Methods (*L. animalis, L. reuteri,* and "L. peromysci") or in Table 2. The other organisms represented are from Reference RNA sequences database of the National Center for Biotechnology Information; the accession numbers are given after the species name. The scale for distance by criterion of observed differences is indicated. Percent bootstrap (100 iterations) support values of ≥ 90% at a node are shown.

"Lactobacillus peromysci" was distant from other sequenced lactobacilli by rMLST (panel A), as well as by its 16S ribosomal RNA gene (figure 8). The nearest taxon in the sequence alignment was *L. intestinalis*, which was first isolated from the intestines of *M. musculus* and other murids (147). The unique ST for the rMLST for strain LL6 of this organism is 115326.

Draft and complete genomes of numerous *L. reuteri* strains have been sequenced, for example, strain Byun-re-01, which was isolated from *M. musculus* small intestine (148). Many of these are utilized in the fermented foods industry, such as production of kimchi, or as dietary supplements,

but others were isolated as constituents of the GI microbiota of several varieties of animals. *L. reuteri* strain LL7 was in a cluster that mainly comprised isolates from *M. musculus*.
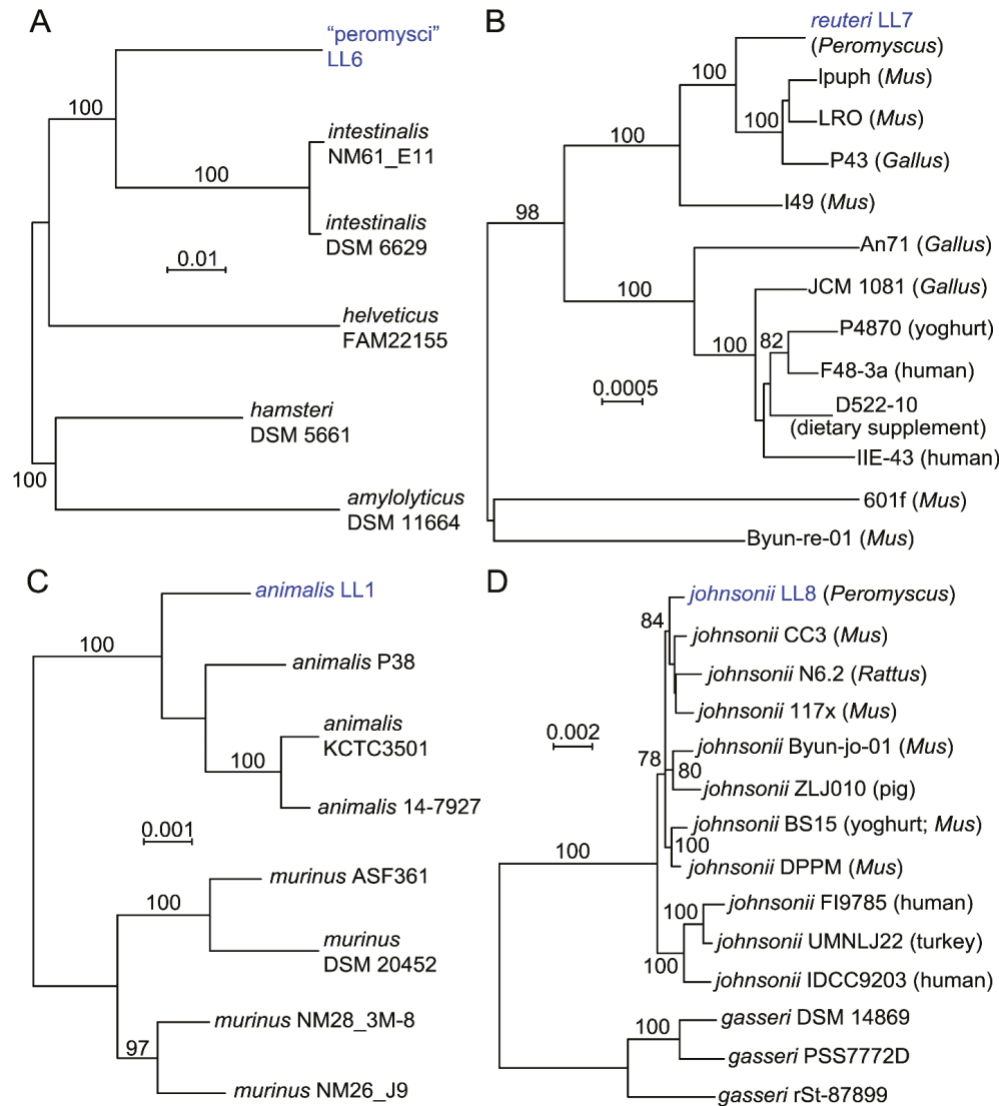
Figure 9 Neighbor-joining distance phylograms Lactobacillus isolated from P. leucopus
Neighbor-joining distance phylograms of codon-aligned, concatenated nucleotide sequences for complete sets of ribosomal proteins of "L. peromysci" (panel A), *L. reuteri* (panel B), *L. animalis* (panel C), and *L. johnsonii* (panel D) of *P. leucopus* compared with *Lactobacillus* spp. (strain identifier) of other sources.The scales for distance by Jukes-Cantor criterion are indicated in each panel. Percent bootstrap (100 iterations) support values of ≥ 75% at a node are shown. In panels B and D the host animal or other origin for a given isolate are given in parentheses.

L. animalis and *L. murinus* are closely related species that primarily have been associated with GI microbiota of rodents and some other mammals. Isolate LL1 grouped with representatives of *L. animalis* in the analysis (panel C) and not *L. murinus* (149). LL1's 16 ribosomal RNA sequence

was identical to that of the type strain ATCC 35046 of *L. animalis* (150) at 1488 of 1489 positions (GCA_000183825) (151). Another pair of closely-related species are *L. johnsonii* and *L. gasseri*, for which there are several sequenced genomes. The LL8 isolate from fecal pellets of *P. leucopus* clustered with *L. johnsonii* strains from mice and rats (panel D). More distant were strains of *L. johnsonii* isolated from humans and a bird; more distant still were representatives of *L. gasseri*.

Plasmids were identified in each of the four species on the basis of a circularly permuted sequence for a contig and presence of coding sequences that were homologous to known plasmid replication or partition proteins (Table 2). Large plasmids of 179 kb and 76 kb were present in *L. reuteri* and *L. johnsonii*, respectively. *L. animalis* and "L. peromysci" had small plasmids of 4 kb and 7 kb, respectively. Megaplasmids of greater than 100 kb have been observed in other *Lactobacillus* spp. (152). In all genomes there was evidence of lysogenic bacteriophages or their remnants. All species except *L. reuteri* discernibly had coding sequences for Class I or Class III bacteriocins or their specific transport and immunity proteins .

Table 3 summarizes differentiating genetic profiles among the four species for 11 selected genes or pathways. Two species, *L. reuteri* and "L. peromysci", had coding sequences for a urease, which could provide for tolerance of acidic conditions, such as in the stomach. A urease had previously been identified in a *L. reuteri* strain that was considered a gut symbiont in rodents (153). The four species had *secY1-secA1* transport and secretion systems. Accessory Sec systems (*secY2-secA2*) were identified in genomes of *L. reuteri*, *L. johnsonii*, and *L. animalis* but not in "L. peromysci". The LL7 strain of *L. reuteri* on its megaplasmid also had coding sequences for a third SecY-SecA system. An accessory Sec system was involved with adhesion and biofilm formation in the Lactobacillales bacterium *Streptococcus pneumoniae* (154). A coding sequence for an IgA protease was identified in *L. johnsonii* but not in the other three species. An IgA protease in another strain of *L. johnsonii* was

associated with long-term persistence in the gut of mice (155). The presence or absence of other genes or pathways that differentiated between the four species were an L-rhamnose biosynthesis pathway in one species, a *luxS* gene associated with a quorum sensing system in *L. reuteri* and *L. johnsonii* (156), a type 1 CRISPR-Cas3 array in "L. peromysci" (155), pathways for thiamine biosynthesis (157) and for reduction of nitrate (158) in three species, an arginine deiminase and its repressor in *L. reuteri* (159), and a type VII secretion system in *L. animalis* (160).

Table 3 Selected genes and pathways in four species of Lactobacillus of the gastrointestinal microbiota of Peromyscus leucopus.

Species	Urease	secY2-secA2	secY3-secA3	IgA protease (pfam07580)	L-rhamnose pathway	luxS	Type 1 CRISPR Cas	Thiamine biosynthesis	Nitroreductase nfnB-nifU	Arginine deiminase/ repressor	Type VII secretion system
reuteri	+	+	+	-	-	+	-	-	-	+	-
johnsonii	-	+	-	+	-	+	-	+	+	-	-
animalis	-	+	-	-	-	-	-	+	+	-	+
"peromysci"	+	-	-	-	+	-	+	+	+	-	-

Of the four species found in *P. leucopus* feces, only *L. johnsonii* and *L. reuteri* have been commonly isolated from human feces (143). While various strains of *L. reuteri*, *L. johnsonii*, and either *L. animalis* or the closely related *L. murinus* have been observed among the GI microbiota of *M. musculus* represenstatives (161), an organism similar to "L. peromysci" has not. Whether this is an indication of a restricted host range or a specific adaptation for this bacterium is considered below. *Lactobacilliaceae* organisms were previously noted in the gut microbiota of *Peromyscus* species but were not further classified as to species or otherwise characterized (138), (123)

Helicobacter

Among the assembled metagenomic contigs were three totaling 666,100 bp of a *Helicobacter* genome (Table 2). The contigs had non-overlapping in genetic content, and blast searches with translated genes from each of the 3 contigs yielded the identical rankings of taxa for homologous proteins. On these bases, we concluded that the contigs represented a single type of *Helicobacter* bacterium, and designated it strain LL4. Using the DNA sequences for 53 ribosomal proteins of this organism, we compared it with similar sets from other *Helicobacter* spp. (panel A of figure 10). This analysis showed that the organism was near-identical to orthologous sequences of *Helicobacter* sp. MIT 05–5293 (accession JROZ02000000), which had been cultivated from a wild *P. leucopus* captured in Massachusetts (162); J.G. Fox, personal communication). This finding

indicated that the organism was autochthonous for *Peromyscus* and had not been acquired from

another rodent housed in the same vivarium. Baxter et al. reported "abundant representatives" of the

genus *Helicobacter* in wild *P. maniculatus* and *P. leucopus* captured in Michigan (138).
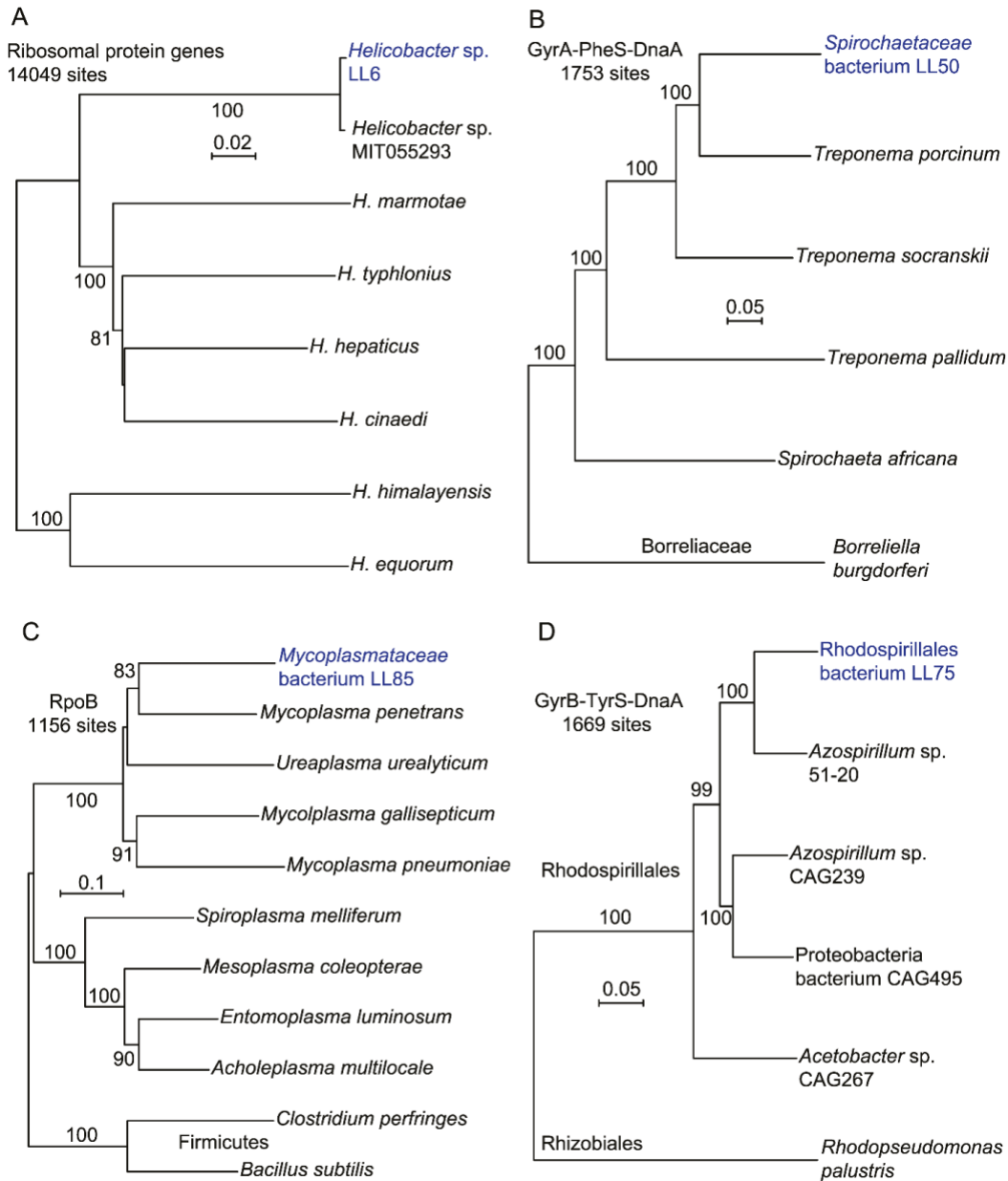
Figure 10 Neighbor-joining distance phylograms of Spirochettes and Helicobacter sp.
Neighbor-joining distance phylograms of concatenated nucleotide (panel A) or amino sequences
(panels B-C) of *Helicobacter* spp. (panel A), *Spirochaetaceae* bacteria (panel B), Mollicutes and
Firmicutes bacteria (panel C) and Rhodospirillales bacteria (panel D) of gut metagenome
of *P. leucopus* and from other sources. The respective phylogenetic analyses used concatenated
sequences of the following: ribosomal protein genes (panel A); the DNA gyrase A (GyrA), phenylalanyl
t-RNA synthase, alpha subunit (PheS), and chromosomal replication iniator protein (DnaA) (panel B);

DNA-directed RNA polymerase, beta-subunit (RpoB) (panel C); and DNA gyrase B (GyrB), tyrosyl t-RNA synthase (TyrS), and DnaA (panel D). The distance criteria were Jukes-Cantor for the codon-aligned nucleotide sequences and Poisson for amino acid sequences. The scales for distance are shown in each panel. Percent bootstrap (100 iterations) support values of ≥ 80% at a node are shown.

Strains LL4 and MIT 05–5293 are in a cluster of species known as "enterohepatic" *Helicobacter* for their primary residence in the intestine rather than the stomach and for their frequent presence in liver tissue (163). These species may not be benign. *H. hepaticus* is associated with hepatitis, bowel inflammation, and carcinoma (164), and *H. typhlonius* is associated with reduced fecundity in mice (165).

Spirochaetaceae

The phylogenetic relationship of the uncultured *Spirochaetaceae* bacterium LL50 to selected other spirochetes is shown in panel B of figure 10. The organism was classified in the genus *Treponema* by the MG-RAST analysis program. Bacteria that have assigned to this genus are highly divergent and include free-living organisms in a variety of environments, symbionts of termites, the agent of syphilis, and gut residents, such as *T. porcinum*, which was isolated from the feces of pigs (166). More distant still was the agent of Lyme disease, *Borreliella burgdorferi*, of the family *Borreliaceae* (167). In our view naming the organism as a "treponeme" on the evidence to date would provide little insight about its role in the microbiome and may even be misleading.

Seven other bacterial taxa

The *Mycoplasmataceae* bacterium LL85 (panel C of figure 10) was unlike any other mollicute represented in the database but was in a cluster with vertebrate-associated species, like *M. pneumoniae*. But there is also deep branching in this tree, as the tree including two Firmicutes as outgroups shows. Panel D is a phylogram of selected alphaproteobacteria and includes the organism LL75. The algorithmic analysis identified this at the genus level as *Azospirillum*, which is a largely uncharacterized taxon with highly divergent members. While assignments as to genus or

family are uncertain at this time, LL75 clustered within the order Rhodospirillales and not with rhizobacteria.

Table 2 lists five other types of novel bacteria that were identified in the *P. leucopus* gut metagenome and partially sequenced and annotated. These were a *Candidatus* Melainabacteria bacterium (isolate LL20), an Elusimicrobia bacteria (isolate LL30), a Clostridiales bacterium (isolate LL40), a *Prevotella* species (isolate LL70), and a *Muribaculaceae* bacterium (isolate LL71). *Candidatus* Melainabacteria is either a non-photosynthetic sister phylum of cyanobacteria or a class within the phylum Cyanobacteria (141). Besides a variety of environmental sources, including hot springs and microbial mats, these poorly-characterized organisms have also been identified in the feces of humans and other animals. The phylum Elusimicrobia, formerly "Termite Group 1" (168), is a strictly-anaerobic, deeply-branched lineage of gram-negative bacteria, representatives of which were first observed in the hindgut of termites (140). The family *Muribaculaceae* (formerly "family S24-7") of the order Bacteroidales were first identified among gut microbiota of mice and subsequently in the intestines of other animals, including humans and ruminants (169).

DNA viruses

Of 112,677,080 reads of the metagenome high-coverage sequencing of the LL stock animals, 97,147 (0.09%) were assigned to one of 28 DNA virus families. Three classifications accounted 92% of the reads: *Siphoviridae* (50%), which are bacteriophages with long contractile tails; *Myoviridae* (21%), which are bacteriophages with contractile tails; and "unclassified viruses" (21%). At the species level, 31,812 (68%) of the 46,904 *Siphoviridae*-matching reads mapped specifically to bacteriophages of *Lactobacillus* spp.

Tritrichomonas protozoan

Intestinal flagellated protozoa named "Trichomonas muris" or "Tritrichomonas muris" had previously been identified in wild *P. leucopus* and *P. maniculatus* (170). While laboratory mice are typically free

of intestinal protozoa, (171) the anaerobic *Tritrichomonas muris* has been reported in some populations of colony *M. musculus* (172). To further investigate the protozoa that were provisionally identified as "Trichomondidae" at the family level in the metagenome analysis, we euthanized 14 healthy adult animals (6 females and 8 males) and examined fresh cecal contents by phase microscopy. Six of the LL stock animals had been born at the PGSC facility, and 8 had been born at U.C. Irvine.

In each of the 14 animals examined there were numerous motile flagellates consistent in morphology with *T. muris* in their ceca (173). These were each at a cell density of ~10^6 per milliliter of unconcentrated cecal fluid (figure 11). The entire ceca and their contents from two adult females and two adult males were subjected to DNA extraction, library preparation from the DNA, sequencing, and de novo assembly of contigs.
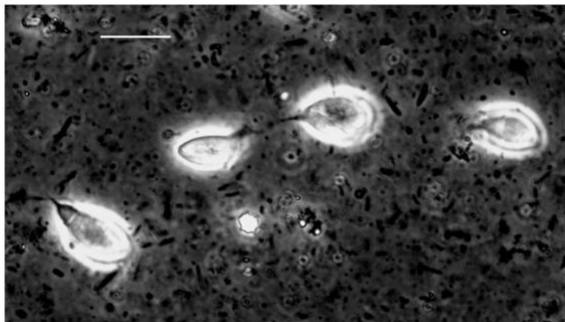
Figure 11 Photomicrograph of live Tritrichomonas flagellated protozoan
Photomicrograph of live *Tritrichomonas* flagellated protozoan in the cecal fluid of *P. leucopus* LL stock. Four organisms against the background of intestinal bacteria were visualized in the wet mount by differential interference microscopy. Bar, 10 μm.

Figure 12 shows phylograms of nucleotide sequence of the small subunit (SSU) ribosomal RNA gene and of the partial amino acid sequence of the iron hydrogenase of the hydrogenosome of anaerobic protozoa (174). The SSU of isolate LL5 indicates that it is probably synonymous with *Tritrichomonas muris*, for which only a SSU sequence was available. The sequence of the iron hydrogenase further supported placement in the genus *Tritrichomonas*. *Histomonas melagridis*, a sister taxon by this analysis, is recognized as a pathogen of poultry. Another sequence of the LL5 organism encodes a

type B DNA polymerase (Table 2), which likewise matched closely with an ortholog in the *Tritrichomonas foetus* genome sequence. *T. foetus* is a sexually-transmitted pathogen of cattle (175) and a cause of chronic diarrhea in domestic cats (176).
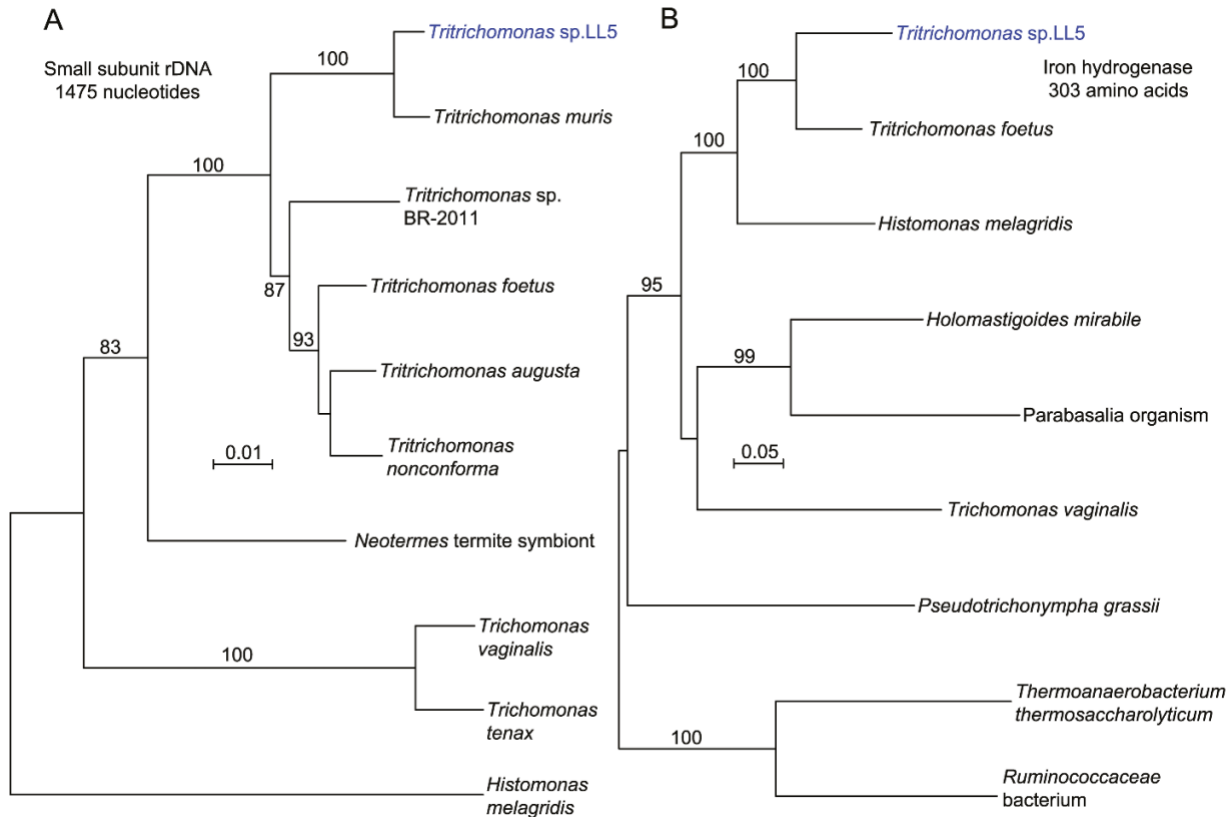
Figure 12 Neighbor-joining distance phylograms of Tritrichomonas sp.
Neighbor-joining distance phylograms of nucleotide (panel A) or amino acid (panel B) partial sequences of small subunit ribosomal RNA gene (rDNA) (panel A) and iron hydrogenase protein (panel B) of *Tritrichomonas* sp. LL5 of *P. leucopus* and selected other parabasilids and other microbes. The distance criteria were observed differences for nucleotide alignment and Poisson for amino acid alignment. The scales for distance are shown in each panel. Percent bootstrap (100 iterations) support values of ≥ 80% at a node are shown.

Whether the *T. muris* is a commensal shared across natural populations of *Peromyscus* or a parasite acquired from another rodent during the colony's history in a vivarium remains to be determined. As related below, there is sequence evidence of the same or related organism in several wild animals. Whatever the case, these organisms may have an effect on immune responses of *P. leucopus*, as has been reported for *T. muris* in *M. musculus* (177), (178), (179), and their presence needs to be

taken into account in interpreting experimental results in the laboratory and in applications for field interventions.

Comparative study of GI microbiota of *P. leucopus* and *M. musculus*

The preceding study revealed several microbes that were either undescribed species or genera, e.g. "L. peromysci" or the *Candidatus* Melainabacteria bacterium, or new strains of known microbial species, e.g. *L. animalis* LL1 and *T. muris* LL4. These novelties notwithstanding, to what extent did the gut microbiota of this deermouse resemble that of the typical laboratory animal, a house mouse that was maintained under similar husbandry conditions, including diet? That question motivated the following experiment.

Fecal pellet samples from 20 adult *P. leucopus* (10 females and 10 males) and 20 adult BALB/c *M. musculus* (10 females and 10 males) were obtained and stored frozen at -80 °C until processing. All animals were approximately 10 weeks old. The animals were housed in the same vivarium facility, though in different rooms. The pellets were subjected to total DNA extractions, and paired-end Illumina sequencing with 250 cycles of indexed libraries were carried out. There were means (95% CI) of 3.4 (3.1–3.7) x 10^6 post-quality control reads for *P. leucopus* samples and 3.4 (3.2–3.6) x 10^6 for *M. musculus* samples.

The reproducibility between replicate library constructions from the same sample was assessed with quantitations of reads assigned by taxonomic family for specimens from seven *P. leucopus* among the 20 total. Pairwise coefficients of determination (R^2) for the 91 possible combinations were calculated. The mean (95% CI) of R^2 values were 0.999 (0.999–1.0) for the 7 pairs of replicates and 0.930 (0.915–0.944) for the 84 non-replicate pairs. We concluded that most of the variation between samples was attributable to inter-specimen differences in the microbiota and not to technical issues in library preparation or sequencing.

The prevalences of different taxonomic families in the *P. leucopus* and *M. musculus* gut metagenomes were similar (left panel of figure 13). But a few families stood out as either more or less common in the deermice. Notable among these were *Lactobacillaceae, Helicobacteriaceae,* and *Spirochaetaceae,* which were approximately 4x, 8x, and 2x, respectively, more prevalent on average among microbiota of *P. leucopus* than in *M. musculus*. There was no evidence of *Tritrichomonas* sp. in the BALB/c mice by this analysis, but direct examination of intestinal contents was not carried out.
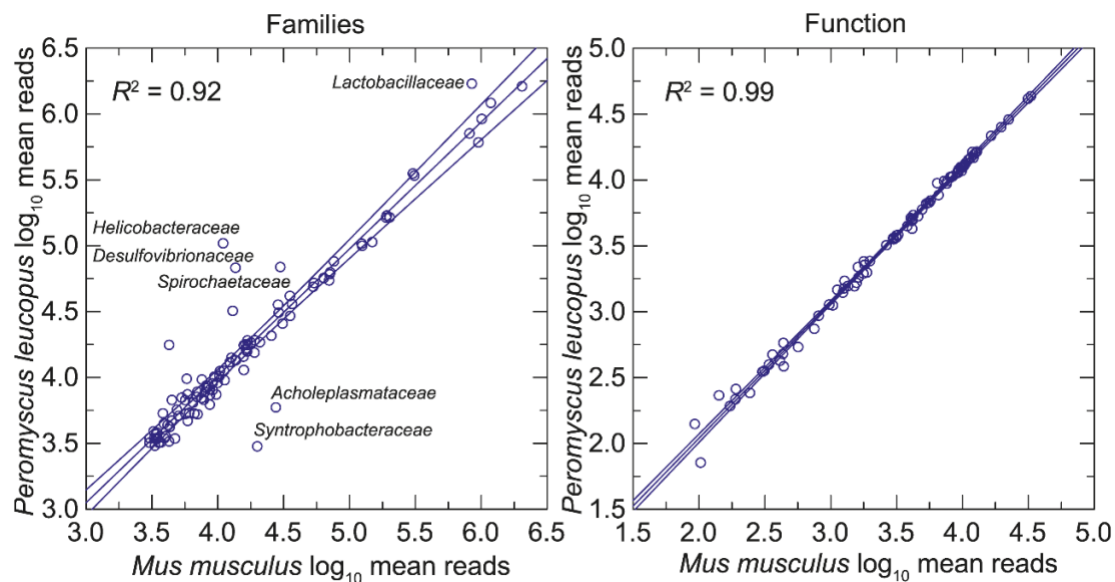
Figure 13 Scatter plots of P. leucopus and M. musculus metagenomes and gene functions
Scatter plots of log-transformed normalized reads of the gut metagenomes of 20 *P. leucopus* on the gut metagenomes of 20 *M. musculus* by bacterial families (left panel) or by function at the pathway level (right panel).The linear regression lines, their 95% confidence intervals, and coefficients of determination (R^2) are shown. Selected families that are comparatively more or less prevalent in *P. leucopus* are indicated.

At the level of 86 operational KEGG pathways, the metagenomes of *P. leucopus* and *M. musculus* were nearly indistinguishable (right panel of figure 13). But, as shown in the heat map of figure 14, at the homologous gene level there were many differences between these two species and also between females and males within each species. Hierarchical clusters 2 and 4 of the analysis

discriminated between mice and deermice regardless of sex, while clusters 1 and 3 signified marked

differences by sex and less so by species.
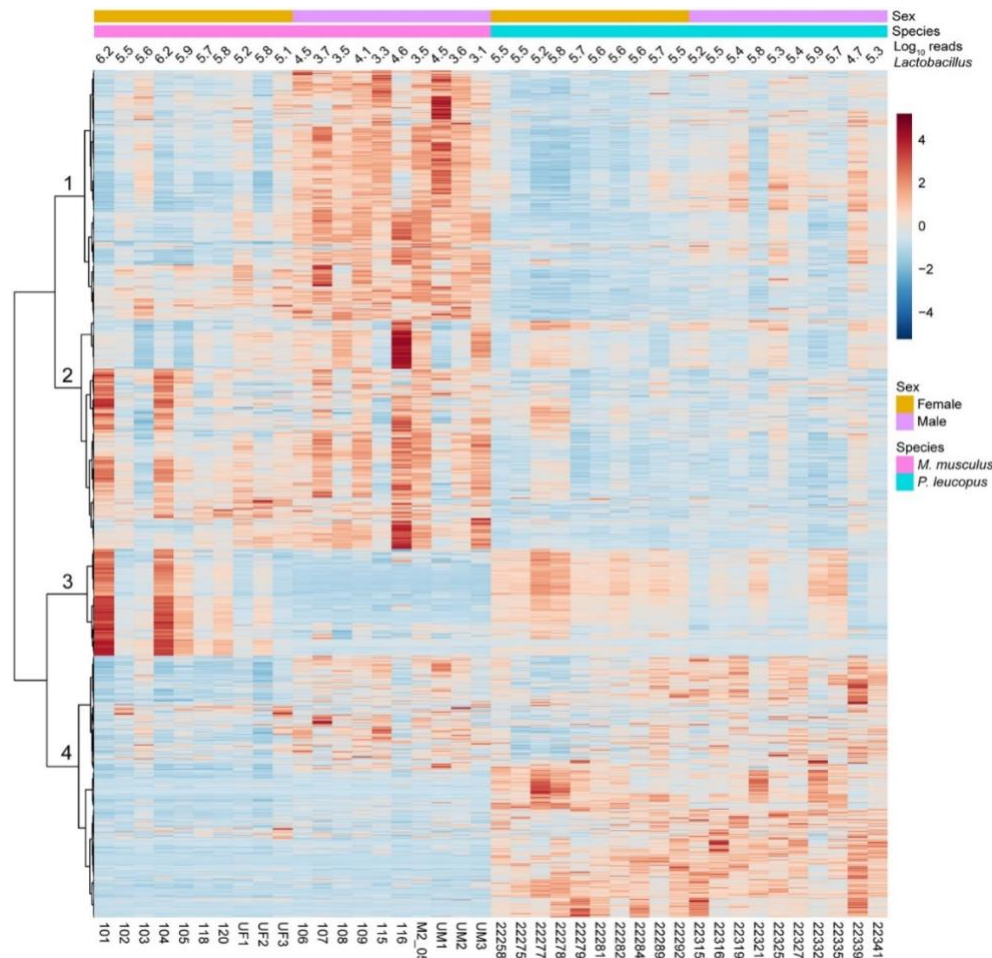
Figure 14 Heat map-formatted shading matrix of KEGG orthology gene level annotations of gut metagenomes of P. leucopus *and* M. musculus.

Heat map-formatted shading matrix of KEGG orthology gene level annotations of gut metagenomes of *P. leucopus* and *M. musculus*. The annotations were generated by MicrobiomeAnalyst (https://www.microbiomeanalyst.ca). Columns are grouped by species and by sex within each species. Individual animal identifications are given on the *x*-axis below the heat map. Above the heat map are the log-transformed reads mapping to the genus Lactobacillus for each animal's fecal sample. Clustering of rows of genes were by Pearson correlation coefficient. Four major clusters are labeled 1–4 on the *y*-axis. Scaling is by relative abundances from low (blue) to high (red).

As one example of differences between species, there was higher representation of genes of the

mevalonate pathway in the gut metagenomes of *P. leucopus*. Beginning with acetyl-CoA and ending

with isopentenyl pyrophosphate, the central intermediate in the biosynthesis of isoprenoids in all

organisms (180), the coding sequences for the following ordered enzymes (with Enzyme Commission [EC] number) in the pathway were comparatively higher in frequency: acetyl-CoA C-acetyltransferase (EC:2.3.1.9), hydroxymethylglutaryl-CoA synthase (EC:2.3.3.10), hydroxymethylglutaryl-CoA reductase (EC:1.1.1.88), mevalonate kinase (EC:2.7.1.36), phosphomevalonate kinase (EC:2.7.4.2), and diphosphomevalonate decarboxylase (EC:4.1.1.33). We further investigated specific differences between *P. leucopus* and *M. musculus* and between individual animals of each species in *Lactobacillus* spp. (181). This was achieved by mapping reads to references of the chromosome sequences of the four species that had been isolated from the feces of LL stock *P. leucopus*. The caveat is that the lactobacilli in the mice would not be expected to be identical to the deermouse strains used as references. Figure 15 shows box plots for *Peromyscus* on the left and for *Mus* on the right. Included in the analysis of *P. leucopus* gut metagenome reads were selected other bacteria that had been frequently identified among the metagenomic contigs and then further characterized by partial genome sequencing (see above).
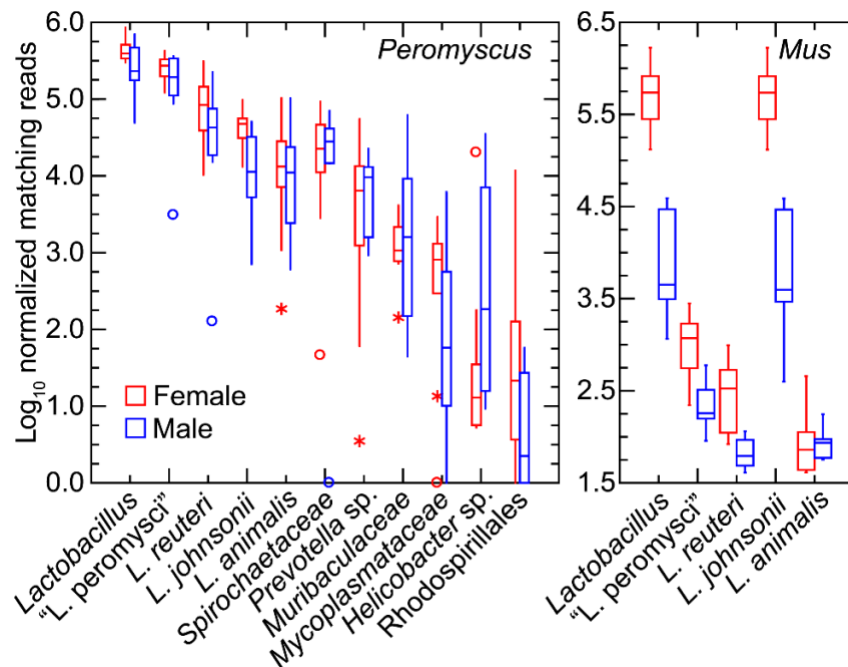
Figure 15 Box-whisker plots of gut metagenomes of P. leucopus *and* M. musculus *that mapped to chromosomes of bacterial species*

Box-whisker plots of log-transformed normalized reads of gut metagenomes of *P. leucopus* (left panel) and *M. musculus* (right panel) that mapped to chromosomes of *Lactobacillus* spp. or other bacteria by host species and grouped by sex. The references to which reads were mapped were complete chromosomes or partial chromosomes of organisms listed in Table 2. "*Lactobacillus*" in the first position of each panel were the cumulative reads for the four individual *Lactobacillus* species in this analysis.

All four species of the lactobacilli were represented in each of the 20 *P. leucopus* metagenomes. "L. peromysci" and *L. reuteri* tended to be the most common and consistently represented, while *L. johnsonii* and *L. animalis* varied more in prevalences between animals. Other bacteria were also identified in the samples of all or most of the individual animals. The *Spirochaetaceae* bacterium was ~10-fold less abundant than the cumulative *Lactobacillus* spp. in the *P. leucopus* samples.

The mean number of lactobacilli in aggregate were ~2-fold more prevalent in *P. leucopus* females than males of the species (*t*-test *p* = 0.03). In *M. musculus* this sex difference for *Lactobacillus* was more pronounced; on average ~100-fold more reads from female mice mapped to *Lactobacillus* genomes than was found for male mice (*t*-test *p* < 0.001). The differences in amounts of fecal lactobacilli in the sample plausibly account for cluster 3 of the heatmap of figure 14. *L. johnsonii* largely accounted for these differences between sexes in *M. musculus*; nearly all of the reads mapping to the *Lactobacillus* genus as a whole were mapping to the *L. johnsonii* genome. The three other species identified in *P. leucopus* were either not present or in much lower numbers in this sampling of *M. musculus*. Strains of *L. johnsonii* have been commonly detected in feces of laboratory mice (182).

A limitation to the study was that the LL stock animals were outbred, and the BALB/c mice were inbred. An inbred lineage derived from the LL stock population was not available. On the other hand, this distinction provided a comparison of microbiome diversities between an outbred and inbred population. As expected, there was greater alpha diversity among the outbred samples than the inbred (figure 16). Another limitation was the dependence on fecal pellets collected at one time

point. The samples were from similar age *P. leucopus* and *M. musculus* and were obtained from the

animals and then processed on the same day, but for this study we did not assess variation within

individuals over time. The gut microbiomes of the mothers of these animals were uncharacterized,

so we could not evaluate maternal effects in this study.
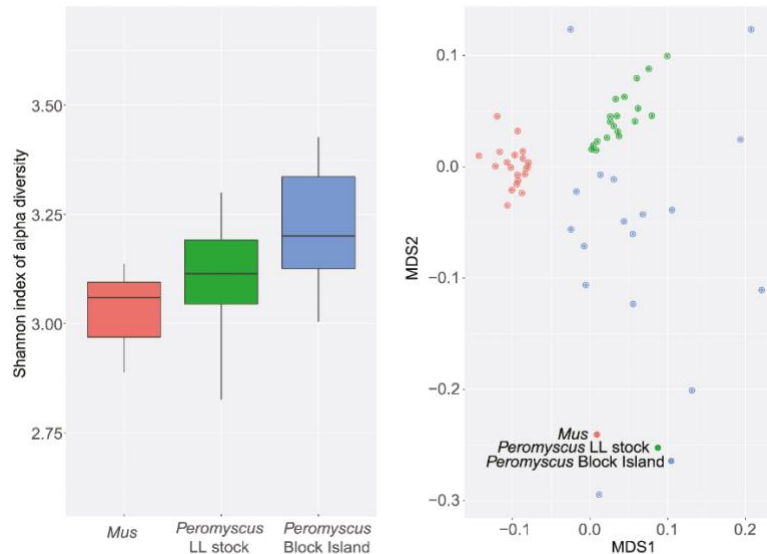
Figure 16. Alpha and beta diversity of gut metagenomes of P. leucopus
Alpha diversity (left) and beta diversity (right) of gut metagenomes of outbred *Peromyscus leucopus* (green), a natural population of *P. leucopus* (blue), and inbred *M. musculus* (red). Left panel, box-whisker plots of Shannon index of alpha diversity for 20 BALB/c *M. musculus*, 20 LL stock colony *P. leucopus*, and 18 *P. leucopus* trapped on Block Island, RI. The 3 pairwise, 2-tailed *t*-test *p* values between the groups were ≤ 0.02. Right panel, beta diversity by Bray-Curtis measure visualized by multi-dimensional scaling (MDS). A fuller description of the samples from the natural population from Block Island is provided below. The greater dispersion of values among these animals in panel B corresponded to the greater alpha diversity of this group (panel A).

Lactobacilli of the stomach of *P. leucopus*

The differences between *P. leucopus* and *M. musculus* in the amount and species richness of the

lactobacilli in their GI microbiota prompted further investigation of *P. leucopus* using histologic,

microbiologic and genomic approaches. Figure 17 shows the gross morphology and histology of the

stomach of representative LL stock *P. leucopus* animals (183). The difference between forestomach

with its stratified squamous epithelium and the discrete region lined with glandular mucosa are

indicated in the dissecting scope and higher magnification light microscope views.
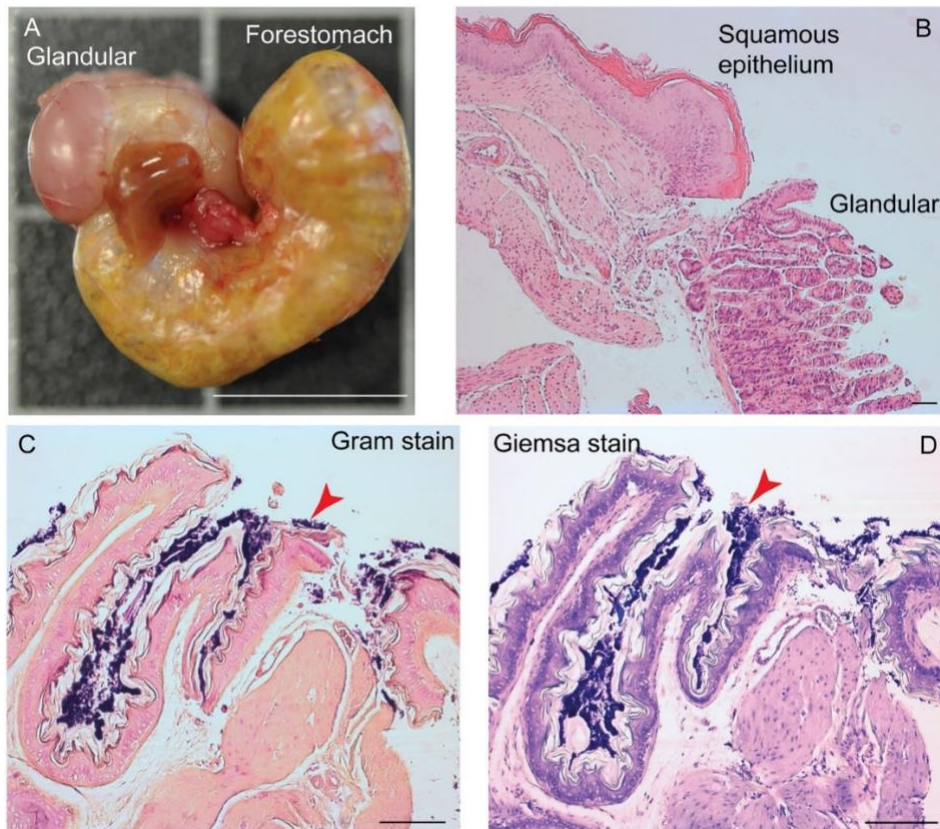
Figure 17 Gross morphology and histology of the stomach of P. leucopus LL stock.
The glandular mucosa portions of the stomach and the forestomach with stratified squamous epithelium are indicated. Panel A, whole stomach after dissection. Portions of the esophagus and small intestine are juxtaposed in the center in this view. Bar, 1 cm. Panel B, histology of hematoxylin and eosin-stained section of junction of glandular and squamous epithelium parts. Bar, 100 μm. Panels C and D, Gram stain (C) and Wright-Giemsa stain (D) of sections of squamous epithelium. Bar, 100 μm. Red arrowheads indicate gram-positive bacteria in a biofilm.

Staining of the sections of the fixed gastric tissue with Gram stain or Giemsa stain show a thick layer of gram-positive bacteria on the non-secretory epithelium portion of the stomach. This is similar to Savage et al. noted in the forestomachs of *M. musculus* (184). The appearance is also consistent with the *Lactobacillus* biofilm that was described by Wesney et al. (185)

Two of the species, "L. peromysci" and *L. animalis*, could reliably be distinguished by their distinctive colony morphologies from the isolated strains of *L. reuteri* and *L. johnsonii*, which had colonies of similar appearance (figure 18). The rough-surfaced, ropy colonies of "L. peromysci" and the compact smooth colonies of *L. animalis* were similar to what Dubos and colleagues described in their study of lactobacilli of the mouse stomach and gut (186).
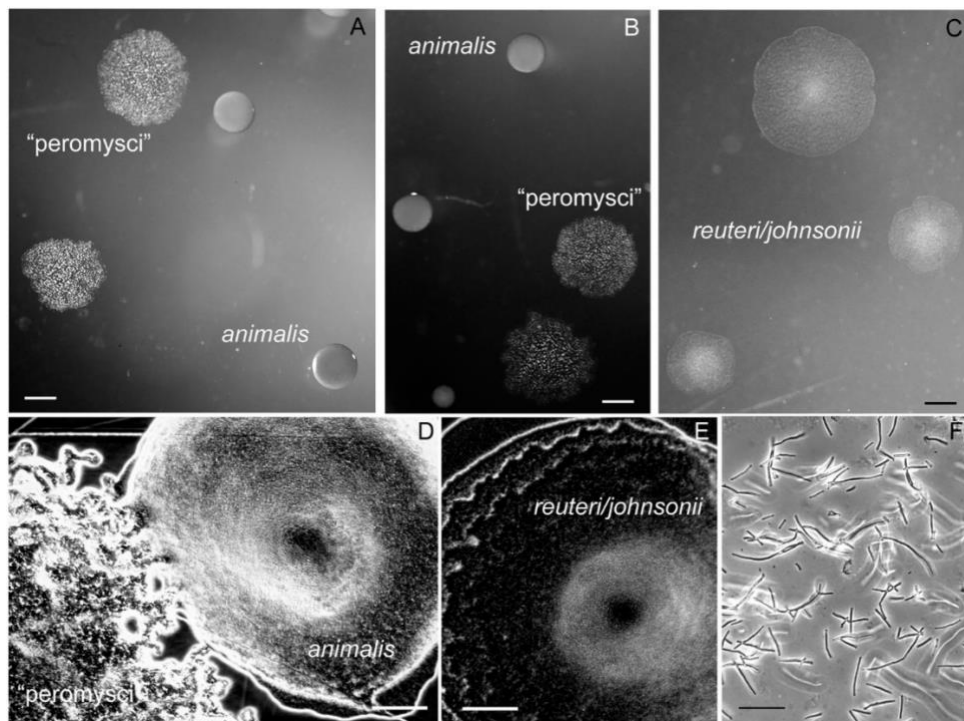
Figure 18 Colonies and cells of lactobacilli of the P. leucopus stomach and gut.
Panelshow representative sizes and morphologies of colonies of "L. peromysci", *L. animalis*, and the less distinguishable *L. reuteri* and *L. johnsonii*. Bars, 1 mm. Panels D and E show magnified view of colonies of "L. peromysci" and *L. animalis* (D) and that of *L. reuteri* and *L. johnsonii* (E). Bar, 100 μm. Panel F, phase microscopy of wet mount of unconcentrated broth culture of "L. peromysci". Bar, 10 μm.

We next used a different set of 20 animals of the LL stock, 6 (2 females and 4 males) of which were born at the PGSC facility and 14 (7 females and 7 males) of which were born at U.C. Irvine. All animals were housed at U.C. Irvine for at least 26 weeks before euthanasia, dissection, and cultivation of the stomach tissue and contents.

Mean (95% CI) colony forming units of lactobacilli per gram of stomach tissue on selective medium plates were ten-fold higher in females at 7.4 (1.1–47.4) x 10^9 than in males at 0.76 (0.40–1.4) x 10^9 (*t*-test *p* = 0.02 for log-transformed values) (Table 4). There was no discernible association with place of birth, and there was no difference between females and males in the proportions of the lactobacilli were identified as "L. peromysci", *L. animalis*, and *L. reuteri/L. johnsonii*. For five animals, whose lactobacilli were subjected to 16S ribosomal RNA gene PCR and sequencing for confirmation, the *L. reuteri/L. johnsonii*-type colonies were predominantly *L. reuteri*. But *L. johnsonii* was

confirmed to be present as well and outnumbered *L. reuteri* in one animal. The results for 3 animals that had been on a 9% fat content diet, which was part of the breeding program, instead of 6% fat content were not distinguishable from those for the other 17.

Table 4 Colony forming units of Lactobacillus spp. in P. leucopus stomach

Animal ID	Sex	Log$_{10}$ total cfu/gm	% colony type		
			animalis	"peromysci"	*reuteri/johnsonii*
22403	F	10.9	3	32	65
22404[a]	F	11.3	61	4	35
25053	F	10.7	37	60	3
25054	F	11.0	96	4	<1 [b]
25055	F	10.4	94	6	<1
25065	F	8.9	50	19	31
25062	F	8.5	60	14	26 (26/0) [c]
25063	F	8.8	52	12	36 (33/3)
25058	F	8.3	82	14	5 (5/0)
22401[a]	M	8.6	33	19	48
22375	M	8.8	58	3	40
22420[a]	M	9.7	48	19	33
22377	M	9.0	81	10	9
26050	M	8.4	87	6	7
25056	M	9.6	60	15	26
25010	M	9.0	84	16	<1
25060	M	8.8	55	23	23
25061	M	8.3	63	38	<1
25059	M	8.5	50	9	41 (36/5)
25011	M	9.0	61	24	14 (4/10)
Mean (95% CI)	n.a.	9.3 (8.9–9.8)	61 (51–70)	17 (11–23)	≤ 22

[a] 9% fat content diet instead of 6% fat content diet.

[b] <1, below limit of detection by serial dilution on plates

[c] (/), *L. reuteri* to *L. johnsonii* ratio by PCR and 16S ribosomal rDNA sequencing

A separate group of 9 adult LL stock *P. leucopus* (5 females and 4 males) were euthanized after withholding food overnight, and the freshly-excised stomachs were subjected to DNA extraction without prior washing of the stomach. A mean (95% CI) of 477,688 (408,988–546,388) PE250 Illumina reads were obtained for the 9 samples. These were mapped to the four *Lactobacillus* genomes as references, as described above, as well as to partial chromosomes for *Prevotella* sp. LL70 and *Helicobacter* sp. LL4 (Table 2). For an estimate of the number of mammalian nuclei represented in the stomach extract the *P. leucopus* genome (accession NMRJ00000000.1) served as the reference. Figure 19 shows the distributions of normalized reads mapping to the references as well as to the *P. leucopus* genome and cumulatively to all *Lactobacillus* spp. Females and males in this sample were indistinguishable by group by these measures. For this group of animals and this

analysis, we confirmed the high prevalence of "L. peromysci" in the stomach as well as the comparatively greater representation of *L. reuteri* over *L. johnsonii*. In this sample *L. animalis* was more variable in numbers between animals. As further evidence that the *Helicobacter* sp. was of the enterohepatic type, it was near undetectable in the stomach extract, while a typically abundnant genus in the intestine, *Prevotella*, was present in small numbers in some samples. The lactobacilli in the stomach were about as numerous as the stomach tissue cells constituting the sample.
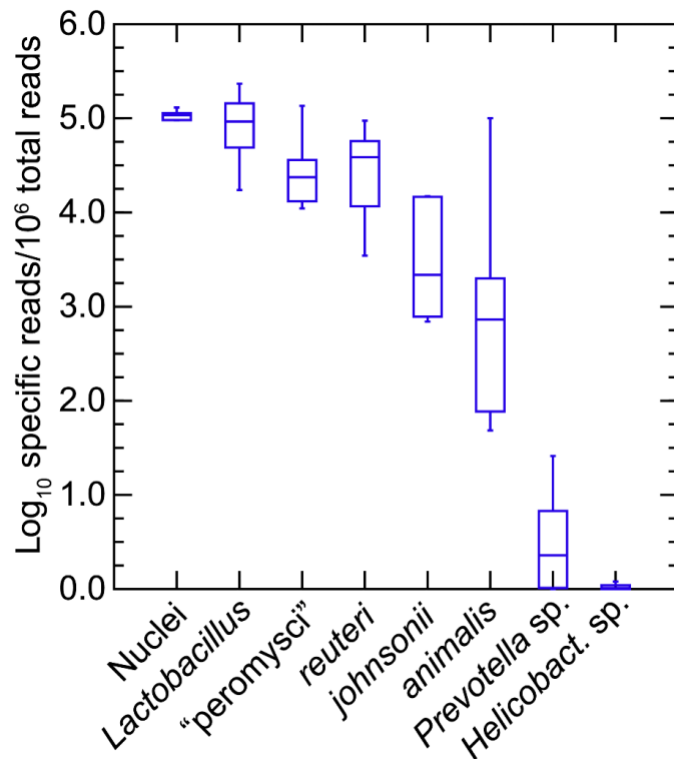
Figure 19 Box-whisker plots of total metagenomes of the stomachs of 9 P. leucopus
Box-whisker plots of log-transformed normalized reads of total metagenomes of the stomachs of 9 *P. leucopus* that mapped to chromosomes of *Lactobacillus* spp. or other bacteria by host species and by sex.The references to which reads were mapped were complete chromosomes or partial chromosomes of organisms listed in Table 2. "*Lactobacillus*" in the first position of each panel were the cumulative reads for the four individual *Lactobacillus* species in this analysis. The number of cell "nuclei" present in the stomach tissue and DNA extract were estimated by mapping reads to the whole genome of *P. leucopus* as described in the text

A strain of *L. reuteri* was shown to be the source of biofilm in the GI tract of mice in one study (187), and *L. murinus*, the sister taxa of *L. animalis* (figure 9), accounted for the biofilm in another study of the upper GI tract of *M. musculus* (188). *L. johnsonii* has also been demonstrated to produce an

exopolysaccharide biofilm (189). In a study of germ-free mice in which bacteria were experimentally introduced, *L. taiwanensis*, which is in the same cluster as *L. johnsonii* and *L. gasseri* (181), formed a mixed-species biofilm with *L. reuteri* (190).

One limitation of this experiment is that we may have overlooked species that were not identified because they were not cultivable by our method and conditions, which were microaerophilic, not strictly anaerobic. That said, if cells of such non-cultivable lactobacilli had been present in the feces or stomach, their numbers did not reach a threshold for assembly into contigs of the de novo assembly of the high coverage sequencing and then detection.

Gut metagenomes of a natural population

The foregoing studies were of animals born and reared under controlled conditions, including the same diet and environmental parameters for all individuals in the group. Infectious diseases and predators were not a variable. The LL stock *P. leucopus* were outbred but the effective population size was small compared to a wild population (5). Which of our findings would hold for animals sampled in their native habitats?

This particular study of a natural population had two specific purposes. The first was to assess the species richness or alpha diversity of microbiota within a given animal and differences in species composition or beta diversity between animals. The second objective's question of whether any of the *Lactobacillus* species we identified in the stock colony were present in natural populations. For this survey we used fecal pellets from *P. leucopus* that were individually captured and then released on Block Island, several miles off-shore from the North American mainland. At time of capture the animals were identified as to species, sex, and stage of maturity.

We analyzed the data from fecal pellets of 18 different animals (10 females and 8 males), the majority of which were adults, collected from *P. leucopus* captured at different locations on Block Island. As expected, there was greater variation between individual animals than was observed with the stock

colony animals maintained under same conditions. Figure 16 compares the alpha diversity by Shannon index and beta diversity by Bray-Curtis dissimilarity of the inbred BALB/c *M. musculus*, outbred LL stock *P. leucopus*, and the natural population of *P. leucopus* of Block Island.

By algorithmic assignment of reads to taxonomic family, *Lactobacillaceae* was one of the most prevalent bacteria with a mean of ~5% of reads, but this was over a range of 0.3% to 20%. As was the case for the stock colony *P. leucopus*, the frequency of *Helicobacteraceae* varied more widely between sampled animals than for comparably-prevalent taxa: a mean of ~1% but over a range of 0.03% to 12%. The frequency of a parabasalid protist in the metagenomes, by the criterion of reads matching to Trichomonadidae, was similar to what we observed in the metagenome of the stock colony *P. leucopus*: the mean was 0.11% with a range of 0.02 to 0.62%. This was an indication that the *T. muris* was autochthonous in *P. leucopus*, but we did not have direct observation of the protozoa to confirm that.

In descending order the two most prevalent DNA viruses in the samples, with the exception of one animal, were bacteriophages of the families *Siphoviridae* and *Myoviridae*. These were the two virus families that were also most prevalent in the feces of the stock colony animals. In the sample from the exceptional animal, the two most prevalent virus types were the genus *Dependovirus*, a group of DNA viruses dependent on another virus in the cell for replication, and the genus *Mastadenovirus*, an adenovirus that can provide that function. Other than the adenovirus in the one sample, there were no discernible representatives of herpesviruses, poxviruses, or parvoviruses, groups of DNA viruses which include pathogens for mammals. The methods used would not have identified RNA viruses other than retroviruses.

Using the chromosome sequences of the four *Lactobacillus* species and partial chromosome sequences of *Spirochaetaceae* bacterium LL50, *Prevotella* sp. LL70, and *Helicobacter* sp. LL4 as references, we mapped and counted reads, as described for the LL stock and *M. musculus* study

above. Figure 20 summarizes results for the 18 animals grouped by sex. Lactobacilli were common but, as seen with family level matching, there was greater variation between samples of the different animals than was observed for colony animals. There was also substantial variation in prevalences of the *Spirochaetaceae* bacterium and the *Prevotella* species. In most of the samples there was scant evidence of the *Helicobacter* species but in two animals, there were higher numbers of this organism, reaching 7% of the total reads in one fecal sample.
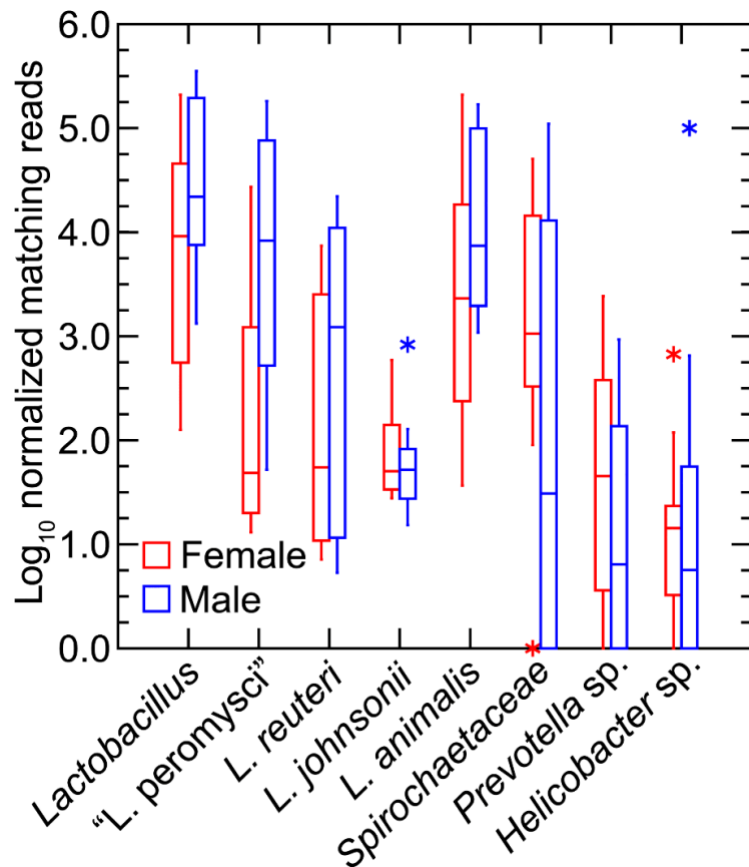
Figure 20 Box plots of fecal metagenomes of P. leucopus *of a natural population*
Box-whisker plots of log-transformed normalized reads of fecal metagenomes of 5 female (red) and 4 male (blue) *P. leucopus* of a natural population of Block Island, Rhode Island. The reference genomes and other sequences were those described for Figure 19 and in addition the partial chromosome sequence of *Spirochaetaceae* sp. LL50. As an estimate of the number of mammalian cells in the extract, "nuclei" corresponded with normalized reads mapped to *P. leucopus* genome. Points at a greater distance from the median than 1.5 times the interquartile range are plotted as asterisks.

Among the lactobacilli used as references for this analysis, the two most prevalent species were "L. peromysci" and *L. animalis*. *L. reuteri* overall was about 10-fold lower in frequency, and *L. johnsonii* was about a hundred-fold lower in frequency. It is likely that reads called as *L. johnsonii* were the result of complete or partial matching to chromosomal loci that were highly conserved across the genus. Unlike the stock colony *P. leucopus* and the *M. musculus*, in this sampling of the Block Island population the samples from female animals had marginally lower representation of lactobacilli in the fecal samples than males.

To better characterize the two predominant *Lactobacillus* species in this set, we assembled contigs of reads mapping to "L. peromysci" or *L. animalis* from a higher coverage sequencing of the DNA of one of the Block Island samples. This yielded 51 of the 53 genes for ribosomal protein genes for a strain of *L animalis*, which was designated 7442BI, and several core or housekeeping genes for the "L. peromysci"-like organism, which was designated BI7442 (Table 2). Figure 21 shows phylograms of DNA sequences for these and related *Lactobacillus* species or strains. The concatenated sequence of the BI7442 isolate was 99.2% identical over the 11,252 nt aligned with the corresponding sequences of the LL6 isolate of "L. peromysci". Isolate 7442BI was comparatively more distant from the LL1 isolate of *L. animalis* in the stock colony but still clustered with it rather than with other examples of *L. animalis*.

The sample size was limited, and we did not attempt culture isolations from the pellets. But the source of samples from *P. leucopus* was notable for its location on an island where *B. burgdorferi* is enzootic (191) and the risk of infection for residents and visitors is high (192). If there were to be future interventions targeting *P. leucopus* to interrupt disease transmission, Block Island would likely be a candidate site for this application.
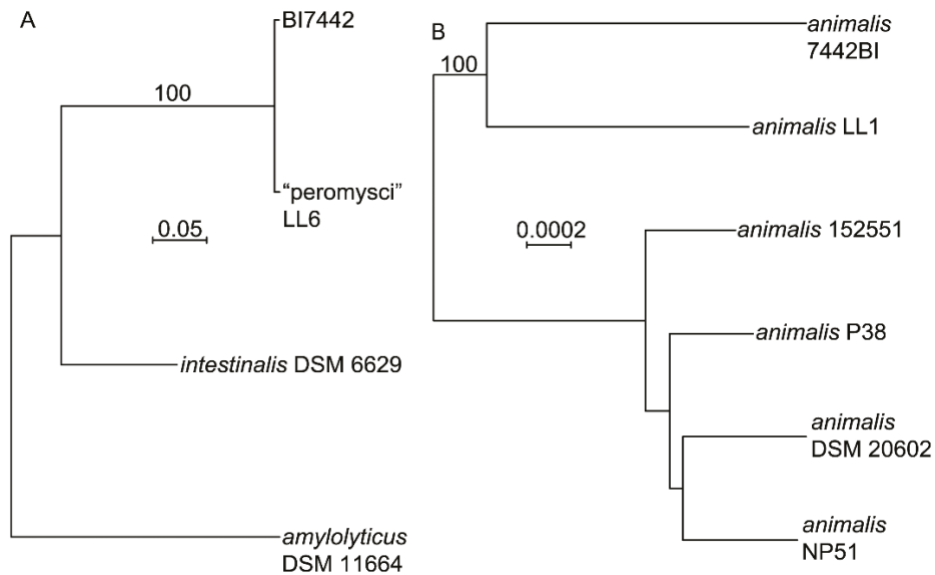
Figure 21 Distance phylograms Lactobacillus spp. of P. leucopus of a natural population
Distance phylograms of concatenated codon-aligned nucleotide sequences of two *Lactobacillus* spp. of *P. leucopus* of Block Island, Rhode Island. Panel A, 10,152 positions of *ftsK*, *ftsZ*, *dnaA*, *dnaN*, *ileS*, and *topA* of "L. peromysci" LL6 and BI7442 and two other *Lactobacillus* species. Panel B, 18,552 positions of 51 of 53 ribosomal protein genes of six *L. animalis* strains, including LL1 and 7442BI. The distance criteria were Jukes-Cantor. The scales for distance are shown in each panel. Percent bootstrap (100 iterations) support values of ≥ 80% at a node are shown.

This survey also documented that a strain or strains of "L. peromysci" and *L. animalis* are present in the native deermice. The high degree of sequence identity between two "L. peromysci" examples, whose origins were North Carolina and a New England island, long separated from the mainland, suggest that this newly-discovered species is autochthonous and plausibly a narrowly host-restricted symbiont of *P. leucopus*. Host-range restrictions of lactobacilli for the stomachs of mice were demonstrated by Wesney and Tannock (185). Supporting an assignment of a symbiont lifestyle was its smaller genome size and lower % GC content of this species in comparison with *L. reuteri* and *L. johnsonii* with their broader host ranges (193).

Description of *Lactobacillus peromysci sp. nov.*, isolated from the gastrointestinal tract of *P. leucopus*, the white-footed deermouse

Microbiome sequencing showed an abundance of *Lactobacillus peromysci sp. nov* (Lp) in *P. leucopus* fecal samples in the colony and in the wild. We isolated the same bacteria from the stomach and the intestines of *P. leucopus*. Histology shows the presence of biofilm in the stomach with gram-positive bacteria.

Since 2020, *Lactobacillus* has been reclassified into 25 genera (194) according to genetic, ecological, and metabolic properties. Before 2020, describing new genera in the *Lactobacillaceae* family was more difficult since the genus *Lactobacillus* was fragmented. Since the new classification was introduced, 37 new species have been described and classified at the genus level by 16S rRNA gene sequence analysis (195). Since then, the genus has been smaller and is now defined as Gram-positive, thermophilic, non-spore-forming rods. Species within the genus are generally host-adapted to a vertebrate host and usually in a commensal relationship with other *Lactobacilli* organisms like *Limolactobacillus reuteri (*classified as *Lactobacillus reuteri* before 2020), mainly due to their role in the formation of biofilms (196).

Lactobacilli are considered safe for humans, animals, and the environment and are health-promoting (197), hence their use as probiotics and in food industry fermentations. General moieties of probiotics are survival in vivo until reaching the gut, metabolizing diverse dietary ingredients, adhering to epithelium as a biofilm, and maintaining their numbers in vivo. Some of the direct benefits for the host are alterations of innate immunity and inflammation and exclusion of pathogens (198).

Genome assembly and annotation

The strain LL6 was isolated from stomach contents of healthy male and female *P. leucopus* on Rogosa SL agar with distinguishing colony morphologies. We selected distinct colonies on Rogosa

selective media for lactobacilli under microaerophilic conditions. Single colonies isolated from each animal's stomach were cultivated and maintained in Difco MRS broth and were supplemented with 10% (v/v) dimethyl sulfoxide (DMSO) at −80 °C for long-term storage.

DNA was extracted from MRS broth, and their 16S ribosomal RNA was amplified using PCR and custom primers specific to the *Lactobacillus* genus. Next, we sequenced DNA using an Oxford Nanopore Technology MinION and PE 150 short reads obtained on a MiSeq sequencer. We carried out hybrid assembly with Unicycler (199). Sequencing obtained 2.73×10^6 ONT reads with 2,601 x coverage on MinION and 278X coverage on Illumina platform. Assembly produced two circular contigs of 1.75 Mb and 0.32 Mb, chromosome and plasmid, respectively. We used the NCBI Prokaryotic Genome Annotation Pipeline for genome annotation (200). Genome size is 2,067,236 bp and plasmid 6,969 with 33.5% and 36.4% G+C, respectively (Table 5). Out of ~2000 genes, 1851 are protein-coding (Table 6).

Table 5 Lactobacillus peromysci sp. nov. assembly statistics

Assembly statistics	
Genome size	2.1 Mb
Total ungapped length	2.1 Mb
Number of chromosomes	1
Number of scaffolds	3
Scaffold N50	1.8 Mb
Scaffold L50	1
Number of contigs	3
Contig N50	1.8 Mb
Contig L50	1
GC percent	33.5
Genome coverage	120.0x
Assembly level	Scaffold

Isolated Lp full-length 16S rRNA blasted on NCBI shows 100% coverage and 96.69% nucleotide identity with *Lactobacillus intestinalis* strain DSM 6629 and 95.6% identity with *Lactobacillus amylolyticus* (figure 21). *L. intestinalis* strain DSM 6629 was isolated from rat intestines. Besides rats,

L. intestinalis is isolated from mice and pigs. In silico genome-to-genome distance comparison (http://ggdc.dsmz.de/distcalc2.php) with *Lactobacillus intestinalis* strain DSM 6629 shows no probability of them belonging to the same species under recommended formula 2 probability that distances and digital DNA:DNA hybridization (dDDH) is more than 70% (i.e., same species) is 0% via logistic regression (201).

Table 6 Genome characteristics of Lactobacillus peromysci sp. nov.

Annotation details genome assembly ASM757093v1		
	RefSeq	GenBank
Provider	NCBI RefSeq	NCBI
Date	8/4/23	19/7/19
Genes	1,958	1,951
Protein-coding	1,851	1,833
Software version	6.5	4.8

Coding sequences for the core genome genes *recA*, *rpoB*, *gyrA*, and *rplB* were no more than 83%, 87%, 81%, and 94%, respectively, with any other *Lactobacillus* sp (139). Based on standards for the description of new cultivable strains that represent novel genera, these distances suggest that Lp represents a previously unknown taxon, provisionally designated *Lactobacillus peromysci sp. nov*. NCBI, genome assembly identification is ASM757093v1, NCBI RefSeq assembly GCF_007570935.1. The full lineage of Lp is cellular organisms; *Bacteria*; *Terrabacteria* group; *Bacillota*; *Bacilli*; *Lactobacillales*; *Lactobacillaceae*; *Lactobacillus*; unclassified *Lactobacillus*.
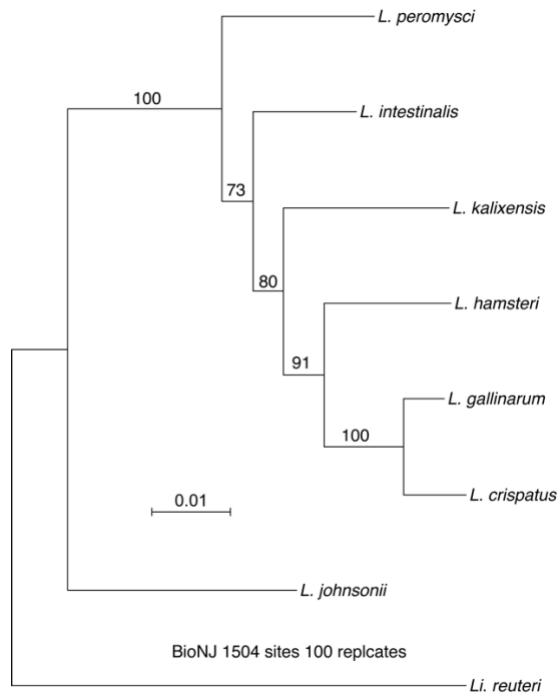
Figure 22 Distance phylogram with observed differences and 100 bootstrap replicates.

The phylogram shows evidence of *Lactobacillus peromysci sp. nov.*, as a novel species. *L. intestinalis* and *L. hamsteri* are isolated from rats and hamsters GIT. *Limosilactobacillus reuteri* is the outgroup. *L. kalixensis* is a *Lactobacillus* isolated from the human intestine.

Lp is similar in chromosome size to the other two bacteria isolated from the stomach of *P. leucopus* from the same *Lactobacillaceae* family- *Limolactobacillus reuteri* strain LL7 and *Ligilactobacillus animalis* strain LL1 (classified as *Lactobacillus animalis* before 2020). However, compared to the other two lactobacilli isolated from the stomach *of P. leucopus*, Lp has a smaller plasmid and G+C% content (Table 6). Lower G+C% content might indicate that Lp is well adapted to the nutrient-rich environment in the stomach epithelium, especially since we detected the same species in the closed colony animals captured near Linville, NC, who have been living in captivity for over 40 years, as well as the natural population of *P. leucopus* on a Block Island, RI natural preserve (202).

Table 7 Comparison of genomes of three different species from Lactobacillaceae family species found in the stomach of P. leucopus

Assembly comparison across species found in the stomach						
Species	Strain	Size bp		G+C%		Accession number	
		Chromosome	Plasmid	Chromosome	Plasmid	Chromosome	Plasmid
Ligilactobacillus animalis	LL1	2,280,577	4,016	41	37.1	CP039849	MK858222
Limolactobacillus reuteri	LL7	2,205,740	178,977	38.9	38.1	CP041676	CP041677
Lp	LL6	2,067,236	6,969	33.5	36.4	VLLR01000001 (contig 1), VLLR01000002 (contig 1)	VLLR01000003

Lp has a total of 1,851/1,958 coding genes out of the total genes. Of those, 5 are 5S rRNA, 16S rRNA, and 23S rRNA each, 65 tRNA, 3 noncoding RNAs, and 35 are pseudo genes. We visualized the genome using the Proksee tool (203) (Figures 22 and 23). Thiamine biosynthesis pathways, type1 CRISPR-Cas arrays, and nitroreductase were annotated on the chromosome.

We further focused on comparing Lp and two other lactobacillus species isolated from the gut of two rodents – rats and hamsters *L. intestinalis* and *L. hamsteri,* respectively (figure 22). *L intestinalis* strain DSM 6629 is closely related to *Lactobacillus hamsteri* DSM 5661 (149). *L. intestinalis* has multiple CRISPR-Cas systems – system I and unidentified, while *L. hamsteri* is also characterized by CRISPR-Cas system I (149). *L. peromysci sp. nov.* is also characterized by CRISPR-Cas system I. Higher G+C% content islands correlate with ribosomal rRNA on a chromosome, pointing out transcription-rich areas of the genome. Lp has the lowest G+C% content than the other two species isolated from rodents (Table 8) (204)

GCA_007570935.1_ASM757093v1_genomic.fna
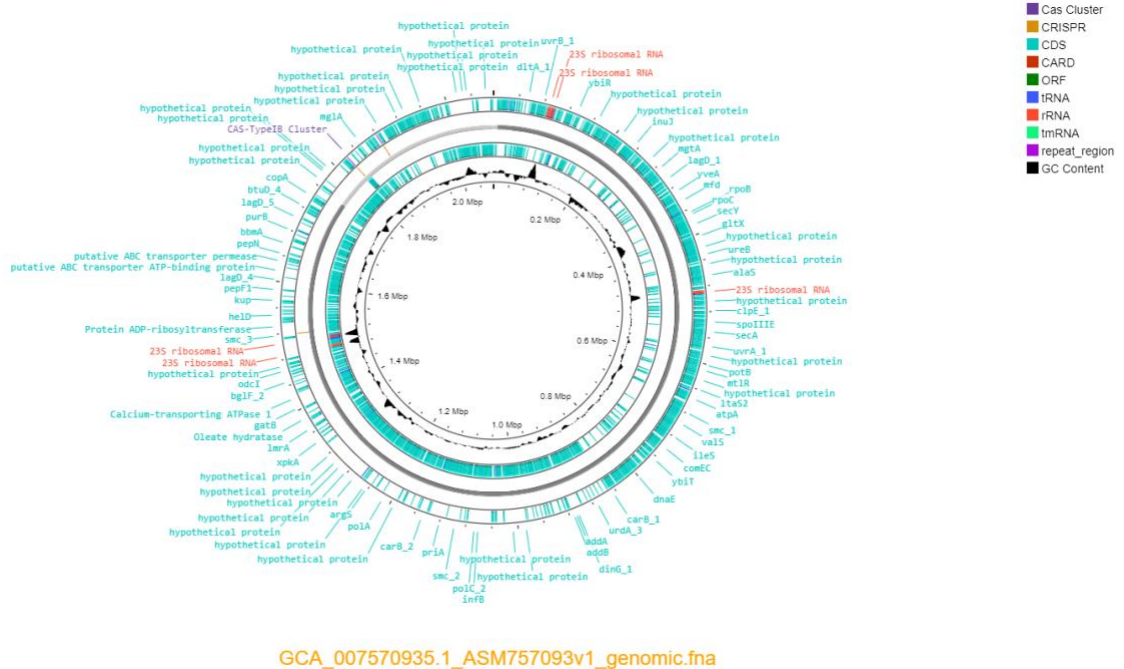
Figure 23 Annotation details of genome characteristics of Lactobacillus peromysci sp. nov. genome assembly

Annotation of a circular chromosome of Lactobacillus peromysci sp. nov. includes the Crisper-Cas cluster (purple) and ribosomal rRNA (red); the rest of the annotated genes are in blue. Inner circle shows G+C content.

plasmid.txt
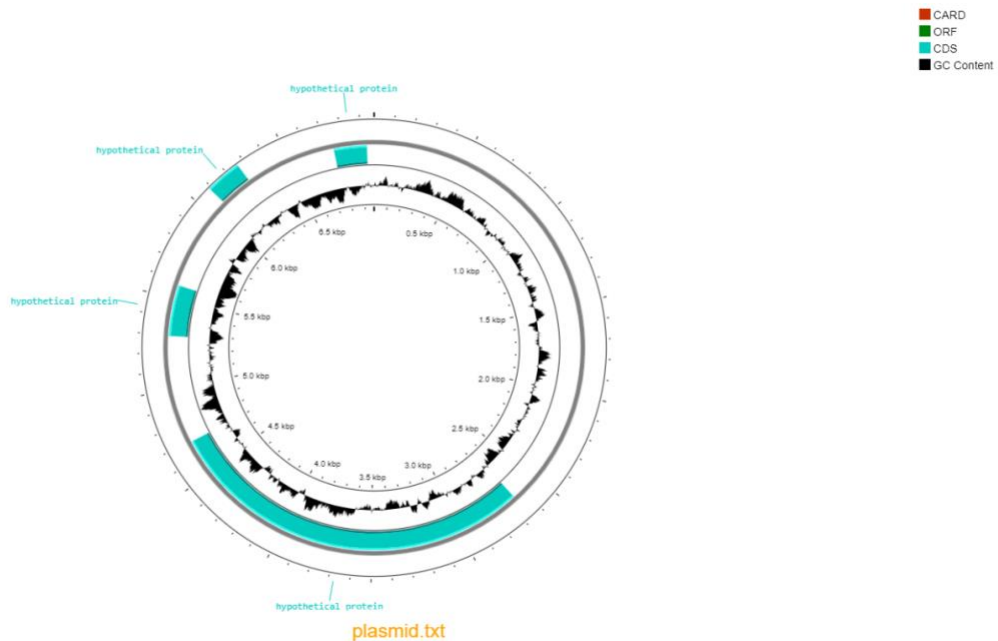
Figure 24 Annotation details of plasmid characteristics of Lactobacillus peromysci sp. nov.
Annotation of a circular plasmid of Lactobacillus peromysci sp. nov. includes 4 hypothetical genes.

60

Phenotypic criteria

Next, we characterized isolated lactobacilli propagated on Rogosa SL agar and MRS agar plates for appearance, metabolism using API 50 CH Strips, urease activity, and antibiotic susceptibility. The colony morphology of Lp is circular and flat; based on the margin, it is undulate to the rhizoid and 1-2 mm in diameter. They are rough-surfaced, ropy colonies. Cells are Gram-positive rods that chain when dividing in one plane (figure 18, panels A and F). They appear as singles, pairs, or chains. Lp as a single isolate deposited in one American (ATCC) and one European (DSMZ) database with accession numbers TSD-309 and DSM 115093, respectively.

Lp is nonmotile and nonspore-forming. It grows on Rogosa solid media or MRS media. In MRS media, it reaches the log phase in 12 hours. It is a microaerophilic organism and successfully grows on Rogosa at pH 5.4. Urease activity in the media was not detected when tested with Escherichia coli and Proteus bacterium as controls. The catalase gene is not annotated in the genome.

Sugar metabolism

We analyzed the metabolism of *L. peromysci sp. nov.* compared to *L. reuteri,* both isolated from the stomach of *P. leucopus,* and read the results at the 24 h and 48 h mark. Lp metabolized 18 different carbon sources and *L. reuteri* 10 with 8 overlappings (L-arabinose, ribose, galactose, D-glucose, maltose, lactose, saccharose, and D-raffinose). Results are represented in Table 9. The isolate with the closest sugar metabolism matching Lp on The Bacterial Diversity Metadatabase BacDive (https://bacdive.dsmz.de) database for standardized bacterial phenotypic information is *Lactiplantibacillus plantarum* CCUG 43738. It is isolated from the human gut with 22 overall carbon sources, out of which 18 are identical to Lp (D-arabitol, D-turanose, melizitose and melibiose) (205). *L. plantarum* is widely prescribed as a probiotic and used in fermented food, like kimchi and dosa batter (206). *L. peromysci sp. nov*. could be another helpful probiotic used in the food industry and fermentation.

L. peromysci sp. nov. is a heterofermentative organism since it metabolizes both pentose and hexose sugars. The closest species to *L. peromysci sp. nov.* based on 16S nucleotide identity – *L. intestinalis* strain DSM 6629 metabolize glucose, mannose, fructose, galactose, sucrose, and mannitol but not trehalose. *L. hamsteri* and Lp metabolize glucose, cellobiose, sucrose, mannitol, raffinose, and ribose, producing dl-lactic acid. In contrast, they do not metabolize melizitose (194).

Comparison in metabolism between three species of lactobacilli isolated from rodent intestines, based on fermentation reactions is presented in table 9 (147), (204).

Table 8 Comparison between three species of lactobacilli isolated from rodent intestines

Characteristics	Species		
	Lp	*L. hamsteri*	*L. intestinalis*
G+C%	33.5	35.1	35.4
CRISPR-Cas system type	I	I	I+ unidentified
Genome size (Mbp)	2.07	1.84	2.01
Trehalose	+	+	-
Arabinose	+	-	-
Xylose	-	-	-
Rhamnose	-	-	-
Melizitose	-	-	NA

Table 9 Carbohydrate metabolic test with a semi-quantitative measure

Performance of carbohydrate metabolism tests				
	Lp 24h	Lp 48h	*L. reuteri* 24h	*L. reuteri* 48h
Control				
Glycerol				
Erythritol				
D-arabinose				
L-arabinose	+	+	+	+
Ribose	+	+	+	+
D-xylose				
L-xylose				
Adonitol				
beta-methyl-xyloside				
Galactose	+	+	+	+
D-glucose	+	+	+	+
D-fructose	+	+		
D-mannose	+/-	+		
L-sorbose				
Rhamnose				
Dulcitol				
Inositol				
Mannitol	+	+		
Sorbitol				
α-methyl-D-mannoside				
α-methyl-D-glucoside				
N-acetyl-glucosamine	+	+		
Amygdalin				
Arbutin	+	+		
Esculin				
Salicin	+	+		
Cellobiose	+	+		
Maltose	+	+	+	+
Lactose	+	+	+	+
Mellibiose			+	+
Saccharose	+	+	+	+
Trehalose	+	+		
Inulin				
Melizitose				
D-raffinose	+/-	+	+	+
Amidon	+/-	+		
Glycogen				
Xylitol				
beta-gentobiose	+/-	+		
D-turanose				
D-lyxose				
D-tagatose				
D-fucose				
L-fucose				
D-arabitol				
L-arabitol				
Gliconate			+/-	+
2-ceto-gluconate				
5-ceto-gluconate				

Antibiotic resistance

Lp is resistant to ciprofloxacin and *L. reuteri* to gentamicin and ciprofloxacin, as shown in Table 10.

Table 10 Antibiogram for Lactobacillus peromysci sp. nov. and L. reuteri

Antibiotic	Lp colonies number	*L. reuteri* colonies number
MH* ampicillin 10µg	sensitive	sensitive
MH gentamicin 10µg	sensitive	**resistant**
MH ciprofloxacin 5 µg	**resistant**	**resistant**
MRS† ampicillin 10µg	sensitive	sensitive
MRS gentamicin 10µg	**resistant**	sensitive
MRS ciprofloxacin 5 µg	**resistant**	**resistant**
* Mueller Hinton agar		
† De Man–Rogosa–Sharpe agar		

Description of Lactobacillus peromysci sp. nov

Lactobacillus peromysci /pɛˈrɒmɪscɪ/ (L. gen. n. "Peromyscus" from where the bacterium was isolated). Gram positive, nonmotile, non-spore-forming, rods bacterium. Colonies are r, circular rough surfaced and flat with 1-2 mm in diameter. They are microaerophilic. Acids are produced from amidon, arbutin, L-arabinose, beta-gentobiose, cellobiose, D-fructose, D-glucose, D-mannose, D-raffinose, galactose, lactose, maltose, mannitol, N-acetyl-glucosamine, ribose, saccharose, salicin, and trehalose. Strain is deposited as *Lactobacillus peromysci* LL6 (=DSM 115093 ᵀ = ATCC TSD-309™). G+C content is 33.5 mol%. The annotated genome data was deposited in GenBank under the accession number VLLR00000000.1.

Discussion

Six decades ago René Dubos (of the epigraph), Russell Schaedler, and their colleagues at what is now Rockefeller University reported in a series of ground-breaking papers on the "fecal flora" of mice and variations in that microflora between mouse strains (207), (208), (209). They associated differences in gastrointestinal flora with growth rates of the mice and the mouse's susceptibility to infection and endotoxin. A featured group of bacteria in their studies were lactobacilli. They showed

that this group of bacteria were present in large numbers in the feces and that they predominated (up to 10^9 cfu per g of homogenate) in the stomachs of the mice (184), similarly to what we observed in *P. leucopus*. As their studies first intimated, the rodent may plausibly owe as much to the genomes of their microbiota as to the nuclei and mitochondria of their somatic cells for either ameliorating or exasperating disease (210).

There are also implications of our findings for development of oral vaccines that target *P. leucopus* to block transmission of pathogens either from tick to the reservoir or from the reservoir to the tick. Two of the candidate vehicles for the bait delivery of recombinant vaccine antigens to rodents have been *E. coli* and a *Lactobacillus* species (145), (211). In neither case were the strains known to be adapted for life in *P. leucopus*. The success rate for achieving a protective response may be enhanced by use as the bacterial vehicle microbes that are adapted to *P. leucopus*. Such organisms presumably would more likely than an allochthonous bacterium to stably colonize and then proliferate to numbers large enough for the recombinant protein to elicit the sought-after immune responses.

Finally, this exploration and curation of microbes in the gut of the white-footed deermouse concludes the third leg of our project on the total genome of representative animals of the species: the nuclear genome (5), the mitochondrial genome (6), and now the GI microbiome (212). This provides a foundation for testing of hypotheses by selective manipulation of the microbiota, for instance, by specifically targeting a certain species with a lytic phage or bacteriocin, to which it is not immune, and then evaluating the phenotype of the animal after this "knock-out". Now that there is an annotated *P. leucopus* genome with millions of SNPs identified (UC Santa Cruz genome browser; http://googl/LwHDr5) it also feasible to investigate through forward genetics the host determinants of particular bacterial associations and for which there is evidence of variation within

a population. An example would be the *Helicobacter* species that was highly variable in prevalence in both the wild animals and the stock colony animals.

Materials and Methods

Colony animals

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. At the University of California Irvine protocol AUP-18-020 was approved by the Institutional Animal Care and Use Committee (IACUC)-approved protocol. Adult outbred *P. leucopus* of the LL stock were purchased from the Peromyscus Genetic Stock Center (PGSC) of the University of South Carolina (213). The closed colony of the LL stock was founded with 38 animals captured near Linville, NC in the mid-1980's. Some of the LL stock animals in the study were bred at the University of California Irvine's animal care facility from pairs originating at the PGSC. Adult BALB/cAnNCrl (BALB/c) *M. musculus* were purchased from Charles River. For the species comparison experiment both the PGSC-bred *P. leucopus* and *M. musculus* animals were housed in Techniplast individual ventilated cages in vivarium rooms with a 12 hour-12 hours light-dark cycle, an ambient temperature of 22 ± 1 °C, stable humidity, and on an ad libitum diet of 2020X Teklad global soy protein-free extruded rodent chow with 6% fat content (Envigo, Placentia, CA). Other animals of PGSC origin, including for the high-coverage gut metagenome study, were also housed under the same conditions and on the same diet. Twenty U.C. Irvine-bred animals were under the housing conditions and on the diet except for three (1 female and 2 males) that were on 2019 Teklad global protein extruded rodent chow with 9% fat content. Fecal samples were 2–3 pellets (100–150 mg) collected from an individual animal in temporary (< 1 h) isolation in a clean cage. The pellets were collected with sterile forceps and pooled in a single sterile 1.5 ml microfuge tube on ice. The pellets for DNA extractions were stored frozen at -80° C. Pellets intended for cultivation, as described below, were processed as fresh specimens.

Before euthanasia with carbon dioxide asphyxiation and cervical dislocation and then dissection of the stomach, food but not water was withheld for 12 h for selected animals. *P. leucopus* studied at the PGSC were under IACUC-approved protocol 2349-101211-041917 of the University of South Carolina and were euthanized by isoflurane inhalation.

Field site and animal trapping

The study was performed under IACUC-approved protocol AC-AAAS6470 of Columbia University (191). Block Island, located 23 km from mainland Rhode Island, is part of the Outer Lands archipelagic region, which extends from Cape Cod, MA through to Staten Island, NY. Block Island is 25.2 sq. km, about 40% of which is maintained under natural conditions. The agent of Lyme disease *Borreliella burgdorferi* is enzootic on the island (17). Animals were trapped at three locations: 1, a nature conservation area (41.15694, -71.58972); 2, private land with woodlots and fields (41.16333, -71.56611); and 3, Block Island National Wildlife Refuge (41.21055, -71.57222). Trapping was carried out during the May-August period with Sherman live traps (H.B. Sherman Traps, Inc. Tallahassee, FL) that were baited with peanut butter, oats, and sunflower seeds. Traps were arranged in nine 200 m transects with one trap placed every 10 m for a total of 180 traps at each location. Animals were removed from traps, weighed, sexed, and assessed as to age (adult, subadult, or juvenile) by pelage. Fecal pellets were collected and kept at -20 °C on site, during shipment and until DNA extraction. The species identification of the source of the fecal pellets as *P. leucopus* was confirmed by sequencing of the D-loop of the mitochondrion as described (6).

Cultivation and enumeration of bacteria

Lactobacilli were initially isolated and then propagated on Rogosa SL agar plates (Sigma-Aldrich) in candle jars at 37° C. Gram-negative bacteria and specifically *Escherichia coli* were isolated and propagated on MacConkey Agar plates (Remel) incubated in ambient air at 37° C. For determinations of colony forming units (cfu) homogenates of stomach, cecum, or fecal pellets were suspended and

the serially diluted in phosphate-buffer saline, pH 7.4, before plating in 100 µl volumes on solid media in 150 mm x 15 mm polystyrene Petri dishes. Colonies were counted manually. Liquid cultures of *Lactobacillus* spp. isolates or *E. coli* were in Difco Lactobacilli MRS Broth (Becton-Dickinson) or LB broth, respectively, and incubated at 37° C on a shaker. Bacteria were harvested by centrifugation at 8000 x g for 10 min. Antibiotic susceptibilities were determined by standard disk testing on Mueller-Hinton Agar (Sigma-Aldrich) plates and ciprofloxacin 5 µg, gentamicin 10 µg, ampicillin 10 µg, and sulfamethoxazole 23.75 –trimethoprim 1.25 µg BBL Sensi-Disc antibiotic disks (Becton-Dickinson) according the manufacturer instructions.

API

API 50 CHL Medium (BIOMERIEUX, Marcy-l'Etoile, France) was used for the identification of the genus Lactobacillus and related genera, which allows the fermentation of the 49 carbohydrates on the API 50 CH strip. A suspension is made in the medium with the microorganism to be tested and each tube of the strip is then inoculated with the suspension. During incubation, the carbohydrates are fermented to acids, which produce a decrease in pH, which is detected by the change in color of the indicator. The results make up the biochemical profile that can be used to identify the strain of Lactobacillus. We grew colonies on MRS agar and then made a colony suspension with a turbidity equivalent to 2 McFarland in the API 50 CHL Medium ampule. Then we used it to inoculate API 50 CHL and covered the blister with mineral oil, and incubated it anaerobically for 24 h. Read the color at the first-time point then read again after 48 h incubation at 37°C. A positive test corresponds to acidification revealed by the bromcresol purple indicator contained in the medium changing to yellow.

Urease

ID-Urease (Microbiologics, 60045) was used to detect urease activity in bacteria. In the presence of urease, the pH of the phenol red indicator changes color from yellow to red. We used *Escherichia coli* and *Proteus* bacterium as controls.

Antibiotic resistance

We used sulfamethoxazole 23.75 µg with trimethoprim 1.5 µg Sensi – Dics (Becton, Dickinson and Company BBL, USA) on MH and MRS media. On the same plate, we used ciprofloxacin 5 µg, gentamicin 10 µg, and ampicillin 10 µg discs.

Histology

After the stomachs were removed from two euthanized *P. leucopus* LL stock adult females, they were opened longitudinally, gently flushed with PBS, and fixed in 10% buffered formalin (Thermo Fisher Scientific). Histological and histochemical analysis was performed on paraffin block sections of the stomach with Hematoxylin and Eosin, Wright-Giemsa and Gram stains (Abcam, Cambridge, UK).

Microscopy, photography, and video

Photographs of colonies on plates were taken with a Nikon Df DSLR camera and 60 mm Nikkor AF-S Micro lens with illumination by incident light above and reflected light below the plates on an Olympus SZ40 dissecting scope. An Olympus BX60 microscope with attached QIClick CCD camera and Q-Capture Pro7 software (Teledyne Photometrics, Tucson, AZ) was used for low-magnification images of colonies under bright light microscopy and 400X images under phase and differential interference microscopy. Histology slides were examined on a Leica DM 2500 microscope equipped with a MC120 HD digital camera (Leica Microsystems, Buffalo Grove, IL).

DNA extractions

DNA from fresh and frozen fecal pellets, from tissue of stomach or cecum, and from bacteria harvested from broth cultures were extracted with ZymoBIOMICS DNA Miniprep or Microprep kits (Zymo Research). Freshly-dissected, unwashed gastric or cecal tissues were first cut into small pieces with sterile instruments before homogenization in the lysis buffer in an Omni Bead Ruptor 4 bead beater (Omni International). DNA concentrations were determined by both NanoDrop spectrophotometer and Qubit fluorometer (Thermo Fisher Scientific).

PCR

The near-complete 16S ribosomal RNA gene for *Lactobacillus* spp. was amplified using PCR using custom primers specific for the genus *Lactobacillus*: forward 5'-CCTAATACATGCAAGTCG and reverse 5'-GGTTCTCCTACGGCTA. The Platinum Taq polymerase and master mix (ThermoFisher Scientific) contained uracil-DNA glycosylase. On a T100 thermal cycler (BioRad) PCR conditions (°C for temperature) The conditions were 37° for 10 min, 94° for min, 40 cycles of 94° for 10 s, 55° for 30 s, and 72° for 45 s. The 1.5 kb PCR product was isolated and purified from agarose gel using the NucleoSpin Gel and PCR Clean-up kit (Takara). The product was subjected to Sanger dideoxy sequencing at GENEWIZ (San Diego, CA).

Whole genome sequencing, assembly, and annotation

Long reads were obtained using an Oxford Nanopore Technology MinION Mk1B instrument with Ligation Sequencing Kit, R9.4.1 flow cell, MinKnow v. 19.6.8 for primary data acquisition, and Guppy v. 3.2.4 for base calling with default settings. Paired-end short reads were obtained on a MiSeq sequencer with paired-end v2 Micro chemistry and 150 cycles (Illumina, San Diego, CA). The library was constructed using the NEXTflex Rapid DNA-Seq kit (Bioo Scientific, Austin, TX), the quality of sequencing reads was analyzed using FastQC (Babraham Bioinformatics), and reads were trimmed of Phred scores <15 and corrected for poor-quality bases using Trimmomatic (214). A hybrid

assembly was carried out with Unicycler v.0.4.7 (199) with default settings and 16 threads on the High Performance Computing cluster of the University of California Irvine. Assembly of short reads alone were performed with the Assembly Cell program of CLC Genomics Workbench v. 11 (Qiagen) Annotation was provided by the NCBI Prokaryotic Genome Annotation Pipeline (200). Putative bacteriocins and their associated transport and immunity functions were identified by BAGEL4 (215), (216) . For other analyses paired-end reads were mapped with a length fraction of 0.7 and similarity fraction of 0.9 to whole genomes sequences or concatenated large contigs representing partial genomes (Table 2). Mapped reads were normalized for length of reference sequence and for total reads after quality control and removal of adapters.

Metagenome sequencing

The library was constructed using NEXTflex Rapid DNA-Seq kit v2 (Bioo Scientific) and the NEXTflex Illumina DNA barcodes after shearing the DNA with a Covaris S220 instrument, end repair and adenylation, and clean-up of the reaction mixture with NEXTFLEX Clean Up magnetic beads (Beckman Coulter, Brea, CA). The library was quantified by qPCR with the Kapa Sybr Fast Universal kit (Kapa Biosystems, Woburn, MA), and the library size was determined by analysis using the Bioanalyzer 2100 DNA High Sensitivity Chip (Agilent Technologies). Multiplexed libraries were loaded on either an Illumina HiSeq 2500 sequencer (Illumina, San Diego, CA) with paired-end chemistry for 250 cycles or a MiSeq Sequencer (Illumina, San Diego, CA) with paired-end v2 Micro chemistry and 150 cycles. The Illumina real time analysis software RTA 1.18.54 converted the images into intensities and base calls. De novo assemblies were performed with De Novo Assembly v. 1.4 of CLC Genomics Workbench v. 11 with the following settings: mismatch, insertion, and deletion costs of 3 each; length fraction of 0.3, and similarity fraction of 0.93.

16S ribosomal RNA gene analysis

The same DNA extract used for the metagenome sequencing at University of California Irvine was submitted to the reference laboratory ID Genomics, Inc. (Seattle, WA) and subjected the company's 16S rRNA Metagenomics service (http://idgenomics.com/our-services), which used the 16S Metagenomics v. 1.01 program (Illumina). Of the 333,358 reads 82%, were classified as to taxonomic family.

Metagenome analysis

Fastq files were uploaded to the metagenomic analysis server MG-RAST (https://www.mg-rast.org) (217). Reads were joined using join paired reads function on the browser and filtered for *M. musculus* v37 genome. Artificial replicate sequences produced by sequencing artifacts were removed by the method of Gomez-Alvarez et al. (218). Low quality reads (Phred score <15 for no more than 5 bases) were removed using SolexQA, a modified DynamicTrim protocol (219). When the term "alpha diversity" was reported in the results and tables for MG-RAST analyses, this stood for species richness or count of the species as calculated by MG-RAST. The output of the MG-RAST protocol was also analyzed in *R* using the *vegan* package (https://cran.r-project.org, https://github.com/vegandevs/vegan). By this means we also represented alpha diversity by the Shannon index, which accounts for evenness as well as richness (220). Beta diversity as expressed as the Bray-Curtis Dissimilarity statistic (221) was calculated using the *avgdist* function with 1000 sample depth, the median as the function, and 100 iterations (https://github.com/vegandevs/vegan/blob/master/man/avgdist.Rd). Data was visualized using Nonmetric multidimensional scaling in two dimensions (222). MicrobiomeAnalyst (https://www.microbiomeanalyst.ca) (223) was used for hierarchical clustering by distance criterion and by means of Pearson correlations. The DFAST prokaryotic genome annotation pipeline (https://dfast.nig.ac.jp) was used for annotation of incomplete chromosomes and large contigs

(224). For *Lactobacillus* spp. and the *Helicobacter* sp. the lactic acid bacteria database and *Helicobacter* database, respectively, options were chosen. Alignments and phylogenetic analysis were carried out with the SeaView v. 4 suite and Muscle (225), (226).

Statistical analysis

Normalized reads and other values whose distributions spanned more than one order of magnitude were log-transformed before parametric analysis by 2-tailed *t*-test. Inverse transformation was carried out to provide nonparametric means and corresponding asymmetric 95% confidence intervals. Nonparametric analysis by rank order was by Kruskal-Wallis test. The *Z*-score was the number of standard deviations below or above the population mean a give raw value was. The False Discovery Rate (FDR) with corrected *p* value was estimated by the method of Benjamini and Hochberg (227). The box-whisker plot graphs were made with SYSTAT v. 13.1 software (Systat Software, Inc.).

CHAPTER THREE: An Infection-Tolerant Mammalian Reservoir for Several Zoonotic Agents Broadly Counters the Inflammatory Effects of Endotoxin

Abstract

Animals that are competent reservoirs of zoonotic pathogens commonly suffer little morbidity from the infections. To investigate mechanisms of this tolerance of infection, we used single-dose lipopolysaccharide (LPS) as an experimental model of inflammation and compared the responses of two rodents: *P. leucopus*, the white-footed deermouse and reservoir for the agents of Lyme disease and other zoonoses, and the house mouse *M. musculus*. Four hours after injection with LPS or saline, blood, spleen, and liver samples were collected and subjected to transcriptome sequencing (RNA-seq), metabolomics, and specific reverse transcriptase quantitative PCR (RT-qPCR). Differential expression analysis was at the gene, pathway, and network levels. LPS-treated deermice showed signs of sickness similar to those of exposed mice and had similar increases in corticosterone levels and expression of interleukin 6 (IL-6), tumor necrosis factor, IL-1β, and C-reactive protein. By network analysis, the M. musculus response to LPS was characterized as cytokine associated, while the P. leucopus response was dominated by neutrophil activity terms. In addition, dichotomies in the expression levels of arginase 1 and nitric oxide synthase 2 and of IL-10 and IL-12 were consistent with type M1 macrophage responses in mice and type M2 responses in deermice. Analysis of metabolites in plasma and RNA in organs revealed species differences in tryptophan metabolism. Two genes in particular signified the different phenotypes of deermice and mice: the Slpi and Ibsp genes. Key RNA-seq findings for *P. leucopus* were replicated in older animals, in a systemic bacterial infection, and with cultivated fibroblasts. The findings indicate that *P. leucopus* possesses several adaptive traits to moderate inflammation in its balancing of infection resistance and tolerance.

Introduction

Peromyscus leucopus, the white-footed deermouse, is a major reservoir for several zoonotic agents (reviewed in reference (3)). The infections include Lyme disease, as well as varieties of anaplasmosis, babesiosis, relapsing fever, ehrlichiosis, and viral encephalitis (228). *P. leucopus* is broadly distributed across the eastern and central United States (229), (2), adapted to a variety of habitats, and an important host for the tick vectors of disease (230). Its immune system and other defenses keep the pathogens at bay, but infections persist (231), (16), thereby increasing the likelihood that a tick acquires the microbe during its blood meal (232). Where enzootic transmission is pervasive, the majority of *P. leucopus* animals live parasitized by one or more of these pathogens ((15), (18), (25). If there is a fitness cost, scrutiny has not revealed it in the field (25) or laboratory (233), (21), (234).

The term for persistent infection with minimal morbidity is infection tolerance (26), (235), (236) and distinguished from the immune system's "tolerance" of self-antigens (237). In both contexts, "tolerance" conveys a moderation of the host response and avoidance of injury. Infection tolerance has been observed mainly at the organismal level, for instance, by signs of illness, disability, and fitness measures. These presumably are explained by events at the cellular and molecular levels (238), (239), but these have not been fully explored.

P. leucopus' tolerance of infection is matched by that of another deermouse, *Peromyscus maniculatus*, a reservoir for a hantavirus (240), (241). *P. maniculatus* has also been experimentally infected with the coronavirus (CoV) disease 2019 (COVID-19) virus and observed to transmit it to cagemates. Yet infected animals displayed only moderate pathology and recovered within a few days (108), (109). Other examples of the tolerance phenomenon are found among bat species implicated as reservoirs for severe acute respiratory syndrome (SARS)-CoV, Ebola virus, Nipah virus, and Hendra virus (242), (243). These deermice and bats exhibit a trade-off between defensive processes that

75

check pathogen proliferation and processes that limit collateral damage from those defenses. The result is a state of persistent infection with limited disability.

The innate and adaptive host defenses that in concert resist and neutralize pathogens are well known. Less understood are mechanisms on the other side of the trade-off (244), namely, those that curb sickness and illness due to maladaptive inflammation (245). *P. leucopus* is well suited for such investigations of infection tolerance (7), (8), (246). Colloquially called "mice," the genus Peromyscus, along with hamsters and voles, belong to the family Cricetidae and not the family Muridae, which includes the house mouse, *Mus musculus* (132). We have previously sequenced the genome of *P. leucopus*, annotated its transcriptome using transcriptome sequencing (RNA-seq), and characterized its gastrointestinal microbiota as described in Chapter 2. (6), (5). There was little difference by RNA-seq between the blood of *P. leucopus* experimentally infected with the Lyme disease agent *Borreliella burgdorferi* and that of uninfected controls (5), a finding consistent with the mildness of deermouse infections with this pathogen (247). Accordingly, we looked to alternatives that would more robustly elicit inflammation.

To that end, we compared the response to a single dose of bacterial lipopolysaccharide (LPS) of *P. leucopus* to that of *M. musculus*. This endotoxin leads to inflammation through its binding to a pattern recognition receptor and complex ensuing cascades (248). In its higher doses, injected LPS elicits a sepsis-like state that shares manifestations, such as fever and shock, with a variety of serious infections, including those caused by some protozoa and viruses, as well as bacteria. Tolerance of LPS is a long-recognized phenomenon (249), but this third usage of "tolerance" refers to a diminution in the severity of the response through cumulative previous exposures of animals or isolated cells to LPS (250), (251), rather than an inherent disposition of a naive animal to survive a toxic dose. Some spontaneous and engineered Mendelian traits do render a mouse less susceptible to LPS through an alteration of Toll-like receptor 4 (TLR4) for LPS or in key downstream mediators,

such as MyD88, in the cascade (252),(253). In these examples, this at the cost of a compromised innate defense against bacterial infection. For this and other reasons, we thought it unlikely that a single gene or locus accounted for the more nuanced interplays between complex contending forces in *Peromyscus* and other reservoirs for a variety of disease agents.

The simple experimental system of *in vivo* LPS exposure is a proxy for acute systemic infection but without the confounding variables of changing numbers of the microbe and the inevitable appearance of acquired immunity. Our working assumption is that early events in the response are determinants of the eventual outcomes of the infection, whether assessed from the perspective of pathogen burden or host disability. In the course of the study, we identified several pathways involved with inflammation, oxidative stress, phagocytosis, and metabolism that distinguished *P. leucopus* and *M. musculus* in their responses to LPS. These have implications for studies of other vertebrate reservoirs of the agents of emerging zoonoses.

Results

Susceptibility of *P. leucopus* to LPS

We found no published reports on the systemic effects of different doses of LPS in any Peromyscus species, but there were several on the susceptibility of *M. musculus* to LPS (254), (255), (256), (257), (258), (259). The 50% lethal dose (LD50) for M. musculus was in the range of 5 to 20 mg of LPS per kg of body weight, with an average across inbred and outbred mice of ~15 mg/kg. Death usually occurred between 12 and 48 h after injection. We previously found that adult BALB/c mice injected intraperitoneally (i.p.) with 10 mg LPS per kg body weight manifested reduced activity and ruffled fur at 4 h, with concomitant elevations in serum concentrations of tumor necrosis factor alpha (TNF-α), interleukin 6 (IL-6), and IL-10 by immunoassay, but all animals survived up to 12 h (260).

We began with a study of the effect of single doses of purified Escherichia coli LPS administered i.p. to adult P. leucopus animals and then monitored them for 7 days. Figure 25A includes the survival

77

curve for animals in groups of six receiving doses of 10, 50, 100, 200, or 300 mg per kg and examined

continuously for the first 8 h and then at 12-h intervals thereafter. Death or a moribund state occurred

in at least one animal in all dosage groups except the 10-mg/kg group. Five of 6 in the 50-mg dose

group survived. At higher doses, the fatality rate was higher, with death occurring between days 2 and

5. Remarkably, 3 of 6 of the animals receiving the highest dose of 300 mg per kg, or a total dose of

6 mg on average per animal, survived. Survivors among the *P. leucopus* animals at that and the 100-

and 200-mg/kg doses appeared to have fully recovered by 7 days after the injection. The variation

between individual animals in outcome at the higher doses was consistent with the genetic

heterogeneity of this outbred population (261). From these data, we could not calculate a precise

LD50, but we estimated it to between 100 and 300 mg/kg from this experiment.
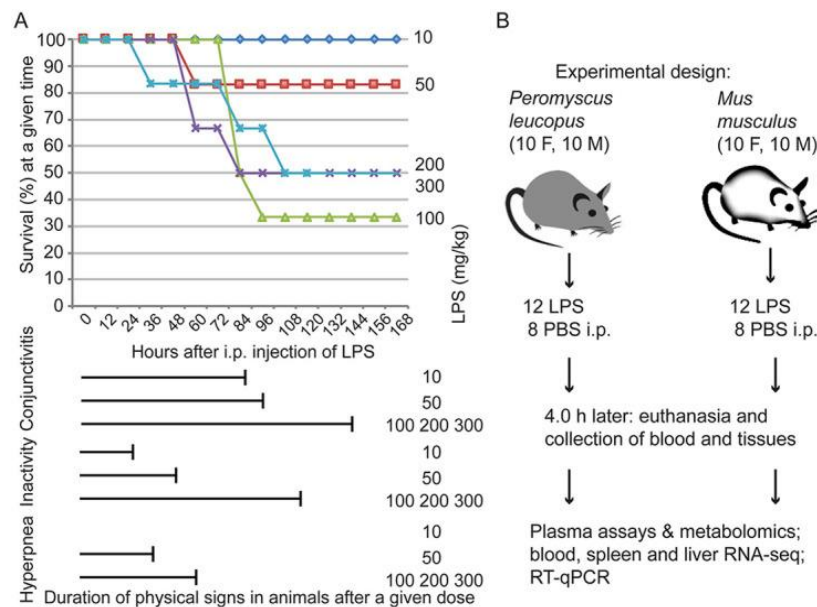
Figure 25 Studies of lipopolysaccharide effects on P. leucopus deermice and experimental design

(A) Dose response of *P. leucopus* to LPS. Groups of 6 adult animals (3 females and 3 males) received different intraperitoneal (i.p.) doses on a milligram/kilogram of body weight basis of *Escherichia coli* LPS at time zero and then monitored for physical signs of sickness (conjunctivitis, inactivity, and hyperpnea) and survival over the succeeding 7 days (168 h). The survival curves by dose are indicated by text to the right and by colors: dark blue (10 mg/kg LPS), red (50 mg/kg), light green (100 mg/kg), blue green (200 mg/kg), and purple (300 mg/kg). The *x* axis scale and label apply to the durations of physical signs as well as to the survival graph. (B) Experimental design of comparative study of short-

term effects of LPS on *P. leucopus* and *Mus musculus*. Animals received 10 mg LPS in saline/kg body weight or saline alone. F, female; M, male.

The deermice in this experiment had an exudative conjunctivitis with eyelid closure at all doses by 12 h, and that persisted for up to 6 days at the highest dose (figure 25A). All animals displayed reduced activity, as defined by the behavior of huddling in groups, moving only for eating and drinking. The durations of this sign, as well as for the sign of hyperpnea, or rapid breathing, in animals receiving doses of 50 mg/kg or higher, correlated with the dose amounts. Thus, even at the lowest dose, all the animals displayed ill effects of this treatment. But in comparison to reported findings in *M. musculus*, a substantial proportion of deermice receiving doses that were 10- to 20-fold higher than the consensus LD50 for the house mouse did not further deteriorate and thereafter recovered.

Experimental design

We compared the short-term responses to a single dose of LPS of the outbred *P. leucopus* mice and the inbred *M. musculus* BALB/c strain mice of both sexes. The comparison of animals from a heterogeneous closed colony with an inbred population provided a gauge of the diversity of responses among the deermice. Figure 25B summarizes the experimental design. Fecal samples obtained from animals a day before the experiment had been subjected to gut microbiome analysis and alpha-diversity values of the microbiota from the earlier study are provided (described in Chapter 2). The animals of each species were sexually mature adults and comparable in size and age, though *P. leucopus* tended to be smaller and older by 1 to 2 weeks. With the exception of two siblings split between treatment and control groups, the *P. leucopus* mice were the offspring of different mating pairs. We alternated the administration of the LPS by species and by sex within a species to control for diurnal effects. Animals were euthanized in the same order, and the intervals between animals were kept within a strict limit. The blood cells and tissues of spleen and liver were subjected to RNA-seq, and the plasma was used for metabolomics. This comparative study was complemented by

experiments with an older set of *P. leucopus* mice similarly exposed to LPS, deermice with a systemic bacterial infection, and cultures of *P. leucopus* fibroblasts.

Comparison of deermice and mice in response to LPS

Both species displayed effects of LPS within an hour of the injection, namely, reduced activity and ruffed fur, but only among 6 LPS-treated *P. leucopus* mice, equally distributed between females and males, did we observe conjunctivitis (P = 0.025). For corticosterone concentrations, the mean for *M. musculus* controls (n = 7) and LPS-treated mice (n = 12) were 77 (95% confidence interval [CI], 34 to 120) and 624 (95% CI, 580 to 667), respectively ($P < 10^{-11}$). Corresponding values for *P. leucopus* control (n = 7) and LPS-treated (n =12) mice were 186 (95% CI, 79 to 293) and 699 (95% CI, 670 to 727), respectively ($P < 10^{-8}$). There were marginally higher baseline levels of corticosterone in *P. leucopus* than in *M. musculus* mice (P = 0.10), but they had similar levels after LPS. The assay for nitric oxide demonstrated higher levels in 11 *M. musculus* mice treated with LPS (mean, 29 [95% CI, 20 to 37]) than in 7 controls (mean, 7 [95% CI, 3 to 12]) (P = 0.008). There was not an elevation in nine LPS-treated *P. leucopus* mice compared with values for six controls: 7 (95% CI, 3 to 11) versus 7 (95% CI, 3 to 12), respectively (P = 0.9).

To estimate the combined contributions of abundance and transcriptional activity of white blood cells in the blood samples at the time of euthanasia, we used the numbers of RNA-seq reads matching the whole mitochondrial genomes of each species in the same analysis. In both species, there were lower total transcriptional activities of mitochondria in LPS-treated animals than in controls at 4 h (lower by an average of 31% in mice and 23% in deermice). Distributions were similar between species for both controls and LPS-treated animals. The mean log10-transformed normalized values were 5.6 (95% CI, 5.5 to 5.7) for *M. musculus* and 5.7 (95% CI, 5.6 to 5.8) for *P. leucopus* controls (Mann-Whitney P = 0.96) and 5.4 (95% CI, 5.4 to 5.5) for *M. musculus* and 5.6 (95% CI, 5.5 to 5.7) for *P. leucopus* LPS-treated animals (P = 0.97).

Metabolomics of plasma

Untargeted metabolomics identified in *M. musculus* plasma 8,125 molecular features (MF), of which 123 (1.5%) differed between LPS-treated and control animals, with a false-discovery rate (FDR) of <0.05 and an absolute fold change of >2.0. In *P. leucopus* plasma, 7,714 MF were identified, of which 215 (2.8%) correspondingly differed between treated and control animals. Pathway enrichment analysis allowed cross-species comparison of identifiable metabolites in the animals (Dryad repository, https://doi.org/10.7280/D1R70J). For *M. musculus* and *P. leucopus*, the numbers of identified pathways with one or more KEGG-defined compounds were 76 and 73, respectively, with 73 in common (figure 26 left). In both species, there was enrichment of the steroid hormone biosynthesis pathway 4 h after LPS injection, which was consistent with the results of plasma corticosterone assays.
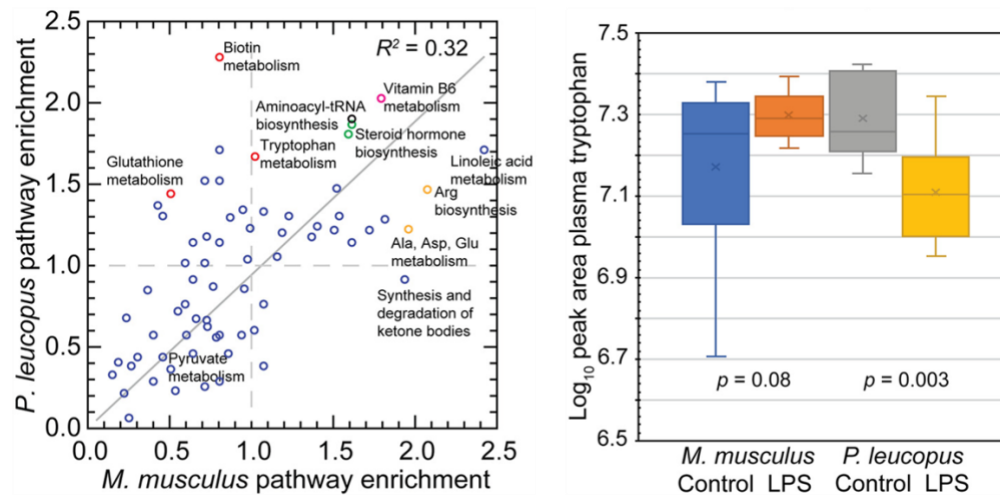
Figure 26 Untargeted metabolomics of plasma of P. leucopus and M. musculus animals with or without LPS treatment 4 h prior
Left panel shows scatterplot of pathway enrichments in LPS-treated P. leucopus (y axis) versus those in LPS-treated M. musculus (x axis). An enrichment value of 1.0 means that there is no difference in number of compounds in a given pathway between treated and untreated conditions for a species. Data, including identified KEGG terms, are provided in the Dryad repository (https://doi.org/10.7280/D1R70J). Selected pathways are labeled. The color of the symbols indicate the following findings for false discovery rate (FDR) P values of <0.05: green, both species; red, *P. leucopus*; and orange, *M. musculus*. The coefficient of determination (R2) shown is for an unspecified intercept. For consistency with the dashed lines indicating enrichment values of 1.0, the regression line for an intercept of 0.0 is shown. Right panel shows box-plots of log-transformed

plasma tryptophan levels estimated as peak areas in LPS-treated and untreated animals of each species. Two-tailed t test P values between the two conditions for each species are shown.

There was overall enrichment of tryptophan metabolism in LPS-treated animals of both species. The magnitude was greatest for *P. leucopus*, which had 30 significant hits out of a possible 41 compounds in the pathway (figure 26 left panel). Tryptophan itself was significantly lower in abundance in plasma of LPS-treated *P. leucopus* animals than in untreated animals, while in *M. musculus*, it was marginally higher after LPS treatment than in controls (figure 26 right panel). Tryptophan depletion activates the amino acid sensor GCN2, which leads to increased production of IL-10 and transforming growth factor beta (TGF-β) by macrophages and dendritic cells (262). Kynurenine, the product of indole 2,3-dioxygenase (Ido1), was elevated in the LPS-treated animals over controls in both species, but the kynurenine/tryptophan ratio was higher, at 1.75, in the LPS-treated deermice than in controls but lower, at 0.79, in the LPS-treated mice (*P* = 0.004). Higher kynurenine/tryptophan ratios are associated with anti-inflammatory effects through kynurenine's binding to the aryl hydrocarbon receptor that drives regulatory T cell differentiation (263).

DEGs in each species

We performed RNA-seq on blood, spleen, and liver extracts and then differential gene expression using 24,295 and 35,805 annotated transcripts for *P. leucopus* or *M. musculus*, respectively, as separate reference sets (Dryad repository, https://doi.org/10.7280/D1VX0C). Figure 27 reveals hundreds of genes that were either up- or downregulated in the blood and two organs of each species by the criteria of a ≥ 4-fold change and an FDR of <0.05. With the exception of the spleens of *P. leucopus* animals, there were more upregulated than there were downregulated genes in each of the comparisons. The highest numbers of differentially expressed genes (DEGs) were in the livers: 1,553 for *P. leucopus* and 3,250 for *M. musculus* animals.
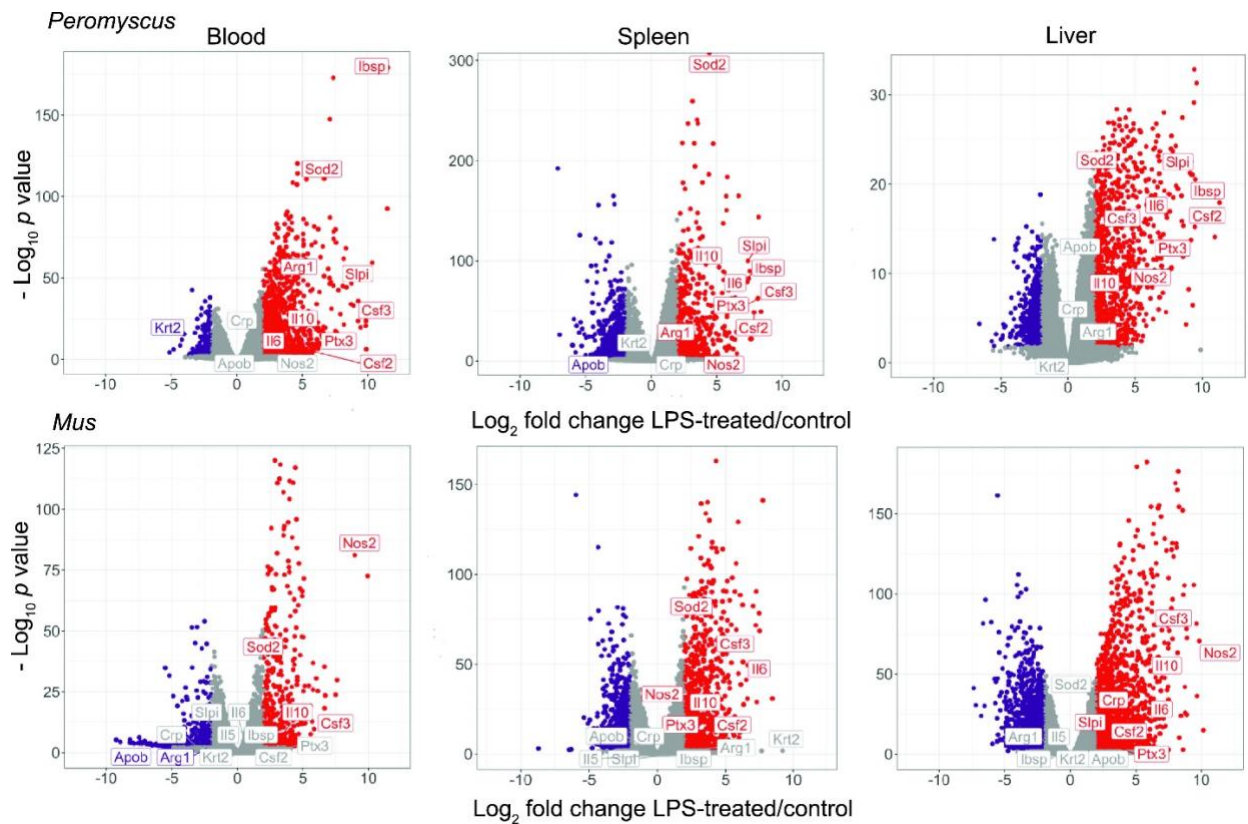
Figure 27 Species- and tissue-specific responses to LPS

Independent differential gene expression analysis of RNA-seq data were performed for blood, spleen, and liver tissues of *P. leucopus* and *M. musculus* collected 4 h after injection with LPS or buffer alone as a control. These are represented as volcano plots with range-adjusted scales for the log2-transformed fold changes on x axes and log10-transformed FDR P values on y axes. Colors of symbols denote the following: red, upregulated gene with an absolute fold change of >4.0 and a P value of <0.05; purple, downregulated gene with an absolute fold change of >4.0 and a P value of <0.05; and gray, all others. Numbers at the top left and right corners in each plot represent numbers of down- and upregulated genes, respectively. Numerical values for each gene in the 6 data sets are provided at the Dryad repository (https://doi.org/10.7280/D1VX0C).

A notable difference in the responses of *P. leucopus* and *M. musculus* was consistent with the results

of the nitric oxide assays of the plasma of these animals. Inducible nitric oxide synthase or nitric

oxide synthase 2 (Nos2) transcript levels were a mean of 493 times higher in the blood of LPS-treated

mice than in the controls ($P = 10^{-78}$). But in the blood of deermice, Nos2 expression was barely

detectable, and expression of Nos2 was indistinguishable between the two conditions ($P = 0.34$). In

contrast, arginase 1 (Arg1), which by its action reduces the amount of arginine available for Nos2 to

produce nitric oxide, was 21 times higher in expression in the blood of LPS-treated *P. leucopus*

animals than in controls ($P = 10-47$), while in *M. musculus* blood, Arg1 expression 4 h after LPS injection was 6 times lower than baseline expression ($P = 0.04$). This reciprocal expression profile for Nos2 and Arg1 between the two species was also observed in the spleen. The products of Nos2 and Arg1 are informative biomarkers for categorizing polarized macrophage responses (264).

Two other genes whose expression profiles distinguished the species in LPS responses were the genes for integrin-binding sialoprotein (Ibsp) and secretory leukocyte peptidase inhibitor (Slpi). These were first and fourth ranked of the upregulated DEGs in the blood for *P. leucopus*, but they ranked numbers 6,232 and 33,541, respectively, among measured transcripts in the blood for *M. musculus* (Dryad repository https://doi.org/10.7280/D1VX0C). The fold differences between Ibsp and Slpi in the LPS-treated deermice and control animals were 2,903 times ($P = 10-174$) and 1,280 times ($P = 10-57$) higher, respectively. In contrast, Ibsp was slightly expressed in mice under both conditions (the number of transcripts per million [TPM] was 0.001), and Slpi in expression was lower by 3× in the treated mice (FDR $P = 0.04$).

Sex-specific responses to LPS in *P. leucopus*

To identify sex-specific responses, the DEG analysis was applied to the RNA-seq of the blood and spleens of *P. leucopus* animals treated with LPS or saline alone. A noncoding RNA (GenBank accession no. XR_003736827) was expressed at orders-of-magnitude-higher levels in females than in males in all tissues. This noncoding RNA (ncRNA) was revealed as the inactive X-specific transcript (Xist), which functions to inactivate genes of one X chromosome of females (265). D1Pas1, an autosomal DEAD box RNA helicase, which is expressed in the testis in mice (266), was markedly lower in expression in all females than in males. Two genes that were expressed by a thousandfold more abundantly in the blood of LPS-treated females than in both control females and the males were those encoding adiponectin receptor 2 (Adipor2) and the X-linked lysine-specific demethylase 6A (Kdm6a). Kdm6a has been implicated in the risk of acquiring autoimmunity and reported to

regulate multiple immune response genes (267). Adiponectin, an adipokine, has an anti-inflammatory effect (268), and expression of its receptor is reportedly affected by the macrophage polarization phenotype (269). Adipor2 was also highly expressed in the spleens of LPS-treated females.

Functional processes distinguishing and shared between species

To further define similarities and difference between the responses of the two species to LPS, we identified gene groups categorized by gene ontology (GO) terms as *P. leucopus* specific, *M. musculus* specific, or shared, meaning, within each pairing, the enrichment groups that were up- or downregulated. For this analysis, 14,685 one-to-one orthologues were used for the comparisons. Figure 28 summarizes this analysis for blood, spleen, and liver.

There were 297 DEGs for the spleen that were shared between *Peromyscus* and *Mus*, but there were distinguishing features as well. Up-regulated GO terms for the *P. leucopus* spleen were tryptophan metabolic and catabolic processes, specifically Ido1. For *M. musculus*, uniquely associated GO terms that were upregulated in the spleen dealt with production and differentiation of myeloid cells, macrophages, and osteoclasts and included the cytokines IL-12 (Il12b) and interleukin 17 members IL-17A (Il17a), Il17f, and IL-23 (Il23a).

In the liver samples, four upregulated GO terms that distinguished the deermice from the mice were phosphatidylserine metabolic process and phosphatidylserine acyl-chain remodeling, which included phospholipase A2, group IIA (Pla2g2a), Pla2g5, and Pla1a; regulation of IL-5 and IL-13 secretion, both of which included the transcription factor GATA binding protein 3 (Gata3) (270); and tumor necrosis factor receptor superfamily, member 21 (Tnfrsf21), which reportedly is a determinant of influenza A virus susceptibility in mice (271). Two distinguishing GO terms for downregulated genes in the livers of LPS-treated *P. leucopus* animals were response to retinoic acid and cellular response to retinoic acid.
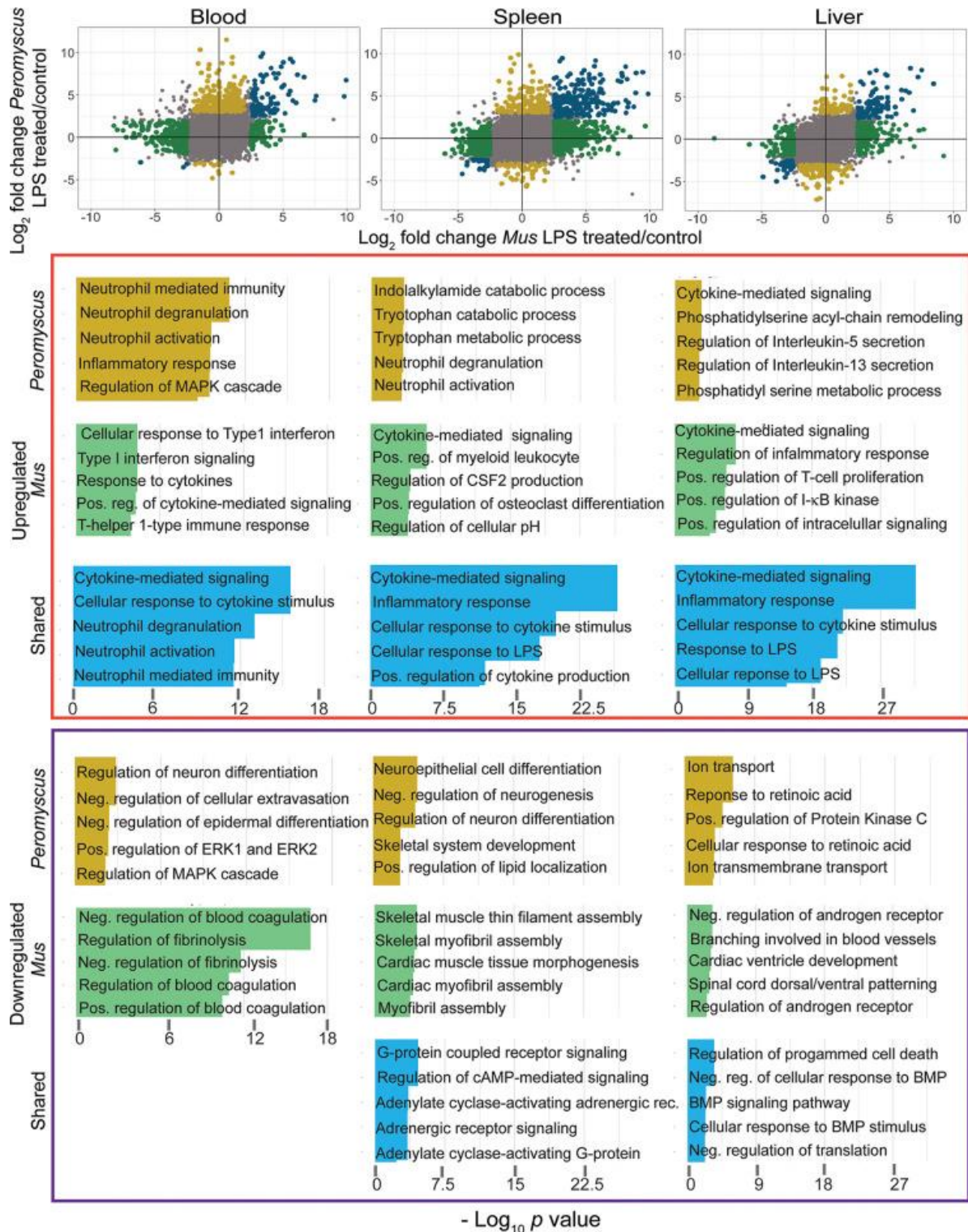
Figure 28 Comparison of P. leucopus and M. musculus animals in their responses to LPS by RNA-seq and categorization of DEGs by gene ontology enrichment

In the three scatterplots for blood, spleen, and liver tissues at the top of the figure, log2 values for the fold change of *P. leucopus* are plotted against corresponding values for *M. musculus* for each gene in the data set. DEGs specific for *P. leucopus* are indicated by gold symbols, while DEGs specific for

M. musculus are green. Genes shared between species among the DEGs are in blue. Gray is for all others. *P. leucopus*-specific, *M. musculus*-specific, and shared upregulated genes are in the upper half, right half, or upper right quadrant, respectively, of the plot. GO term enrichment was performed for each group of genes, separating upregulated and downregulated genes for each one of the tissues, accordingly. The colors in the horizontal bar graphs correspond with the colors indicated above. Pos., positive; reg., regulation; Neg., negative; MAPK, mitogen-activated protein kinase; cAMP, cyclic AMP; rec., receptor.

A distinguishing GO term for upregulated genes in *M. musculus* liver was positive regulation of I-kB kinase/NF-κB signaling, which included the NLR family, apoptosis inhibitory protein 5 or Naip5 (Birc), Ccl19, inhibitor of kappaB kinase epsilon (Ikbke), Il12, Il1α, LPS-induced TN factor (Litaf), Myd88, receptor tyrosine-kinase orphan receptor 1 (Ror1), Tlr2, and tumor necrosis factor (ligand) superfamily member 10 (Tnfsf10). Distinguishing GO terms for downregulated genes in LPS-treated mouse liver concerned the androgen receptor signaling pathway and included the transcription factor forkhead box H1 (Foxh1) and hairy/enhancer-of-split related with YRPW motif-like (Heyl), which is associated with repression of TGF-β signaling (272).

In summary, this layering of established GO terms over the DEG analysis provided consolidation and further evidence that *P. leucopus* and *M. musculus* have much in common in how they respond to LPS in the blood, spleen, and liver in the first few hours. But there were also several distinguishing functional processes. One feature of the *P. leucopus* response that particularly stood out in this analysis was the activation of neutrophils and other phagocytes. Included in the lists of specific genes under all the neutrophil-associated GO terms was the gene for Slpi, which was first identified as an inhibitor of serine proteases, such as neutrophil elastase (273). In *M. musculus*, notable distinguishing GO terms concerned blood coagulation and fibrinolysis, processes that figure in the pathophysiology of sepsis (274).

Gene network patterns across species

To further delineate the gene expression networks without a prior categorization, we performed weighted gene correlation network analysis (WGCNA) to empirically identify groups of genes

(modules) that distinguished specific species and/or treatments in response to LPS. Gene network analysis allowed identification of patterns across the three tissues. Since blood and the two organs inherently differed in their expression profiles across the genome and one tissue may overshadow the patterns in other tissues, we made a matrix in which each column represented a sample and each row represented a gene per tissue. Each gene had a suffix to identify the tissue of origin. Twenty-four modules were identified, and GO term enrichment was assessed for unique genes for each module. Figure 29 summarizes the results for four modules: darkorange2, darkseagreen3, brown4, and lightblue.

The darkorange module comprised genes that were upregulated in LPS-treated *P. leucopus* and not in either control *P. leucopus* or LPS-treated or control *M. musculus*. The blood made the largest contribution, with 1,472 genes to this module, followed by liver and spleen, with 933 and 776, respectively, each. The highest three ranked GO terms related to neutrophils: neutrophil-mediated immunity, neutrophil activation involved in immune response, and neutrophil degranulation. The three GO terms share several neutrophil-associated genes that were upregulated, including those for lysosomal cysteine proteases cathepsin S (Ctss) and cathepsin B (Ctsb) and lysosomal membrane protein 2 (Lamp2). The Slpi gene was also among the contributing genes for each of the three neutrophil GO terms.

The lightblue module comprised genes that were upregulated in LPS-treated *M. musculus* but not in either control mice or LPS-treated or control *P. leucopus* animals. In contrast to the darkorange module, the liver made the largest contribution to the lightblue module, with 1,170 genes, followed by spleen, with 566 genes, and blood, with 33 genes. The three highest-ranked GO terms for this module were cytokine-associated: cytokine-mediated signaling pathway, cellular response to cytokine stimulus, and response to cytokine. While the darkorange module also included two cytokine-associated GO terms at a lower rank than the three neutrophil-associated GO terms, only

a minority of genes (41 of 170) in the combined list overlapped between the two species for the

cytokine-mediated signaling pathway GO term.
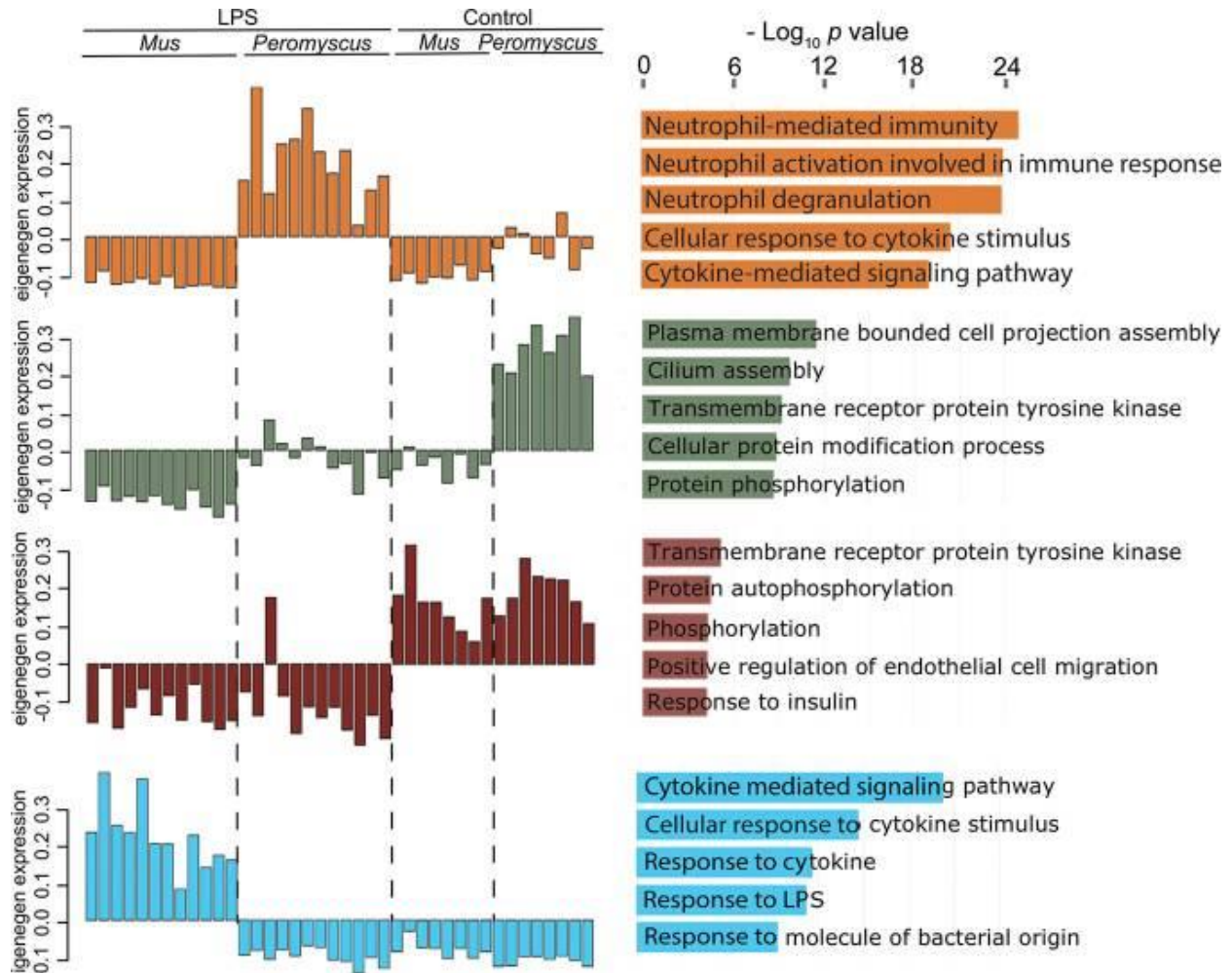
Figure 29 Four selected Eigengene modules by network analysis of differential responses of P. leucopus or M. musculus animals to LPS

The different modules are distinguished by the color of the hexadecimal scheme: dark orange 2 for upregulated in LPS-treated *P. leucopus*, dark sea green 4 for comparatively higher expression in untreated *P. leucopus* than in the other three groups, brown 4 for downregulated in both species after LPS treatment, and light blue for upregulated in LPS-treated *M. musculus*. The top 5 GO terms by adjusted *P* value are shown for each module. The modules from this analysis are available at the Dryad repository (https://doi.org/10.7280/D1B38G).

Selected upregulated genes unique to LPS-treated *M. musculus* under this GO term included those

for the following: cytokines Il1a, Il12, Il17, Il22, Il27, Il33, and Ifng; the chemokines Ccl5 (RANTES),

Cxcl9, and Cxcl13; macrophage markers Cd86 and Cd80; and transcription factors Jak3 and Stat4.

Selected DEGs associated with LPS-treated *P. leucopus* for this GO term included those for the following: Alox5, annexin 1 (Anxa1), Casp3, leptin (Lep), Sod2, Tgfb1, and tissue inhibitor of metalloproteinases (Timp1); the cytokines Il1b, Il10, Il19, and interferon beta (Ifn1b); and the transcription factors Gata3, Irf8, Stat1, and Stat3.

The two other modules highlighted in figure 29 are the darkseagreen module, which distinguished control *P. leucopus* from the other 3 groups, and the brown module, which featured GO terms that were downregulated in comparison to controls in both sets of LPS-treated animals. For the dark sea green module, the spleen was the major contributor, with 4,791 genes, followed by liver, with 3,189 genes, and blood, with 73 genes. Two GO terms with many genes in common were plasma membrane-bounded cell projection assembly and cilium assembly. For the brown module, the liver RNA-seq with a contribution of 876 genes far out-numbered the contributions of spleen, with 6 genes, or blood, with 3 genes. Genes common to the top three GO terms in this module were all genes for kinases: tyrosine-protein kinase ABL (Abl1), tyrosine-protein kinase CSK (Csk), mitogen-activated protein kinase 3 (Mapk3), and megakaryocyte-associated tyrosine kinase (Matk).

RT-qPCR of Nos2, Arg1, and Slpi

In analyses of RNA-seq data at the gene, pathway, or network level, the Slpi, Nos2, and Arg1 genes were three of the genes that distinguished the species. We confirmed the markedly higher expression of Slpi in *P. leucopus* blood after LPS by reverse transcriptase quantitative PCR (RT-qPCR) of replicate cDNA libraries from the RNA extracts (Table 11). The assays also confirmed the absent-to-low expression of Nos2 in *P. leucopus* blood in both LPS-treated and control animals, as well as the high baseline expression of Arg1 in the deermice and even higher in the LPS-exposed animals (Table 11).

Table 11 RT-qPCR of selected transcripts in blood of P. leucopus animals with and without LPS treatment

mRNA	Control (*n* = 8) mean no. of copies (95% CI)	LPS (*n* = 12) mean no. of copies (95% CI)	*t* test *P* value
Glyceraldehyde 3-phosphate dehydrogenase (Gapdh)	55,888 (32,512–96,070)	135,970 (75,934–243,471)	>0.05
Nitric oxide synthase 2 (Nos2)	0 (0–0)	5 (1–9)	>0.05
Arginase 1 (Arg1)	310 (199–481)	8,779 (5,185–14,864)	4×10^{-8}
Secretory leukocyte peptidase inhibitor (Slpi)	5 (3–9)	43,524 (29,039–65,235)	3×10^{-16}

Comparative expression of other genes

We examined other selected coding sequences in a cross-species analysis. Those chosen either typified one of the species' responses to LPS or were shared between species. Some genes were identified through gene-level DEG analyses (figure 27). Others were highlighted by GO term analysis (figure 28) or module-based analysis (figure 29). There was a single reference set that contained the pairs of orthologous protein-coding sequences (CDS) of the mRNAs for both *P. leucopus* and *M. musculus*. For each of the 40 animals, there were 15 genes for blood, 40 genes for spleen, and 12 genes for liver samples. Density distributions of the individual values in the each of the four experimental groups were plotted for 11 genes in blood, 12 genes in spleen, and 9 genes in liver). Figure 30 consolidates subsets of these as plots of 9 pairs of genes and presents results for blood, spleen, and liver.
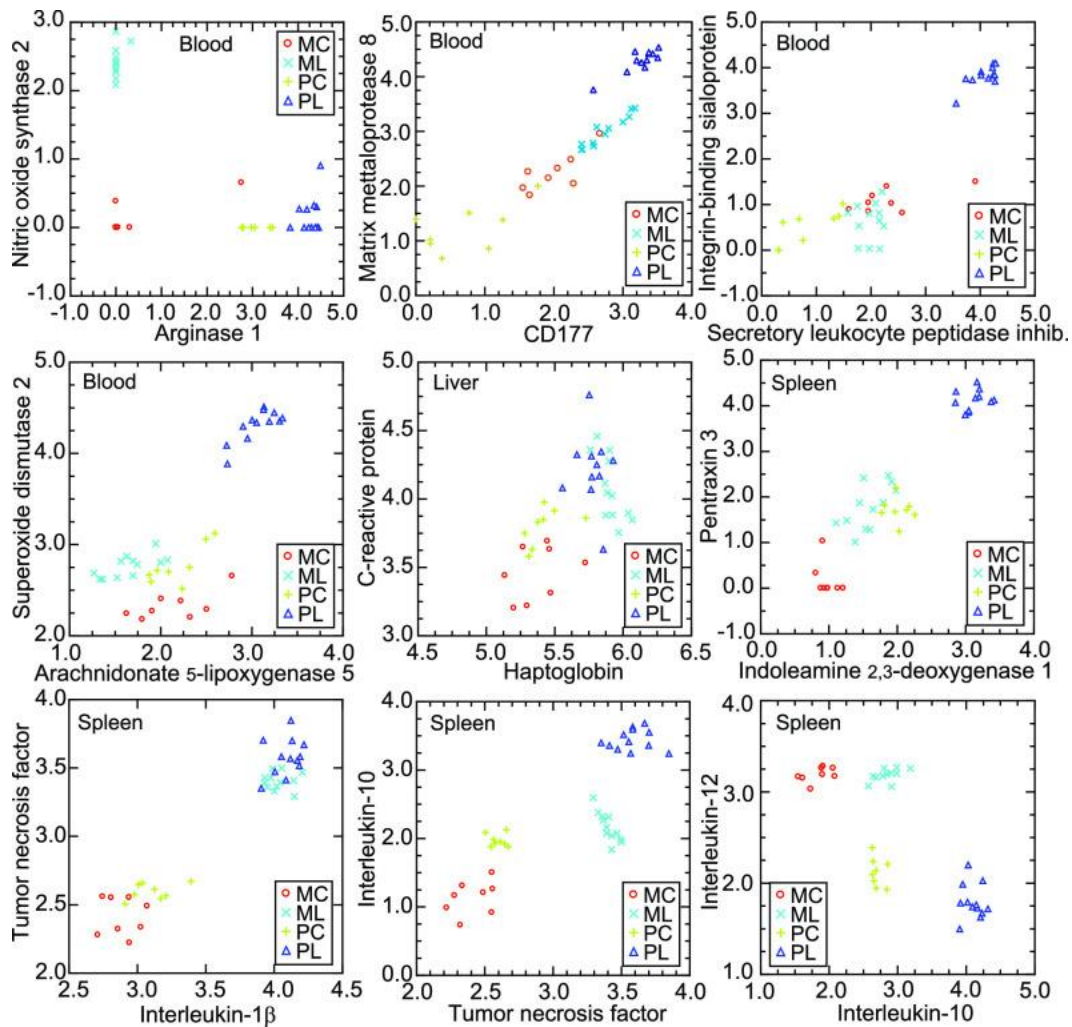
Figure 30 Correlations of pairs of selected genes of P. leucopus and M. musculus from the RNA-seq analysis

The 9 scatterplots are log10 values of normalized unique reads of one coding sequence against another for each of the four groups: control *M. musculus* (MC), LPS-treated M. musculus (ML), control *P. leucopus* (PC), and LPS-treated P. leucopus (PL). Each group is represented by a different symbol.

Up to a certain point, events after exposure to LPS appear similar in both species. Responses for both deermice and mice were highly correlated between the genes for TNF and IL-1β in the spleen and between the genes for the acute-phase reactants C-reactive protein, haptoglobin, and serum amyloid A (figure 30) (275). A couple of representative genes that were upregulated in both species but to a much greater extent in deermice were the Ptx3 and Ido1 genes. Pentraxin 3 is a member of a pattern recognition family involved in innate responses, and IDO1 is the rate-limiting enzyme of

tryptophan catabolism by the kynurenine pathway. Untreated *P. leucopus* and *M. musculus* displayed similar expression levels for the genes for the mitochondrial iron/manganese superoxide dismutase 2 (Sod2), an antioxidant, and arachidonate 5-lipoxygenase 5 (Alox5), which transforms essential fatty acids into leukotrienes. Both Alox5 and Sod2 were 10-fold higher in expression in the blood of LPS-treated deermice but about the same in all mice (figure 30). The difference between untreated animals and LPS-treated *P. leucopus* animals was even more striking for Slpi and Ibsp in the blood.

The previously noted discordance between the species in Nos2 and Arg1 expression was observed in this analysis as well. The profile of high Nos2-low Arg1 expression in response to LPS indicated a type M1 macrophage response in the mouse, while the low-Nos2–high-Arg1 profile in *P. leucopus* was more typical of the type M2 macrophage response. Further evidence of a dichotomy corresponding to polarized macrophage categories was the Il10 to Il12 ratio in the spleen (figure 30) (276). In control *M. musculus* animals, the mean ratio was 0.036 (95% CI, 0.026 to 0.046), rising 10-fold to 0.361 (95% CI, 0.251 to 0.471) in LPS-treated animals. The baseline Il10 to Il12 ratio of 3.47 (95% CI, 2.32 to 4.63) was a hundredfold higher in control deermice than in their mouse counterparts. For LPS-treated *P. leucopus* animals, the ratio was 191 (95% CI, 123 to 259), many hundredfold higher than in LPS-treated *M. musculus* mice.

Gauges of neutrophil or other white blood cell activity in the blood were Itgam (also known as CD11b), Fcgr1 (CD64), Fcgr2b (CD32), Cd177, and Mmp8 (figure 30). In mice treated with LPS, Itgam and Fcgr2b expression was either marginally lower than in controls or unchanged, while both genes were upregulated by severalfold in the LPS-treated deermice. Fcgr1 expression was higher in treated mice than in controls but in deermice was undetectable under both conditions. The pair Cd177 and Mmp8 provided the best discrimination between the two species (figure 30). Values for Cd177 and Mmp8 were tightly associated ($R2 = 0.87$) over the entire range for both species.

One of the genes associated with responses to LPS was growth differentiation factor 15 (Gdf15), which, among its ascribed functions, mediates tissue tolerance through triglyceride metabolism (277), but this gene was not annotated in the current genome assembly for *P. leucopus*; consequently, a corresponding transcript was absent from the reference set. Using the liver RNA-seq data, we performed a targeted analysis with the *M. musculus* coding sequence for Gdf15 and, as a substitute for *P. leucopus*, the orthologue in *P. maniculatus*. The log10-transformed normalized reads were higher in LPS-treated mice than in controls: 2.7 (95% CI, 2.5 to 2.9) versus 1.9 (95% CI, 1.8 to 2.1) (P = 0.0004). This is an average fold change of 5.2. For the deermice, the reads in log10 values for controls were nearly identical to those for control mice: 1.9 (95% CI, 1.8 to 2.1). In contrast to mice, numbers were marginally lower in LPS-treated deermice: 1.8 (95% CI, 1.7 to 1.8). The difference between the species in the Gdf15 response was largely attributable to higher values among treated male mice.

Diversity of responses within the populations

We investigated whether the outbred deermice were more heterogeneous in their responses than the inbred BALB/c mice with a pairwise analysis of coefficients of determination ($R2$) among the 24 animals that received LPS. We used spleen because of 150-nucleotide (nt) reads for this tissue and similar numbers of up- and downregulated DEGs in both species (figure 27). With self-pairings excluded, there were 66 pairs each for *P. leucopus* or *M. musculus* and 144 cross-species pairs. The genes were those for Alox5, Ccl2, Csf2, Csf3, Cxcl11, Gapdh (glyceraldehyde-3-phosphate dehydrogenase), Hmox1, Il1b, Il1rn, Il6, Lcn2, MT-Co1, Nfe2l2, Slpi, Sod2, Tgfb1, and Tnf. The median fold difference between *P. leucopus* and *M. musculus* for these genes was 1.2. Figure 31 shows the distributions of $R2$ values of the sets of pairs. As expected, correlations were lower between mixed-species pairs than between intraspecies pairs (upper panel); the median $R2$ was 0.701. The seven highest intraspecies pairwise $R2$ values (0.987 to 0.994) were observed among the *M. musculus*

mice, and the median of the 66 *R2* values was marginally higher, at 0.954, for the mice than the 0.948

for the deermice, but, overall, *P. leucopus* animals in this experiment were not notably more diverse

in their responses than were the inbred mice under the same conditions (P = 0.68; Mann-Whitney

P = 0.19), an indication that the discriminating DEGs, enrichment GO terms, and modules were not

wholly or partially attributable to greater variances among the deermice than the mice.
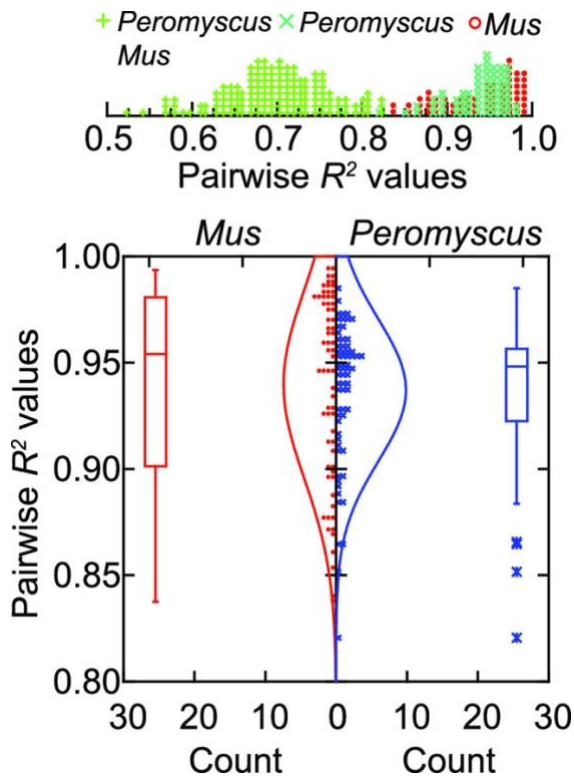
Figure 31 Assessment of diversity among individual animals by species in transcriptional responses to LPS for 17 genes

Pairwise coefficients of determination (R2) were calculated for the 66 intraspecies pairs for LPS-treated *P. leucopus* animals, the 66 intraspecies pairs for LPS-treated *M. musculus* animals, and the 144 interspecies pairs for all LPS-treated animals. The top panel is a frequency distribution of R2 values for all pairwise determinations. The bottom panel shows just the distributions of intraspecies pairwise determinations.

RNA-seq of older *P. leucopus* animals in response to LPS

We found similar profiles in blood and spleen samples of older animals in an experiment that was

carried out before the deermouse-mouse comparison but under the same conditions: a 10-µg/g dose

or buffer alone as a control and samples taken after 4 h. The 16 animals (12 females) had a median age of 81 weeks, with a range of 54 to 94 weeks. Nine animals received LPS and seven buffer alone. We observed again in this second experiment low baseline and after-treatment expression of Nos2 in the blood but high baseline expression of Arg1, with a further increase in the LPS-treated animals. Both Slpi and Ibsp increased more than a hundredfold in expression in the blood after LPS. In the spleen, Il1b, Tnf, Il6, Il10, Ccl2, Cxcl11, Csf2, and Csf3 all were increased in the LPS-treated animals, as were two genes that notably were more upregulated in the samples from young *P. leucopus* deermice than in comparably aged *M. musculus* mice: the Ido1 and Ptx3 genes.

Systemic bacterial infection

The simplicity of the LPS model of inflammation was a strength, but how representative was it of a bacterial infection in terms of the findings to this point? We chose the relapsing fever agent *Borrelia hermsii*, because *Peromyscus* species are natural reservoirs for the species (278), and because it has lipoproteins, instead of LPS, that are the ligands for Toll-like receptor signaling (279), (280). For this question, five *P. leucopus* were infected with the relapsing fever agent *Borrelia hermsii* on day 0. Another 3 animals received buffer alone. Infection of the blood was directly confirmed by microscopy on day 4. The animals were euthanized on day 5, just before the appearance of neutralizing antibody was anticipated (95). The infected animals had enlarged spleens, as well as large numbers of bacteria in the spleens, as assessed by quantitative PCR. We used previously collected but unanalyzed PE100 reads for the blood of these animals (6). As noted in the two LPS experiments, Nos2 expression was hardly detectable in either the controls or the infected animals. On the other hand, Arg1 expression was high at baseline and further elevated by day 5 of infection. Two genes of *P. leucopus* that clearly distinguished this species from *M. musculus* in its short-term response to LPS were the Slpi and Ibsp genes; both of these were a hundredfold more highly expressed in infected animals than in controls.

RNA-seq was carried out with the same reference set and settings as for the samples in a comparative response study. Given the smaller sample size of the infection study, we focused on more highly expressed genes. Of the 46,154 transcripts in the full reference set, 1,773 had mean TPM values of ≥50 in at least one of the two groups in each experiment. The latter set was used as the basis for determining and then comparing fold changes between the LPS condition with a duration in hours and the systemic bacterial infection condition with a duration in days. Figure 32 summarizes the pairwise fold changes for individual genes. Two-thirds of the DEGs in the infection experiment were also represented among the DEGs in the LPS experiment. We noted again the substantial upregulation of Slpi and Ibsp. Other coding sequences upregulated under both conditions were Acod1, Alox5, Arg1, Cxcl2, Il1rn, Il1b, Lcn2, and Mmp8. A discordant DEG was dual-specificity phosphatase (Dusp), which has been implicated in regulating lipid metabolism during sepsis (281).
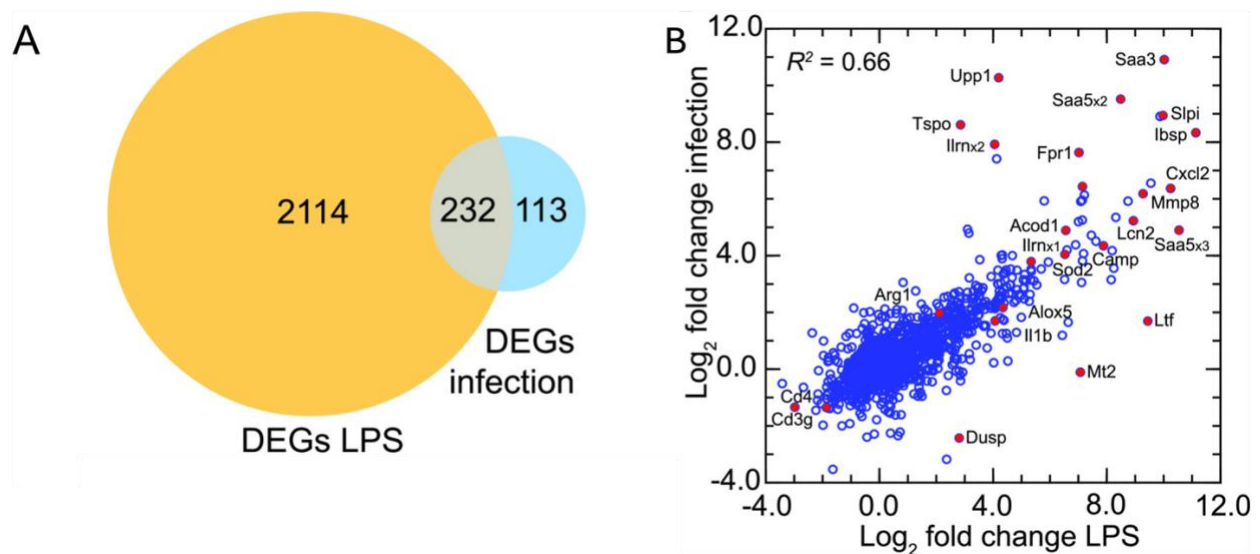
Figure 32 Comparison of DEGs of RNA-seq of P. leucopus deermice treated with LPS and P. leucopus animals systemically infected with the bacterial agent Borrelia hermsii
There were 12 LPS-treated animals with 8 controls and 5 infected animals with 3 controls. (A) Venn diagram of numbers of DEGs in each experiment and the overlap between them (B) scatterplot of log2-transformed fold changes between study and control conditions for the infection experiment (*y* axis) versus the short-term LPS experiment (*x* axis). Selected genes are indicated with a label adjacent to a red symbol for the data point.

LPS treatment of *P. leucopus* fibroblasts

Low-passage-number cultures of fibroblasts, a type of stromal cell isolated from ear tissues from five LL stock deermice, were split into pairs. Then one member of each pair was treated with LPS for 4 h, and the other member was treated with saline alone. Of the 46,141 transcripts in the reference set, 18,462 had a mean TPM of >1 in either the control or the LPS group, and these were used in the DEG analysis (Dryad repository, https://doi.org/10.7280/D1MD69). For protein coding sequences, the 10 highest numbers of TPM, in order, among control samples were for ferritin heavy chain (Fth), Slpi, secreted protein acidic and cysteine rich (Sparc), eukaryotic translation elongation factor 1 alpha (Eef1a1), ferritin light chain (Ftl), vimentin (Vim), serpin family H member 1 (Serpinh1), ribosomal protein lateral stalk subunit P1 (Rplp1), collagen type I alpha 1 chain (Col1a1), and Gapdh. There were 324 genes that were upregulated by the criterion of a fold change of ≥4 and an FDR of <0.05, and 17 genes that were downregulated, with an additional 80 downregulated by the criterion of a fold change of ≥2 and an FDR of <0.5 (figure 33). Among those displaying the marked increases in expression between the control and LPS conditions were two subunits of nuclear factor kappa B (Nfkb), two forms of Saa, Nos2, Csf3, Sod2, and Pla2g2a, which is associated with inflammation during systemic bacterial infection (282). Slpi transcripts were not only surprisingly abundant in the fibroblast cultures under usual cultivation conditions, but expression further increased by 1.5-fold (95% CI, 1.2- to 1.9-fold) in those pair members exposed to LPS ($P = 0.02$). Expression of Nos2 increased 792-fold (95% CI, 124- to 5,073-fold) in the LPS-treated fibroblasts between paired specimens ($P = 10^{-5}$). This was evidence that the absent or scant expression of inducible nitric oxide synthase that we observed in both control and LPS-treated *P. leucopus* animals was not attributable to a genotypic incapacity to transcribe the Nos2 gene.
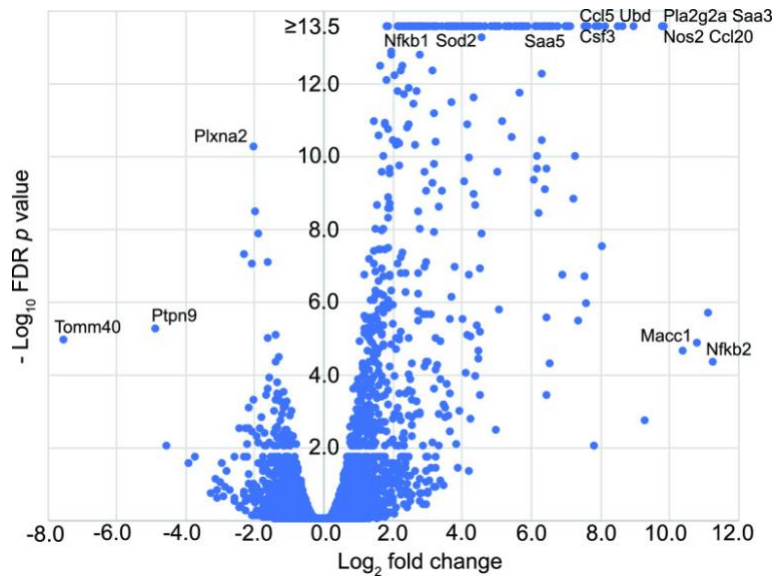
Figure 33 Volcano plot of RNA-seq results for pairs of P. leucopus fibroblast cultures with or without exposure to LPS

Fold changes for 543 genes are given on the x axis, and false discovery rate P values are given on the y axis. For conciseness, the upper limit for the –log10 values for this graph was 13.5. The exact or approximate locations of selected differentially expressed genes are shown. Numerical values for each gene in the data set are provided in the Dryad repository (https://doi.org/10.7280/D1MD69).

Conjunctivitis

After identification of a variety of DEGs under conditions of LPS treatment of whole animals and isolated cells or of systemic infection, we returned to the disease sign that we observed in *P. leucopus* but not in *M. musculus*, namely, conjunctivitis severe enough to cause eyelid closure within 2 to 4 h of the LPS injection. Six of 12 animals receiving the LPS overtly manifested conjunctivitis. DEGs of note, in the blood of deermice with conjunctivitis, we found 21-fold-higher expression of Cxcl13, 12-fold-higher expression of Ccl6, but 46-fold-lower expression of lysozyme (with an FDR of <10−4), an antimicrobial enzyme and constituent of tears that coat and protect the conjunctiva (283). In spleen samples, the conjunctivitis animals was distinguished by 6- to 7-fold-higher expression (FDR, <0.005) of two forms of carbonic anhydrase, an enzyme of relevance for the eye disorder glaucoma, because of its role in producing the aqueous humor in the eye (284).

Discussion

The study yielded abundant information about a small mammal that is emerging as an informative model organism, not only for the study of pathogenesis and immunology of infectious diseases but also for the fields of aging, behavior, ecology, and reproductive biology. The question driving this investigation was how do deermice largely avoid morbidity and mortality when infected with pathogens that otherwise are disabling if not fatal for humans? "Sickness" has various definitions, including one operationally based on behavior for animal studies (285). *P. leucopus* and *M. musculus* animals became equivalently sick by the behavioral criteria of reduced activity, mutual huddling, and lower food intake. Nevertheless, under conditions that frequently led to death or the moribund state in various strains of *M. musculus* in several studies, *P. leucopus* animals receiving the same or even much higher doses of LPS pulled through and recovered.

To address this central question, we used RNA-seq of whole blood, spleen, and liver. The study did not specifically examine differences in mRNA isoforms or small noncoding RNAs. These results were complemented by analysis of a limited set of coding sequences, specific RT-qPCR assays, direct detection of the compounds in the blood, and metabolomics. The most extensive experiment was on the responses among several young adult animals of both sexes to a single dose of LPS at approximately the LD50 for mice, but we obtained similar results with *P. leucopus* of considerably older age under the same conditions and with animals that had systemic bacterial infections of 5 days' duration. While we noted some differences between the sexes of *P. leucopus* animals in the experiment, for the most part, females and males responded the same to LPS, at least in the short term and at these doses. The experiment with *P. leucopus* fibroblast cells in culture exposed to LPS showed that the some of the findings, such as high expression of Slpi, are reproduced *in vitro* and, thus, feasibly exploited by transgenic and silencing technologies.

Integration of the metabolomics results for plasma with the transcriptomics results for the three tissues was necessarily limited. The sources of some of the metabolites in the blood would likely have been organs, e.g., the adrenal glands, or gut microbiota not subjected to RNA-seq. Nevertheless, metabolomics confirmed that *P. leucopus* responded similarly to *M. musculus* in many respects to the same dose of LPS, as well as in corticosteroid biosynthesis. The evidence of greater catabolism of tryptophan by the kynurenine pathway in the LPS-treated deermice than in mice was associated with comparatively higher expression of Ido1 in the deermice.

One of the striking differences between the two species in their responses to LPS was overall predominance in the GO term and *de novo* module analyses of neutrophil-associated genes for *P. leucopus* and cytokine-associated genes for *M. musculus* (figures 28 and 29). The blood was the greatest contributor to the set of genes that distinguished LPS-treated *P. leucopus* animals from treated and untreated *M. musculus* animals and from control deermice, while the liver provided the majority of genes that distinguished LPS-treated *M. musculus* mice from deermice and from control mice. A limitation of the study was that complete blood cell counts were not terminally performed; this was to ensure sufficient RNA for high-coverage sequencing and plasma for untargeted metabolomics. It is possible that the prominence of neutrophil-associated GO terms in LPS-treated deermice was accounted for by higher numbers and proportions of neutrophils in the blood of deermice than in the blood of mice. But this is not likely the sole explanation. Healthy animals of the LL stock population of *P. leucopus* at the Peromyscus Genetic Stock Center (PGSC) had lower absolute numbers and proportions of neutrophils than what was reported for the BALB/c population from which our animals were drawn (286), (287) (see Materials and Methods). A study of a separate breeding colony of LL stock *P. leucopus* animals also reported comparatively low absolute numbers and proportions of neutrophils among healthy adult deermice of both sexes (288). In addition, if there was a difference between species in overall white cell activity 4 h after injection of LPS, this was not

reflected in total mitochondrion reads, which represented the number of mitochondria and transcription by individual mitochondria in the blood sample. Values declined in both species from baseline by about the same proportion.

The comparatively heightened neutrophil transcriptional activity in *P. leucopus* paralleled the appearance of conjunctivitis with purulent exudates in the eyes of deermice but not mice. While increased activity of neutrophils may lead to tissue damage from elastase and other proteases and from reactive oxygen species, this may be ameliorated in the deermice by such factors as the leukocyte protease inhibitor Slpi and Sod2, a defense against oxygen radicals.

Slpi is a nonglycosylated 12-kDa cationic, cysteine-rich protein, which is known to be increased in expression in response to LPS and lipoteichoic acids and under the stimulus of TNF and IL-1β (289). It has an inhibitory effect on the formation of neutrophil extracellular traps (290). Slpi inhibits neutrophil proteases like cathepsins and elastase but also prevents degradation of the NF-κB inhibitory proteins IκBα and IκBβ through its antiprotease activity (291). Slpi knockout or deficient animals had impaired wound healing and increased inflammation (292), elevated Nos2 activity in macrophages (293), and increased susceptibility to LPS-induced shock (294).

While LPS-treated *P. leucopus* animals exhibited increased expression of the TLR signaling adapter MyD88 and downstream inflammatory cytokines, such as IL-6, TNF, and IL-1β, to approximately the same degree as in *M. musculus*, this was countered by greater expression of cytokines associated with M2 polarization, such as IL-10 and TGFβ, as well as other anti-inflammatory mediators, like annex A1 (295), in the context of lesser expression of IL-12 and IFN-γ. The GO term cytokine-mediated signaling pathway served to distinguish both LPS-treated *P. leucopus* and LPS-treated *M. musculus* animals from their untreated counterparts (figure 29) but the sets of constituent genes that populated the GO term for each species overlapped for only a minority.

Another notable difference was the contrasting expression of arginase 1 and inducible nitric oxide synthase, or Nos2, in the two species. In this phenotype, which distinguishes deermice from mice under the same conditions, *P. leucopus* resembled another cricetine rodent, the golden hamster (*Mesocricetus auratus*), during infection with *Leishmania donovani* (296), (297). This phenotype can also be mapped to macrophage polarization, where type M1 macrophages generally feature high Nos2 and low Arg1 expression and where most varieties of type M2, or alternatively activated macrophages, are characterized by low Nos2 and high Arg1 (264). The differences between species in their IL-10/IL-12 ratios before and after LPS were also consistent with a greater disposition of *P. leucopus* to a type M2 response (figure 30) (276). Further categorization into the M2 subtypes recognized in mice was inconclusive at this point. Nos2 gene knockout mice are less likely to succumb to the toxicity of LPS, but at the cost of greater susceptibility to *Listeria monocytogenes* or to *Listeria major* infections (259), (298). While experimental infections with these particular intracellular pathogens have not been reported for *P. leucopus*, this species has not been notably at risk of morbidity or mortality from the obligate intracellular bacterial pathogen *Anaplasma granulocytophilia* (299) or the Powassan encephalitis virus (300), which utilizes *P. leucopus* as a reservoir.

Integrin-binding sialoprotein (Ibsp), also known as bone sialoprotein 2 (BSP2), is a revelatory case in this study. There were marked increases in the expression of its gene in *P. leucopus* but not in *M. musculus* upon exposure to LPS. The Ibsp protein has predominantly been associated with bone and tooth morphogenesis and not with inflammation or innate immunity. Consequently, it is not expected to be included in innate immunity and inflammation pathways or GO terms that distinguish LPS-treated from untreated animals. However, there is justification for viewing Ibsp as *P. leucopus'* functional substitute for osteopontin or bone sialoprotein 1 (BSP1), another small integrin-binding glycoprotein that is heavily modified posttranslationally (301). Both BSP1 and BSP2 bind to alpha v

beta 3 integrins, and their genes are within 100 kb of each other on chromosome 4 of humans and chromosome 22 of *P. leucopus*. Osteopontin is a biomarker of sepsis in humans (302) and reportedly acts through Stat1 degradation to inhibit Nos2 transcription (303).

Could observed differences between the two species in these experiments be attributable in part to differences in their microbiomes? The gut metagenomes were determined from preexperiment fecal pellets from the animals in this study in Chapter 2. In general, the deermice and mice had similar distributions and frequencies of bacteria at the taxonomic level of the family and in the representation and proportions of different biosynthetic, metabolic, catabolic, and regulatory functions, but there were also substantive differences in the gut microbiotas between the species that plausibly account for some distinguishing responses to LPS. The first was the presence in the intestine in *P. leucopus*, but not in *M. musculus*, of a protozoan, provisionally a new species of the parabasalid genus *Tritrichomonas*. The presence of *Tritrichomonas muris* in some populations of the same inbred strains of *M. musculus* mice altered their cellular immune responses to other microorganisms and antigens (177). Further distinguishing the gut microbiota was the greater abundance and diversity in *P. leucopus* of *Lactobacillus* spp as described in Chapter 2. In two studies on *Lactobacillus* spp. as probiotics, feeding of mice with *Lactobacillus paracasei* or *Lactobacillus plantarum* resulted in less inflammation in controls when they were challenged with an influenza virus or *Klebsiella pneumoniae* (304), (177). In the present study, a tryptophan metabolite in the plasma that distinguished the two species was indolepropionic acid which in controls was ~4-fold-greater in abundance in deermice than in mice ($P = 0.003$). This compound is the metabolic product of certain intestinal bacteria (305) and reportedly has anti-inflammatory and antioxidant properties (306).

In summary, *P. leucopus* differed in many respects from *M. musculus* in these experiments in its response to LPS. Under conditions in which about half the mice were expected to die within 48 h, the

deermice survived the acute insult and recovered, a phenomenon that modeled the tolerance of these animals to infection. Simple pairwise variables characterizing the two species' different responses to LPS were the ratios of Arg1 to Nos2 transcripts, IL-10 to IL-12 transcripts, and kynurenine to tryptophan metabolites. In their outcomes after LPS exposure, the naive *P. leucopus* animals resembled *M. musculus* animals that had become accustomed to LPS by prior exposure to a low dose of LPS (307). A characteristic profile of these pretreated mice is comparatively lesser expression of proinflammatory mediators and greater expression of genes of phagocytes (308), similar to what we observed in naive *P. leucopus*.

Whether this points to a single pathway or even a single gene remains to be determined, but we doubt that the phenomenon of infection tolerance can be reduced to a simple explanation. The evidence rather is of multiple adaptive traits representing different aspects of innate immunity, metabolism, oxidative stress management, and perhaps the microbiome that serve to sustain *P. leucopus* populations amid the varied infectious agents that they face. This is also to the benefit of these agents in this trade-off, because it renders *P. leucopus* a competent vertebrate reservoir for them and one of impact for human populations in areas of endemicity.

Materials and Methods

Animals

Adult outbred *P. leucopus* of the LL stock were obtained from the Peromyscus Genetic Stock Center (PGSC) of the University of South Carolina (http://stkctr.biol.sc.edu, accessed 10 February 2021). The LL stock colony was founded with 38 animals captured near Linville, NC, between 1982 and 1985 and has been closed since 1985. Sib-matings are avoided, and complete pedigree records are kept. Animals of the LL stock have mitochondria with the same genome sequence with little or no heteroplasmy (33). Adult BALB/cAnNCrl (BALB/c) and C.B.17 strain severe combined immunodeficiency (SCID) *M. musculus* mice were purchased from Charles River. The SCID mice

were used for propagation of bacteria for infections (see below). The *P. leucopus* deermice in the study were drawn from an LL stock population at the PGSC with a mean total number of white blood cells per microliter of blood of 8,600 and absolute numbers (percentages of total white blood cells) of neutrophils, lymphocytes, and monocytes of 510 (5.9%), 8,000 (93%), and 90 (1.0%), respectively, per µl (286). The BALB/c mice were obtained from a population with a reported mean of 9,250 white blood cells per µl of blood and corresponding values for neutrophils, lymphocytes, and monocytes of 1,880 (20%), 6,600 (71%), and 640 (7%), respectively, for 8- to 10-week-old animals (287). All animals in the LPS experiments spent at least 2 weeks at the UCI facilities before the experiment.

Animals were maintained in the AAALAC-accredited UC Irvine vivarium, with 2 to 5 animals per cage according to sex and on a 12-h-light/1-h-darkness lighting schedule, at a temperature of 21 to 23°C and a humidity of 30 to 70%, with water *ad libitum*, and with a diet of 8604 Teklad Rodent (Harlan Laboratories). Prior to injections, animals were lightly anesthetized with 2.5% isoflurane in the presence of 2 liters/min oxygen. The rodents were euthanized by carbon dioxide overdose and intracardiac exsanguination at the termination of the experiment or if they were moribund, unable to eat or drink, or otherwise distressed. Dissection was carried out immediately. Instruments were cleaned first and then sterilized between dissections.

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 (309). University of California Irvine protocol AUP-18-020 was approved by the Institutional Animal Care and Use Committee (IACUC). The protocol for the comparative study of *P. leucopus* and *M. musculus* animals for responses to LPS after 4 h was in addition approved by the Animal Care and Use Review Office of the U.S. Army Medical Research and Materiel Command. *P. leucopus* animals studied at the PGSC were under IACUC-approved protocol 2349-101211-041917 of the University of South Carolina.

The gut metagenomes from feces collected from the 20 *P. leucopus* and 20 *M. musculus* animals 1 to 2 days before the comparative experiment have previously been described in Chapter 2 and are available from the MG-RAST database (https://www.mg-rast.org) under accession numbers mgm4832531 to mgm4832578.

LPS susceptibility and dose responses

Thirty adult *P. leucopus* deermice, divided into five groups of three females and three males, were each injected intraperitoneally (i.p.) on day 0 with a 50-μl volume of *Escherichia coli* O111:B4 LPS purified by ion exchange and with <1% protein and <1% RNA (Sigma-Aldrich; catalog L3024), which was diluted in sterile, endotoxin-free 0.9% saline (Sigma-Aldrich) to achieve the following doses in milligrams per kilogram of body weight: 10, 50, 100, 200, and 300. The doses were administered in randomized order over the period from 1400 to 1700 h of a single day. Animals were returned to their cages with *ad libitum* food and water and then monitored every 12 h for the following signs: reduced activity by criterion of huddling with little or no movement for >5 min, ruffled fur (piloerection), hyperpnea or rapid respiration rate, and conjunctivitis by the criterion of closed eyes with crusting observable on the eyelids. The primary endpoint was death during the period between observations or the moribund state (immobility, rapid respiration, and inability to feed or drink) at the scheduled monitoring time or when notified in the interim by vivarium attendants.

Single-dose LPS comparison

Animals were anesthetized with isoflurane and injected i.p. with a single dose of *E. coli* O111:B4 LPS at a concentration of 10 mg/kg body weight in a 50-μl volume as described above. The control group was anesthetized and then injected with the 0.9% saline alone. The experiment started at 0800 h, with 10-min intervals between animals and with alteration of LPS and control injections. At 4.0 h after their injection, the animals were euthanized. After the chest was opened, exsanguination was performed by cardiac puncture and blood was transferred to a heparin sulfate-coated tube (Becton,

Dickinson Microtainer). Anticoagulated blood was centrifuged to pellet blood cells for 3 min at 4,600 × g at 4°C. Plasma and blood pellets were kept separately at −80°C until further analysis. Liver and spleen were extracted, flash-frozen in liquid nitrogen, and stored at −80°C until RNA extraction.

Experimental infection

Infection of a group of adult *P. leucopus* deermice of the LL stock with the relapsing fever agent *Borrelia hermsii* strain MTW of genomic group II was described by Barbour et al. (6). *Peromyscus* species are natural hosts of genomic group II strains of *B. hermsii* (278). In brief, animals were anesthetized and then injected on day 0 with 103 bacteria divided between the i.p. and subcutaneous routes in 50-μl volumes of PBS and diluted plasma from infected SCID mice, as described previously (20). On day 4, a drop of tail vein blood was mixed with PBS and examined as a wet mount by phase-contrast microscopy to confirm infection. On day 5, animals were euthanized with carbon dioxide and terminal exsanguination. Whole blood was dispensed into heparin-coated tubes, and the spleens were removed by dissection, weighed, and then flash frozen in liquid nitrogen. Confirmation of infection and quantitation of bacterial burdens in the spleens were carried by quantitative PCR, as described previously (310).

Fibroblast culture and LPS treatment

Fresh ear punches were collected from five LL stock *P. leucopus* animals (2 females and 3 males) during routine marking procedures at the time of weaning at ~3 weeks of age. The PGSC identification numbers (and mating pair for each animal) were 22608 (H-1075), 22609 (H-1075), 22610 (H-1121), 22611 (H-1121), and 22614 (H-1127). After the ear punch tissue was bathed in 70% ethanol for 2 min, it was placed in RPMI 1640 medium (HyClone FetalClone II; Thermo Scientific) supplemented with 10% fetal bovine serum (Gibco), 100 U/ml penicillin, 100 μg/ml streptomycin, and 0.292 μg/ml l-glutamine (HyClone) as previously described (311). Ear punches were minced and then treated with 2 mg/ml collagenase type I (Millipore) for 1 h. Undigested debris was removed once

cells were visible. The disassociated cells were cultivated in the same medium at 37°C and in 5% CO2. Cells were passed when adherent layers reached 90% confluence and for no more than 7 passages. For the experiment, individual cultures were split into pairs, and they were incubated at initial concentrations of 3 × 105 cells per well for 24 h. *E. coli* O111:B4 LPS or saline alone was added to the medium for a final LPS concentration of 1 μg/ml, and then the incubation was continued for 4 h. After disassociation of the fibroblast layer with trypsin and then addition of RNAlater (Thermo Scientific), the cells were harvested and stored at a concentration of ~106/ml in −80°C until RNA extraction.

Analyses of blood

Nitric oxide was analyzed from frozen plasma samples using a nitric oxide assay kit (Invitrogen) according to the manufacturer's instructions; the nitrite concentration was calculated based on standard curve measuring optical density at 540 nm on a plate reader (BioTek Synergy2). Corticosterone was measured colorimetrically in plasma using DetectX corticosterone enzyme immunoassay kit K014-H1 (Arbor Assays). After incubation, the reaction was read at 450 nm on a BioTek Synergy2 plate reader.

Metabolomics

Untargeted detection and analysis of metabolites in plasma of *P. leucopus* and *M. musculus* were carried out essentially as described previously (312). In brief, 40-μl volumes of plasma, which had been stored frozen at −80°C, were extracted with 120 μl methanol containing as internal standards phenylalanine d5 (175 ng/ml), 1-methyl tryptophan (37.5 ng/ml), and arachidonoyl amide (30 ng/ml). After precipitated proteins were removed by centrifugation, the supernatant was dried under vacuum and then suspended in 50% methanol. Aliquots of 10 μl were subjected to high-pressure liquid chromatography (HPLC) and quadrupole time-of-flight mass spectrometry (Q-TOF-MS) with periodic inclusion of pooled samples for quality control. Metabolites were separated on an Agilent

Technologies Poroshell C8 column (100 by 2.1 mm, 2.7 µm) with a gradient of acetonitrile and water, both containing 0.1% formic acid, with an Agilent 1260 HPLC pump under conditions previously described (105). The eluent was introduced into an Agilent Technologies 6520 Q-TOF-MS instrument equipped with an electrospray ionization source. The parameters for the analysis were a capillary voltage of 4,000 V, a fragmentor voltage of 120 V, gas at 310°C, a gas flow of 10 liters/min, and a nebulizer pressure of 45 lb/in2 (gauge). Data were acquired in the positive ion mode at a scan range of 75 to 1,700 for the mass-to-charge ratio (*m/z*) and at a rate of 1.67 spectra per second. The raw data were deposited at the Metabolomics Workbench repository (https://www.metabolomicsworkbench.org) (https://doi.org/10.21228/M8PH55).

The raw data files were converted to the XML-based mzML format with ProteoWizard (313). The files were then processed for peak picking, grouping, and retention time correction with the XCMS suite of software (314). The centWave algorithm was applied to detect chromatographic peaks with the following parameter settings: ≤30 ppm for *m/z* deviation in consecutive peaks, a signal-to-noise ratio of 10, a prefilter setting of 3 scans with peak intensities of ≥750, and 10 to 45 s for peak width. Molecular features (MF), defined by *m/z* and retention time, were grouped across samples using a bandwidth of 15 and an overlapping *m/z* slice of 0.02. Retention time correction was performed with the retcor algorithm of XCMS. Finally, peak area was normalized by the median fold normalization method as previously described (315).

Pathway enrichment analysis was performed with Mummichog network analysis software v. 2 (316), which predicts functional activity from spectral feature tables without *a priori* identification of metabolites. This was implemented online at MetaboAnalyst (https://www.metaboanalyst.ca/). The parameter settings were a mass accuracy of 20 ppm, a positive ion mode, and a cutoff *P* value of 0.05 for the Fisher's exact test of observed and expected hits of a given pathway. Pathway enrichment in Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg/) terms was

determined separately for the *P. leucopus* and *M. musculus* sets of LPS-treated and control animals. A cross-species comparison of enrichments for pathways in common was plotted. Identifiable metabolites constituting the pathways that distinguished between LPS-treated and control animals were listed for each species. As a measure of abundance, the peak areas of selected metabolites were extracted from raw data set using Skyline software (317). Molecular features data for the 40 individual plasma samples was deposited within the Dryad data repository (https://doi.org/10.7280/D12M4N).

RNA extractions

Cells of freshly obtained heparinized blood were pelleted by centrifugation for 3 min at 7,000 × *g*. The plasma supernatant and cell pellet were rapidly frozen on dry ice and stored at −80°C. For extraction of RNA from blood cell pellets in LPS experiments, we used a NucleoSpin RNA blood minikit (Macherey-Nagel). Lysis buffer and proteinase K were added to the frozen pellet and shaken while the pellet thawed. RNA from liver and spleen was extracted from 30 mg of frozen tissues, mechanically homogenized, and further lysed in Buffer RLT (Qiagen) with 2-mercaptoethanol in a TissueLyser (Qiagen) instrument with 3-mm stainless steel beads. Extraction was carried out using an RNeasy minikit (Qiagen). Extraction of total RNA from blood of *P. leucopus* animals infected with *B. hermsii* was as described by Long et al. (5). Frozen suspensions of dissociated, cultured fibroblasts in RNAlater (ThermoFisher Scientific) were extracted with an RNeasy minikit (Qiagen). Nucleic acid concentrations were determined with a Qubit fluorometer (ThermoFisher Scientific). The quality of the extracted RNA assessed with an Agilent 2100 Bioanalyzer with the Nano RNA chip. The RNA was stored in RNase-free distilled water at −80°C.

RT-qPCR

Reverse transcriptase quantitative PCR (RT-qPCR) assays were developed and implemented for measurement of transcripts of genes for the following proteins of *P. leucopus*: nitric oxide synthase

2 (Nos2), arginase 1 (Arg1), secretory leukocyte peptidase inhibitor (Slpi), and glyceraldehyde 3-phosphate dehydrogenase (Gapdh). cDNA was synthesized from extracted RNA using an iScript reverse transcription kit and iScript Supermix (Bio-Rad) for qPCR, according to the manufacturer's instructions. The reaction mix was incubated in a thermal cycler for 5 min at 25°C, 20 min at 46°C, and 1 min at 95°C. PCR of cDNA was carried out with qPCR PowerUp SYBR green master mix (Applied Biosystems) and performed in 96-well plates in a StepOnePlus real-time PCR system (Applied Biosystems) instrument. The forward and reverse primers, synthesized by Integrated DNA Technologies (San Diego, CA), were, respectively, the following:

for Arg1, 5'-*TCCGCTGACAACCAACTCTG* and 5'-*GACAGGTGTGCCAGTAGATG*;

for Nos2, 5'-*GACTGGATTTGGCTGGTCCC* and 5'-*GAACACCACTTTCACCAAGAC*;

for Slpi, 5'-*TCCCATCAGCAGACCAGTG* and 5'-*TTGGGAGGATTCAGCATCATACA*;

and for Gapdh, 5'-*TCACCACCATGGAGAAGGC* and 5'-*GCTAAGCAGTTGGTGGTGCA*. The product sizes were 352, 192, 81, and 169 bp, respectively. For all assays, the initial step was 95°C for 10 min, followed by 40 cycles. The cycle conditions for Arg1 and Nos2 were 95°C for 15 s, 52°C for 30 s, and 72°C for 60 s. For Slpi and Gapdh, they were 95°C for 15 s and 52°C for 30 s. Standards were the corresponding PCR products cloned into the *E. coli* plasmid vector pUC57 and stored frozen at −80°C in single-use aliquots after plasmid purification, as described above.

RNA-seq

Library preparation with the Illumina TruSeq mRNA stranded kit was carried out as described previously (34). The libraries were normalized and then multiplexed to achieve 12 samples per flow cell on an Illumina HiSeq 4000 instrument and 100 or 150 cycles of paired-end read chemistry at the UC Irvine Genomic High Throughput Facility. The quality of sequencing reads was analyzed using FastQC (Babraham Bioinformatics). The reads were trimmed of low-quality reads (Phred score of <15) and adapter sequence, and corrected for poor-quality bases using Trimmomatic (214). After

these steps, there were overall for all the studies between 45 and 100 million reads for each of the samples. For the comparative study of 40 animals, the mean and median numbers, respectively, of reads (× 106) for the three tissues were 62.9 and 62.9 for blood, 65.2 and 63.8 for spleen, and 60.5 and 59.4 for liver. Blood and liver RNA-seq outputs were 100 paired-end reads (PE100), and spleen and fibroblast RNA-seq outputs were PE150. Paired-end reads in fastq files were quantified using kallisto v. 0.46.1 (318) with the GENCODE annotation (v. 21; https://www.gencodegenes.org) for mouse and GCF_004664715.1_Pero_0.1 _rna from the NCBI for *P. leucopus* (34). The mean (95% confidence interval [95% CI]) coefficient of determination ($R2$) between paired replicates of 4 RNA extracts (blood of 4 LPS-treated *P. leucopus* deermice), but with independent cDNA libraries and sequencing lanes, was 0.994 (0.990 to 0.997) for 8,055 transcripts, with mean numbers of transcripts per million (TPM) being >1, while the mean $R2$ for the 24 discordant pairs for these samples was 0.975 (0.973 to 0.977). Sequencing reads were deposited with the Sequence Read Archive Table 12.

Table 12 BioProject, BioSample, and Sequence Read Archive accession numbers for RNA-seq

Subject	BioProject accession no.	BioSample accession no.	Sequence read archive accession no.
Blood, spleens, and livers of 2- to 3-mo-old *Peromyscus leucopus* LL stock deermice after injection of lipopolysaccharide (LPS) or buffer alone (control)	PRJNA643534	SAMN15445639–SAMN15445698 (20 animals × 3 tissues = 60 samples)	SRR12781556–SRR12781615
Blood, spleens, and livers of 2- to 3-mo-old *Mus musculus* BALB/c mice after treatment with LPS or buffer alone	PRJNA643535	SAMN15445905–SAMN15445964 (20 animals × 3 tissues = 60 samples)	SRR12782328–SRR12782387
Blood and/or spleens of 1 to 2-yr-old *P. leucopus* LL stock deermice after injection of LPS or buffer alone	PRJNA644403	SAMN15469591–SAMN15469598; SAMN16439936–SAMN16439945	SRR15769470–SRR15769487
Blood of *P. leucopus* LL stock deermice infected with *Borrelia hermsii* strain MTW	PRJNA508222	SAMN10522571–SAMN10522573; SAMN10522575–SAMN10522578	SRR8283810–SRR8283816
Fibroblasts from the ears of 5 *P. leucopus* LL stock deermice cultivated with or without LPS	PRJNA672217	SAMN16563388–SAMN16563397	SRR13021354–SRR13021363

Differential expression

We used edgeR v. 3.28.1 for DEG analysis (319). Genes were called differentially expressed if their absolute fold change between conditions was >4.0 and the false-discovery rate (FDR) was <0.05. To compare fold changes across species, we merged the output tables from the DEG analyses and retained 14,685 orthologous genes that were synonymously annotated between both species, out of a total of 24,295 annotated genes for *P. leucopus* and 35,805 for *M. musculus*. To screen for DEGs that varied in magnitude by species, we required an absolute fold change of >5.0 (log2 = 2.5) in one species and <5.0 in the other. For DEGs designated "shared" between the species, the absolute fold change was >5.0 in both. Enrichment analysis was done for each one of the gene groups, separated by up- or downregulation. Enrichment of gene ontology (GO; http://geneontology.org) terms for biological processes was computed using EnrichR (https://amp.pharm.mssm.edu/Enrichr) (320)and plotted using ggplot (v. 3.3.2) of the R package (321). The GO terms were sorted by ascending *P* value. RNA-seq of a limited set of protein coding sequences (CDS) of both species in the same DEG analysis was carried out using CLC Genomics Workbench v. 20 (Qiagen). Paired-end reads were aligned with a length fraction of 0.4, similarity fraction of 0.9, and penalties of 3 for mismatch, insertion, or deletion to the CDS of sets of corresponding to orthologous mRNAs of *P. leucopus* and *M. musculus*. Accession numbers are given in Dryad repository, https://doi.org/10.7280/D1R70J. Counted reads were those that were species specific. For example, for Arg1, the coding sequences for both *P. leucopus* and *M. musculus* were in the reference set, but for *P. leucopus*, only the unique reads for Arg1 for that species were tabulated. The cross-hybridization of reads, e.g., number of *M. musculus* reads mapping to *P. leucopus* Arg1, was no more than 5% of that of the homologously mapped gene, e.g., *M. musculus* Arg1, and usually <1%. Expression values were unique reads normalized for total reads across all the samples without adjustment for reference sequence length. These were log10 transformed.

WGCNA

We used weighted gene correlation network analysis (WGCNA) to identify densely interconnected genes (modules) (322). This was applied for the 6 data sets across species and tissues by building a matrix of gene expression with genes with numbers of TPM of >1 in one or more individual. We selected a power (β) of 13 for a soft threshold for the weighted network and specified a minimum of 100 genes per module. To merge modules with similar gene expression profiles, we carried out a dynamic tree cut with an Eigengene dissimilarity threshold of 0.2 that generated the final Eigengene profiles, where Eigengene is the first principal component of the module expression matrix (323). The inferred modules were distinguished by assigned hexadecimal color code (https://www.color-hex.com/color-names.html), e.g., darkorange2. Modules, associated GO terms, and constituent genes that are not presented in this paper have been deposited with Dryad (http://datadryad.org) under the data set name "Peromyscus_WGCNA_supplement" (https://doi.org/10.7280/D1B38G).

Statistics

Means are presented with 95% confidence intervals. Parametric (*t* test) and nonparametric (Mann-Whitney) tests of significance were two-tailed. Unless otherwise stated, the *t* test *P* value is given. Adjustment of *P* values for multiple testing was by the Benjamini-Hochberg method (227), as implemented in edgeR, CLC Genomics Workbench (see above), or the False Discovery Rate Online Calculator (https://tools.carbocation.com/FDR). For categorical data, an exact likelihood ratio test was performed with StatXact v. 6 (Cytel statistical software). Other methods are given in software programs or suites cited above.

Data availability

The National Center for Biotechnology Information BioProject, BioSample, and Sequence Read Archive (SRA) identification and accession numbers for the nucleic acid sequences obtained in this study are listed in Table 12. The raw data for the metabolomics study is at

https://doi.org/10.21228/M8PH55 of the Metabolomics Workbench repository (https://www.metabolomicsworkbench.org) under project PR001060.

CHAPTER FOUR: The white-footed deermouse, an infection-tolerant reservoir for several zoonotic agents, tempers interferon responses to endotoxin in comparison to the mouse and rat

Abstract

The white-footed deermouse *Peromyscus leucopus*, a long-lived rodent, is a key reservoir for agents of several zoonoses, including Lyme disease. While persistently infected, this deermouse is without apparent disability or diminished fitness. For a model for inflammation elicited by various pathogens, the endotoxin lipopolysaccharide (LPS) was used to compare genome-wide transcription in blood by *P. leucopus, Mus musculus* and *Rattus norvegicus* and adjusted for white cell concentrations. Deermice were distinguished from the mice and rats by LPS response profiles consistent with non-classical monocytes and alternatively-activated macrophages. LPS-treated *P. leucopus,* in contrast to mice and rats, also displayed little transcription of interferon-gamma and lower magnitude fold-changes in type 1 interferon-stimulated genes. This was associated with comparatively reduced transcription of endogenous retrovirus sequences and cytoplasmic pattern recognition receptors in the deermice. The results reveal a mechanism for infection tolerance in this species and perhaps other animal reservoirs for agents of human disease.

Introduction

How does the white-footed deermouse *Peromyscus leucopus* continue to thrive while sustaining infections with disease agents it serves as reservoir for (3)? The diverse tickborne pathogens (and diseases) for humans include the extracellular bacterium *Borreliella burgdorferi* (Lyme disease), the intracellular bacterium *Anaplasma phagocytophilum* (anaplasmosis), the protozoan *Babesia microti* (babesiosis), and the Powassan flavivirus (viral encephalitis). Most deermice remain persistently infected but display scant inflammation in affected tissues (20), (5), (233), and without apparent consequence for fitness (21), (25).

A related question—conceivably with the same answer--is what accounts for the two-to-three fold longer life span for *P. leucopus* than for the house mouse, *Mus musculus* (324), (10)? The abundance of *P. leucopus* across much of North America (325), (326) and its adaptation to a variety of environments, including urban areas and toxic waste sites (327), (328), (329), indicates successful adjustment to changing landscapes and climate. *Peromyscus* species, including the hantavirus reservoir *P. maniculatus*, are more closely related to hamsters and voles in family Cricetidae than to mice and rats of family Muridae (330).

As a species native to North America, *P. leucopus* is an advantageous alternative to the Eurasian-origin house mouse for study of natural variation in populations that are readily accessible (7), (5). A disadvantage for the study of any *Peromyscus* species is the limited reagents and genetic tools of the sorts that are applied for mouse studies. As an alternative, we study *P. leucopus* with a non-reductionist approach that is comparative in design and agnostic in assumptions as presented in Chapter 3. The genome-wide expression comparison for *P. leucopus* is with *M. musculus* and, added here, the brown rat *Rattus norvegicus*. Given the wide range of pathogenic microbes that deermice tolerate, we use the bacterial endotoxin lipopolysaccharide (LPS) as the primary experimental treatment because the inflammation it elicits within a few hours has features common to different kinds of serious infections, not to mention severe burns and critical injuries (331).

We previously reported in Chapter 3 that a few hours after injection of LPS, *P. leucopus* and *M. musculus* had distinguishing profiles of differentially expressed genes (DEG) in the blood, spleen, and liver. In brief, the inflammation phenotype of deermice was consistent with an "alternatively activated" or M2-type macrophage polarization phenotype instead of the expected "classically activated" or M1-type polarization phenotype that was observed for *M. musculus* (264). The deermice also differed from mice in displaying evidence of greater neutrophil activation and degranulation after LPS exposure. The potentially damaging action from neutrophil proteases and

reactive oxygen species appeared to be mitigated in part in *P. leucopus* by proteins like secretory leukocyte protease inhibitor (Slpi) and superoxide dismutase 2 (Sod2).

Here, we first address whether the heightened transcription of neutrophil-associated genes in *P. leucopus* is attributable to differences in numbers of white cells in the blood. To better match for genetic diversity, we substituted outbred *M. musculus* for the inbred BALB/c mouse of the previous study. We retained the experimental protocol of short-term responses to LPS. This main experiment was supplemented by a study of rats under the similar conditions, by an investigation of a different dose of LPS and duration of exposure in another group of deermice, and by analysis of deermice infected with a bacterium lacking LPS. The focus was on the blood of these animals, not only because the distinctions between species in their transcriptional profiles were nearly as numerous for this specimen as for spleen and liver, as described in Chapter 3, but also because for ecological and immunological studies of natural populations of *Peromyscus* species blood is obtainable from captured-released animals without their sacrifice.

The results inform future studies of *Peromyscus* species, not only with respect to microbial infections and innate immunity, but conceivably also determinants of longevity and resilience in the face of other stressors, such as toxic substances in the environment. The findings pertain as well to the phenomenon of infection tolerance broadly documented in other reservoirs for human disease agents, such as betacoronaviruses and bats (242). Less directly, the results provide for insights about maladaptive responses among humans to microbes, from systemic inflammatory response syndrome (SIRS) to post-infection fatigue syndromes.

Results

LPS experiment and hematology studies

Twenty adult animals each for *P. leucopus* and *M. musculus* and equally divided between sexes received by intraperitoneal injection either purified E. coli LPS at a dose of 10 µg per g body mass or

saline alone (Table 13). Within 2 h LPS-treated animals of both species displayed piloerection and sickness behavior, i.e. reduced activity, hunched posture, and huddling. By the experiment's termination at 4 h, 8 of 10 *M. musculus* treated with LPS had tachypnea, while only one of ten LPS-treated *P. leucopus* displayed this sign of the sepsis state (p = 0.005).

Table 13 Characteristics and treatments of M. musculus CD-1 and P. leucopus LL stock.

Animal	Genus	Sex	Age (d)	Mass (g)	Treatment	Tachypnea	Hct * (%)	MCV *	WBC *	Neutrophils	Lymphocytea	Monocytes	Eosinophils	Neutrophil/ Lymphocyte
MM19	*Mus*	female	149	60.4	control	no	71	62	3800	570	2926	190	114	0.19
MM21	*Mus*	female	149	51.4	control	no	51	59	8600	826	4897	177	0	0.17
MM23	*Mus*	female	149	30.6	control	no	51	.	1500	135	1305	60	0	0.10
MM25	*Mus*	female	149	42.1	control	no	65	61	7900	1106	6557	237	0	0.17
MM27	*Mus*	female	149	39.7	control	no	58	62	3700	962	2257	333	148	0.43
MM1	*Mus*	male	149	50.9	control	no	67	59	3300	726	2376	198	0	0.31
MM17	*Mus*	male	149	40.8	control	no	51	.	7000	4970	2030	0	0	2.45
MM3	*Mus*	male	149	45.8	control	no	58	59	4300	344	3655	301	0	0.09
MM5	*Mus*	male	149	41.2	control	no	60	57	5400	810	4266	324	0	0.19
MM7	*Mus*	male	149	44.1	control	no	47	58	3800	798	2736	266	0	0.29
MM31	*Mus*	female	149	65.5	LPS	yes	53	58	1900	209	1539	114	38	0.14
MM33	*Mus*	female	149	48.4	LPS	yes	62	62	3200	256	2784	96	64	0.09
MM35	*Mus*	female	149	40.5	LPS	yes	49	64	2200	528	1518	88	66	0.35
MM37	*Mus*	female	149	38.5	LPS	no	54	65	1900	152	1634	57	57	0.09
MM39	*Mus*	female	149	40.2	LPS	yes	61	63	600	60	510	12	18	0.12
MM11	*Mus*	male	149	57.4	LPS	yes	59	58	2600	572	1898	104	0	0.30
MM13	*Mus*	male	149	58.5	LPS	yes	66	59	3700	1258	2183	259	0	0.58
MM15	*Mus*	male	149	51.0	LPS	yes	33	56	1800	90	1602	90	0	0.06
MM29	*Mus*	male	149	39.1	LPS	no	53	59	1800	306	1368	72	54	0.22
MM9	*Mus*	male	149	44.1	LPS	yes	58	.	1300	260	858	182	0	0.30
24841	*Peromyscus*	female	162	19.9	control	no	58	48	5100	204	4590	102	204	0.04
24842	*Peromyscus*	female	164	18.3	control	no	47	48	3900	663	3003	117	117	0.22
24843	*Peromyscus*	female	162	20.5	control	no	45	46	4200	942	3150	42	84	0.30
24845	*Peromyscus*	female	161	18.7	control	no	49	48	9100	2002	6916	182	0	0.29
24853	*Peromyscus*	female	160	22.8	control	no	42	.	4800	1008	2880	864	1	0.35
24852	*Peromyscus*	male	162	19.4	control	no	44	46	7100	994	4970	284	852	0.20
24861	*Peromyscus*	male	157	20.8	control	no	28	50	1300	104	1053	91	52	0.10
24863	*Peromyscus*	male	157	17.1	control	no	26	58	7200	720	5904	288	288	0.12
24869	*Peromyscus*	male	143	29.0	control	no	48	52	6000	1260	4260	180	240	0.30
24876	*Peromyscus*	male	142	16.1	control	no	54	48	9700	2716	6208	194	485	0.44
24846	*Peromyscus*	female	162	22.7	LPS	no	23	47	1100	231	718	44	44	0.32
24847	*Peromyscus*	female	162	16.7	LPS	no	39	.	1500	570	675	180	75	0.84
24848	*Peromyscus*	female	166	16.2	LPS	no	43	49	2700	918	1701	27	54	0.54
24850	*Peromyscus*	female	157	19.4	LPS	no	46	.	3300	1551	1683	66	0	0.92
24851	*Peromyscus*	female	161	25.4	LPS	no	24	47	2300	552	1242	437	69	0.44
24855	*Peromyscus*	male	165	27.1	LPS	no	51	51	2100	1281	714	42	63	1.79
24860	*Peromyscus*	male	160	17.8	LPS	yes	42	54	13200	6996	3696	1320	1056	1.89
24865	*Peromyscus*	male	160	16.7	LPS	no	49	46	1800	396	1080	288	0	0.37
24873	*Peromyscus*	male	145	23.2	LPS	no	43	48	2200	550	1430	110	110	0.38
24879	*Peromyscus*	male	145	22.0	LPS	no	40	.	1100	220	538	352	0	0.41

* Abbreviations: Hct, hematocrit; MCV, mean cellular volume of erythrocytes; WBC, white blood cell count

Within a given species there was little difference between LPS-treated and control animals in values for erythrocytes. But overall the deermice had lower mean (95% confidence interval) hematocrit at 42 (36-48) %, hemoglobin concentration at 13.8 g/dL (12.1-15.5), and mean corpuscular volume for erythrocytes at 49 fL (47-51) than *M. musculus* with respective values of 56 (51-62)%, 16.1 g/dL (14.6-

17.7), and 60 fL (58-62) (*p* < 0.01). These hematology values for adult CD-1 *M. musculus* and LL stock *P. leucopus* in this study were close to what had been reported for these colony populations (287), (286).

In contrast to red blood cells, the mean numbers of white blood cells in the LPS groups in both species were lower than those of control groups (figure 34). Controls had a mean 4.9 (3.5-6.4) x 103 white cells per μl among *M. musculus* and 5.8 (4.2-7.4) x 103 white cells per μl among *P. leucopus* (p = 0.41). For the LPS-treated animals the values were 2.1 (1.5-2.7) x 103 for mice and 3.1 (0.9-5.4) x 103 for deermice (*p* = 0.39). However, there was difference between species among LPS-treated animals in the proportions of neutrophils and lymphocytes in the white cell population. The ratios of neutrophils to lymphocytes were 0.25 (0.14-0.45) and 0.20 (0.13-0.31) for control *M. musculus* and *P. leucopus,* respectively (*p* = 0.53). But under the LPS condition. the neutrophil-to-lymphocyte ratio was 0.18 (0.11-0.28) for mice and 0.64 (0.42-0.97) for deermice (*p* = 0.0006). The regression curves for plots of neutrophils and lymphocytes for LPS-treated and control *P. leucopus* and LPS-treated *M. musculus* had similar slopes, but the *y*-intercept was shifted upwards towards a higher ratio of neutrophils to lymphocytes for blood from the LPS group of deermice. Control group mice and deermice and LPS-treated mice had similar percentages (~5%) of monocytes in their blood; the mean monocyte percentage rose to 10% in LPS treated deermice (*p* = 0.12). Eosinophil percentages tended to be higher in deermice at a mean 3.4 (2.1-4.7) % than mice at 1.2 (0.5-1.9)% under either condition (*p* = 0.004).

In the *P. leucopus* experiment with a 10-fold lower dose of LPS and a 12 hour duration, the mean (95% confidence interval) white blood cell count (x 103) at termination 3.5 (2.5-4.5) in controls and 7.9 (6.0-9.7) in the LPS-treated (p = 0.01). Even with the higher overall white blood cell count the increase white cells was proportionately higher for neutrophils than for lymphocytes, as was seen in the

deermice in the higher dose LPS experiment. The ratio of neutrophils-to-lymphocytes was 0.20 (0.07-0.32) in the controls and 0.38 (0.26-0.50) in the LPS-treated (p = 0.10).
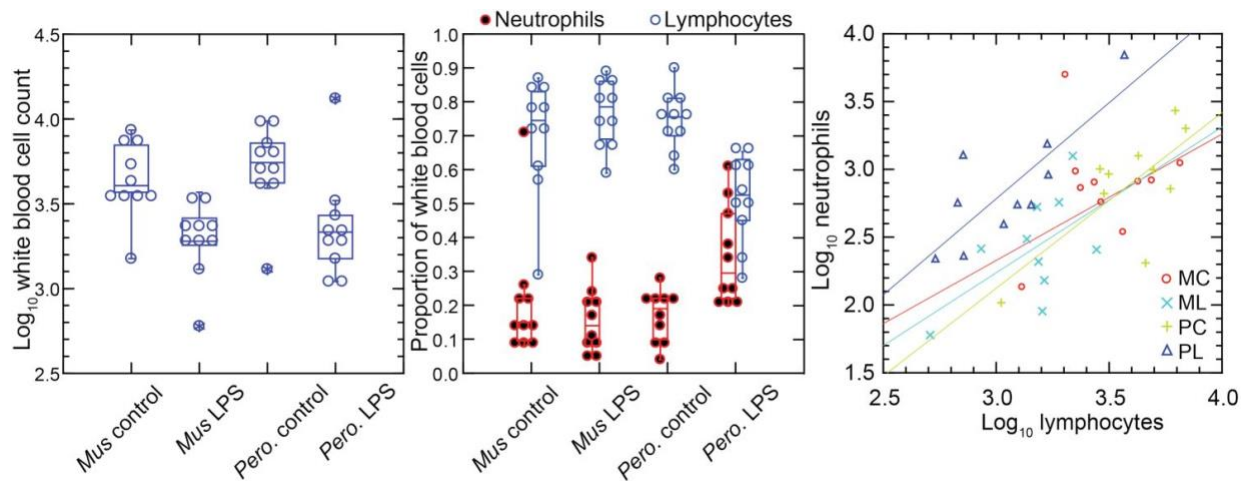
Figure 34 *Total white blood cells, neutrophils, and lymphocytes of M. musculus and P. leucopus with or without treatment with LPS*
Total white blood cells, neutrophils, and lymphocytes of Mus musculus (M) and Peromyscus leucopus (P) with or without (control; C) treatment with 10 µg lipopolysaccharide (LPS; L) per g body mass 4 h previous. The box plots of left and center panels show values of individual animals and compiled median, quartiles, and extreme values. The linear regressions of the right panel are color-coded according to the species and treatment designations. The outlier value for a M. musculus control (MM17) was excluded from the linear regression for that group.

The higher neutrophil to lymphocyte ratio in the deermice exposed to LPS was consistent with the greater neutrophil activation noted by transcriptional analysis as described in Chapter 3. But many individual genes that constitute this and related gene ontology (GO) terms had transcription levels in the deermice that far exceeded a three-fold difference in neutrophil counts. For some genes the differences were a hundred or more-fold, which suggested that the distinctive LPS transcriptional response profile between species was not attributable solely to neutrophil counts.

Genome-wide expression in blood of deermice and mice

We used the respective transcript sets from the reference genomes for *P. leucopus* and *M. musculus* for deep coverage RNA-seq with paired-end ~150 nt reads. Principle component analyses (PCA) of the P. leucopus data and M. musculus data revealed that untreated controls had coherent profiles

within each species (figure 35). With the exception of one mouse, the LPS-treated *M. musculus* were

also in a tight PCA cluster. In contrast, the LPS-treated deermice displayed a diversity of genome-

wide transcription profiles and limited clustering.
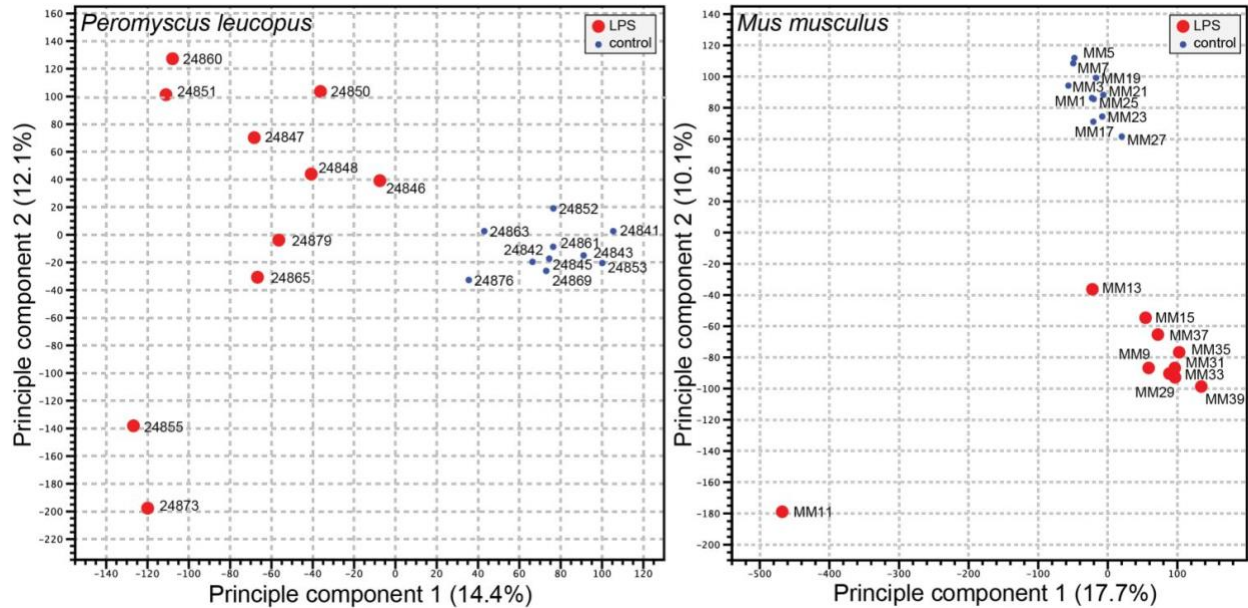
Figure 35 Principal component analysis of genome-wide RNA-seq data of P. leucopus or M. musculus

Principal component analysis of genome-wide RNA-seq data of *P. leucopus* or *M. musculus* with or without (blue dot) treatment with LPS 4 h previous. The individual animals listed in Table 13 are indicated on the graphs. The insets indicate the size and color of the symbol for the experimental condition (LPS-treated or control).

For both species the number of genes with higher expression with LPS exposure exceeded those with

lower or unchanged expression. For *P. leucopus* and *M. musculus* the mean fold-changes were 1.32

(1.29-1.35) and 1.30 (1.24-1.36), respectively (p = 0.31). For GO term analysis the absolute fold-

change criterion was ≥ 2. Because of the ~3-fold greater number of transcripts for the *M. musculus*

reference set than the *P. leucopus* reference set, application of the same false-discovery rate (FDR)

threshold for both datasets would favor the labeling of transcripts as DEGs in *P. leucopus*.

Accordingly, the FDR *p* values were arbitrarily set at <5 x 10-5 for *P. leucopus* and <3 x 10-3 for *M.*

musculus to provide approximately the same number of DEGs for *P. leucopus* (1154 DEGs) and *M.*

musculus (1266 DEGs) for the GO term comparison.

Figure 36 shows the GO terms for the top 20 clusters by ascending *p*-value for up-regulated and down-regulated in *P. leucopus* and the corresponding categories for *M. musculus*. The up-regulated gene profile for *P. leucopus* featured terms associated with "neutrophil degranulation", "myeloid leukocyte activation", "leukocyte migration", and "response to molecule of bacterial origin". Other sets of up-regulated genes for the deermice were "negative regulation of cytokine production" and "regulation of reactive oxygen species metabolic process". None of these were among the top 20 up-regulated clusters for *M. musculus*. Indeed, "leukocyte activation" and "leukocyte migration" were GO terms for down-regulated DEGs in *M. musculus*. Distinctive GO terms for up-regulated genes distinguishing mice from deermice were "response to virus", "response to interferon-beta", "response to interferon-gamma", "response to protozoan", and "type II interferon signaling".
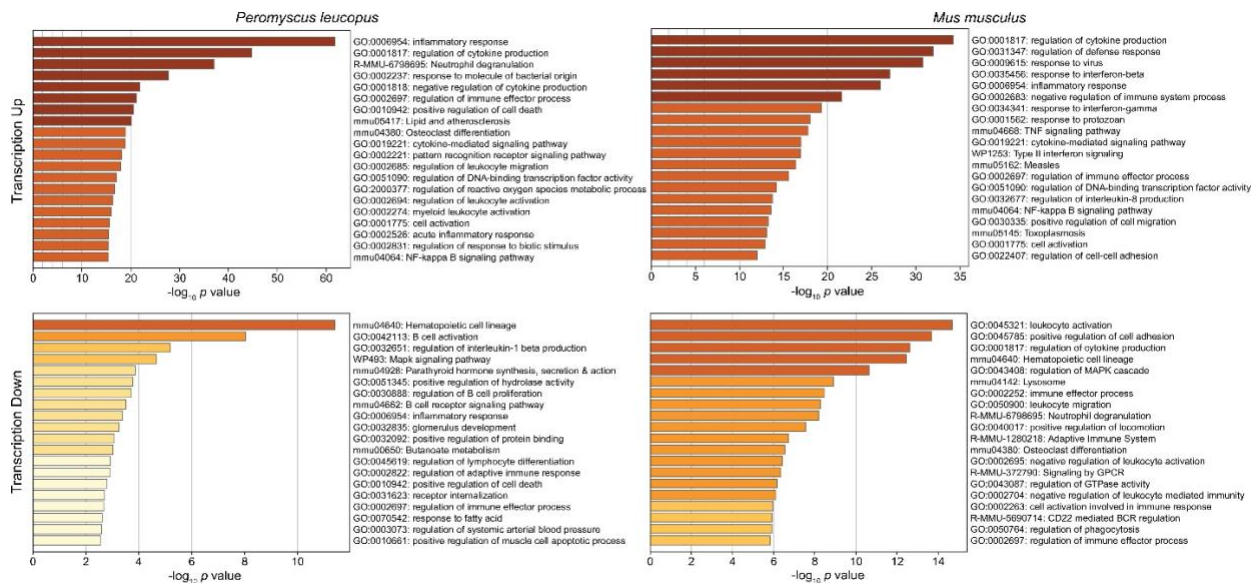
Figure 36 Gene Ontology term clusters of P. leucopus and M. musculus

Gene Ontology (GO) term clusters associated with up-regulated genes (upper panels) and down-regulated genes (lower panels) of Peromyscus leucopus (left panels) and Mus musculus (right panels) treated with LPS in comparison with untreated controls of each species The scale for the x-axes for the panels was determined by the highest -log10 p values in each of the four sets. The horizontal bar color, which ranges from white to dark brown through shades of yellow through orange in between, is a schematic representation of the -log10 p values.

By arbitrary criterion of 100 for the top DEGs by ascending *p* value for each species, 24 genes were shared between species. These included up-regulated Bcl3, Ccl3, Cxcl1, Cxcl2, Cxcl3, Cxcl10, Il1rn, and Sod2. Among the 100 mouse DEGs, 20 were constituents of GO terms "response to virus" or "response to interferon-beta" and only 3 were members of GO term sets "response to molecule of bacterial origin" or "response to lipopolysaccharide". In contrast, among the top 100 deermouse DEGs, there were only 2 associated with the virus or type 1 interferon GO terms, but 12 were associated with either or both of the bacterial molecule GO terms.

We confirmed the sex identification for each sample with sex-specific transcripts of Xist for females and Ddx3y for males, data presented in Chapter 3. For female and male *P. leucopus* there were 5012 transcripts out of 54,466 in the reference set for which there were TPM values of ≥10 in at least one animal in each of the sexes under either condition. The comparison of females to males by fold-changes between LPS-treated and control animals revealed transcripts that were differentially expressed between sexes under LPS treatment (figure 37). Some were down-regulated in one sex while unchanged in expression in the opposite sex. Of note in this category were different isoforms or variants of Lilra6 (leukocyte immunoglobulin-like receptor, subfamily A, member 6), one of a family of orphan receptors of myeloid cells (332). The opposite case was exemplified by the Dnajc15 and Hspa8 genes for two chaperones: DnaJ heat shock protein family (Hsp40) member C15 and heat shock protein 8, respectively. These were substantially lower in transcription in the LPS-treated females than in untreated animals, but little changed in LPS-treated males. Coordinates for some other genes, e.g. Saa5 and Cxcl2, fell outside the prediction limits at the extreme end of up-regulation, but their vectors were within 20-25° of each other. While these and other sex-associated differences merit attention for future studies, overall, they were not of sufficient number or magnitude in our view to warrant division by sex for the subsequent analyses, which had the aim of identifying differences applicable for both females and males.
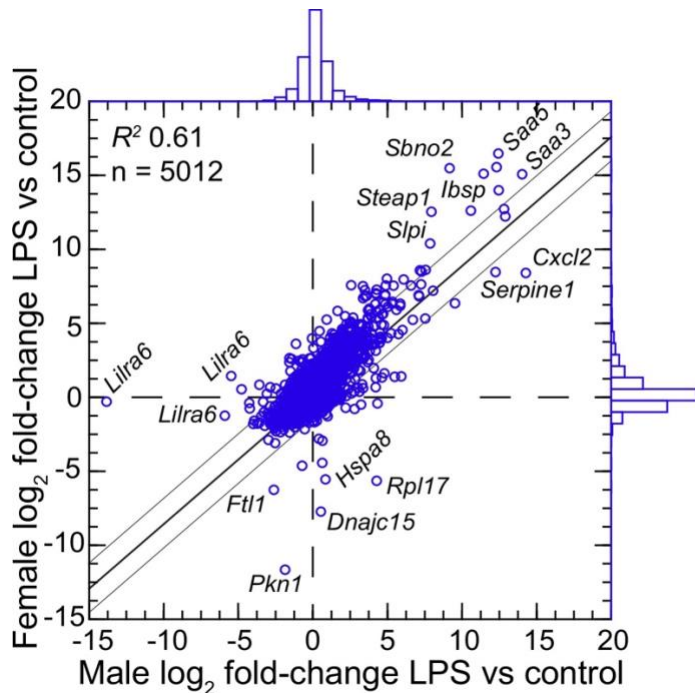
Figure 37 Scatter plot with linear regression of pairs of log2-transformed mean fold-changes between LPS-treated and control P. leucopus by male and female sex

Scatter plot with linear regression of pairs of log2-transformed mean fold-changes between LPS-treated and control *P. leucopus* by male and female sex. The 5012 reference transcripts from the genome reference set are defined in the text. The coefficient of determination (R^2), the 95% upper and lower prediction limits for the regression line, and distributions of the values on the *x*-and *y*-axes are shown. Selected genes for which their *x-y* coordinates fall outside the limits of prediction are labeled. Cxcl2, Ibsp, Saa3, Saa5, Sbno2, Serpine1, Slpi, and Steap1 were noted as up-regulated DEGs for the groups with both sexes.

Targeted RNA seq analysis

The emerging picture was of *P. leucopus* generally responding to LPS exposure as if infected with an extracellular bacterial pathogen, including with activated neutrophils. While *M. musculus* animals of both sexes shared with *P. leucopus* some features of an antibacterial response, they also displayed type 1 and type 2 interferon type response profiles associated with infections with viruses and intracellular bacteria and parasites.

Going forward, the challenge for a cross-species RNA-seq was commensurability between annotated transcripts of reference sets. Orthologous genes can be identified, but mRNA isoforms and their 5' and 3' untranslated regions may not fully correspond. Accordingly, we limited targeted

RNA-seq to protein coding sequences of mRNAs for the corresponding sets of *P. leucopus* and *M. musculus* sequences.

The 113 mRNA coding sequences, which are listed in Methods, were drawn from the identified DEGs for *P. leucopus* and *M. musculus* from the genome-wide RNA-seq. For cross-species normalization we first evaluated three methods: (1) normalization using the ratio of mean total reads for all samples to total reads for a given sample, (2) the ratio of reads mapping to a given target gene (i.e. the numerator) to the reads to the mitochondrial 12S rRNA gene (i.e. the denominator), or (3) when the denominator instead was the myeloid cell marker CD45 (protein tyrosine phosphatase, receptor type C) encoded by the Ptprc gene (figure 38). In humans, mice, and hamsters, Ptprc is expressed by nucleated hematopoietic cells, and CD45 is commonly used as a white cell marker for flow cytometry (333). The coefficients of determination (R^2) between comparison pairs (e.g. normalized total reads vs. Ptprc ratio) within a species were ≥ 0.95. There was also little difference between the choice of Ptprc or 12S rRNA transcripts as denominator with respect to cross-species comparisons of LPS-treated to control fold changes. Given the widespread adoption of CD45 for flow cytometry, we chose Ptprc as denominator and as an adjustment for white cell numbers in the samples. Pearson correlation between log-transformed total white blood cell counts and normalized reads for Ptprc across 40 animals representing both species, sexes, and treatments was 0.40 ($p = 0.01$).

The normalization options were total reads for the same sample, unique reads for the mitochondrial 12S rRNA, and unique reads for the Ptprc (CD45) transcript are presented in figure 38. The variable for comparison was, first, the mean fold-change between LPS-treated and control animals of each sex and by each of the three methods. The coefficients of determination (R^2) were calculated for each of the pairs and for within each species and across species. The results of this analysis are in the matrix of the upper panel of figure 38.

Genus Normalization	Mus reads R^2	Mus 12S rDNA R^2	Mus Ptprc R^2	Peromyscus reads R^2	Peromyscus 12S R^2	Peromyscus Ptprc R^2
Mus reads	1.000					
Mus 12S rDNA	0.994	1.000				
Mus Ptprc	0.969	0.945	1.000			
Peromyscus reads	0.394	0.371	0.426	1.000		
Peromyscus 12S	0.416	0.396	0.446	0.986	1.000	
Peromyscus Ptprc	0.424	0.401	0.457	0.996	0.990	1.000
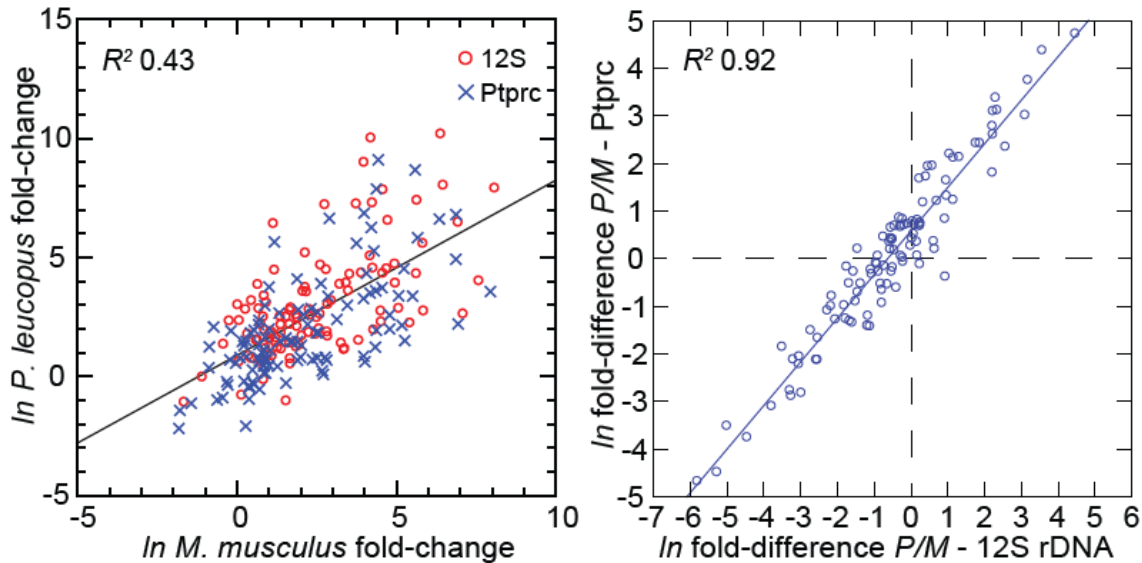
Figure 38 Comparison of three different methods for normalization for cross-species targeted RNA-seq

Within each species the outputs from the 3 different methods were highly similar. Since the aim was to identify a normalization procedure suitable for use on blood sample with differing concentrations of nucleated cells and different states of activation, we next compared normalization by 12S rRNA with normalization by Ptprc. The lower left panel compares in the same scatterplot the LPS to control fold-changes by each method and the *P. leucopus* (P) result regressed on the *M. musculus* (M) for the same gene. As expected, the two species differed overall in many of their responses, and this was more pronounced for genes with the higher magnitudes of fold-change. But areas of distributions of datapoints over the space for each of the methods overlapped. The next question was whether there would a difference in outcome for a cross-species comparison. For this we used the metric of the log-transformed ratio of the LPS:control fold- change for *P. leucopus* to the LPS:control fold-change for *M. musculus*. To avoid confusion with the "fold-change" variable, we termed this cross-species

variable "fold-difference". If the two methods yielded commensurate results, they would be highly correlated. The lower right panel of the figure is a scatterplot with linear regression and the R2 value. The slope (beta) was 0.92. The results indicated that the two normalization methods, one based on a mitochondrion gene and other on a chromosome gene, were commensurate.

Figure 39 comprises plots of the log-transformed mean ratios for the 10 *P. leucopus* controls and 10 *M. musculus* controls and for the 10 *P. leucopus* and 10 *M. musculus* treated with LPS. For untreated animals (left panel) there was high correlation and a regression coefficient of ~1 between the paired data for deermice and mice. The mitochondrial cytochrome oxidase 1 gene (MT-Co1) and S100a9, a subunit of calprotectin, were comparably transcribed. But, there were other coding sequences that stood out for either their greater or lesser transcription in untreated deermice than mice. Two examples of greater expression were Arg1 and MX dynaminin-like GTPase 2 (Mx2), an ISG, while two examples of lesser expression were matrix metalloprotease 8 (Mmp8) and Slpi. There was low to undetectable transcription of Nos2 and interferon-gamma (Ifng) in the blood of controls of both species.
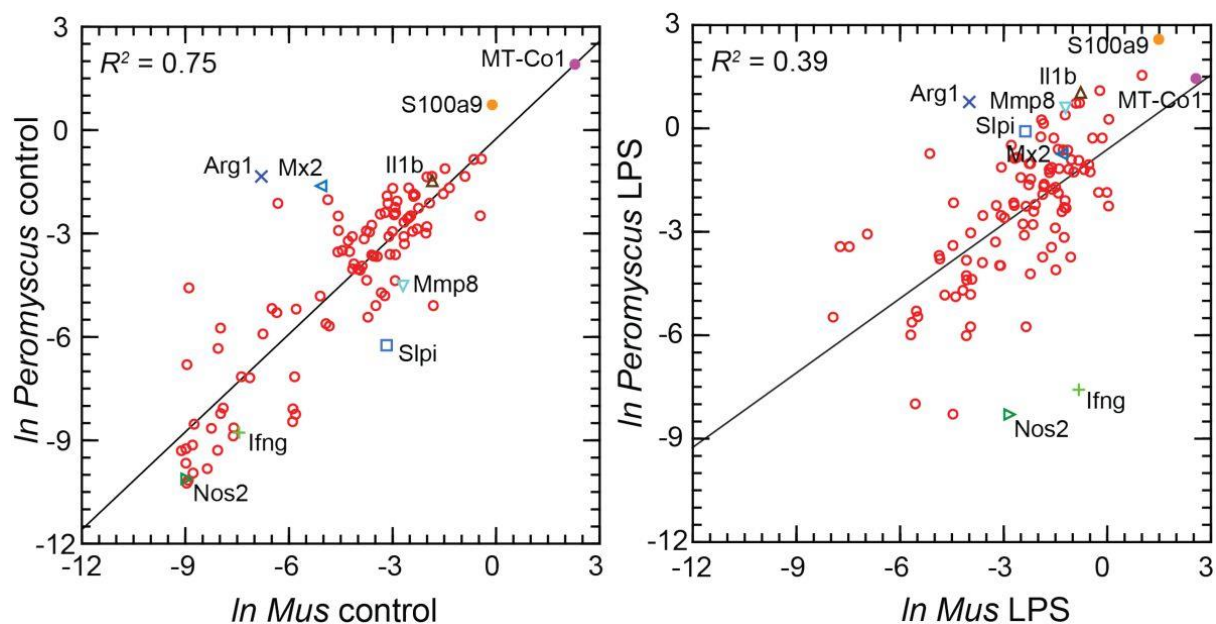
Figure 39 Scatter plots with linear regression of pairs of log-transformed P. leucopus and M. musculus sequences

Scatter plots with linear regression of pairs of log-transformed (*ln*) normalized RNA-seq reads for selected coding sequences for control *P. leucopus* and *M. musculus* (left panel) and LPS-treated *P. leucopus* and *M. musculus* (right panel). The R^2 values and selected genes (each with a different symbol) are indicated in each graph.

For the LPS-treated animals (right panel figure 39) there was, as expected for this selected set, higher expression of the majority genes and greater heterogeneity among *P. leucopus* and *M. musculus* animals in their responses for represented genes. In contrast to the findings with controls, Ifng and Nos2 had higher transcription in treated mice. In deermice the magnitude of difference in the transcription between controls and LPS-treated was less. A comparatively restrained transcriptional response in deermice was also noted for Mx2. On the other hand, there were greater fold-change from baseline in *P. leucopus* than in *M. musculus* for interleukin-1 beta (Il1b), Mmp8, Slpi, and S100a9.

Table 14 lists all the selected targets with the means and confidence intervals for the normalized values for controls and LPS-treated *M. musculus* and controls and LPS-treated *P. leucopus*. The fold-changes within each species and between treatments across species are given. The final column is the ratio of the fold-change between LPS to control in *P. leucopus* to the corresponding value for *M. musculus*. This along with the derived heat-map of these ratios, presented in the second column, indicates the genes for which there was little difference between species in their responses to LPS--either up or down--as well as those that were comparatively greater or lesser in one species or the other. Several of these genes are considered in other specific contexts below. Of note here are the places of Nos2 and Ifng at the bottom of the table, and Il1b near the top at position 20.

Table 14 Targeted RNA-seq of blood of Peromyscus leucopus and Mus musculus 4 hours after intraperitoneal injection of lipopolysaccharide (LPS) or saline control

Gene *	Heatmap †	Mus control mean (95% CI) ‡	Mus LPS mean (95% CI)	Mus FC LPS vs control**	Mus FDR p value**	Peromyscus control mean (95% CI)	Peromyscus LPS mean (95% CI)	Peromyscus FC LPS v control	Peromyscus FDR p value	Peromyscus v Mus control FC	Peromyscus v Mus LPS FC	Peromyscus v Mus LPS/control †
Ibsp		0.51 (0.28-0.93)	5.99 (2.04-17.6)	11.7	2E-03	0.17 (0.06-0.47)	471 (305-728)	2720	3E-10	0.34	78.6	233
Slpi		42.0 (32.6-54.0)	93.2 (44.5-195)	2.22	7E-02	1.94 (0.96-3.94)	921 (636-1333)	474	1E-10	0.05	9.88	213
Steap1		0.13 (0.10-0.16)	0.58 (0.21-1.60)	4.52	1E-02	0.09 (0.05-0.17)	31.7 (16.3-61.4)	335	1E-09	0.73	54.3	74.2
Saa3		2.82 (0.89-8.95)	458 (165-1267)	162	7E-04	0.21 (0.12-0.37)	2052 (1273-3306)	9934	2E-13	0.07	4.48	61.2
Lpo		0.13 (0.09-0.20)	0.37 (0.19-0.72)	2.78	2E-02	0.03 (0.03-0.04)	4.10 (1.44-11.7)	118	1E-07	0.26	11.2	42.4
Mmp8		67.4 (34.3-133)	298 (187-477)	4.43	3E-03	10.9 (4.26-27.8)	1823 (1311-2534)	168	3E-08	0.16	6.11	37.9
Lrg		40.8 (22.1-75.3)	107 (52.2-219)	2.62	7E-02	8.01 (2.98-21.5)	353 (268-464)	44.1	2E-06	0.20	3.30	16.8
Mmp9		222 (147-335)	153 (90.1-258)	0.69	3E-01	154 (96.8-245)	1258 (903-1753)	8.17	2E-06	0.69	8.25	11.9
Csf3		0.14 (0.10-0.19)	7.85 (3.07-20.1)	57.1	6E-07	0.04 (0.03-0.05)	24.5 (8.95-67.1)	648	2E-09	0.27	3.12	11.4
Serpine1		0.51 (0.27-0.93)	11.7 (3.04-45.2)	23.2	9E-04	0.14 (0.06-0.32)	32.8 (4.82-223)	239	9E-05	0.27	2.80	10.3
Fpr2		36.9 (25.8-52.7)	246 (168-362)	6.68	2E-06	8.74 (6.48-11.8)	535 (399-717)	61.2	5E-12	0.24	2.17	9.16
Cxcl3		0.32 (0.17-0.59)	19.6 (12.3-31.3)	60.8	2E-08	0.09 (0.04-0.21)	47.2 (15.9-140)	523	2E-07	0.28	2.41	8.59
Retn		0.35 (0.18-0.68)	0.45 (0.26-0.77)	1.28	6E-01	3.13 (1.80-5.46)	31.8 (22.0-45.9)	10.1	4E-06	8.96	71.1	7.93
Csf3r/Cd114		669 (455-982)	302 (221-411)	0.45	7E-03	423 (318-563)	1445 (1024-2040)	3.42	6E-05	0.63	4.79	7.58
Fcgr3/Cd16		45.7 (31.7-65.8)	70.1 (35.1-140)	1.53	3E-01	44.7 (34.1-58.6)	409 (338-497)	9.17	1E-09	0.98	5.84	5.97
Ptx		0.24 (0.17-0.33)	4.06 (0.94-17.6)	17.0	2E-03	0.05 (0.03-0.09)	4.88 (1.10-21.6)	92.2	4E-05	0.22	1.20	5.42
Fcgr2a		78.8 (59.1-105)	149 (93.8-237)	1.89	4E-02	75.0 (62.0-90.8)	766 (597-983)	10.2	2E-10	0.95	5.14	5.40
Cd177		25.0 (11.1-56.1)	158 (80.2-310)	6.31	4E-03	4.29 (2.12-8.69)	144 (107-194)	33.5	1E-07	0.17	0.91	5.31
Ncf4		134 (112-160)	204 (172-242)	1.52	5E-03	49.1 (38.4-62.9)	317 (258-389)	6.46	5E-09	0.37	1.55	4.24
Il1b		158 (102-244)	464 (341-631)	2.94	1E-03	230 (146-362)	2845 (1998-4052)	12.4	2E-07	1.45	6.14	4.22
Fcgr2b		32.7 (27.4-39.0)	111 (78.2-156)	3.38	1E-05	24.9 (22.4-27.8)	352 (302-409)	14.1	2E-14	0.76	3.18	4.17
Hk3		35.8 (25.7-50.0)	63.6 (46.6-86.8)	1.78	3E-02	85.5 (66.8-109)	602 (468-773)	7.04	1E-08	2.39	9.46	3.96
Fgr		138 (110-174)	161 (125-207)	1.17	4E-01	251 (205-308)	1131 (987-1295)	4.50	3E-09	1.82	7.02	3.85
Tlr4		11.9 (8.62-16.4)	11.9 (6.23-22.8)	1.00	1E+00	29.7 (22.9-38.6)	113 (90.7-141)	3.81	1E-06	2.50	9.49	3.80
Glrx		54.8 (43.1-69.7)	122 (65.4-227)	2.22	4E-02	12.4 (11.1-13.9)	88.4 (71.6-109)	7.11	5E-11	0.23	0.73	3.20
Slc11a1/Nramp		54.9 (41.6-72.5)	75.3 (60.0-94.6)	1.37	1E-01	104 (88.0-123)	424 (308-582)	4.07	1E-06	1.89	5.62	2.97
Ccl3		3.07 (2.20-4.30)	227 (161-320)	73.9	2E-11	0.26 (0.14-0.48)	54.6 (33.2-89.8)	213	1E-09	0.08	0.24	2.88
Tlr1		71.8 (63.3-81.5)	52.2 (37.4-72.8)	0.73	1E-01	36.0 (29.8-43.4)	73.3 (61.6-87.1)	2.04	5E-05	0.50	1.40	2.81
Hmox1		16.0 (12.5-20.6)	27.8 (13.7-56.7)	1.74	2E-01	17.4 (14.9-20.4)	78.2 (51.5-119)	4.49	6E-06	1.09	2.81	2.58
Tnfrsf1a		70.8 (57.2-87.7)	69.3 (45.5-106)	0.98	9E-01	44.6 (34.1-58.4)	105 (89.0-123)	2.34	6E-05	0.63	1.51	2.40
Olfm4		7.41 (3.43-16.0)	162 (98.8-266)	21.9	6E-06	3.56 (1.96-6.44)	181 (101-326)	51.0	1E-07	0.48	1.12	2.33
Sod2		58.1 (48.3-69.9)	412 (262-650)	7.10	8E-07	124 (80.5-192)	2030 (1666-2475)	16.3	5E-09	2.14	4.92	2.30
Il2ra/Cd25		12.3 (3.95-38.3)	8.23 (3.06-22.2)	0.67	6E-01	4.33 (3.80-4.94)	5.77 (4.38-7.61)	1.33	9E-02	0.35	0.70	1.99
Lcn2		29.6 (14.9-58.6)	2797 (1832-4271)	94.5	1E-08	25.2 (11.4-55.9)	4571 (3268-6393)	181	3E-09	0.85	1.63	1.92
Sting		51.2 (42.7-61.3)	47.8 (38.5-59.3)	0.93	7E-01	181 (160-204)	317 (277-363)	1.76	2E-05	3.53	6.64	1.88
Jak2		166 (150-185)	355 (267-472)	2.13	2E-04	6.02 (5.43-6.66)	23.4 (16.5-33.2)	3.89	2E-06	0.04	0.07	1.82
Lrrk2		93.9 (82.2-107)	83.5 (66.1-105)	0.89	4E-01	153 (125-188)	236 (163-340)	1.54	6E-02	1.63	2.82	1.73
Il1rn		18.6 (12.1-28.6)	325 (185-572)	17.5	7E-07	17.4 (13.0-23.3)	523 (408-670)	30.0	2E-11	0.94	1.61	1.72
Csf1		8.18 (5.53-12.1)	17.5 (12.0-25.5)	2.14	2E-02	3.32 (2.43-4.55)	12.1 (8.04-18.1)	3.63	1E-04	0.41	0.69	1.70
Cfb		6.28 (4.26-9.25)	111 (46.0-269)	17.7	2E-05	7.93 (3.52-17.9)	225 (140-362)	28.4	3E-06	1.26	2.03	1.60
Itgam/Cd11b		265 (207-339)	361 (283-461)	1.37	1E-01	183 (135-250)	396 (278-564)	2.16	5E-03	0.69	1.10	1.58
Timp1		108 (94.0-124)	192 (117-316)	1.78	5E-02	102 (91.0-114)	272 (226-326)	2.67	1E-07	0.94	1.41	1.50
Socs3		19.9 (12.8-30.8)	254 (213-304)	12.8	2E-08	16.8 (12.1-23.4)	304 (224-413)	18.1	1E-09	0.84	1.20	1.42
Il6		0.27 (0.17-0.42)	9.27 (5.27-16.3)	34.5	5E-08	0.17 (0.09-0.33)	7.77 (1.74-34.6)	45.5	3E-04	0.64	0.84	1.32
Ltf		10.5 (5.82-18.8)	824 (445-1526)	78.8	3E-08	28.5 (12.0-68.0)	2930 (2069-4151)	103	5E-08	2.73	3.56	1.30
S100a9		912 (539-1543)	4540 (2426-8497)	4.98	2E-03	2036 (1240-3344)	12972 (10079-16695)	6.37	6E-06	2.23	2.86	1.28
Cxcl1		0.13 (0.09-0.16)	20.3 (9.03-45.7)	158	9E-09	0.06 (0.03-0.13)	12.3 (3.08-49.0)	198	6E-06	0.48	0.60	1.25
Il10		0.11 (0.10-0.13)	3.57 (1.86-6.87)	31.7	9E-08	0.09 (0.05-0.15)	3.53 (1.13-11.0)	39.6	3E-05	0.79	0.99	1.25
Gapdh		646 (591-707)	1014 (558-1845)	1.57	2E-01	667 (590-754)	1295 (923-1818)	1.94	2E-03	1.03	1.28	1.24
Csf1r/Cd115		233 (201-270)	47.3 (31.7-70.7)	0.20	2E-06	320 (262-390)	78.3 (44.7-137)	0.24	2E-04	1.37	1.66	1.21
Nfkb1		149 (133-168)	336 (283-399)	2.25	1E-06	118 (114-122)	300 (209-431)	2.54	1E-04	0.79	0.89	1.13
Nr3c1		124 (105-148)	119 (101-139)	0.95	7E-01	47.0 (41.5-53.1)	49.9 (38.3-65.2)	1.06	7E-01	0.38	0.42	1.11
Akt1		98.1 (82.8-116)	68.1 (54.1-85.7)	0.69	3E-02	148 (133-164)	113 (88.8-144)	0.76	6E-02	1.51	1.66	1.10
Pkm		538 (469-618)	884 (506-1544)	1.64	1E-01	417 (364-479)	741 (558-983)	1.77	2E-03	0.78	0.84	1.08
Il4ra/Cd124		91.2 (80.2-104)	247 (198-309)	2.71	1E-06	51.7 (47.2-56.7)	150 (122-184)	2.90	8E-08	0.57	0.61	1.07
Fcer2		104 (70.7-153)	15.7 (10.0-24.5)	0.15	1E-05	55.7 (47.6-65.1)	8.93 (4.96-16.1)	0.16	2E-05	0.54	0.57	1.06
Camp		3.12 (1.76-5.54)	285 (145-558)	91.2	3E-08	5.42 (2.26-13.0)	524 (354-776)	96.7	8E-08	1.74	1.84	1.06
Hif1a		53.4 (46.5-61.5)	191 (144-254)	3.58	8E-07	83.6 (77.0-90.8)	299 (203-440)	3.58	1E-05	1.56	1.57	1.00
Padi4		15.9 (8.70-29.0)	41.2 (22.3-76.0)	2.59	5E-02	44.6 (34.1-58.4)	105 (89.0-123)	2.34	6E-05	2.81	2.54	0.90
Bcl3		21.1 (14.8-30.0)	209 (168-260)	9.92	1E-08	18.9 (15.1-23.8)	167 (127-219)	8.80	3E-09	0.90	0.80	0.89
Mt2		0.38 (0.19-0.77)	89.1 (33.8-235)	235	1E-07	0.31 (0.13-0.75)	61.2 (20.5-182)	199	2E-06	0.81	0.69	0.84
Myd88		50.8 (39.5-65.4)	127 (96.0-168)	2.50	2E-04	51.5 (46.6-56.8)	108 (99.7-116)	2.09	5E-09	1.01	0.85	0.84
Thy1		35.0 (22.6-54.2)	25.2 (13.3-48.0)	0.72	3E-01	503 (457-553)	297 (240-368)	0.59	4E-04	14.4	11.8	0.82
Jun		31.4 (23.9-41.2)	157 (105-234)	5.01	6E-06	6.02 (5.43-6.66)	23.4 (16.5-33.2)	3.89	2E-06	0.19	0.15	0.78
Icam		24.1 (20.3-28.6)	301 (257-352)	12.5	3E-12	12.6 (10.5-15.0)	121 (67.3-217)	9.61	2E-06	0.52	0.40	0.77
Cx3cr1		27.9 (21.8-35.7)	3.46 (1.99-6.00)	0.12	4E-06	26.1 (19.0-35.9)	2.45 (1.44-4.15)	0.09	1E-06	0.94	0.71	0.76
Cxcl2		0.16 (0.09-0.26)	119 (76.0-186)	762	6E-12	0.10 (0.04-0.26)	60.2 (19.6-185)	574	2E-07	0.67	0.51	0.75
Akt2		138 (117-163)	275 (182-415)	1.99	9E-03	59.4 (57.1-61.7)	87.6 (71.5-107)	1.47	2E-03	0.43	0.32	0.74
Jak1		415 (386-447)	623 (467-830)	1.50	2E-02	256 (232-281)	277 (214-359)	1.09	6E-01	0.61	0.45	0.72
ERV Gag-pol		3210 (2406-4283)	13633 (10083-18433)	4.3	4E-06	233 (195-278)	692 (548-873)	2.97	2E-06	0.07	0.05	0.70
Pbib		93.7 (82.3-107)	97.2 (58.5-161)	1.04	9E-01	143 (134-153)	102 (71.2-147)	0.71	9E-02	1.53	1.05	0.69
Nox1		0.14 (0.11-0.18)	0.98 (0.65-1.48)	7.01	8E-07	10.1 (8.12-12.5)	45.6 (34.4-60.5)	4.53	3E-07	71.8	46.4	0.65
Tnf		2.99 (2.26-3.96)	19.5 (14.0-27.1)	6.52	3E-07	0.76 (0.42-1.40)	3.10 (1.62-5.95)	4.06	7E-03	0.26	0.16	0.62
Tlr2		26.1 (18.1-37.5)	165 (126-215)	6.31	6E-07	50.8 (37.2-69.5)	194 (128-294)	3.83	1E-04	1.95	1.18	0.61
Cd8		23.8 (19.6-29.0)	17.3 (13.6-22.1)	0.73	7E-02	52.6 (41.6-66.5)	21.3 (10.2-44.7)	0.41	4E-02	2.21	1.23	0.56
Stat1		160 (132-193)	449 (247-814)	2.81	6E-03	256 (221-297)	387 (323-465)	1.51	3E-03	1.60	0.86	0.54
Mapk1		71.1 (65.5-77.0)	92.6 (65.0-132)	1.30	2E-01	67.5 (63.9-71.3)	44.2 (37.4-52.3)	0.65	2E-04	0.95	0.48	0.50
Arg1		1.12 (0.36-3.42)	18.4 (5.99-56.7)	16.5	4E-03	261 (154-441)	2161 (1709-2732)	8.28	2E-06	234	117	0.50
Ccl2		0.16 (0.10-0.27)	12.7 (4.87-33.0)	77.7	8E-07	0.19 (0.07-0.50)	7.42 (3.23-17.0)	38.5	3E-05	1.18	0.58	0.50
Rigi/Ddx58		47.3 (41.9-53.3)	293 (175-491)	6.20	0.00	26.6 (20.5-34.4)	98.9 (76.9-127)	3.72	0.00	0.60	0.29	0.49
MT-Co1		10015 (7976-12576)	13316 (6322-28047)	1.33	5E-01	6608 (4976-8775)	4164 (2563-6765)	0.63	1E-01	0.66	0.31	0.47
Cd4		55.6 (45.7-67.6)	27.5 (19.6-38.8)	0.50	4E-03	88.5 (75.5-104)	19.9 (10.4-38.3)	0.23	5E-04	1.59	0.72	0.45
Socs1		0.82 (0.59-1.13)	45.4 (25.6-80.4)	55.6	3E-09	0.74 (0.53-1.04)	18.4 (9.19-36.9)	24.8	5E-07	0.91	0.41	0.45
Cd3d		81.9 (71.8-93.4)	46.7 (32.8-66.5)	0.57	1E-02	77.1 (64.1-92.7)	18.3 (8.91-37.6)	0.24	2E-03	0.94	0.39	0.42
Mtor		27.9 (24.3-31.9)	40.2 (34.9-46.3)	1.44	2E-03	62.1 (55.6-69.4)	36.3 (27.4-48.1)	0.58	3E-03	2.23	0.90	0.41
Il18		1.78 (1.18-2.70)	4.24 (2.84-6.32)	2.38	1E-02	4.89 (4.22-5.67)	4.15 (2.40-7.18)	0.85	6E-01	2.74	0.98	0.36
Stat2		41.0 (36.6-45.9)	229 (144-364)	5.59	2E-06	89.1 (75.3-105)	177 (153-206)	1.99	2E-05	2.17	0.77	0.36
Il7r/Cd127		36.8 (29.2-46.3)	39.6 (28.4-55.1)	1.08	7E-01	59.2 (50.5-69.3)	22.3 (15.5-32.0)	0.38	2E-04	1.61	0.56	0.35
Tgfb		0.13 (0.09-0.19)	8.02 (3.82-16.8)	60.6	4E-08	1.08 (0.78-1.50)	22.1 (11.5-42.5)	20.4	5E-07	8.18	2.76	0.34
Acod1		1.20 (0.60-2.37)	388 (289-521)	324	2E-10	2.64 (1.72-4.05)	276 (202-378)	105	2E-11	2.21	0.71	0.32
Csf2		0.16 (0.11-0.23)	3.94 (2.09-7.43)	24.9	8E-07	0.05 (0.03-0.07)	0.33 (0.11-0.98)	7.10	4E-03	0.29	0.08	0.28
Stat4		55.4 (50.4-60.8)	110 (95.3-128)	1.98	8E-07	26.4 (21.8-32.0)	14.4 (10.4-19.9)	0.55	6E-03	0.48	0.13	0.27
Dhx58/Lgp2		11.3 (9.76-13.2)	156 (77.1-317)	13.8	2E-06	20.2 (14.3-28.4)	73.1 (55.9-95.5)	3.63	3E-05	1.78	0.47	0.26
Isg15		81.5 (65.6-101)	600 (219-1646)	7.37	2E-03	183 (104-322)	342 (175-668)	1.87	2E-01	2.24	0.57	0.25
Oas1		10.6 (8.90-12.6)	106 (49.4-225)	9.98	3E-05	81.1 (51.9-127)	191 (127-285)	2.35	1E-02	7.67	1.81	0.24
Cgas		1.55 (1.19-2.02)	17.3 (14.3-21.0)	11.2	4E-10	5.53 (4.52-6.77)	13.7 (10.8-17.4)	2.47	3E-05	3.57	0.79	0.22
Gbp4		86.1 (73.9-100)	795 (523-1208)	9.23	4E-08	81.5 (70.5-94.1)	153 (112-209)	1.88	2E-03	0.95	0.19	0.20
Il12		2.85 (2.19-3.71)	11.7 (7.75-17.8)	4.12	4E-05	0.30 (0.17-0.53)	0.25 (0.14-0.43)	0.82	6E-01	0.11	0.02	0.20
ERV Env		913 (661-1261)	3398 (2240-5155)	3.7	2E-04	48.1 (42.9-53.9)	33.1 (25.7-47.5)	0.69	2E-02	0.05	0.01	0.18
Irf7		43.5 (34.1-55.5)	528 (276-1010)	12.1	2E-06	145 (83.0-253)	296 (186-469)	2.04	7E-02	3.33	0.56	0.17
Fos		16.6 (11.9-23.2)	17.3 (9.38-31.8)	1.04	9E-01	20.2 (14.5-28.0)	2.39 (1.61-3.55)	0.12	5E-07	1.21	0.14	0.11
Ifih1/Mda5		8.39 (6.56-10.7)	181 (82.4-395)	21.5	2E-06	43.3 (35.5-52.8)	92.1 (75.8-112)	2.13	6E-05	5.16	0.51	0.10
Cxcl10		0.35 (0.21-0.58)	231 (81.4-655)	659	1E-08	0.26 (0.12-0.59)	16.2 (5.83-44.9)	61.7	1E-05	0.75	0.07	0.094
Mpo		0.33 (0.19-0.57)	19.5 (8.56-44.6)	59.9	6E-07	1.73 (0.71-4.26)	7.91 (2.61-24.0)	4.56	5E-02	5.32	0.41	0.076
Myc		22.4 (18.0-27.7)	294 (209-413)	13.1	3E-09	41.6 (35.9-48.3)	41.4 (25.2-68.1)	1.00	1E+00	1.86	0.14	0.076
Ngp		7.82 (4.03-15.2)	1071 (580-1977)	137	1E-08	130 (77.8-216)	1275 (1022-1589)	9.83	5E-07	16.6	1.19	0.072
Ifit1		10.7 (8.18-13.9)	309 (96.6-991)	29.0	5E-03	53.1 (32.8-85.8)	96.7 (59.0-158)	1.82	1E-01	4.98	0.31	0.063
Mx2		6.29 (5.53-7.16)	270 (158-460)	42.9	1E-09	198 (121-323)	476 (300-757)	2.41	2E-02	31.5	1.76	0.056
Cd69		14.1 (11.5-17.3)	203 (155-266)	14.4	1E-10	39.3 (34.6-44.6)	31.1 (17.3-55.9)	0.79	5E-01	2.78	0.15	0.055
Ccl4		1.84 (1.10-3.09)	217 (126-376)	118	3E-09	117 (75.4-181)	740 (487-1126)	6.33	2E-05	63.5	3.40	0.054
Gzmb		14.7 (9.75-22.1)	1063 (603-1873)	72.4	3E-09	28.8 (15.8-52.6)	103 (56.2-188)	3.57	1E-02	1.96	0.10	0.049
Tnfrsf9/Cd137		0.64 (0.46-0.89)	97.5 (47.3-201)	152	3E-09	0.76 (0.42-1.40)	3.10 (1.62-5.95)	4.06	7E-03	1.19	0.03	0.027
Cd14		44.5 (25.5-77.7)	667 (408-1091)	15.0	2E-06	88.5 (75.5-104)	19.9 (10.4-38.3)	0.23	5E-04	1.99	0.03	0.015
Nos2		0.13 (0.10-0.17)	58.9 (22.1-157)	459	4E-09	0.04 (0.03-0.05)	0.25 (0.09-0.69)	6.16	3E-03	0.31	0.00	0.013
Ifng		0.58 (0.33-1.02)	438 (178-1080)	751	3E-09	0.15 (0.08-0.29)	0.51 (0.14-1.89)	3.30	1E-01	0.26	0.00	0.004

* Gene: Mus musculus reference genome GenBank name/other name in literature. The list includes the 113 coding sequences listed in Methods and two endogenous retrovirus (ERV) proteins, Env and Gag-pol.
† Heatmap (red-white-blue scale) of log-transformed "Peromyscus v Mus LPS/control" values (last column). This is ratio of "Peromyscus v Mus LPS FC" (2nd column from right) to "Peromyscus v Mus control FC" (3rd column from right)
‡ Mean unique reads for given gene normalized for reads for Ptprc (Cd45) gene for a sample. The 95% confidence intervals (CI) are asymmetric. Actual [gene]/Ptprc ratios are X 10^-3.
** Other abbreviations: Peromyscus, Peromyscus leucopus; Mus, Mus musculus FC, fold-change; FDR, false discovery rate

"Alternatively-activated" macrophages and "nonclassical" monocytes in *P. leucopus*

While we could not type single cells using protein markers, we could assess relative transcription of established indicators of different white cell subpopulations in whole blood. The present study, which incorporated outbred *M. musculus* instead of an inbred strain, confirmed the previous finding of differences in Nos2, the gene for inducible nitric oxide synthase and Arg1, the gene for arginase 1, expression between *M. musculus* and *P. leucopus* (figure 39, Table 15). Results similar to the RNA-seq findings were obtained with specific RT-qPCR assays for Nos2 and Arg1 transcripts for P. musculus and *M. musculus* (Table 15).

Table 15 RT-qPCR of blood of LPS-treated and control P. leucopus and M. musculus

Blood mRNA source	Control mean copies (95% CI)	LPS mean copies (95% CI)	Fold difference LPS/control	*t* test *p* value/Mann-Whitney *p* value
P. leucopus				
Gapdh	$1.22 (0.40\text{-}3.74) \times 10^5$	$2.38 (1.10\text{-}5.16) \times 10^5$	1.95	0.35/0.49
Nos2	191 (141-260)	138 (61-315)	0.72	0.47/0.53
Arg1	$4.62 (3.09\text{-}6.92) \times 10^3$	$12.3 (3.22\text{-}47.1) \times 10^3$	2.66	0.18/0.55
M. musculus				
Gapdh	$6.10 (2.32\text{-}16.0) \times 10^6$	$1.77 (0.80\text{-}3.89) \times 10^6$	0.29	0.06/0.02
Nos2	101 (68-151)	1891 (866-4130)	18.6	<0.00001/0.002
Arg1	27 (15-20)	16 (8-34)	0.59	0.29/0.45

Low transcription of Nos2 in both in controls and LPS-treated P. leucopus and an increase in Arg1 with LPS was also observed in another experiment for present study where the dose of LPS was 1 µg/g body mass instead of 10 µg/g and the interval between injection and assessment was 12 h instead of 4 h (Table 16).

In addition to the differences in Nos2 and Arg1 expression for typing macrophage and monocyte subpopulations, there are also the relative expressions of three other pairs of genes: (1) IL12 and IL10, where a lower IL12/IL10 ratio is more characteristic of alternatively activated or M2 type (334), (269); (2) the proto-oncogene kinases Akt1 and Akt2, where the associations are Akt1 with M2-type and Akt2 with M1-type macrophages (335), (336); and (3) CD14 and CD16 (low affinity

immunoglobulin gamma Fc region receptor III; Fcgr3), where low expression of CD14 and high expression of CD16/Fcgr3 is associated with "non-classical" monocytes (337). There is evidence that nonclassical monocytes can change to M2-type macrophages (338).

Table 16 Targeted RNA-seq of P. leucopus blood in 12 h and 1 ug/g LPS experiment

Gene	Control (n=3) mean (95% CI)	LPS (n=3) mean (95% CI)	Fold change	FDR *p* value
Akt1	220 (12-4202)	514 (321-825)	2.3	0.05
Akt2	145 (97-217)	336 (236-478)	2.3	0.04
Arg1	146 (58-367)	2812 (273-28925)	19	0.018
Cd14	82 (12-569)	914 (161-5197)	11	0.05
Cd69	165 (87-310)	68 (14-329)	0.42	0.15
ERV Env	25 (3-242)	40 (7-224)	1.6	0.86
ERV Gag-pol	3085 (132-14695)	2768 (533-1278)	0.9	0.66
Fcgr3/Cd16	40 (25-66)	841 (533-1278)	21	0.008
Gapdh	7176 (3504-15142)	23811 (5827-97306)	3.3	0.07
Gbp4	97 (6-1551)	439 (33-5819)	4.5	0.05
Ifit1	367 (51-2663)	1373 (374-5047)	3.7	0.05
Ifng	0 (0-0)	0 (0-0)	.	.
Il1b	258 (18-3652)	1432 (183-11220)	5.6	0.1
Irf7	121 (93-157)	11405 (530-245616)	94	0.003
Isg15	429 (184-1001)	19505 (11140-34152)	45	0.005
Mx2	157 (73-341)	1310 (323-5315)	8.3	0.04
Nos2	0 (0-0)	0 (0-0)	.	.
Oas1	65 (31-138)	1458 (367-5795)	22	0.03
Rigi/Ddx58	38 (3-504)	173 (35-845)	4.5	0.07
Saa3	2 (0-250)	6683 (1494-29896)	3372	0.03
Slpi	6 (1-46)	779 (326-1864)	123	0.03
Sod2	180 (54-607)	3406 (698-16633)	19	0.03

These four relationships, which are presented as log-transformed ratios of Nos2/Arg1, IL12/IL10, Akt1/Akt2, and CD14/CD16, are shown in figure 40. We confirmed the difference between *P. leucopus* and *M. musculus* in the ratios of Nos2/Arg1 and IL12/IL10 (3) with outbred mice and normalization for white cells. In both species the Akt1/Akt2 ratio declined in LPS-treated animals, but for *P. leucopus* the ratio remained > 1.0 even among LPS-treated animals, while in the blood of *M. musculus* the ratio was < 1.0 at baseline and declined further in the LPS-treated animals.

An orthologous gene for Ly6C (339), a protein used for typing mouse monocytes and other white cells, has not been identified in Peromyscus or other Cricetidae family members. Therefore, expression of Cd14 was compared with expression of the Ly6c alternative Fcgr3, which deermice and other cricetines do have. In mice, the Cd14/Fcgr3 transcription ratio increased from baseline in the LPS group. In the deermice, the ratio in control animals was midway between the two groups of mice but there was a marked decrease in the LPS-treated deermice (figure 40). This was not associated with a fall in the absolute numbers or percentages of monocytes in the blood of these animals (Table 13).
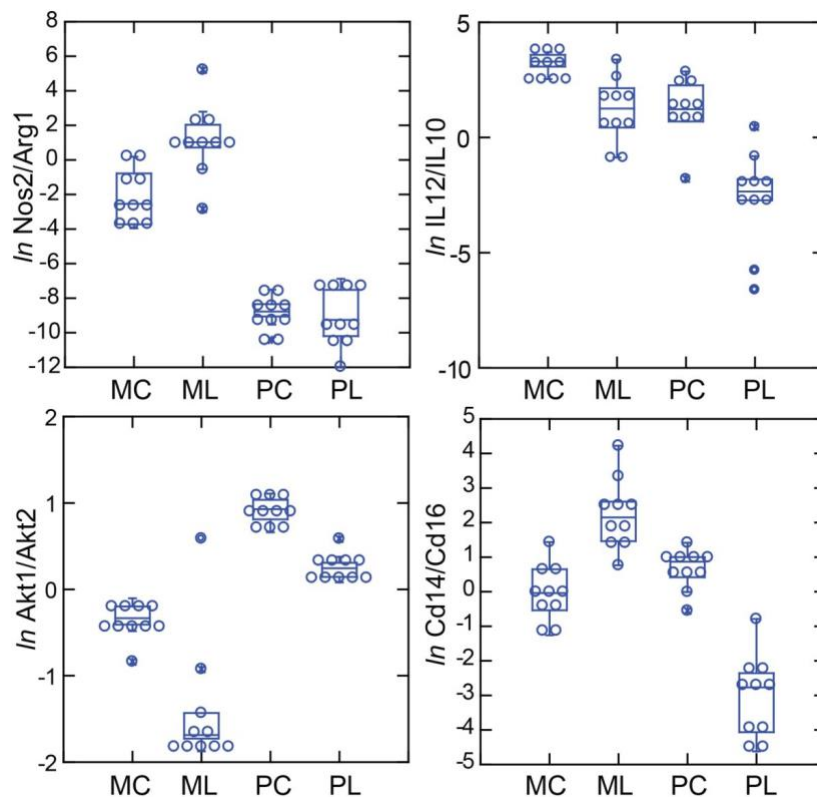
Figure 40 Box plots of log-transformed ratios of four pairs of gene transcripts from targeted RNA-seq analysis
Box plots of log-transformed ratios of four pairs of gene transcripts from targeted RNA-seq analysis of blood of P. leucopus (P) or M. musculus (M) with (L) or without (C) treatment with LPS. Upper left, Nos2/Arg1. Upper right, IL12/IL10. Lower left, Akt1/Akt2. Lower right, Cd14/Cd16.

Taken together, the Nos2-Arg1, Il12-Il10, Akt1-Akt2, and CD14-CD16 relationships document a disposition toward alternatively-activated macrophages and nonclassical monocytes in *P. leucopus*

both before and after exposure to LPS. This contrasts with profiles consistent with a predominance of classically-activated macrophages and classical monocytes in mice.

Interferon-gamma and interleukin-1 beta dichotomy between deermice and murids

For mice the Ifng transcript was one of the top ranked DEGs by both fold-change and adjusted p value by genome-wide RNA-seq (Table 14). In contrast, for *P. leucopus* Ifng was far down the list, and the comparably ranked DEG instead was Il1b. This inversion of relationships between two pro-inflammatory cytokines was confirmed by analysis of the individual animals of both species (figure 41). There was little or no detectable transcription of Ifng in the blood of deermice in which Il1b expression was high. There was also scant to no transcription of Ifng in the blood of *P. leucopus* 12 h after injection of LPS (Table 16).
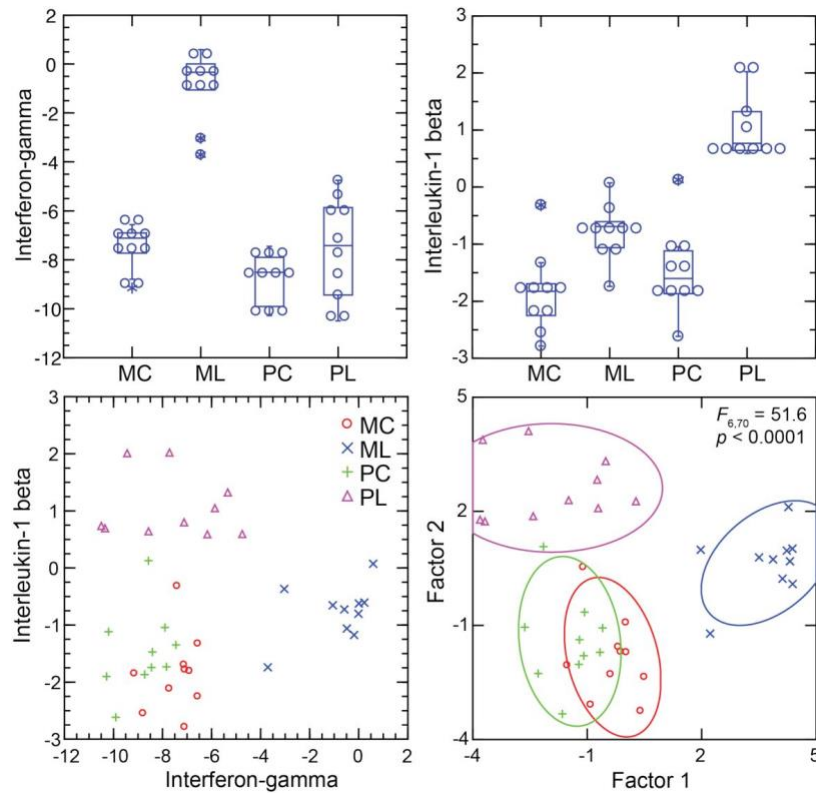
Figure 41 Transcripts of interferon-gamma and interleukin-1 beta by targeted RNA-seq of the blood of two rodents

Transcripts of interferon-gamma and interleukin-1 beta by targeted RNA-seq of the blood of *P. leucopus* (P) or *M. musculus* (M) with (L) or without (C) treatment with LPS. The top panels are box plots of the individual values. The lower left panel is a scatter plot of interleukin-1 ý on interferon-ψ values. The lower right panel is a Discriminant Analysis of these pairs of values where Factor 1 corresponds to interferon-gamma, and Factor 2 corresponds to interleukin-1 beta.

The up-regulation of Ifng within 4 hours of exposure to LPS was not limited to the species *M. musculus*. In an experiment with the rat *R. norvegicus*, we used two different LPS doses (5 µg/g and 20 µg/g), but the same 4 h endpoint and whole blood as the sample. Both groups of LPS-treated rats had lowered total white blood cells, and, like the mice, lower neutrophil-to-lymphocyte ratios compared to controls (Table 17). There were also elevations of interferon-gamma, interleukin-6 and interleukin-10 proteins from undetectable levels in the blood of the treated rats. The values for rats receiving 5 µg/g or 20 µg/g doses were similar, so these groups were combined. By targeted RNA-seq there were 24x fold-changes between the LPS-treated rats and control rats for Ifng and Nos2 but only ~3x fold-change for Il1b (Table 17).

Given these findings, we asked why the interferon-gamma response observed in CD-1 mice and rats here was not as pronounced in BALB/c mice as described in Chapter 3. Accordingly, we used the RNA-seq reads obtained from the prior study described in Chapter 3 in combination with the reads of the present study and carried out targeted RNA-seq (figure 42). The BALB/c inbred mice had, like the CD-1 mice, modest elevations of Il1b transcription. Ifng expression was also elevated in the BALB/c animals but not to the degree noted in CD-1 mice or rats. One explanation is an inherent difference of BALB/c mice from other strains in their lower interferon-gamma response to LPS.
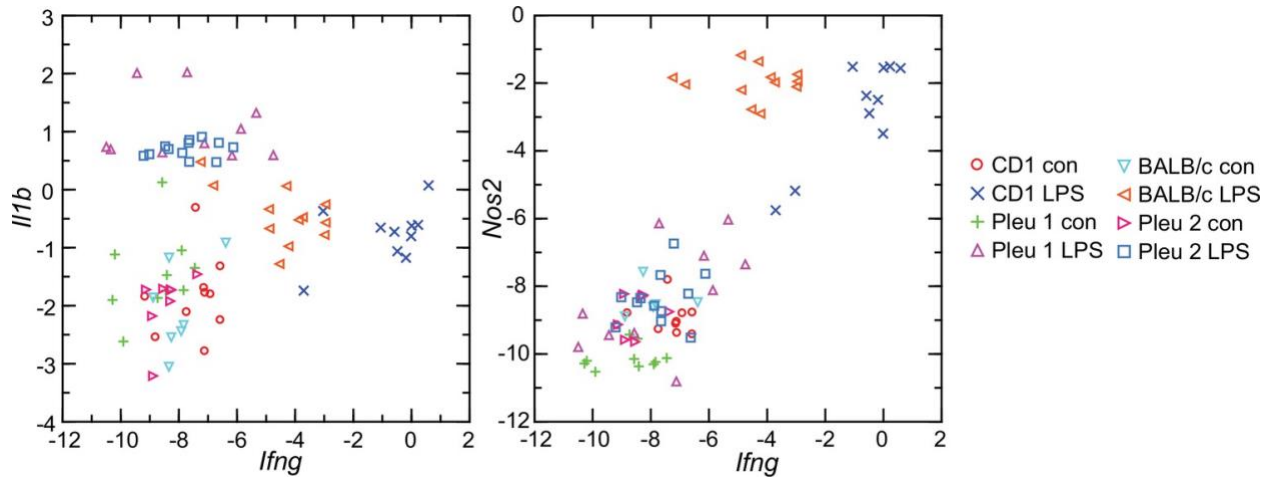
Figure 42 Scatter plot of interleukin-1 beta and nitric oxide synthase by interferon-gamma targeted RNA-seq of the blood of P. leucopus or M musculus with or without treatment with LPS

Scatter plots of log-transformed normalized transcripts of genes for interleukin-1 beta (Il1b; left panel) or nitric oxide synthase 2 (Nos2; right panel) on interferon-gamma (Ifng) of blood of *P. leucopus* (Pleu) or *M. musculus* (outbred CD-1 and inbred BALB/c) with (LPS) or without (con) treatment with lipopolysaccharide 4 hr previously. The data are from the present study of *P. leucopus* and *M. musculus* (Pleu 2 and CD-1) and from the study presented in Chapter 3 of *P. leucopus* and *M. musculus* (Pleu 1 and BALB/c).

Table 17 Hematology, cytokines, and targeted RNA-seq of LPS-treated and control rats

Variable	Control (n=5) mean (95% CI)*	LPS (n=11) mean (95% CI)*,†	Fold change	FDR *p* value
Hematology				
Hematocrit (%)	48 (46-50)	48 (46-49)	1.0	1E+00
White blood cells	7660 (7200-8310)	4980 (2020-7940)	0.75	2E-01
Neutrophils	3680 (3260-4090)	1410 (520-2290)	0.38	4E-03
Lymphocytes	3170 (3010-3330)	2830 (1230-4430)	0.89	8E-01
Neutrophil/lymphocyte	1.17 (1.01-1.33)	0.49 (0.44-0.55)	0.42	7E-09
Plasma cytokines (pg/ml)				
Interleukin-6	0 (0-0)	36933 (21676-52190)	.	5E-15
Interleukin-10	9 (1-17)	640 (477-802)	71	2E-09
Interferon-gamma	0 (0-0)	9091 (7126-11056)	.	9E-20
Targeted RNA-seq				
Akt1	35.8 (29.8-42.9)	24.3 (21.5-27.5)	0.68	4E-03
Akt2	50.8 (44.9-57.4)	109 (95.2-125)	2.2	9E-06
Arg1	0.04 (0.02-0.08)	0.21 (0.10-0.44)	4.8	2E-02
Cd14	7.7 (5.3-11.2)	43.4 (29.6-63.6)	5.6	9E-05
Cd177	0.89 (0.43-1.8)	190 (143-251)	213	2E-10
Cd3	32.3 (27.7-37.6)	21.9 (18.5-25.8)	0.68	1E-02
Cd69	17.9 (16.7-19.2)	58.9 (49.1-70.7)	3.3	9E-07
Cgas	1.5 (1.3-1.6)	12.8 (10.8-15.1)	8.7	3E-10
Cxcl10	0.43 (0.34-0.56)	130 (93.7-181)	302	1E-11
Dhx58	7.1 (6.7-7.6)	113 (97.3-130)	15.8	3E-12
ERV Env	8.4 (7.5-9.3)	713 (624-815)	85.3	1E-14
ERV gag-pol	6.0 (5.4-6.7)	506 (449-570)	84.3	7E-15
Fcgr2a	114 (87.0-150)	764 (631-925)	6.7	4E-08
Fcgr2b	32.5 (24.6-43.1)	161 (127-204)	4.9	2E-06
Fcgr3/Cd16	15.2 (13.6-17.0)	13.9 (11.8-16.3)	0.91	5E-01
Gapdh	327 (237-451)	1643 (1385-1949)	5.0	2E-07
Gbp4	35.7 (33.6-38.0)	269 (237-306)	7.5	3E-11
Ibsp	0.41 (0.31-0.55)	0.61 (0.47-0.80)	1.5	1E-01
Ifih1	18.7 (17.1-20.4)	165 (149-184)	8.8	2E-12
Ifit1	102 (70.0-147)	756 (677-844)	7.4	3E-09
Ifng	0.47 (0.32-0.67)	10.3 (6.4-16.5)	22.1	1E-06
Il10	0.12 (0.07-0.21)	4.5 (3.4-5.8)	38.2	5E-09
Il12	0.07 (0.04-0.11)	3.2 (2.0-4.9)	45.9	7E-08
Il1b	58.6 (39.7-86.4)	618 (503-760)	10.6	2E-08
Il6	0.06 (0.05-0.08)	4.4 (2.9-6.6)	70.9	6E-09
Irf7	44.8 (36.9-54.3)	443 (372-528)	9.9	6E-10
Isg15	15.6 (13.1-18.7)	624 (534-729)	39.9	6E-13
Itgam?Cd11b	66.3 (52.5-83.7)	208 (161-269)	3.1	9E-05
Mmp8	75.2 (50.4-112)	519 (438-615)	6.9	7E-08
Mx2	40.7 (35.9-46.2)	900 (780-1039)	22.1	1E-12
Nos2	32.4 (17-60.6)	2990 (2491-3589)	92.4	1E-10
Oas1	23.1 (18.8-28.4)	151 (140-164)	6.6	3E-11
Rigi/Ddx58	8.6 (7.8-9.4)	151 (135-168)	17.6	1E-13
S100a9	298 (190-466)	2884 (2269-3666)	9.7	2E-07
Saa1	0.60 (0.49-0.73)	699 (552-884)	1167	4E-14
Slpi	20.2 (13.1-31.3)	262 (197-347)	12.9	1E-07
Sod2	63.8 (51.1-79.6)	901 (759-1070)	14.1	2E-10
Tlr4	5.3 (4.6-6.0)	20.1 (17.2-23.6)	3.8	8E-08
Tnf	1.1 (0.63-1.9)	78.8 (62.6-99.1)	72.8	2E-10
Ratios				
Akt1/Akt2	0.70 (0.65-0.76)	0.22 (0.19-0.25)	0.31	2E-08
Cd14/Fcgr3	0.51 (0.34-0.76)	3.14 (2.26-4.36)	6.16	1E-05
IL12/IL10	0.54 (0.22-1.37)	0.70 (0.51-0.95)	1.30	5E-01
Nos2/Arg1	741 (403-1362)	14244 (7615-26646)	19.2	5E-05

* For targeted RNA-seq it is mean unique reads for given gene normalized for reads for Ptprc (Cd45) gene for a sample. The 95% confidence intervals (CI) are asymmetric. Actual [gene]/Ptprc ratios are X
† The results for rats receiving 5 µg/g (n=6) and 20 µg/g (n=5) were combined

Interferon-gamma and inducible nitric oxide synthase

Interferon-gamma is a determinant of Nos2 expression (340), (341). So, the scant transcription of Ifng in *P. leucopus* conceivably accounted for the low expression of Nos2 in that species. The analysis shown in upper left panel of figure 43 shows a tight correlation between the levels of transcription of Ifng and Nos2 for both species and both experimental conditions. A significant correlation was also observed for the combined set of animals between the ratios of Nos2 to Arg1 and Ifng to Il1b (upper right panel figure 43), an indication of co-variation between Ifng expression and macrophage polarization.
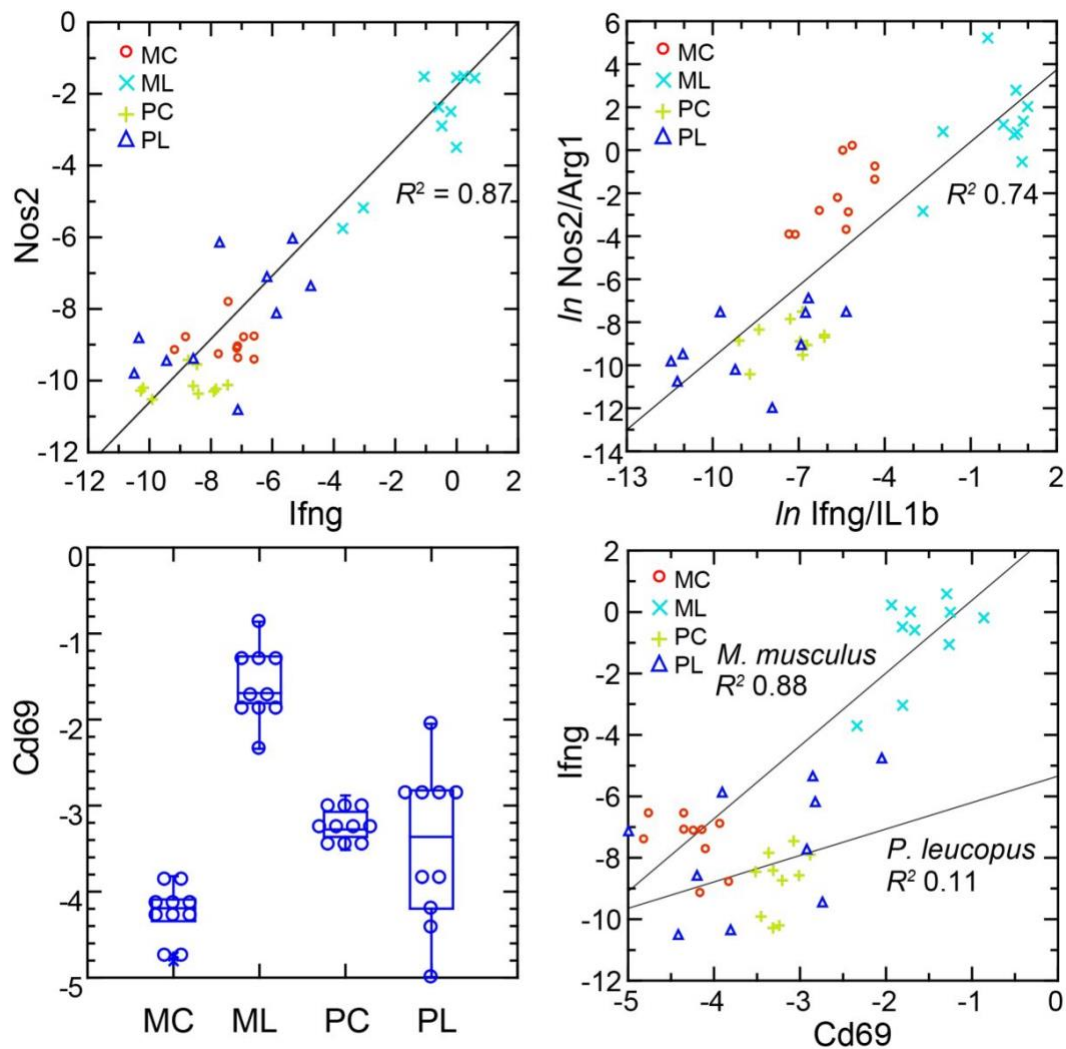
Figure 43 Normalized transcripts of Nos2, Ifng, and Cd69 in targeted RNA-seq analysis of blood of
P. leucopus or *M. musculus* with or without treatment with LPS

Normalized transcripts of Nos2, Ifng, and Cd69 in targeted RNA-seq analysis of blood of *P. leucopus* (P) or *M. musculus* (M) with (L) or without (C) treatment with LPS. Upper left: scatter plot of individual values for Nos2 on Ifng with linear regression curve and coefficient of determination ($R2$). Upper right: natural logarithm (*ln*) of ratios of Nos2/Argi1 on Ifng/IL1b with regression curve and $R2$. Lower left: Box plots of individual values of normalized transcripts of Cd69. Lower right: Scatter plot of Ifng on Cd69 with separate regression curves and R2 values for *M. musculus* and *P. leucopus*.

The plausible sources of Ifng mRNA in whole blood are T-cells, Natural Killer cells, and Type 1 Innate Lymphoid Cells (342). A DEG for *M. musculus* by both genome-wide and targeted RNA-seq (Table 14) was Cd69, a C-type lectin protein and an early activation antigen for these cells (343). In *P. leucopus* transcription Cd69 occurred in the blood of control *P. leucopus* , but it was the same or only marginally different for the LPS-treated animals (lower left panel figure 43). In contrast, in *M. musculus* the baseline transcription of Cd69 was below that of *P. leucopus,* but in the LPS-treated mice it was many fold higher. In mice transcripts for Cd69 correlated tightly with Ifng transcription, while in the deermice there was little correlation between Cd69 and Ifng expression at those low levels (lower right panel figure 43).

The findings are consistent with CD69-positive cells being a source of Ifng in mice. Cd69 transcription was comparatively higher in control deermice than in control mice, so we presume that deermice have CD69-positive cells at baseline. One explanation then for the comparatively few Ifng transcripts in the deermice after LPS is a diminished responsiveness of these cells. Tlr4 expression increased ~3 -fold more in *P. leucopus* than in *M. musculus* after LPS (Table 14), but the magnitude of the decline in expression of Cd14 in deermice than mice was even greater. CD14 is required for LPS-stimulated signaling through surface TLR4 (344), and, as such, its decreased availability for this signaling pathway is a possible explanation for the moderated response to LPS in *P. leucopus*.

Interferon-stimulated genes and RIG-I-like receptors

As noted, GO terms differentiating mice from deermice included "response to interferon-beta" and "response to virus" (figure 36). There was also the example of Mx2, an ISG with anti-viral activity on

its own, that showed a greater fold-change from baseline in mice than in deermice (figure 39). Five other ISGs--guanylate binding protein 4 (Gbp4), interferon-induced protein with tetratricopeptide repeat (Ifit1), interferon regulatory factor 7 (Irf7), ubiquitin-type modifier ISG15 (Isg15), and 2'-5' oligoadenylate synthase 1A (Oas1a)—had higher transcription in all the LPS-treated animals. But the magnitude of fold change was less in the deermice, ranging from 6-25% of what it was in the LPS group of mice (Table 14).

The up-regulation of these ISGs was evidence of an interferon effect, but transcripts for interferon-1 beta (Ifnb) or -alpha (Ifna) themselves were scarcely detectable in deermice or mice in the blood under either condition. We then considered pattern recognition receptors (PRR) that might be part of a signaling pathway leading to ISG expression. Among the DEGs from the genome-wide analyses were four cytoplasmic PRRs: (1) Rigi (formerly Ddx58), which encodes the RNA helicase retinoic acid-inducible I (RIG-I); (2) Ifih1, which encodes interferon induced with helicase C domain 1, also known as MDA5 and a RIG-I-like receptor; (3) Dhx58, also known LGP2 and another RIG-I-like receptor; and (4) cGAS (cyclic GMP-AMP synthase), which is part of the cGAS-STING sensing pathway.

All four cytoplasmic PRRs were upregulated in the blood of LPS-treated mice and deermice (Table 14). But, again, for each of them the magnitude of fold change was less by 50-90% in treated *P. leucopus* than in *M. musculus*. The coefficients of determination for the 6 ISGs and the 4 PRRs are provided in figure 44. For most of the pairs there was evidence of covariation across all 40 animals. When the correlation was low across all the data, e.g. between the ISG Mx2 and the PRR Rigi or the ISGs Mx2 and Gbp4, it was high within a species.

These findings were evidence that pathways in *P. leucopus* for PRR signaling and ISG expression functioned similarly to those in *M. musculus* but differed under these experimental conditions in magnitude of the changes, being more moderate in the deermice.

	Cgas	Dhx58	Ifih1	Rigi	Gbp4	Irf7	Isg15	Mx2	Oas1
Cgas	1.00	0.70	0.78	0.40	0.46	0.63	0.39	0.77	0.61
Dhx58		1.00	0.86	0.80	0.76	0.79	0.61	0.58	0.62
Ifih1			1.00	0.54	0.59	0.79	0.61	0.77	0.75
Rigi				1.00	0.83	0.52	0.43	0.20	0.26
Gbp4					1.00	0.52	0.43	0.24	0.24
Irf7						1.00	0.81	0.71	0.77
Isg15							1.00	0.48	0.59
Mx2								1.00	0.87
Oas1									1.00
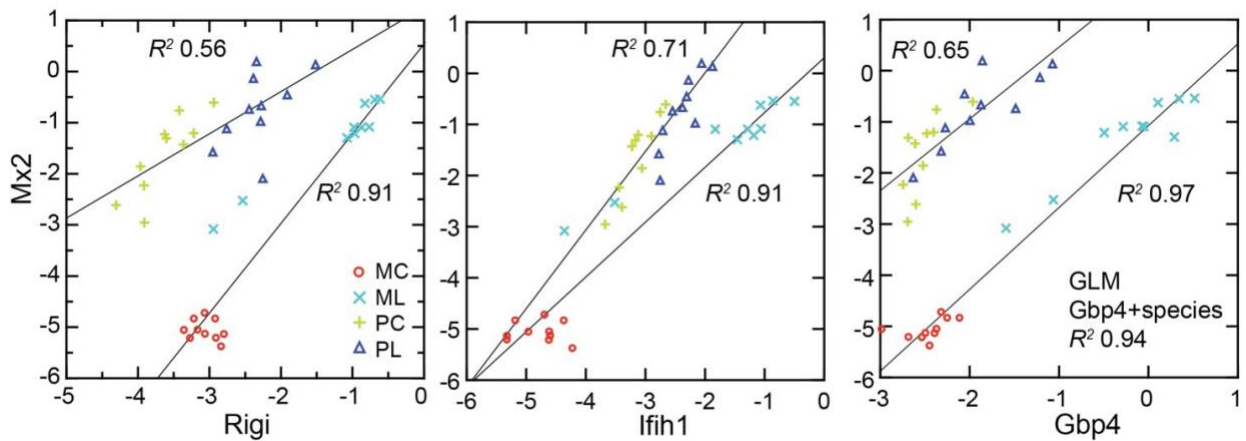
Figure 44 Co-variation between transcripts for selected PRRs and ISGs in the blood of P. leucopus or M. musculus with or without LPS treatment

Co-variation between transcripts for selected PRRs and ISGs in the blood of *P. leucopus* or *M. musculus* (M) with (L) or without (C) LPS treatment. Top panel: matrix of coefficients of determination (R2) for combined *P. leucopus* and *M. musculus* data. PRRs are indicated by yellow fill and ISGs by blue fill on horizontal and vertical axes. Shades of green of the matrix cells correspond to R2 values, where cells with values less than 0.30 have white fill and those of 0.90-1.00 have deepest green fill. Bottom panels: scatter plots of log-transformed normalized Mx2 transcripts on Rigi (left), Ifih1 (center), and Gbp4 (right). The linear regression curves are for each species. For the right-lower graph the result from the General Linear Model (GLM) estimate is also given.

Endogenous retroviruses in deermice, mice, and rats after LPS exposure

The six ISGs are nonexclusive consequences of activity of type I interferons. What we could document was the association of transcription of the gene for the cytoplasmic PPRs, including RIG-I, and the ISGs in both species, as well as the distinction between deermice in the magnitude of the responses of both PRRs and ISGs. These findings led us to ask could be a pathogen-associated molecular pattern (PAMP) for signaling pathways leading to expression of type 1 interferons.

One of these is endogenous retroviruses (ERV). The activity of these diverse, abundant, and pervasive elements have been recognized as one of the drivers of innate immune responses to a microbe (345), (346), (347). Our attention was drawn to ERVs by finding in the genome-wide RNA-seq of LPS-treated and control rats. Two of the three highest scoring DEGs by FDR p value and fold-change were a gag-pol polyprotein of a leukemia virus with 131x fold-change from controls and a mouse leukemia virus (MLV) envelope (Env) protein with 62x fold-change.

We returned to the mouse and deermouse data. There were four MLV or other ERV Env proteins among the 1266 genome-wide RNA-seq DEGs for *M. musculus*. But there was no ERV Env protein identified as such among the 1154 DEGs identified for *P. leucopus*. One possible explanation for the difference was an incomplete annotation of the *P. leucopus* genome. We took three approaches to rectify this. The first was to examine the DEGs for *P. leucopus* that encoded a polypeptide ≥ 200 amino acids and was annotated for the genome as "uncharacterized". A search with these candidates of both the virus and rodent proteins databases identified two that were homologous with gag-pol polyproteins of ERVs, mainly leukemia viruses, of mammals.

For a second approach we carried out a de novo transcript assembly of mRNA reads from blood of LPS-treated and control *P. leucopus* and used the resultant contigs as the reference set for RNA-seq analysis. This identified two contigs that were measurably transcribed in the blood, differentially expressed between conditions, and were homologous to ERV sequences. One was revealed to be an Env protein that was identical to a *P. leucopus* protein annotated as "MLV-related proviral Env protein" (XP_037065362). The second was a gag-pol protein. The latter was near-identical to the gag-pol protein identified by the first approach.

The third approach was to scan the *P. leucopus* genome for nonredundant sequences, defined as <95% identity, that were homologous with ERV gag-pol sequences, which are not typically annotated because of masking for repetitive sequences. This analysis yielded 615 unique sequences. These

were used in turn as a reference set for RNA-seq. There were 4 sequences that met the criterion of FDR p value <0.01. Three were transcribed at 5-to 40-fold higher levels in LPS-treated deermice than in controls. But all three, as well as the fourth, a down-regulated DEG, were ERV relics with truncations, frame shifts, and in-frame stop codons. These were assessed as non-coding RNAs and not further pursued in this study.

To represent *P. leucopus* in a targeted RNA-seq comparison with mice and rats we settled on the Env protein and gag-pol coding sequences identified present in the blood mRNA and as DEGs. Representing *M. musculus* were highest ranked MLV Env and gag-pol protein sequences among the DEGs. For rats we chose Env and gag-pol proteins that were second and third ranked DEGs identified in the genome-wide RNA-seq. Because of length differences for the coding sequences, the unit used for cross-species analysis was reads per kilobase before normalization for Ptprc transcription.

The left panel of figure 45 shows the striking transcriptional fold-change in LPS-treated rats of both Env and gag-pol sequences over controls. Of lesser magnitude but no less significant was the fold-change observed *M. musculus* for both Env and gag-pol sequences. In both mice and rats Env and gag-pol read values were highly correlated across conditions. In contrast, in *P. leucopus* the magnitudes of fold-change upwards for gag-pol was less than in mice or rats, and transcription of the Env protein sequence was actually lower in LPS-treated animals than in controls. While there was a tight association between Env protein and the PRR Rigi transcription in the *M. musculus*, this was not observed in *P. leucopus*. Rigi transcription was moderately higher at the time that Env protein's transcription was lower in the LPS group.
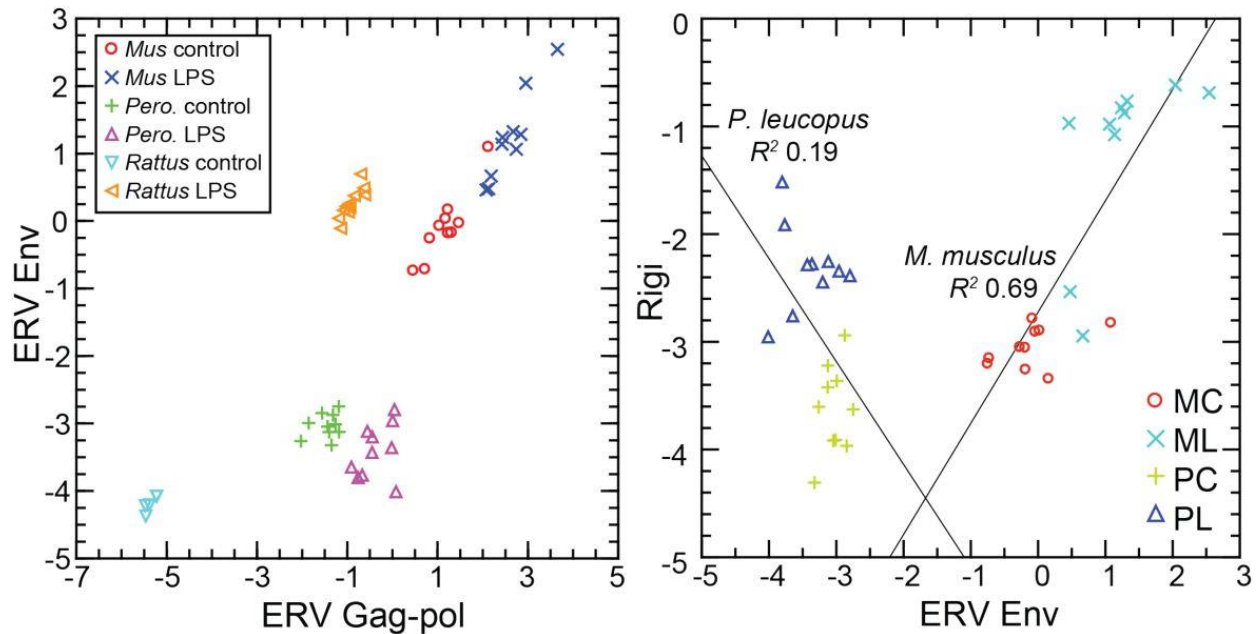
Figure 45 Scatter plots of endogenous retrovirus
Scatter plots of endogenous retrovirus (ERV) Env and Gag-pol protein gene transcription (left) and association of ERV Env with Rigi transcription (right) in the blood of *P. leucopus* (*Pero*; P), *M. musculus* (*Mus*; M), or *R. norvegicus* (*Rattus*) with (L) or without (control; C) treatment with LPS. In right panel the linear regression curve and coefficients of determination (R^2) for *P. leucopus* and *M. musculus* are shown.

Borrelia hermsii infection of *P. leucopus*

The phenomena reported so far were consequences of exposures to a particular PAMP--bacterial lipopolysaccharide with its hallmark lipid A moiety--recognized by a particular PRR, TLR4. While the focus was primarily on events downstream from that initial signaling, we asked in a concluding study whether the profile observed in *P. leucopus* applied in circumstances when the PAMP or PAMPs did not include LPS. This question is germane, given *P. leucopus'* role as a natural host for B. burgdorferi. This organism and other members of the spirochete family *Borreliaceae* do not have LPS (348), (279), but they have abundant lipoproteins, which are agonists for TLR2 in a heterodimer with TLR1 (349). B. burgdorferi is transiently blood-borne at low densities in *P. leucopus,* but in its life cycle B. burgdorferi is mainly tissue-associated in vertebrate hosts (310). We previously observed that the blood of B. burgdorferi-infected *P. leucopus* manifested few DEGs in comparison to skin (52). More

145

comparable to the LPS experimental model is infection of *P. leucopus* with a relapsing fever Borrelia species, which commonly achieve high densities in the blood. *P. leucopus* is a reservoir for Borrelia miyamotoi, which causes hard tick-borne relapsing fever (6), and the related P. maniculatus is a natural host for the soft tick-borne relapsing fever agent B. hermsii (278).

Accordingly, we used blood RNA-seq reads, which were taken from a prior study of B. hermsii infection of *P. leucopus,* described in Chapter 3*,* for targeted analysis with the same reference set employed for the LPS analyses (Table 18). The blood samples were taken from infected and uninfected animals on day 5, when bacteremia was at its peak, as documented by microscopy of the blood, qPCR of the spleen, and transcripts of a B. hermsii plasmid in the RNA extracts of the blood. As expected for B. hermsii infection (95), the spleen was enlarged in infected animals.

Similarities in the profiles for the LPS-treated and B. hermsii-infected deermice were as follows: (1) low levels of transcription of Nos2 and Ifng that contrasted with the high levels for Arg1 and Il1b in the same animals, (2) maintenance of Akt1/Akt2 > 1.0 under both conditions, (3) reduction of the Cd14/Cd16 ratio, (4) decreased transcription of Cd69, and (5) stable, low transcription of ERV Env and gag-pol loci with only marginal increases in transcription of ISGs and RIG-I-like receptors. Other equivalences under the two experimental conditions included increases in expression of genes for superoxide dismutase 2, low-affinity Fc gamma receptors, and secretory leukocyte protease inhibitor. Thus, the responses that distinguish deermice are not confined to the singular case of LPS as the elicitor.

Table 18 Targeted RNA-seq of Peromyscus leucopus with and without Borrelia hermsii infection.

Variable	Uninfected (n=3) mean (95% CI)*	Infected (n=4) mean (95% CI)*	Fold change	FDR *p* value
B. hermsii qPCR spleen	.	13615 (1882-98476)	.	.
B. hermsii reads blood	.	3487 (743-16362)	.	.
% spleen/body mass	0.15 (0.12-0.19)	0.36 (0.26-0.51)	2.4	1E-02
Targeted RNA-seq				
Akt1	184 (135-253)	347 (191-630)	1.88	3E-01
Akt2	92.9 (72.4-119)	141 (89.4-222)	1.52	3E-01
Arg1	247 (96.9-630)	1375 (848-2230)	5.57	4E-02
Cd14	308 (119-799)	598 (357-1002)	1.94	3E-01
Cd177	2.11 (0.71-6.27)	35.9 (11.9-108)	17.0	4E-02
Cd3	86.9 (77.7-97.2)	56.2 (16.4-192)	0.65	6E-01
Cd69	133 (114-155)	65 (33.6-128)	0.49	2E-02
Cgas	8.87 (5.40-14.6)	16.8 (8.99-31.4)	1.89	3E-01
Cxcl10	0.33 (0.14-0.76)	4.85 (2.74-8.59)	14.7	2E-02
Dhx58	22.0 (9.26-52.4)	37.0 (13.3-103)	1.68	6E-01
ERV Env	26.8 (22.4-32.1)	29.2 (14.8-57.8)	1.09	9E-01
ERV gag-pol	1853 (1351-2542)	1578 (820-3037)	0.85	8E-01
Fcgr2a	44.8 (24.0-83.6)	715 (303-1689)	16.0	2E-02
Fcgr2b	47.3 (30.2-74.3)	578 (167-2000)	12.2	5E-02
Fcgr3 (Cd16)	30.6 (18.6-50.4)	392 (171-897)	12.8	2E-02
Gapdh	1985 (1142-3447)	5366 (2383-12081)	2.70	2E-01
Gbp4	126 (82.3-193)	289 (130-644)	2.30	3E-01
Ibsp	0.81 (0.56-1.16)	358 (95-1350)	444	1E-02
Ifih1	48.5 (31.9-73.7)	49.1 (26.8-89.7)	1.01	1E+00
Ifit1	223 (107-465)	604 (303-1203)	2.71	2E-01
Ifng	0.58 (0.09-3.92)	2.28 (0.86-6.06)	3.94	3E-01
Il10	0.25 (0.07-0.87)	1.49 (0.31-7.28)	5.94	3E-01
Il12	0.43 (0.23-0.81)	1.13 (0.45-2.88)	2.63	3E-01
Il1b	477 (174-1308)	2828 (1325-6034)	5.93	7E-02
Il6	0.51 (0.18-1.39)	1.80 (0.61-5.28)	3.55	3E-01
Irf7	93.3 (13.4-65)	626 (196-1998)	6.71	2E-01
Isg15	302 (30.1-3030)	1922 (623-5934)	6.36	3E-01
Itgam (Cd11b)	72.2 (44.5-117)	322 (211-492)	4.45	2E-02
Mmp8	7.1 (2.74-18.6)	537 (148-1952)	75.2	2E-02
Mx2	152 (48.6-476)	167 (48.0-582)	1.10	9E-01
Nos2	0.16 (0.08-0.30)	0.32 (0.13-0.80)	2.01	4E-01
Oas1	51.3 (6.75-390)	159 (39.2-643)	3.10	5E-01
Rigi (Ddx58)	38.7 (18.9-79.2)	55.7 (33.9-91.5)	1.44	5E-01
S100a9	1739 (657-4596)	18430 (6546-51883)	10.6	5E-02
Saa3	0.49 (0.13-1.87)	212 (25.9-1733)	431	2E-02
Slpi	0.41 (0.18-0.95)	166 (49.0-566)	401	1E-02
Sod2	104 (51.4-211)	2011 (804-5028)	19.3	2E-02
Tlr1	21.4 (15.0-30.4)	111 (53.9-230)	5.21	4E-02
Tlr2	83.2 (59.7-116)	371 (234-587)	4.46	2E-02
Tlr4	44.8 (26.1-77.0)	256 (138-474)	5.71	3E-02
Tnf	2.86 (1.64-4.96)	78.8 (38.8-160)	27.6	1E-02
Ratios				
Akt1/Akt2	2.0 (1.7-2.3)	2.5 (2.1-2.9)	1.25	1E-01
Cd14/Fcgr3	11.3 (4.8-17.9)	1.7 (0.72-2.8)	0.15	2E-02
IL12/IL10	2.3 (0.00-5.0)	2.8 (0.0-7.6)	1.21	9E-01
Nos2/Arg1	0.001 (0.0-0.002)	0.0001 (0.0-0.0004)	0.28	2E-01

* For targeted RNA-seq it is mean unique reads for given gene normalized for reads for Ptprc (Cd45) gene for a sample. The 95% confidence intervals (CI) are asymmetric. Actual [gene]/Ptprc ratios are X 10^-3.

Discussion

Study limitations

The approach was forward and unbiased, looking for differences between species broadly across their transcriptomes. The findings lead to hypotheses, but reverse genetics in service of that testing was not applied here. In selective cases we could point to supporting evidence in the literature on *M. musculus* and the phenotypes of relevant gene knockouts, but there are no such resources for Peromyscus as yet. The resource constraint also applies to the availability of antibodies for use with Peromyscus for immunoassays for specific proteins, e.g. interferon-gamma, in serum, or for cell markers, e.g. CD69, for flow cytometry of white blood cells.

While a strength of the study was use of an outbred population of *M. musculus* to approximate the genetic diversity of the *P. leucopus* in the study, this meant that some genes of potential relevance might have gone undetected, i.e. from type II error. The variances for a sample of genetically diverse outbred animals, like the LL stock of *P. leucopus* (5), (5), would be expected to be greater than for the same sized sample of inbred animals. For some traits, especially ones that are complex or under balancing selection, even sample sizes of 10 in each group may not have provided sufficient power for discrimination between deermice and mice. For the same reason differences between sexes of a species in their responses might have been undetected. The interpretations applied to mixed-sex groups of deermice and mice. Expression strongly associated with female or male sex could have yielded an average fold change for the whole group that fell below the screen's threshold.

The parameters for the experiment of LPS dose, the route, and duration of experiment each might have had different values under another design. Those particular choices were based on past studies of deermice and mice as described in Chapter 3 and Langeroudi et al, (260). In another experiment we found that with doses twice or half those given the deermice the responses by rats to the different doses were indistinguishable by hematology, cytokine assays, and RNA-seq. Thus, there seems to

be some latitude in the dose and still achieving replication. We obtained similar results for *P. leucopus* when we looked at a replicate of the experiment with the same conditions as in Chapter 3 when the dose was lower and duration lengthened to 12 h (this study). The analysis here of the B. hermsii infection experiment also indicated that the phenomenon observed in *P. leucopus* was not limited to a TLR4 agonist.

While the rodents in these experiments were housed in the same facility and ate the same diet, we cannot exclude inherent differences in gastrointestinal microbiota between species and individual outbred animals as co-variables for the experimental outcomes. We reported differences between the LL stock *P. leucopus* and BALB/c *M. musculus* of the same age and diet in their microbiomes by metagenomic analysis and microbiologic means in Chapter 2. This included a commensal Tritrichomonas sp. in *P. leucopus* but not in the *M. musculus* in the study. The presence of these protozoa affects innate and adaptive immune responses in the gastrointestinal tract (350), (179), but it is not clear whether there are systemic consequences of colonization by this flagellate.

LPS, ERVs, and interferons

The results confirm previous reports of heightened transcription of ERV sequences in mice or mouse cells after exposure to LPS (351), (352), (353). Here we add the example of the rat. The LPS was administered in solution and not by means of membrane vesicles. The sensing PRR presumably was surface-displayed, membrane-anchored TLR4 (344). It follows that a second, indirect of LPS on the mouse is through its provocation of increased ERV transcription intracellularly. ERV-origin RNA, cDNA and/or protein would then be recognized by a cytoplasmic PRR. RIG-I was one associated with ERV transcription in this study. Kong et al. reported that LPS stimulated expression of Rigi in a mouse macrophage but did not investigate ERVs for an intermediary function in this phenomenon (354). As was demonstrated for LINE type retrotransposons in human fibroblasts, intracellular PRR signaling can trigger a type 1 interferon response (355). The combination of these two signaling events, i.e. one

through surface TLR4 by LPS itself and another through intracellular PPR(s) by to-be-defined ERV products, manifested in mice and rats as a response profile that had features of both a response to a virus with type 1 interferon and ISGs and a response to a bacterial PAMP like LPS with acute phase reactants such as calprotectin and serum amyloid.

This or a similar phenomenon has been observed under other circumstances. In humans there was heightened transcription of retrotransposons in patients with septic shock (356), as well as in peripheral blood mononuclear cells from human subjects experimentally injected with LPS (357). Bacteria like Staphylococcus epidermidis that express TLR2 agonists, such as lipoteichoic acid, promoted expression of ERVs, which in turn modulated host immune responses (346). A synthetic analog of a B. burgdorferi lipoprotein activated human monocytic cells and promoted replication of the latent HIV virus in cells that were persistently infected (358).

P. leucopus does not fit well with this model. Instead of the prominent interferon-gamma response observed in mice and rats, there were prominent responses of interleukin-1 beta and genes associated with neutrophil activation. Instead of the much-heightened expression of ISGs, like Mx2 and Isg15, in mice treated with LPS, the deermice under the same condition had a more subdued ISG transcription profile. Instead of increased expression of ERV Env protein sequences in blood of mice and rats treated with LPS, there was decreased transcription of the homologous MLV Env gene in like-treated *P. leucopus*.

This suppression in the deermice may be attributable to defensive adaptations of Peromyscus to repeated invasions of endogenous retroviruses, as Gozashti et al. has proposed for P. maniculatus (359). This includes expanding the repertoire of silencing mechanisms, such as Kruppel-associated box (KRAB) domain-containing zinc finger proteins (360). Like P. maniculatus, *P. leucopus* has an abundance of Long Terminal Repeat retrotransposons, several named for their endogenous retrovirus heritages (5). Our initial analysis of the *P. leucopus* genome reported a depletion of KRAB

domains compared to Muridae (5). But a subsequent annotation round identified several genes for KRAB domain zinc finger proteins in *P. leucopus,* including Zfp809 (XP_006982432), which initiates ERV silencing (361), and Zfp997 (XP_037067826), which suppresses ERV expression (362). Another possible adaptation in *P. leucopus* is the higher baseline expression of some ISGs as noted here (figure 44).

Reducing differences between *P. leucopus* and murids *M. musculus* and R. norvegicus to a single attribute, such as the documented inactivation of the Fcgr1 gene in *P. leucopus* (348), may be fruitless. But the feature that may best distinguish the deermouse from the mouse and rat is its predominantly anti-inflammatory quality. This characteristic likely has a complex, polygenic basis, with environmental (including microbiota) and epigenetic influences. An individual's placement is on a spectrum or, more likely, a landscape rather than in one or another binary or Mendelian category.

One argument against a purely anti-inflammatory characterization is the greater neutrophil numbers and activity in *P. leucopus* compared to *M. musculus* in the LPS experiment. The neutrophil activation, migration, and phagocytosis would be appropriate early defenses against a pyogenic pathogen. But if not contained, they bring local and systemic risks for the host. This damage would not likely be from nitric oxide and reactive nitrogen species, given the minimal Nos2 transcription. But deermice showed heightened expression of proteases, such as Mmp8, enzymes for reactive oxygen species, such as NADPH oxidase 1 (Nox1), and facilitators of neutrophil extracellular traps, such as PAD4 (Padi4) (Table 14). We had previously identified possible mitigators, such as the protease inhibitor Slpi and superoxide dismutase 2 presented in Chapter 3. These findings were replicated here. The topic of neutrophil activation and these and other possible counters is considered in more detail elsewhere.

An anti-inflammatory disposition but at what cost?

An assignment of infection tolerance to a host and pathogen pairing assumes sufficient immunity against the microbe to keep it in check if elimination fails. *P. leucopus* and P. *maniculatus*, are in this sense "immunocompetent" with respect to the microbes they host and with which they may share a long history (363). Yet, has this balance of resistance and tolerance for certain host-associated microbes been achieved in a trade-off that entails vulnerabilities to other types of agents?

The selection of LPS as the experimental model was meant to cover this contingency, at least for the common denominator of acute inflammation many types of infections elicit. But LPS studies revealed potential weaknesses that some pathogens might exploit. One of these is the low expression of inducible nitric oxide. Although Nos2 gene knockouts in *M. musculus* had lower LPS-induced mortality than their wild-type counterparts, the mutants were more susceptible to the protozoan Leishmania major and the facultative intracellular bacterium Listeria monocytogenes (298), (259). While there are no known studies of either of these pathogens in *P. leucopus* , the related species P. yucatanicus is the main reservoir for Leishmania mexicana in Mexico (17). Compared with *M. musculus*, which suffer a high fatality rate from experimental infections with L. mexicana, P. yucatanicus infections are commonly asymptomatic (364).

Given the restrained interferon and ISG response shown by *P. leucopus,* another plausible vulnerability would be viral infections. But other studies suggest that neither RNA nor DNA viruses pose an inordinately high risk for Peromyscus. Both tolerance of and resistance to the tickborne encephalitis flavivirus Powassan virus by *P. leucopus* were demonstrated in an experimental model in which mice, by contrast, were severely affected (300). P. maniculatus has been successfully infected with the SARS-CoV-2 virus by the respiratory route, but the infected animals displayed only mild pathology, manifested little if any disability, and recovered within a few days (108), (109). Among natural populations and in the laboratory P. maniculatus is noted for its tolerance of hantavirus,

which commonly is fatal for infected humans (240), (365). P. maniculatus was permissive of infection with monkeypox virus, but the infection was mild and transient (366).

A distinguishing *P. leucopus* characteristic, which was not expressly examined here, is its aforementioned 2-3 fold greater life span than that of *M. musculus*. While deermice may not be in the same longevity league as the naked mole-rat (*Heterocephalus glaber*), which can live for over 30 years (367), some features of naked mole-rat immunology are intriguingly similar to what we have observed for *P. leucopus*. These include macrophages and blood myeloid cells with low to absent transcription of Nos2 or production of nitric oxide in response to LPS, even in the presence of added inteferon-gamma (368). Like *P. leucopus* and in distinction to *M. musculus*, naked mole-rats showed an increase in the proportion of neutrophils in the blood 4 hours after intraperitoneal injection of LPS (369). In another comparative study, the hematopoietic stem and progenitor cells of these rodents had a lower type 1 interferon response than mice to a TLR3 agonist (370).

In summary, if there is a vulnerability that Peromyscus accepts in return for relief from inflammation (and perhaps a longer life), it has not been identified yet. However, potential threats and stressors are many, and the number assessed either in the field or laboratory has been limited to date.

Implications for Lyme disease and other zoonoses

Our studies of *P. leucopus* began with a natural population and documented a >80% prevalence of infection and high incidence of re-infections by B. burgdorferi in the area's white-footed deermouse, the most abundant mammal there (15). This was a Lyme disease endemic area (371), where residents frequently presented for medical care for a variety of clinical manifestations, from mild to serious, of B. burgdorferi infection (372). Subclinical infections in humans occur, but most of those who become infected have a definable illness (373). The localized or systemic presence of the microbe is a necessary condition for Lyme disease, but the majority of the symptoms and signs are attributable to inflammation elicited by the organism's presence and not from virulence properties

per se or the hijacking of host cells (374). Since humans are transmission dead-ends for B.

burgdorferi and many other zoonotic agents in their life cycles, it is not surprising that human

infections are generally more debilitating if not fatal than what adapted natural hosts experience.

It is in the space between the asymptomatic natural host and symptomatic inadvertent host where

there may be insights with basic and translational application. With this goal, we consider the ways

the results inform studies of the pathogenesis of Lyme disease, where "disease" includes lingering

disorders akin to "long Covid" (375), and where "pathogenesis" includes both microbial and host

contributions. Plausibly-germane deermouse-mouse differences identified in our studies are

summarized in figure 46. Two are highlighted here.

	Mus musculus	Peromyscus leucopus
Interferon-gamma	+++	+
Interleukin-1 beta	+	+++
NO synthase 2	+++	+
Arginase 1	+	+++
Akt1/Akt2	< 1.0	> 1.0
Cd14/Cd16	> 1.0	< 1.0
Cd69	++	±
ISGs	+++	++
RIG-I, MDA5	+++	+
Env-ERV	++	-

Figure 46 Summary of distinguishing features of transcriptional responses in the blood between P. leucopus and M. musculus 4 h after treatment with LPS
Summary of distinguishing features of transcriptional responses in the blood between *P. leucopus* and *M. musculus* 4 h after treatment with LPS. There is semi-quantitative representation of relative transcription of selected coding sequences or ratios of transcription for selected pairs of genes in the blood.

The first is macrophage polarization (334). By the criteria summarized above, the response to LPS by *P. leucopus* is consistent with the alternatively-activated or M2 type, rather than the expected classical or M1 type. But it was not only LPS-treated deermice that had this attribute, the blood of untreated animals also displayed M2 type polarization features. This included a comparatively high Arg1 expression level and a Akt1/Akt2 ratio of more than 1 at baseline. This suggests that studies of other mammals, including humans, need not administer LPS or other TLR agonist to assess disposition toward M1 or M2-type polarization. This reading could serve as a prognostic indicator of the inflammatory response to infection with *B. burgdorferi* or other pathogen and the long-term outcome.

The second difference we highlight is the activation of ERV transcription that was prominent in the LPS-treated mice and rats but not in similarly-treated deermice. A paradoxical enlistment of antiviral defenses, including type 1 and type 2 interferons, for an infection with an extracellular bacterium, like *B. burgdorferi*, may bring about more harm than benefit, especially if the resultant inflammation persists after antibiotic therapy. There are various ways to assess ERV activation in the blood, including assays for RNA, protein, and reverse transcriptase activity. A xenotropic MLV-related retrovirus has been discounted as a cause of chronic fatigue syndrome (376). However, production of whole virions need not occur for there to be PRR signaling in response to cytoplasmic Env protein, single stranded RNA, or cDNA (377).

Materials and Methods

Animals

The study was carried out in accordance with the Guide for the Care and Use of Laboratory Animals: Eighth Edition of the National Academy of Sciences. The protocols AUP-18-020 and AUP-21-007 were approved by the Institutional Animal Care and Use Committee of the University of California Irvine.

Peromyscus leucopus, here also referred to as "deermice", were of the outbred LL stock, which 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 (246). The LL stock animals for this study were bred and raised at the vivarium of University of California Irvine, an AAALAC approved facility. Outbred Mus musculus breed CD-1, specifically Crl:CD1(ICR) IGS, and here also referred to as "mice", were obtained from Charles River. Fischer F344 strain inbred Rattus norvegicus, specifically F344/NHsd, and here also referred to as "rats", were obtained from Charles River.

For the combined *P. leucopus* -*M. musculus* experiment the 20 *P. leucopus* were of a mean (95% confidence interval) 158 (156-159) days of age and had a mean 21 (19-22) g body mass. The 20 *M. musculus* were all 149 days of age and had a mean body mass of 47 (43-50) g. The ratio of average male to average female body mass was 1.04 for *P. leucopus* and 1.03 for *M. musculus*. The 6 female *P. leucopus* for the 12 h duration experiment were of a mean 401 (266-535) days of age and mean body mass of 20 (17-23) g. The 16 adult 10-12 week old female R. norvegicus had a mean 139 (137-141) g body mass. The 7 male *P. leucopus* for the infection study were of a mean 107 (80-134) days and mean body mass of 21 (18-24) g.

Animals were housed in Techniplast-ventilated cages in vivarium rooms with a 16 h-8 h light-dark cycle, an ambient temperature of 22 °C, and on ad libitum water and a diet of 2020X Teklad global soy protein-free extruded rodent chow with 6% fat content (Envigo, Placentia, CA).

For injections the rodents were anesthetized with inhaled isoflurane. The rodents were euthanized by carbon dioxide overdose and intracardiac exsanguination at the termination of the experiment. No animals died or became moribund before the 4 hour or 12 h termination time points in the LPS experiments or before the 5 d termination point of infection study.

LPS experiments

For the *P. leucopus* and *M. musculus* combined experiment, sample sizes replicated the specifications of the previous study, in which there were 20 *P. leucopus* and 20 *M. musculus*, equally divided between females and males and equally allotted between conditions described in Chapter 3. The treatments were administered in the morning of a single day. At 15 min intervals and alternating between species, sex, and treatments, animals were intraperitoneally (ip) injected 50 µl volumes of either ion-exchange chromatography-purified Escherichia coli O111:B4 LPS (Sigma-Aldrich L3024) in a dose of 10 µg per g body mass or the diluent alone: sterile-filtered, endotoxin-tested, 0.9% sodium chloride (Sigma-Aldrich). The animals were visually monitored in separate cages continuously for the duration of the experiment. We recorded whether there was reduced activity by criterion of huddling with little or movement for > 5 min, ruffled fur or piloerection), or rapid respiration rate or tachypnea. At 4.0 h time after injection animals were euthanized as described above, and sterile dissection was carried out immediately.

Lower dose and longer duration experiment. In an experiment with 6 *P. leucopus,* the animals were administered the same single dose of LPS but at 1.0 µg/g and the same control solution. The animals were euthanized 12 h after the injection the following day.

Rat LPS experiment. The same experimental design was used for the rats as for the combined deermice-mice experiment, with the exception that the formulation of the E. coli O111:B4 LPS was "cell culture grade" (Sigma-Aldrich L4391), and the groups were sterile saline alone (n = 5), 5 µg LPS per g body mass (n = 6), or 20 µg LPS per g (n = 5).

Experimental infection

The infection of a group of *P. leucopus* LL stock with the relapsing fever agent B. hermsii and the processing of blood and tissues for RNA extraction 5 days into the infection were described previously in Chapter 3. In brief, animals were infected intraperitoneally on day 0 with either

phosphate-buffered saline alone or 103 cells of B. hermsii MTW, a strain that is infectious for Peromyscus species (278). Bacteremia was confirmed by microscopy on day 4, and the animals were euthanized on day 5. For that prior study the RNA-seq analysis was limited to the genome-wide transcript reference set. For the present study we used the original fastq format files for targeted RNA-seq as described below.

Hematology and plasma analyte assays

For the combined *P. leucopus -M. musculus* experiment automated complete blood counts with differentials were performed at Antech Diagnostics, Fountain Valley, CA on a Siemens ADVIA 2120i with Autoslide hematology instrument with manual review of blood smears by a veterinary pathologist. For the 12 h duration *P. leucopus* experiment hematologic parameters were analyzed on an ABCVet Hemalyzer automated cell counter instrument at U.C. Irvine. For the rat experiment complete blood counts with differentials were performed at the Comparative Pathology Laboratory of the University of California Davis. Multiplex bead-based cytokine protein assay of the plasma of the rats was performed at Charles River Laboratories using selected options of the Millipore MILLIPLEX MAP rat cytokine/chemokine panel.

RNA extraction of blood

After the chest cavity was exposed, cardiac puncture was performed through a 25 gauge needle into a sterile 1 ml polypropylene syringe. After the needle was removed, the blood was expelled into Becton-Dickinson K2E Microtainer Tubes, which contained potassium EDTA. Anticoagulated blood was split into a sample that was placed on ice for same-day delivery to the veterinary hematology laboratory and a sample intended for RNA extraction which was transferred to an Invitrogen RiboPure tube with DNA/RNA Later and this suspension was stored at −20 °C. RNA was isolated using the Invitrogen Mouse RiboPure-Blood RNA Isolation Kit]. RNA concentration was determined on a

NanoDrop microvolume spectrophotometer (ThermoFisher) and quality was assessed on an Agilent Bioanalyzer 2100.

RNA-seq of blood

The chosen sample sizes and coverage for the bulk RNA-seq were based on empirical data from the Chapter 3, which indicated that with 10 animals per group and a two-sided two sample t-test we could detect with a power of ≥ 0.80 and at a significance level of 0.05 a ≥1.5 fold difference in transcription between groups for a given gene. We also were guided by the simulations calculations of Hart et al. (378), which indicated for a biological coefficient of variation of 0.4 within a group, a minimum depth of coverage of ≥ 10, and a target of ≥ 2x fold change that a sample size of 7-8 was sufficient for 80% power and type I error of 5%. For the *P. leucopus* and *M. musculus* samples production of cDNA libraries was with the Illumina TruSeq Stranded mRNA kit. After normalization and multiplexing, the libraries were sequenced at the University of California Irvine's Genomic High Throughput Facility on a Illumina NovaSeq 6000 instrument with paired-end chemistry and 150 cycles to achieve ~100 million reads per sample for the combined *P. leucopus -M. musculus* experiment. The same method for producing cDNA libraries was used for the R. norvegicus RNA and the *P. leucopus* in the infection study, but these were sequenced on a Illumina HiSeq 4000 instrument with paired-end chemistry and 100 cycles. The quality of sequencing reads was analyzed using FastQC (Babraham Bioinformatics). The reads were trimmed of low-quality reads (Phred score of <15) and adapter sequences, and corrected for poor-quality bases using Trimmomatic (214).

For the combined species experiment the mean (95% CI) number of PE150 reads per animal after trimming for quality was 1.1 (1.0-1.2) x 108 for *P. leucopus* and 1.1 (1.0-1.2) x 108 for *M. musculus* (p = 0.91). For *P. leucopus* of this experiment a mean of 83% of the reads mapped to the genome transcript reference set of 54,466; mean coverages for all transcripts and for the mean 62% of reference transcripts with ≥ 1x coverage were 97x and 157x, respectively. For *M. musculus* of this

experiment a mean 91% of the reads mapped to the genome transcript reference set of 130,329; mean coverages for all transcripts and for the mean of 21% of reference transcripts with ≥ 1x coverage were 103x and 568x, respectively. For the lower dose-longer duration experiment with *P. leucopus* the mean number of PE150 reads was 2.5 (2.3-2.6) x 107. For the rat experiment the mean number of PE100 reads was 2.4 (2.2-2.5) x 107. For the B. hermsii infection experiment the mean number of PE100 reads was 4.9 (4.5-5.3) x 107.

Batched fastq files were subjected to analysis with CLC Genomics Workbench version 23 (Qiagen). Library size normalization was done by the TMM (trimmed mean of M values) method of Robinson and Oshlack (77). The reference genome transcript sets on GenBank were the following: GCF_004664715.2 for *P. leucopus* LL stock, GCF_000001635.27_GRCm39 for *M. musculus* C57Bl/6, and GCF_015227675.2_mRatBN7.2 for R. norvegicus. The settings for PE150 reads were as follows for both strands: length fraction of 0.35, similarity fraction of 0.9, and costs for mismatch, insertion, or deletion of 3. For PE100 reads the settings were the same except for length fraction of 0.4. Principal Component Analysis was carried with the "PCA for RNA-Seq" module of the suite of programs.

For the *P. leucopus* RNA-seq analysis there were 54,466 reference transcripts, of which 48,164 (88%) were mRNAs with protein coding sequences, and 6,302 were identified as non-coding RNAs (ncRNA). Of the 48,164 coding sequences, 40,247 (84%) had matching reads for at least one of the samples. The five most highly represented *P. leucopus* coding sequences among the matched transcripts of whole blood among treated and control animals were hemoglobin subunits alpha and beta, the calprotectin subunits S100A8 and S100A9, and ferritin heavy chain. For the *M. musculus* analysis there were available 130,329 reference transcripts: 92,486 (71%) mRNAs with protein coding sequences and 37,843 ncRNAs. Of the coding sequences, 59,239 (64%) were detectably transcribed in one or both groups by the same criterion. The five most highly represented coding sequences of mRNAs of identified genes for *M. musculus* were hemoglobin subunits alpha and beta,

aminolevulinic synthase 2, ferritin light polypeptide 1, and thymosin beta. For R. norvegicus there were 99,126 reference transcripts, of which 74,742 (75%) were mRNAs. The five most highly represented coding sequences of mRNAs of identified genes for R. norvegicus were hemoglobin subunits alpha and beta, beta-2 microglobulin, ferritin heavy chain, and S100a9.

Genome-wide differential gene expression

Differential expression between experimental conditions was assessed with an assumption of a negative binomial distribution for expression level and a separate Generalized Linear Model for each (379). Fold changes in TPM (transcripts per million) were log2-transformed. The False Discovery Rate (FDR) with corrected p value was estimated by the method of Benjamini and Hochberg (227). To assess the limit of detection for differentially expressed genes between 10 animals treated with LPS and 10 with saline alone, we took the data for 4650 reference transcripts for which the mean TPM across 20 *P. leucopus* was >10 and randomly permuted the data to achieve another 9 sets and calculated the fold-change of sub-groups of 10 and 10 with one random group serving as the proxy of the experimental treatment and other second as the control for each of the sets. The expectation was that mean fold-change of the 9 permuted sets and the 4650 reference sequences would be ~1. The result was a mean and median of 1.08 with a 99.9% asymmetric confidence interval for the mean of 0.81-1.49. This was an indication that the choices for sample sizes were realistic for achieving detection of ≥ 1.5x fold changes.

Gene Ontology term analysis

M. musculus was selected as the closest reference for the *P. leucopus* data. The analysis was implemented for data for differentially expressed genes meeting the criteria of a FDR p-value ≤ 0.01 and fold-change of ≥ 1.5. The analysis was implemented with the tools of Metascape (https://metascape.org) (380). Functional enrichment analysis was carried out first with the hypergeometric test and FDR p-value correction (227). Then pairwise similarities between any two

enriched terms were computed based on a Kappa-test score (381). Similarity matrices were then hierarchically clustered, and a 0.3 similarity threshold was applied to trim resultant trees into separate clusters. The lower the p-value, the less the likelihood the observed enrichment is due to randomness (382). The lowest p-value term represented each cluster shown in the horizontal bar graph. Besides the terms beginning with "GO" and referring to the Gene Ontology resource (http://geneontology.org) (2), others 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.

Targeted RNA-seq

RNA-seq of selected set of protein coding sequences (CDS), which are listed below, was carried out using CLC Genomics Workbench v. 23 (Qiagen). Paired-end reads were mapped with a length fraction of 0.35 for ~150 nt reads and 0.40 for ~100 nt reads, a similarity fraction of 0.9, and costs of 3 for mismatch, insertion, or deletion to the CDS of sets of corresponding orthologous mRNAs of *P. leucopus, M. musculus*, and R. norvegicus. Preliminary expression values were unique reads normalized for total reads across all the samples without adjustment for reference sequence length, as described in Chapter 3. Exceptions were the endogenous retrovirus coding sequences which differed in lengths between species. For within- and cross-species comparisons we initially normalized three different ways after quality filtering and removing vector and linker sequence: for total reads for the given sample, for unique reads for 12S ribosomal RNA for the mitochondria of nucleated cells in the blood, and for unique reads for the gene Ptprc, which encodes CD45, a marker for both granulocytes and mononuclear cells in the blood. This is described in more detail in Results. Following the recommendation of Hedges et al. we used the natural logarithm (ln) of ratios (383).

The target CDS were as follows: Acod1, Akt1, Akt2, Arg1, Bcl3, Camp, Ccl2, Ccl3, Ccl4, Cd14, Cd177, Cd3d, Cd4, Cd69, Cd8, Cfb, Cgas, Csf1, Csf1r, Csf2, Csf3, Csf3r, Cx3cr1, Cxcl1, Cxcl10, Cxcl2,

Cxcl3, Dhx58, Fcer2, Fcgr2a, Fcgr2b, Fcgr3 (CD16), Fgr, Fos, Fpr2, Gapdh, Gbp4, Glrx, Gzmb, Hif1a, Hk3, Hmox1, Ibsp, Icam, Ifih1, Ifit1, Ifng, Il10, Il12, Il18, Il1b, Il1rn, Il2ra, Il4ra, Il6, Il7r, Irf7, Isg15, Itgam, Jak1, Jak2, Jun, Lcn2, Lpo, Lrg, Lrrk2, Ltf, Mapk1, Mmp8, Mmp9, Mpo, MT-Co1, Mt2, Mtor, Mx2, Myc, Myd88, Ncf4, Nfkb1, Ngp, Nos2, Nox1, Nr3c1, Oas1, Olfm4, Padi4, Pbib, Pkm, Ptx, Ptprc (CD45), Retn, Rigi (Ddx58), S100a9, Saa3, Serpine1, Slc11a1 (Nramp), Slpi, Socs1, Socs3, Sod2, Stat1, Stat2, Stat4, Steap1, Sting, Tgfb, Thy1, Timp1, Tlr1, Tlr2, Tlr4, Tnf, Tnfrsf1a, and Tnfrsf9. The sources for these coding sequences were the reference genome transcript sets for *P. leucopus, M. musculus*, and R. norvegicus listed above. If there were two or more isoforms of the mRNAs and the amino acid sequences differed, the default selection for the coding sequence was the first listed isoform. The lengths of the orthologous pairs of *P. leucopus* and *M. musculus* coding sequences were either identical or within 2% of the other. Fcgr1, the gene for high-affinity Fc gamma receptor I or CD64, was not included in the comparison because in P. leucopus, it is an untranscribed pseudogene (7). For the targeted RNA-seq of blood of deermice infected with B. hermsii (Table 18), the reference sequence was the cp6.5 plasmid of B. hermsii (NZ_CP015335).

Quantitative PCR assays

Reverse transcriptase (RT)-qPCR assays and the corresponding primers for measurement of transcripts of genes for nitric oxide synthase 2 (Nos2) and glyceraldehyde 3-phosphate dehydrogenase (Gapdh) were those described previously in Chapter 3. These primers worked for *M. musculus* as well as *P. leucopus* using modified cycling conditions. For the arginase 1 transcript assays different primer sets were used for each species. The forward and reverse primer sets for the 352 bp Arg1 product for *P. leucopus* were 5'-TCCGCTGACAACCAACTCTG and 5'-GACAGGTGTGCCAGTAGATG, respectively. The corresponding primer pairs for a 348 bp Arg1 of *M. musculus* were 5'-TGTGAAGAACCCACGGTCTG and 5'-ACGTCTCGCAAGCCAATGTA. cDNA synthesis and qPCR were achieved with a Power Sybr Green RNA-to-Ct 1-Step Kit (Applied Biosystems) in 96

163

MicroAmp Fast Reaction Tubes using an Applied Biosystems StepOne Plus real-time PCR instrument. The initial steps for all assays were 48 °C for 30 min and 95 °C for 10 min. For Arg1 and Nos2 assays, this was followed by 40 cycles of a 2-step PCR of, first, 95°C for 15 s and then, second, annealing and extension at 60 °C for 1 min. The cycling conditions for Gapdh were 40 cycles of 95 °C for 15 s followed by 60 °C for 30 s. Quantitation of genome copies of B. hermsii in extracted DNA was carried out by probe-based qPCR as described (310).

Additional statistics

Means are presented with asymmetrical 95% confidence intervals (CI) to accommodate data that was not normally distributed. Parametric (t test) and non-parametric (Mann-Whitney) tests of significance were 2-tailed. Unless otherwise stated, the t test p value is given. Categorical variables were assessed by 2-tailed Fisher's exact test. FDR correction of p values for multiple testing was by the Benjamini-Hochberg method (227), as implemented in CLC Genomics Workbench (see above), or False Discovery Rate Online Calculator (https://tools.carbocation.com/FDR). Discriminant Analysis, linear regression, correlation, coefficient of determination, and General Linear Model 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.

CHAPTER FIVE: The Peromyscus leucopus - model of SARS-CoV-2 virus infection of lungs and brain

Abstract

The white-footed deermouse *Peromyscus leucopus* is a long-lived rodent and a key reservoir in North America for agents of several zoonoses including Lyme disease, babesiosis, anaplasmosis, and viral encephalitis. While persistently infected, this deermouse avoids apparent disability or diminished fitness. Its tolerance to infection with sometimes more than one pathogen makes *P. leucopus* comparable to bats. This study uses *P. leucopus*, LL colony stock, as a genetically diverse animal model for viral infection with SARS-CoV-2. We infected *P. leucopus* with SARS-CoV-2, collected plasma, lungs, and brain three and six days post-infection, and compared to control animals. *P. leucopus* mount an appropriate immune response against viral pathogens through production of neutralizing antibodies and genome-wide transcription of type I interferon stimulated genes in lungs compared to naïve animals. We report that viral RNA detection correlates with gene expression of type I interferon stimulated genes in response to viral infection in the brain. These results show that *P. leucopus* is a viable animal model for SARS-CoV-2, particularly in research of viral infection of the brain.

Introduction

Since the declaration of the COVID-19 pandemic until March 2024, there have been 774,699,366 reported cases and 7,033,430 deaths due to confirmed infection of SARS-CoV-2 (384), (385). The disease is characterized on a spectrum from asymptomatic (60%), mild (20%), and moderate/severe (10%) with hospitalization in one percent of cases (386). Symptoms vary from fever, cough with sputum production, dyspnea, myalgia, and fatigue, diarrhea and involve the nervous system (387), (388). The outbreak trajectory was not initially predictable, but then cases of exacerbated pneumonia resulting in acute respiratory distress syndrome were reported in combination with systemic

inflammatory response and multiorgan failure, which caused an increased number of fatalities (389), (390), (391), (392). Laboratory findings predictive of poor prognosis included elevated markers of dysregulated coagulation, leucopenia, with high neutrophil count and proinflammatory cytokine release in lungs (bronchial lavage specimen): increased interleukin-6 (IL-6), IL-8, and lower interferon alpha (IFN-α) - infamous cytokine storm (393), (394), (395), (396). About 5 percent of individuals infected with SARS-CoV-2 do not fully recover from acute disease and develop long COVID characterized by plasma markers suggestive of myeloid inflammation and complement activation (397). The science is not set on viral involvement in the brain, although many neurological symptoms are associated with COVID-19, including loss of smell, headaches, and memory problems (398). Now, new research adds to the evidence that inflammation in the brain might underlie these symptoms (399), (400).

Neuropathogenesis of COVID-19 in humans is associated with cognitive function decline and inflammation signature in postmortem frontal cortex samples, leading some groups to conclude that symptoms are independent of the virus, mainly since the viral receptor angiotensin-converting enzyme 2 (ACE2) expression is segregated in the brain – it is not expressed in the prefrontal cortex, but it is highly expressed in the olfactory bulb (401), (402), (403), (400). Both ACE2 docking protein for viral spike protein and transmembrane protease serine 2 (TMPRSS2) necessary for viral protein processing are expressed on blood vessels, and the blood-brain barrier is the proposed viral point of entry. Epithelial cells of the olfactory bulb express viral docking proteins and are suggested as another viral entry route along with infected immune cells (403), (404), (405). This data supports both viral invasion and inflammation in the brain, causing neurological symptoms, but in the case of long COVID with no resolution and prolonged inflammation. More studies using appropriate animal models could provide more information on brain involvement of SARS-CoV-2.

Animal models are crucial to understanding the host's viral pathogenesis, including neural involvement in the infection, as well as vaccine and drug development. Widely used animal models for COVID-19 are transgenic K18-hACE2 mice or mouse-adapted strains of SARS-CoV-2 (406). Ideally, an animal model would recapitulate mild to lethal forms of infection that resemble human disease; it would be a genetically diverse small mammal (407), (408), (409).

It was experimentally established that other Peromyscus species like *P. maniculatus* (later reclassified as possibly *Peromyscus sonoriensis*) infected with SARS-CoV-2 could infect naïve deermice through direct contact since the virus was detected in nasal, oral, and rectal swabs, and viral RNA was detectable in feces and urine (109). *A*nother species within the *Peromyscus* genus *(Peromyscus sonoriensis)* was successfully infected with SARS-CoV-2 and had detectable viral replication in the lungs, intestines, and brain (108). After two passages were attempted, a mutation was detected diverging from the original viral stock, demonstrating the potential for adaptation in the natural population (108). *P. maniculatus*, *P. sonoriensis*, *P. polionotus*, and *P. californicus* had detectable neutralizing antibodies, viral RNA, and viral particles in oral swabs and lung tissue. In contrast, *some P. californicus* developed severe clinical disease (107). Infection of *P. leucopus* with SARS-CoV-2 has not been attempted so far in laboratory conditions.

For this study, we used the white-footed deermouse (*Peromyscus leucopus*), the primary reservoir for several zoonoses in North America, notably Lyme disease, as an outbred, small animal model for COVID-19. Unlike mice, this rodent does not require a modified virus (a species-adapted strain of the virus, as is the case in mice) or transgenic ACE2 receptor for successful infection with SARS-CoV-2 (as in K18-hACE2 mice) (410), (407), (406). Another advantage is its relatively small size compared to Syrian hamsters, also used as animal models for COVID-19 (406), (411), (412), (413), (414). In addition, *P. leucopus* used in research are genetically diverse, while Syrian hamsters used in research, although not inbred, are genetically isogenic (415). Just like for hamsters, an annotated

genome for *P. leucopus* is available, and immunity to infection has been investigated in response to bacteria, and endotoxins were described in Chapters 3 and 4 (6), (5). *P. leucopus* has demonstrated remarkable tolerance during infection. Some of the tolerance hallmarks in *P. leucopus* were identified as: moderate reaction to bacteria or bacterial endotoxin; macrophage polarization showing alternative polarization including upregulated enzyme arginase (Arg1) utilizing substrate for nitric oxide synthetase 2 (Nos2); and reducing reactive oxide species utilizing superoxide dismutase 2 (Sod2). An additional mitigating factor in *P. leucopus* infection is its microbiome (Chapter 2). Although we report that *P. leucopus* is infection-tolerant, it responds appropriately to pathogens. Transcriptome analysis showed upregulated genes related to neutrophil activation gene ontology (GO) terms, and this was associated with the higher neutrophil to lymphocyte ratio cell counts in the deermice in response to LPS compared to mice described in Chapters 3 and 4, and Barbour (3). The damaging effect of neutrophil proteases on local tissue in infection is modulated by a 1,280-fold increase of secretory leukocyte peptidase inhibitor (Slpi) expressed in reaction to endotoxin described in Chapter 3. Our next question was, can *P. leucopus*, as a genetically diverse animal from a closed colony, get infected with SARS-CoV-2, and would it also show hallmarks of tolerance in this case?

Recent evidence suggests that *Peromyscus* species can be successfully infected with SARS-CoV-2. Furthermore, *P. leucopus* is a comparable animal model to the Syrian hamster, but unlike hamsters, *Peromyscus* are genetically diverse. Thus, *P. leucopus* is a more informative and practical alternative for COVID-19 research. Here, we infected *P. leucopus* with a high titer SARS-CoV-2 and analyzed the lung and brain transcriptional profile and the level of humoral response in the plasma in this genetically diverse and infection-tolerant animal model. The lungs and brains of each animal were analyzed using histopathology, RNAscope in situ hybridization, genome-wide RNAseq, and RT-qPCR

analysis for the virus. Blood plasma was used to test *P. leucopus*' antibodies specific to SARS-CoV-2 spike (S) and nucleocapsid (N) protein and virus neutralization.

Results

Experimental infection of P. leucopus with SARS-CoV-2

Genetically diverse *Peromyscus leucopus* from a closed colony was used as an animal model for infection with SARS-CoV-2. We performed two experiments and performed experimental procedures in the same way. This study aimed to describe the response of *P. leucopus* to SARS-CoV-2 and compare it to a group of control animals. Thirty *P. leucopus* were used over both experiments (figure 47). The age of animals varied with an average age of 613 days (219-998 days range). Since deermouse species live 6-8 years, two-year-old animals are not considered aged adults (10). Animals used in experiment 2 were older, with a mean age of 645 (468-735) days, than the animals used in experiment one, with a mean age of 561(518-603) days, (p = 0.3).

As a control for intranasal SARS-CoV-2 virus inoculum, we performed a mock infection with a tissue culture medium. The first experiment included animals (n=14) of both sexes. Out of those, n=8 were intranasally challenged with the virus. We intranasally inoculated (n=6) animals with VeroE6 growth medium as a control group. Animals were euthanized either 3 or 6 days post-inoculation in both groups. The second experiment was repeated 12 months later. This time, we used animals (n=16) of Animals from each group were euthanized either 3 or 6 days post-challenge. All of the animals and their characteristics, along with selected results of analyses, are described in Table 19.

Animals were observed daily for clinical signs of disease or sickness behavior. Thirty *P. leucopus* in both experiments receiving either virus or medium did not show overt signs of sickness over periods of observation of 3 days or 6 days. The exceptions were two animals in experiment 2: 25255, which by day 3 had ruffled fur, hunched posture, lethargy, and tachypnea, and 25165, which had periorbital edema that lasted for more than a day (Table 19). These animals were euthanized on day 3 due to

concerns that they were not capable of eating and drinking, although they were intended for a day 6

group. No clinical signs were observed in controls. One female *P. leucopus* (ID 25288) lost 14% of its

weight, while others lost or gained less than 5% (Table 19). No animals succumbed to infection

before euthanasia. The lungs of each animal were collected and analyzed for histopathology, the

presence of the virus using RNAscope in situ hybridization, and bulk RNA was sequenced. The RNA

derived from the brain was also sequenced and in addition analyzed for the viral RNA. Blood plasma

was used to detect *P. leucopus'* antibodies specific to SARS-CoV-2 S- and N- proteins and to conduct

a virus neutralization assay (figure 47).

16 deermice (8M, 8F)	14 deermice (7M, 7F)
SARS-CoV2 WA strain	DMEM media
Euthanasia on day 3 n=8 and day 6 n=8	Euthanasia on day 3 n=7 and day 6 n=7

Plasma ELISA N- and S- titer and neutralization assay. Lungs, brain
RNAseq and histopathology, RT-qPCR of viral copy numbers.
Experiment 1 included 14 animals (n=8 infected and n=6 controls
and experiment 2 had 16 animals (n=8 infected and n=8 controls)

Figure 47 Schematic representation of experimental design
Animals of both sexes and a wide range of ages received intranasal doses of SARS-CoV-2 or media
independent of weight, and then post-euthanasia were processed the same way.

Table 19 Characteristics of P. leucopus infection with SARS-CoV-2

Characteristics of *Peromyscus leucopus* treatment with SARS-CoV-2 and media

Animal ID	Treatment Intranasal	Euthanasia day	Sex	Age (d)	Weight loss (g)	Viral CN lungs log10 (95% CI)*	Viral CN brain log10 (95% CI)*	Viral CN feces log10 (95% CI)*	Anti CoV-2 N titer	Anti CoV-2 S titer	Neutralization assay ID50	Clinical disease	Lung histology score	Brain RNA-seq reads mapped to SARS
Experiment 1														
25160	Media	3	female	501	.	<2.94	<3.43	<2.1	381	<50	1		0	1
25169	Media	3	male	483	.	<2.94	<3.43	<2.1	100	<50	1		0	3
25124	Media	3	male	578	.	<2.94	<3.43	<2.1	113	<50	1		0	5
25167	Media	6	female	486	1.0	<2.94	<3.43	<2.1	120	598	1		0	4
25018	Media	6	female	627	-1.0	<2.94	<3.43	<2.1	<50	<50	1		0	3
25159	Media	6	male	499	0.0	<2.94	<3.43	<2.1	458	<50	1		0	3
25013	Sars-CoV-2	3	female	662	-2.0	6.5(6.3-6.6)	4.1(4.0-4.1)	6.2(6.2-6.2)	736	<50	1	1	1	38
25025	Sars-CoV-2	3	female	761	0.0	7.1(6.9-7.2)	<3.43	2.2(1.1-3.4)	156	<50	1	1	1	1
25117	Sars-CoV-2	3	male	583	0.0	6.8(6.6-6.9)	4.1(4.0-4.2)	5.6(5.6-5.6)	205	<50	1	1	2	48
25130	Sars-CoV-2	3	male	578	-1.0	7.7(7.5-7.8)	4.7(4.6-4.7)	3.4(3.4-3.4)	300	<50	1	1	2	334
25128	Sars-CoV-2	6	female	581	0.0	4.9(4.8-4.9)	4.4(4.3-4.4)	3.1(3.0-3.1)	643	38,632	459	1	1	219
25158	Sars-CoV-2	6	female	504	-1.0	3.1(2.5-3.7)	<3.43	1.7(0.9-2.5)	161	30,203	186	1	1	74
25157	Sars-CoV-2	6	female	504	1.0	5.1(4.8-5.4)	<3.43	4.1(4.1-4.1)	730	35,470	108	1	1	4
25161	Sars-CoV-2	6	male	509	0.0	5.6(4.9-6.3)	<3.43	2.3(2.2-2.4)	1,914	56,519	334	1	1	2
Experiment 2														
25225	Media	3	female	287	0.0	<2.94	<3.43	.	182	<50	1		0	-
25226	Media	3	female	223	1.0	<2.94	<3.43	.	474	96	1		0	5
25281	Media	3	male	644	0.0	<2.94	<3.43	.	160	<50	1		0	-
25152	Media	3	male	949	0.0	<2.94	<3.43	.	253	79	1		0	3
25219	Media	6	female	317	0.0	<2.94	<3.43	.	376	<50	1		0	4
25313	Media	6	female	373	1.0	<2.94	<3.43	.	477	<50	1		0	8
25287	Media	6	male	675	0.0	<2.94	<3.43	.	268	<50	1		0	10
25280	Media	6	male	647	0.0	<2.94	<3.43	.	167	<50	1		0	4
25298	Sars-CoV-2	3	female	626	-0.2	3.8(3.5-4.0)	6.7(6.6-6.7)	.	2,133	121	1	+	1	686
25304	Sars-CoV-2	3	female	469	-0.5	6.8(6.8-6.9)	7.1(7.0-7.1)	.	2,660	220	1	+	2	5,302
25255	Sars-CoV-2	3	male	932	0.2	<2.94	7.2(7.2-7.2)	.	577	126	1	+++	1	8,447
25165	Sars-CoV-2	3	male	919	-0.3	5.4(5.3-5.4)	6.7(6.6-6.7)	.	2,364	460	1	+++	2	507
25306	Sars-CoV-2	6	female	472	-1.3	5.6(5.5-5.6)	6.5(6.5-6.5)	.	2,515	27,066	136		2	1,403
25228	Sars-CoV-2	6	female	219	3.4	3.4(3.3-3.5)	7.8(7.8-7.8)	.	2,693	117,372	607		1	16,840
25136	Sars-CoV-2	6	male	952	1.0	4.2(4.1-4.2)	7.8(7.8-7.8)	.	2,844	11,500		+	1	50,596
25163	Sars-CoV-2	6	male	922	-0.9	4.2(4.1-4.3)	6.1(6.1-6.2)	.	761	62,542	164	+	1	322
						control mean +2SD = 2.94‡	control mean +2SD = 3.43‡	control mean +2SD= 2.1‡						

‡ Test limit of detection was determined as the mean of control plus two standard deviations

SD - standard deviation

* Values in parenteses represent 95% confidence interval

Viral CN - viral copy number in the tissue by RT-qPCR

Anti CoV-2 N titer - Endpoint binding of SARS-CoV-2 N titer calculating the highest serum dilution to vieal N protein, which gives a reading above the blank, including three standard deviations

Anti CoV-2 S titer - Endpoint binding of SARS-CoV-2 S titer calculating the highest serum dilution to viral S protein, which gives a reading above the blank, including three standard deviations

Neutralization assay ID50 - the reduction of the median fluorescence intensity compared to the virus control was used to determine the ID_{50} of the tested plasma samples

Clinical disease - observed and recorded disease presentation: lethargic and unable to eat or drink (+++), ruffled fur or periorbital edema but able to move and recovered within 24h (+))

Lung histology score - perfused and sliced, stained lung tissues were scored based on the American Thoracic Society Report (Matute-Bello, 2011 #306)

Brain RNAseq reads mapped to SARS - brain RNA depleted of rRNA mapped against the whole viral genome, and the number of hits normalized

Antibody response

The antibody response *of P. leucopus* to SARS-CoV-2 is highest on day 6 post-infection, measured by an S protein-specific binding ELISA. S-specific antibody titers significantly increased from day 3 to day 6 (p=0.0011 Student t-test) (figure 48). We also detected IgG titers against the SARS-CoV-2 N protein on days 3 and 6 post-infection, with no differences between cohorts by day (p = 0.96). Both experiments detected neutralizing antibodies on day 6 post-infection (Table 19). Focus forming assay of flash-frozen lung tissue isolated from animals in experiment 1 showed live virus counts of 30 PFU/mg of tissue for animals ID 25025 and 4PFU/mg for animal 25117. The other two day 3 animals had the not detectable virus for a given dilution.
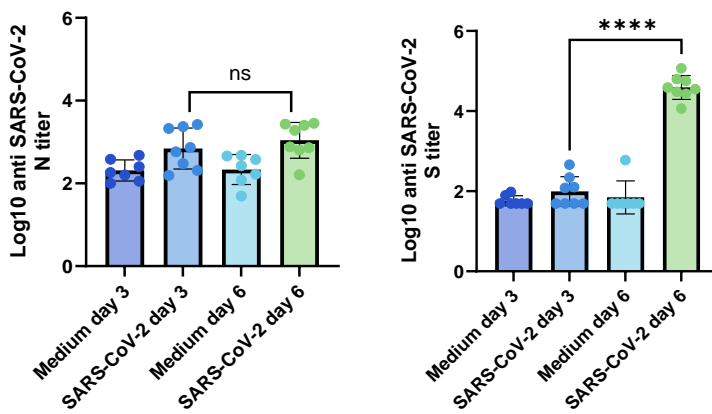
Figure 48 P. leucopus plasma antibody response to virus
Peromyscus IgG antibody response against SARS-CoV-2 mixed spike (N) - left panel and nucleoprotein (S) - right panel antigens from both experiments were assessed by ELISA using plasma collected on the indicated days post-infection. Titer is represented as a Log10 value of reciprocal titration value. Solid lines and bars indicate the mean; error bars indicate 95% CI.

RT-qPCR assays for virus in lungs, brain, and feces

Viral RNA was detected in the lungs of all animals in both experiments, except in animal 25298, in which case it was below the calculated detection limit. On day 3 post-infection, we observed a higher viral copy number in the lung; the mean of log10 transformed copy number is 5.8 (4.4-7.3) compared to day 6, where the mean was 4.5 (3.7-5.3). We compared the viral load between day 3 and day 6 cohort using an unpaired t-test and found that it was not significantly different with a p-value of

0.085, t=1.855, and df=14 (Table 19). The cutoff for detection was determined as the mean copy number of the control group plus two standard deviations.

Viral RNA was also detected in the brains of all animals (8/8) in experiment 2 and half (4/8) of them in experiment 1. In experiment 2, we included the olfactory bulb, while in experiment 1, we only used the cerebrum and cerebellum for RNA isolation.

There was no difference between sexes in viral copy numbers in either the brain or lungs based on a Student t-test (figure 49).

In the first experiment, we tested the animals for viral shedding in the feces collected at euthanasia. RT-qPCR detected virions in the animals inoculated with SARS-CoV-2 on days 3 and 6, while controls were below the test detection limit. There was no difference in animal copy number on both days, p=0.19, with a log10 mean of 3.6 (2.4-4.7) of both day cohorts combined.
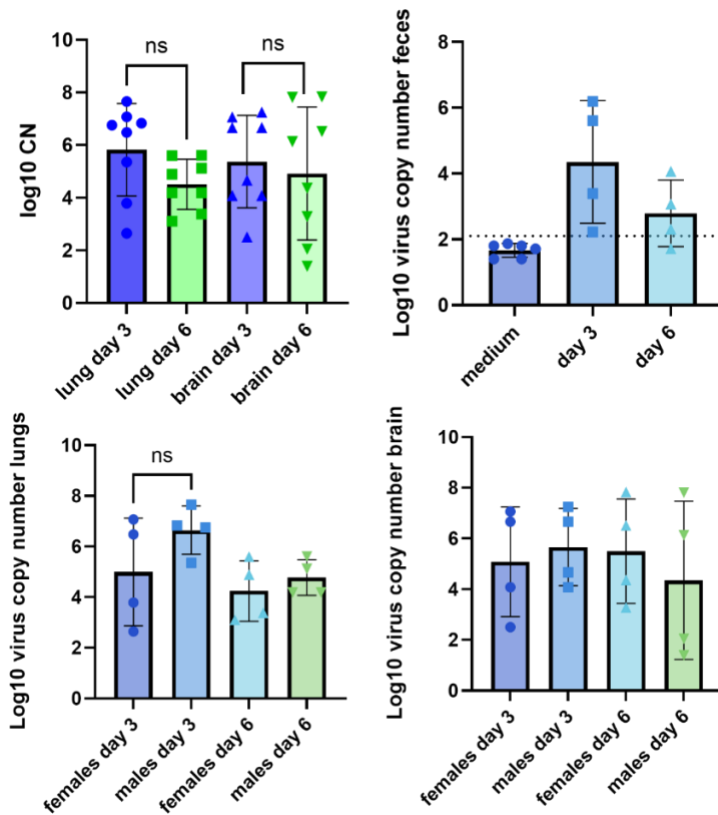
Figure 49 Viral copy numbers in lungs and brain of P. leucopus infected with SARS-CoV-2

The upper left panel shows viral log10 copy numbers by RT-qPCR on days 3 (blue) and 6 (green). The upper right panel shows the log10 viral copy number in feces. The dotted horizontal line shows a calculated detection limit (mean of medium-treated samples plus two standard deviations). Lower panels show bar graphs of log10 viral copy numbers in the lungs and brain, respectively, of infected animals stratified by sex. Solid lines and bars indicate the mean; error bars indicate 95% CI.

Histopathology of the lungs

Infection with SARS-Cov-2 caused inflammation with histological changes that indicate the induction of acute lung injury (ALI). Hallmarks of acute inflammation were most prominent at 3 days after the infection. Histology showed that the changes in acute inflammation were more prominent 3 days after the infection. Trichrome staining demonstrating fibrosis following acute inflammation was more prominent in day 6 cohort than day 3. The evaluation was based on the American Thoracic Society Report (416). Control animals that inhaled media presented with mild intraseptal inflammation with neutrophils and focal foreign body reaction compared to normal tissue.

Additionally, focal peribronchial inflammatory infiltrates and areas with mild interseptal inflammation were observed, as shown in figure 50 upper panel. Animals infected with SARS-CoV-2 presented with neutrophils in the interstitium of the alveolar septae, septal thickening, exudate in the alveolar sacs, and hyaline membranes covering the alveoli (figure 50). In 4/8 of the animals in the day 3 infected group, the lungs were marked as having scored 2 on the lung injury scoring system in animals, and the rest scored 1. One animal scored 2 in the day 6 group (ID 20306), and the rest scored 1 (Table 19). Combined hematoxylin and eosin-stained lung sections with probe-based SARS-CoV-2 RNAscope *in situ* hybridization to determine sites of viral presence. In *P. leucopus* day 3 group, SARS-CoV-2 is localized near neutrophil infiltrates (figure 51). SARS-CoV-2 presence was not detected in the day 6 cohort as infection in the lungs was cleared by that time. In the K18-hACE2 mouse, SARS-CoV-2 localizes to epithelial cells in the alveolar region as previously reported (417).
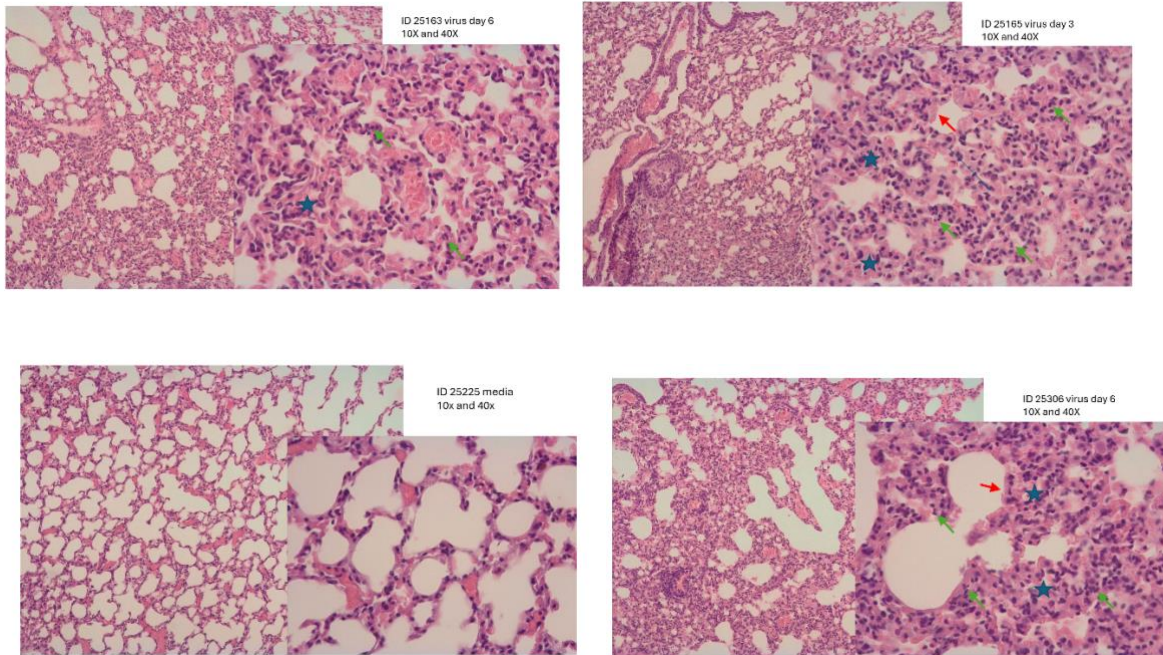
Figure 50 Lung histology of P. leucopus infected with SARS-CoV-2

Adult *P. leucopus* of both sexes were inoculated with $2x10^4$ PFU in 20 µl or DMEM by an intranasal route. Lung pathology slides were stained with hematoxylin and eosin. A summary of all pathology scores and observations is provided in Table 19. Top panel lung tissue infected with SARS-CoV2 day 3 (ID 25165) and 6 day (ID 25163) both from experiment two. The bottom is control with media (ID 25225) and day 6 (ID 25306). Arrowheads indicate neutrophil infiltration in the alveolar septae (green), hyaline membrane (red), and alveolar exudate (white). Blue star indicates alveolar septal thickening.
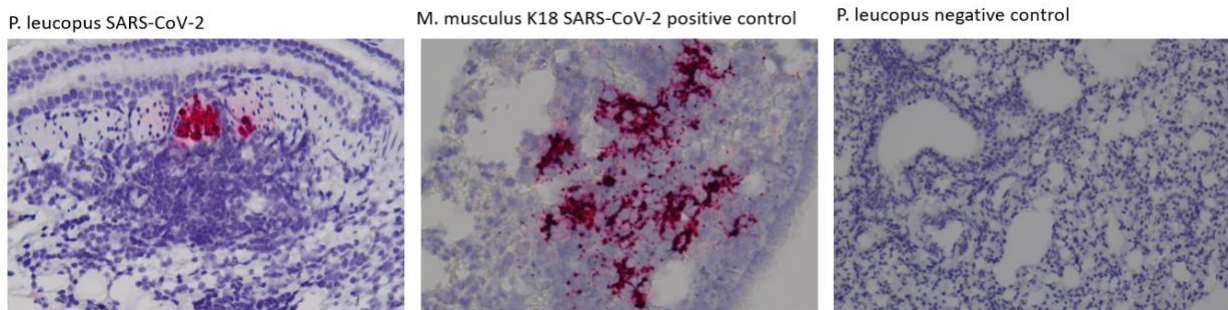
Figure 51 In situ viral RNA hybridization of SARS-CoV-2

Deermouse lung fixed sections were analyzed using RNAscope. Deparaffinized slides were hybridized with probes targeting the SARS-CoV-2 spike, positive-control Hs-PPIB (Homo sapiens control probe for peptidyl isomerase B) for the housekeeping gene, or negative control DapB. Tissues were counterstained with Gill's hematoxylin. Deermouse lung infected with SARS-CoV-2 day 3 shows a positive reaction to the probe. Magnification is 10x.

Transcriptional analysis of early responses to SARS-CoV-2 infection in the lungs

We used high-quality RNA paired-end 150 nt reads and the latest NCBI annotation for *P. leucopus* LL with 54475 transcripts. Since we currently don't have isoform data, transcripts post RNAseq were concatenated and analyzed for differential gene expression. Overall, we have 26570 genes for analysis.

Lung transcripts were analyzed for differential gene expression on day 3 and day 6 for both experiments. To increase the power of analysis, we combined controls on both days 3 and 6. With a stringent cutoff of |log2 FC|>1.5 and FDR <0.05 and p<0.05 in both day 3 and 6 clusters, the number of genes was still higher on day 3, as shown in the volcano plot (figure 52). The total number of differentially expressed genes in each direction showed 61 upregulated and 2 downregulated genes in the day 3 group and 14 upregulated and 1 downregulated in the day 6 group, presented in the heatmap in figure 52.

The Edwards-Venn diagram of differentially upregulated shows an overlap between the day 3 and 6 groups, with genes overlapping Cxcl9, Cxcl10, Gbp6, Gzmb, Gzmk, and Iigp1. Cxcl9 is produced by activated macrophages and recruits Th1-polarized CD4$^+$ CXCR3$^+$ cells. Cxcl10 is expressed upon interferon gamma activation early in infection with RNA viruses in humans, and it tends to persist in infected individuals since it is used to bridge viral elimination until the adaptive immune system is activated. In plasma and bronchial lavage, it correlates with disease severity (418). There are two significantly downregulated genes based on the same criteria in the day 3 cohort (Ciart and St8sia5) while day 6 downregulated gene is Asgr2.

Lungs in day 3 cohort have upregulated genes highlighted in Table 20. *P. leucopus* responds to infection with SARS-CoV-2 in the lungs by expressing type I interferon stimulated genes and effectors of negative stimulation of inflammation, striking a balanced response.
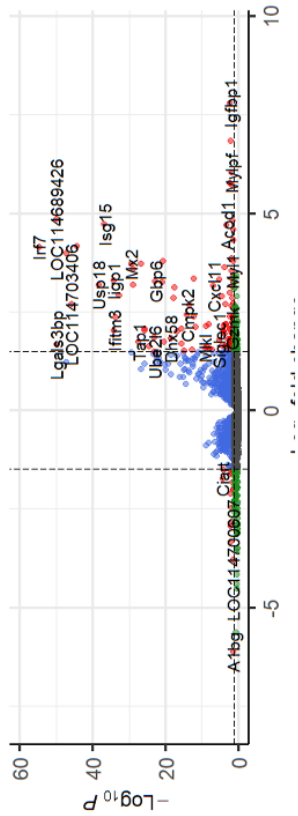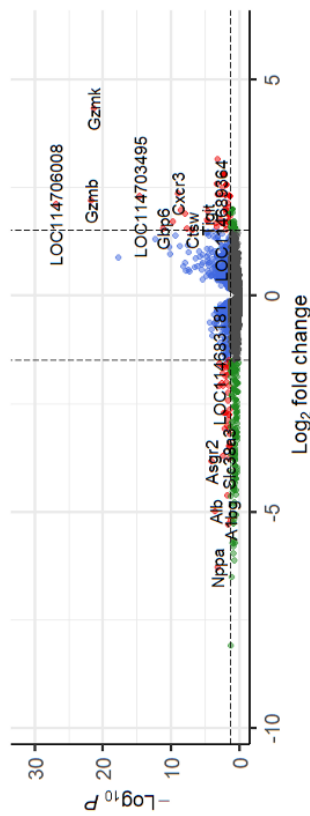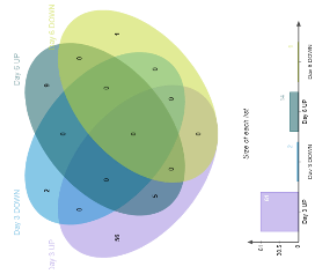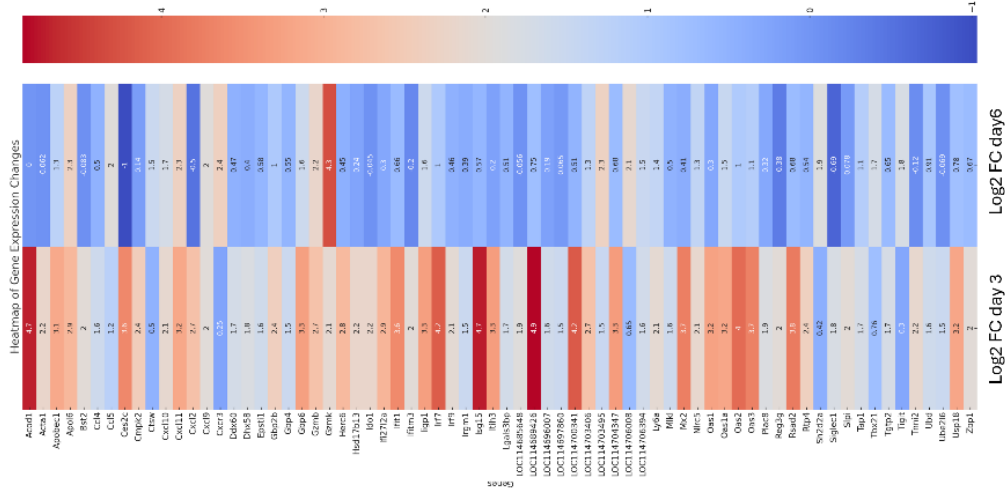
Figure 52 Volcano plot, Venn diagram and heat map of differential gene expression in the lungs

Volcano plot comparing differentially expressed genes at day 3 versus media and day 6 versus media. Dots indicate upregulated and downregulated genes. Colors denote: red |log2 FC| > 1.5 and p-value < 0.05, blue is |log2 FC| <1.5 and p-value < 0.05, green is |log2 FC| > 1.5 and p-value > 0.05 and black is |log2 FC| < 1.5 and p-value > 0.05. The tables next to each plot show significantly upregulated genes with FDR < 0.05. with |log2 FC|>1.5 and FDR <0.05. Edwards-Venn diagram shows upregulated and downregulated genes in day 3 and day cohorts. Heatmap of differentially expressed genes on day 3 and day 6 with the log2 FC>1.5 and FDR <0.05.

Type I interferon-stimulated genes

Day 3 group features type I interferon upregulated genes: Mx2, Oas, Ifit, Itim, Isg15, Rsad2, Herc6, Apobec, Rtp4, and Gbp (table 20). Apobec gene is primarily known for its role in editing apolipoprotein B mRNA, which involves changing cytidine to uridine in the mRNA. Deamination can cause mutation in viral DNA and hider replication (419). Rsad2, whose gene product is viperin - an enzyme belonging to the radical S-adenosylmethionine superfamily, exhibits antiviral activity against a wide range of viruses by perturbing lipid microdomains rich in glycolipids at the sites for viral release (420). Mx genes block the nuclear import of viral cDNA. This action disrupts the virus's life cycle, preventing it from integrating into the host genome and further replicating. Mx genes in bats are under constant positive selection, indicating structural flexibility, especially in surface-exposed variable regions indicating changes under viral pressure (421). We detected three possible isoforms, Mx2, LOC114689426, LOC114700341, using NCBI blast and orthoDB to detect orthologs (422). Rtp4 gene binds to viral replicase and suppresses amplification. Like the Mx gene, it undergoes positive selection in bats, usually achieving patterns of specificity against certain viruses (423).

Ifitm3 gene product inhibits viral entry into host cells. It adds to the cellular defense by incorporating itself into host cell membranes, which can block the fusion of viral membranes with host membranes, effectively preventing the virus from entering and replicating the cell. Lower levels of IFITM3 are linked to an increased risk of severe influenza infections. Studies in human COVID-19 patients suggest that higher expression of Ifitm3 is protective against severe disease, although results are at odds in some studies (424), (425), (426).

Ifi27l2a has been implicated in promoting apoptosis and shows strong anatomical variation (427).

Isg15 is a ubiquitin-like protein, an intracellular post-translational modifier conjugating to viral protein and modulating its function. Isg15 contributes to immunomodulation (428).

Regulation of immune response

Ifr7 is a transcription factor crucial for regulating the production of type I interferons. Ifr7 is activated by PRR and shortlived, regulated by numerous posttranslational modifications and cellular concentration of type I interferons by feedback amplification (429) (430).

Negative regulation of type I interferons

The Igrm gene product is a master negative regulator of the interferon response, suppressing cGas and Rigi. IGRM is responsible for the mitophagy of dysfunctional mitochondria, preventing oxidative damage (431). We detected two possible isoforms, Igrm and LOC114706394, using NCBI blast and orthoDB to detect orthologs (422).

Anti-inflammatory response

IDO1 is an anti-inflammatory enzyme in kynurenine pathway depleting Triptophan in plasma, detected in COVID-19 patients (432). Ang (LOC114696007) gene expressing angiogenin is found in various tissues, and its expression can be induced by hypoxia. Ang has a role in tissue repair, where new blood vessels are needed. Reg3g gene product is an anti-microbial peptide typically associated with regeneration and inflammation in various tissues, particularly the gastrointestinal tract. Lactobacillus in the gut increased expression and circulating levels of Reg3g while antibiotics suppress it (433). Studies suggest it has beneficial effects on epithelial injury, neuron regeneration, and cardiac inflammation (434)

Lungs in the day 6 cohort feature cell-mediated immune response and anti-inflammatory genes; however, the overall number of genes and the fold change is lower than in the day 3 cohort.

Type I interferon stimulated genes

There is a lingering effect of type I interferon with GPB6 and Oas1a, which are still differentially upregulated.

T cell response

Gzmk, Gzmb, Prf1, and Cxcr3 are significantly expressed in the day 6 cohort, which indicates activation of cytotoxic T-cells for removal of virus-infected cells.

Anti-inflammatory response

T-cell immunoglobulin and immunoreceptor tyrosine–based inhibitory motif domain (TIGIT) is a novel target for acute inflammation. TIGIT is expressed in NK cells and T cells at high levels, but long exposure to this product can lead to progression of tumors (435). $CD8^+$ $Tigit^-$ T cells are associated with long covid, while subjects expressing Tigit in $CD8^+$ cells recovered (400). Tigit expressed on $CD4^+$ T cells induces Il10 production while suppressing Il12 production, inhibiting $CD4^+$ T cell proliferation and Ifng production and favoring alternative macrophage polarization (436), (437), (438). Cd52 (LOC114706008) or campath-1 antigen, expressed on T cells, suppresses other T cells. Soluble CD52 binds to the inhibitory receptor Siglec-10 and impairs T cell activation. This mechanism of T cell regulation may be used in translational studies of autoimmune diseases (439). Soluble CD52 also suppresses the production of inflammatory cytokines by macrophages, monocytes, and dendritic cells in response to LPS, signifying its potential as an immunotherapeutic agent (440).

Table 20 Select differentially upregulated genes in the lungs three days post infection with SARS-CoV-2

Gene	Function	log2 fold change	False discovery rate p-value
Select upregulated genes in *P. leucopus* lungs three days post infection with SARS-CoV-2			
Type I interferon stimulted genes			
Isg15	Ubiquitin-like protein that is conjugated to target proteins and can modulate immune responses.	4.7	4.0E-34
Acod1	Involved in the immune response to bacterial infection and is induced by interferons.	4.7	4.1E-02
Rsad2	Encodes Viperin, a protein with antiviral properties.	3.8	1.5E-18
Ifit1	Proteins that bind to viral RNA and inhibit translation.	3.6	1.5E-26
Ifi27l2a	Interferon Alpha-Inducible Protein 27 Like 2A, inhibiting viral replication.	2.9	2.9E-30
Ifitm3	Interferon Induced Transmembrane Protein 3 inhibiting viral entry into host cells.	2.0	1.8E-31
Herc6	An E3 ubiquitin-protein ligase that could be implicated in antiviral response.	2.8	1.1E-15
Mx	GTPases involved in antiviral activities against a range of viruses.		
Mx2		3.7	6.3E-27
Mx-isoform		4.9	9.6E-47
Mx-isoform		4.2	1.3E-41
Oas	Enzymes that synthesize 2'-5'-oligoadenylates which activate RNase L, leading to degradation of viral RNA.		
Oas2		4.0	2.2E-44
Oas1		3.2	1.1E-18
Oas1a		3.2	1.5E-26
Oas3		3.7	1.9E-24
Ifr	Transcription factors that play central roles in IFN signaling.		
Irf7		4.2	3.9E-51
Irf9		2.1	2.1E-23
Gbp	Guanylate binding proteins, which have roles in immune responses to pathogens.		
Gbp6		3.3	1.6E-20
Gbp2b		2.4	1.6E-10
Gbp4		1.5	6.1E-11
Cxcl	Chemokines that recruit immune cells to the site of infection or inflammation.		
Cxcl11		3.2	2.8E-05
Cxcl2		2.7	3.1E-02
Cxcl10		2.1	5.2E-10
Anti-inflammatory			
Ido1	Indoleamine-2,3-dioxygenase oxydizing tryptophane via kynurenine pathway.	2.2	1.7E-06
Slpi	Secretory leukocyte protease inhibitor controls excessive protease activity in inflammation.	2.0	8.3E-03
Ang	Angiogenin expressed under hypoxia promoting repair or regeneration where new blood vessels are needed.	1.6	2.8E-02
Irgm1	Immunity Related GTPase M negative regulator of the interferon response.	1.5	5.6E-13
Irgm-isoform		1.6	3.2E-03
Anti-viral			
Reg3g	Regenerating family member gamma, antimicrobial peptide involved in wound reepithelialization.	2.0	1.3E-02
Rtp4	Receptor transporting protein 4, found in bats to suppress flavivirus genome amplification.	2.4	4.7E-31

Table 21 Select differentially upregulated genes in the lungs six days post infection with SARS-CoV-2

Gene	Function	log2 fold change	False discovery rate p-value
Select upregulated genes in *P. leucopus* lungs six days post infection with SARS-CoV-2			
Ifng stimulated genes			
Cxcl9	Chemokine attracting activated T cells to sites of inflammation.	2.0	1.5E-02
Cxcl10	Chemokine recruiting and activation T and NK cells.	1.7	4.6E-07
Type I interferon stimulated genes			
Gbp6	Guanylate binding proteins, which have roles in immune responses to pathogens.	1.6	2.5E-08
Oas1a	Enzymes that synthesize 2'-5'-oligoadenylates which activate RNase L, leading to degradation of viral RNA.	1.5	1.4E-02
CD8+ T cell			
Gzmk	Serine protease recognizes and enters virus-infected or transformed cancer causing direct DNA damage apoptosis.	4.3	3.8E-18
Gzmb	Serine protease direct DNA damage and triggering of alternative apoptotic pathways.	2.2	2.2E-18
Prf1	Pore-forming protein facilitating delivery of cytotoxic granules, such as granzymes.	1.4	1.5E-02
Cxcr3	T-cell receptor	2.4	2.2E-06
Anti-inflammatory			
Tigit	Expressed on CD4+ T cells and induces IL-10 production while suppressing IL-12 production.	1.8	1.4E-02
Cd52	Suppresses T-cell activation and production of inflammatory cytokines.	2.1	3.4E-23

Gene ontology analysis of the lungs

To further define the response to SARS-CoV-2 in the lungs of *P. leucopus*, we used identified

categories of upregulated and downregulated genes based on the same cutoff. In a day 3 cohort, 61

genes are upregulated, and 53 are recognized by enriched biological pathways and protein complexes analysis as *M. musculus* genes. With a p-value of 0.01 cutoff for enrichment, we got 15 gene ontology (GO) terms, and 6 had a P-value less than 10^{-10}. The most significant GO term in day 3 cohort is Response to virus enrichment (p-value 10^{-27}) constituted of: Acod1, Apobec1, Bst2, Cxcl10, Cxcl9, Ddx60, Dhx58, Gbp2b, Ifi27l2a, Ifit1, Ifitm3, Irf7, Irgm1, Isg15, Mlkl, Mx2, Oas1a, Oas1b, Oas2, Oas3, Rsad2, Rtp4, Zbp1. The rest of the GO terms are related to early viral infection (GO: 0009815) and activation of innate immune response (GO: 0045087). They also include Response to interferon-beta (GO: 0035456), Negative regulation of viral genome replication (GO: 0045071), Regulation of innate immune response (GO: 0045088), and Response to interferon-alpha (GO: 0042742) (figure 7).

Differentially expressed genes in the lungs 6 days post-challenge featured Chemokine receptors bind chemokine made up of genes: Ccl5, Cxcl9, Cxcl10, Cxcr3, Gbp6, Gzmb, Iigp1, Tbx21 while genes driving second GO term: regulation of leukocyte cell-cell adhesion are: Ccl5, Tbx21, Tigit, Sh2d2a. Based on the same cutoff, downregulated genes in the day 3 cohort are Ciart and St8sia5, and in the day 6 cohort Asfr2.

Upregulated genes in day 3 cohort when compared to day 6 as a control in differential gene expression analysis are: Alb, Apobec1, Arg1, Bst2, Cd209d, Cmpk2, Galnt15, Gbp6, Gp2, Herc6, Ido1, Ifi27, Ifi27l2a, Ifitm3, Ifit1, Iigp1, Irf7, Irf9, Isg15, Il1r2, Mx2, Oas1, Oas1a, Oas2, Oas3, Plac8, Reg3g, Rsad2, Rtp4, Saa3, Saa5, Siglec1, Ube2l6, Usp18. Day 6 differentially upregulated genes in the same analysis are: Cxcr3, Dbp, Gzmk, Lratd1, and Per3.
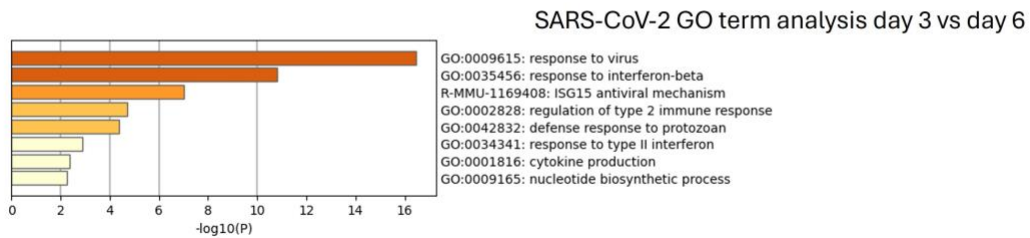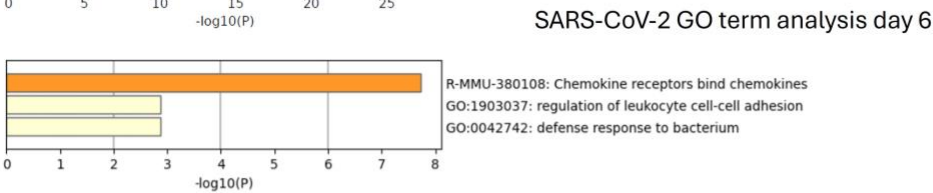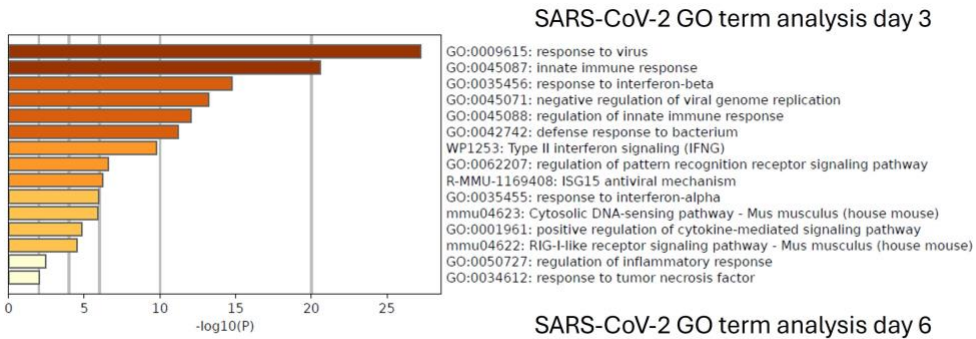
Figure 53 Gene ontology of DEGs in the lungs
GO term analysis of *P. leucopus* lungs with upregulated genes on day 3 (upper panel) and day 6 (middle panel) compared to animals treated with media inhalation as controls and sacrificed on days 3 and 6 post-treatment. The lower panel shows GO terms analysis of genes upregulated on day 3 compared to the treated cohort on day 6X-axes are -log10 P-value and bars, while bars represent a range of the -log10 p depicted by shades from yellow to dark brown.

Targeted RNA seq analysis in the lungs

To confirm RNAseq analysis, we used a set of genes identified as significant in this and previous analyses of *P. leucopus* immunity and made a list of target 156 mRNA coding sequences used in Chapters 3 and 4. Out of 156 transcripts we included in the targeted analysis, we further normalized, transformed, and analyzed data using ANOVA and Tukey's multiple comparison tests to test which relationships with groups (day 3, day 6, and controls) are marked as significant.
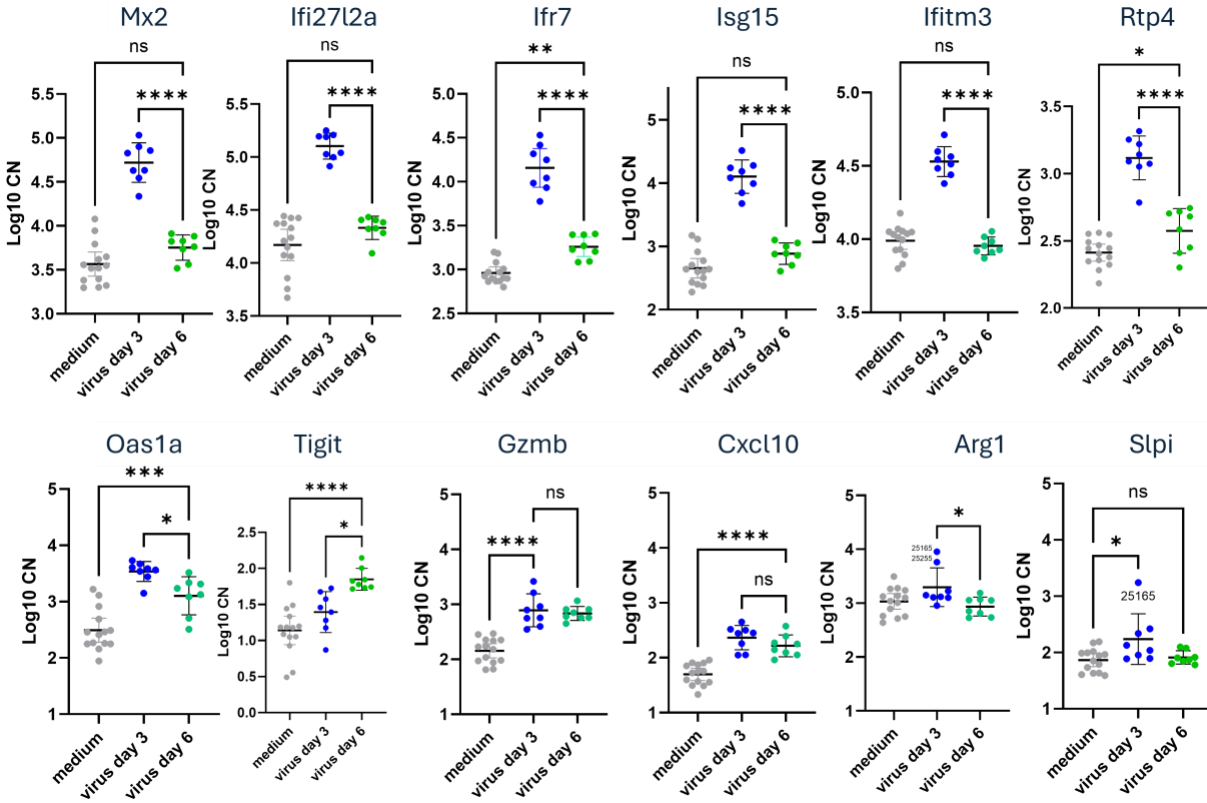
Figure 54 Targeted RNAseq analysis in the lungs
Plot with mean and 95% CI of lung viral copy number and targeted RNAseq. Genes selected based on 156 select genes mapping ANOVA between groups with $p<0.001$. RNAseq reads are normalized based on all reads per sample and transformed as log10. Green dots are day 6, blue are day 3, and gray are control with media. ANOVA was followed by a Tukey multiple comparison test to compare groups. "ns" on the bar means comparison has a p-value >0.05, one star means p-value < 0.05-, and four-stars p-value < 0.0001.

Interferon-stimulated genes like Mx2, Irf7, and Isg15 are significantly upregulated on day 3 post-infection, while the innate response to viruses is dominant (figure 54). Chemokine Cxcl10, Gzmb - a serine protease expressed by cytotoxic T lymphocytes and Oas1a - RNA sensing ISG, are all upregulated in lungs and stay upregulated over 6 days (441), (442), (83). Targeted gene analysis unveiled some extreme cases where animals presented with strong disease phenotypes. Those are associated with the expression of Slpi, Acod1, and Arg1. Mx2 is constitutively expressed higher than other ISGs like Isg15 and Irf7, although Ifi27l2a virus sensing is expressed highest in controls. In sepsis, Arg1 is associated with inflammation, as noted in Chapters 3 and 4. In viral infection, Arg1

increases 5 and 10 fold in animals with a disease phenotype compared to other infected animals. As seen previously, Nos2 did not change significantly during infection, nor did Ifng and Il6.

Transcriptional analysis of early responses to SARS-CoV-2 infection in the brain

Since there is evidence of infection with SARS-CoV-2 and inflammation in the brain in other *Peromyscus species* (108) and Syrian hamsters (413), ecologically important, widespread, and disease-tolerant *P. leucopus* would be a valuable model to simulate disease progress and describe pathogenesis in the brain during viral infection.

Our experiment used a stranded rRNA depletion brain RNA library paired-end ~150 nt and got 40 million reads per sample. Since we are interested in the brain's immunity, we included the cerebrum and cerebellum in the first experiment, while the olfactory bulb was analyzed together with the whole brain in the second experiment. Hence, the two experiments were analyzed separately. To determine the differences between experiments, we analyzed differential gene expression on day 3 and day 6 versus all control animals and used the Quantile-Quantile Plot Linear Fit (QLF) test to identify genes differentially expressed across two experiments (figure 55). In the day 3 cohort, there are 19 significantly upregulated genes (FDR<0.05 and log FC > 1.5) in experiment 2; out of those, 8 are downregulated in experiment 1 (Bpifa6, Ctse, Deg21, Ifi27, Irs4, Obp2a, Sptlc3, and Sulta1), 8 are upregulated but not significantly log2 FC < 1.5 (Bst2, Ifi27l2a, Irf7, Mx2,ncRNA, Pla1a, Usp18) and 3 are upregulated in both experiments log2 FC > 1.5 (Isg15, Mx2 isoform, and Oas1a). In the same analysis for day 6 animals, there are only 2 significantly upregulated genes in experiment 2 (Cxcl10 and Isg15), and in experiment 1, Cxcl10 is downregulated, and Isg15 is upregulated but not significantly (figure 55). Data show that DEGs between experiments 1 and 2 differ, and the olfactory bulb included in experiment 2 could be the driver of inflammation.
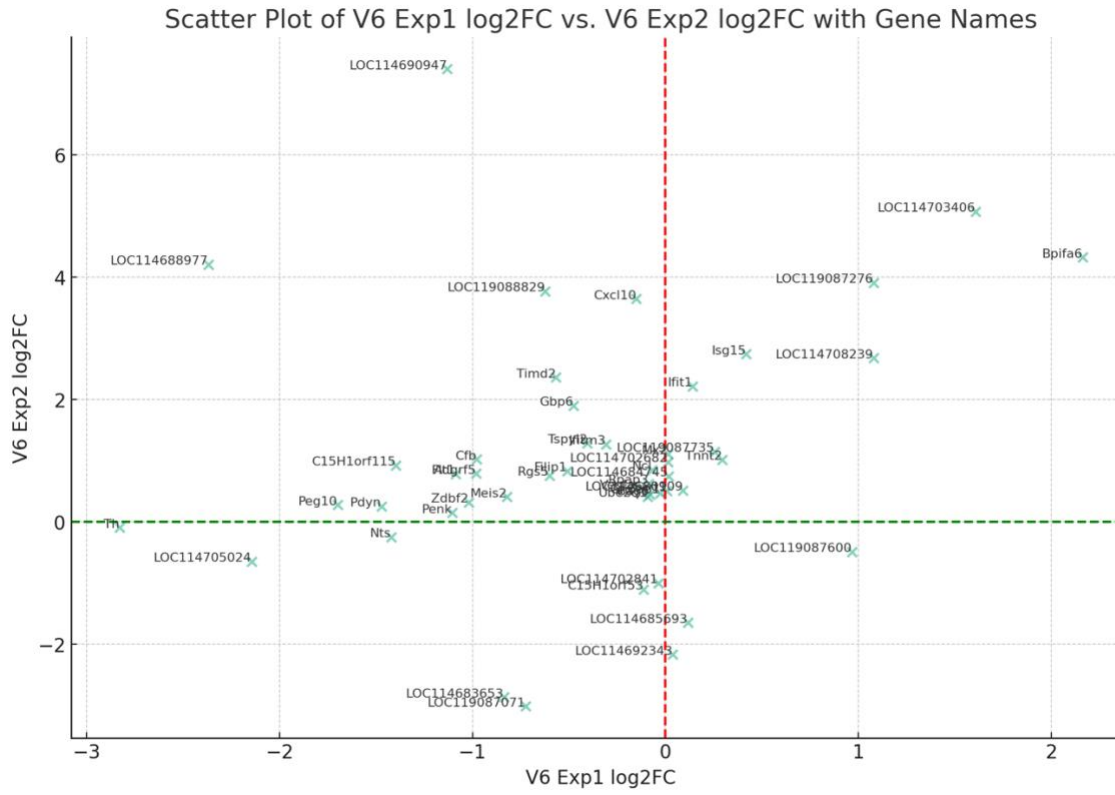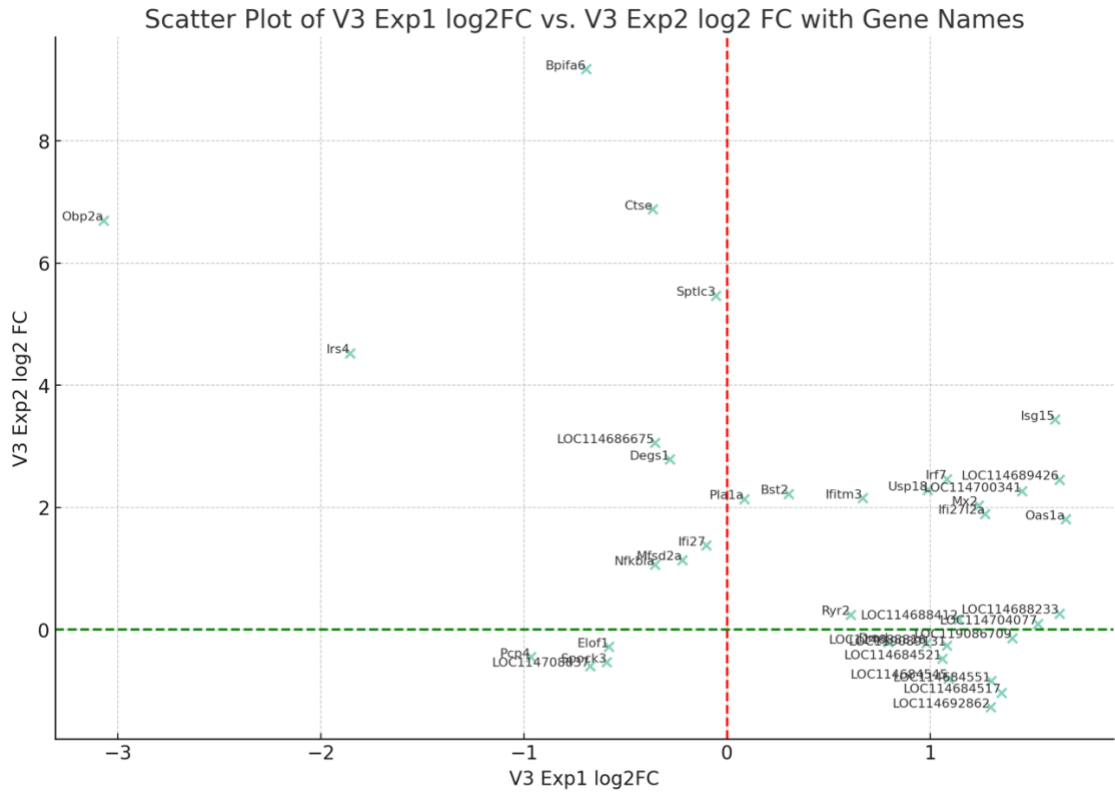
Figure 55 Scatter plot of differential gene expression correlated between two experiments on day 3 and day 6

Scatter plot of differential gene expression between experiments on day 3 in the upper panel and day 6 in the lower panel. Upregulated genes in experiment 1 are above the green dotted line, while the upregulated genes in both experiments are in the top right quadrant.

There is a correlation between the RT-qPCR analyzed copy number in the brain and targeted RNAseq against the whole SARS-CoV-2 genome (figure 56). Hence, in this analysis of the brain, we use normalized viral copy numbers in the brain for the analysis and experiment 2 animals as they show significant inflammation in the brain.
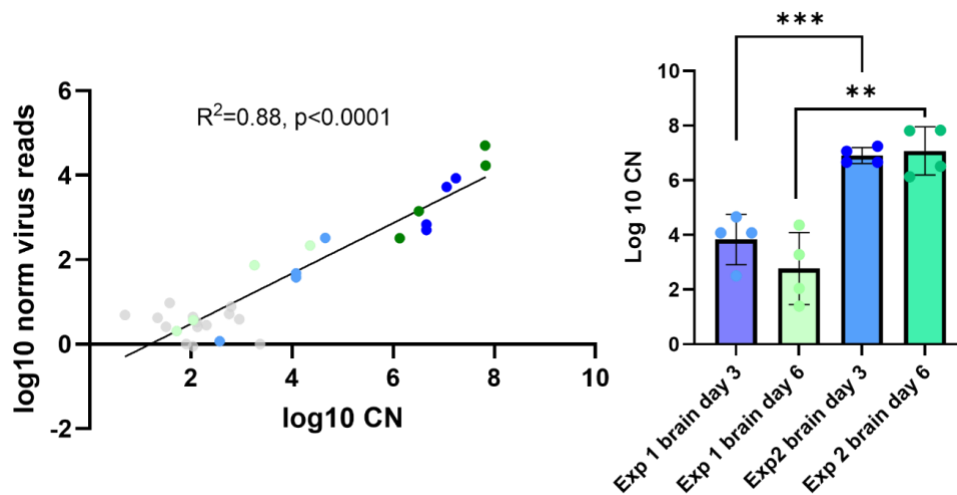
Figure 56 Viral copy numbers in the brain
Left panel shows a correlation between log10 viral copy number in the brain by RT-qPCR and log10 RNAseq targeted against the SARS-CoV-2 genome, normalized for the number of reads. The right panel shows the difference in copy number between the two experiments. Blue dots are day 3, and green are day 6. Lighter colors are experiment 1, and darker colors are experiment 2. Gray are controls. Animal IDs below the calculated test detection limit for RT-qPCR of 3.43 are highlighted on the graph as a dotted line.

The overlapping genes between upregulated brain activity on days 3 and 6 are Aadacl2, Aadacl4 (both expressed on olfactory region), Bpifa6 (antimicrobial protein presented in the olfactory bulb and nasal epithelia (443)), Ifit1 (induced as a response to type I interferon, binding ssRNA and acting as a viral sensor single and inhibiting expression of viral RNA (444)), Isg15 and Mx2 (type I interferon response genes (445)), innate response Mogat3 (involved in triacylglycerol metabolism, not known to be present in mice (446)), ncRNA, Oas1a (type I interferon response gene limiting viral spread in

the brain (447)), Olah (neutrophil related sepsis biomarker (448)), and Tmem29 (associated with programmed cell death in the brain post-ischemic shock and glucose insufficiency (449)). These genes are expressed in both datasets, indicating that they are active on both days or through the course of infection. There is no overlap in downregulated genes. There are nearly ten times more upregulated genes in the day 3 cohort than downregulated, contributing to the notion of increased transcriptional activity during inflammation.

RNAseq of the brain used 26570 annotated transcripts. Figure 57 shows either up- or downregulated genes in the brain with criteria of |log2 FC| >1.5 and FDR < 0.5, and there are 57 upregulated genes in the day 3 cohort and 3 downregulated. There are 13 upregulated genes in the day 6 cohort and 1 downregulated.
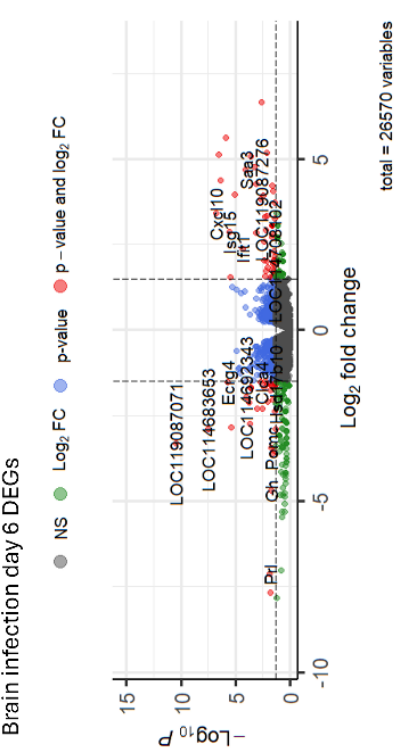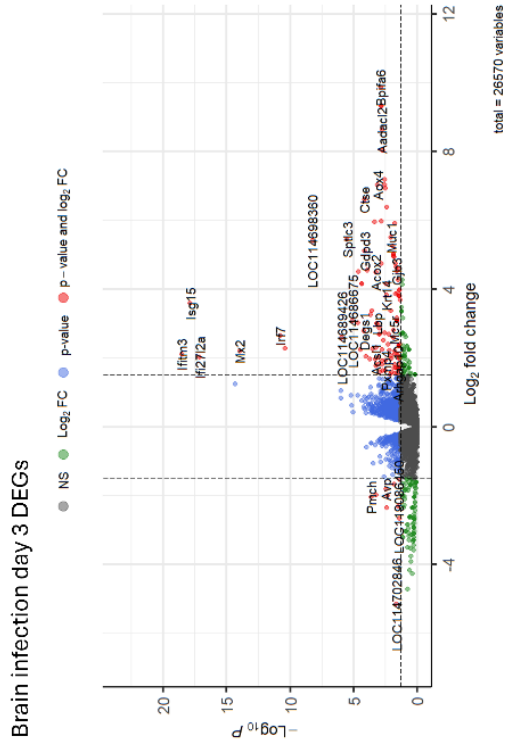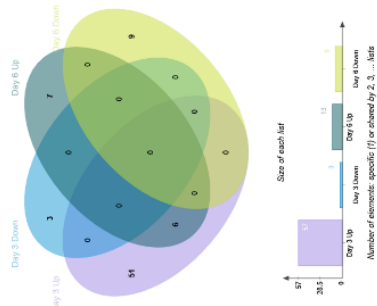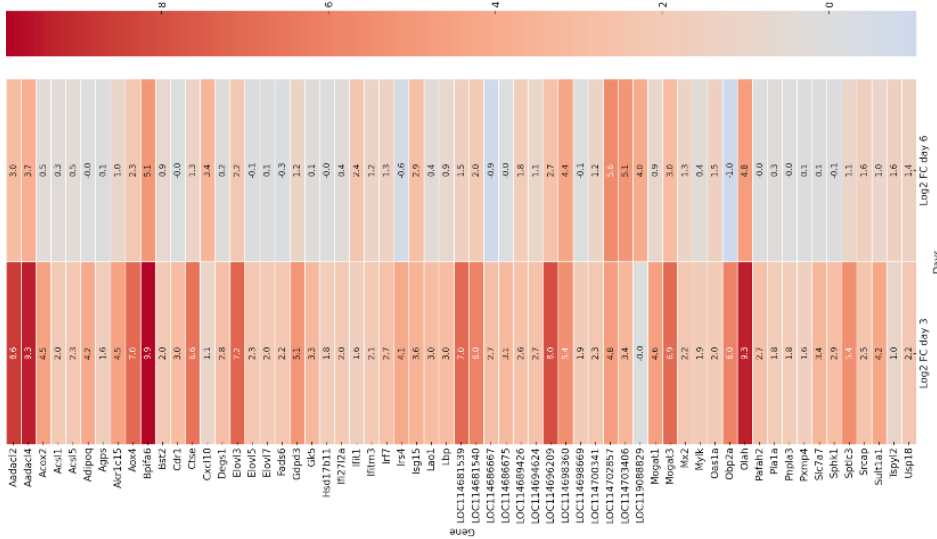
Figure 57 Volcano plot, Venn diagram and heat map of differential gene expression in the brain

Volcano plot of log2 FC and log10 p-value based on differential gene expression analysis comparing differentially expressed genes at day 3 versus media and day 6 versus media. Dots indicate upregulated and downregulated genes. Colors denote: red |log2 FC| > 1.5 and p < 0.05, blue is |log2 FC| <1.5 and p-value < 0.05, green is |log2 FC| > 1.5 and p-value > 0.05 and black is |log2 FC| < 1.5 and p-value > 0.05. The tables next to each plot show significantly upregulated genes with p < 0.05 and |log2 FC|>1.5. Edwards-Venn diagram shows upregulated and downregulated genes in day 3 and day cohorts. Heatmap shows Log2 FC>1.5 data of the brain with FDR<0.5.

Gene ontology analysis of the brain

Gene ontology (GO) analysis uses defined categories of genes (figure 58). The most significant GO

term in the day 3 upregulated cohort was GO:0008610 lipid biosynthesis process. The second one is

GO:0034340: response to type I interferon: Acsl1, Adipoq, Bst2, Ifi27l2a, Ifit1, Ifitm3, Irf7, Isg15, Lbp,

Mx2, Oas1a, Sphk1; driven by innate response and Interferon I induced genes. Sphk1 gene

sphingosine kinase phosphorylates sphingosine SP1 is shown to influence proliferation, survival,

migration, angiogenesis, and differentiation, especially during neuroinflammation (450), (451). Day 3

downregulated gene Avp with log2 FC=-1.8 is responsible for the regulation of the body's osmotic

system, body water content, blood pressure, and plasma volume, and its upregulation is correlated

with poor clinical outcomes in Covid-19 patients (452). The brain day 6 cohort had 13 upregulated

genes and 1 unique GO term. Genes driving GO term GO:0051607 cellular Response to interferon-

beta (Cxcl10, Ifit1, Isg15).
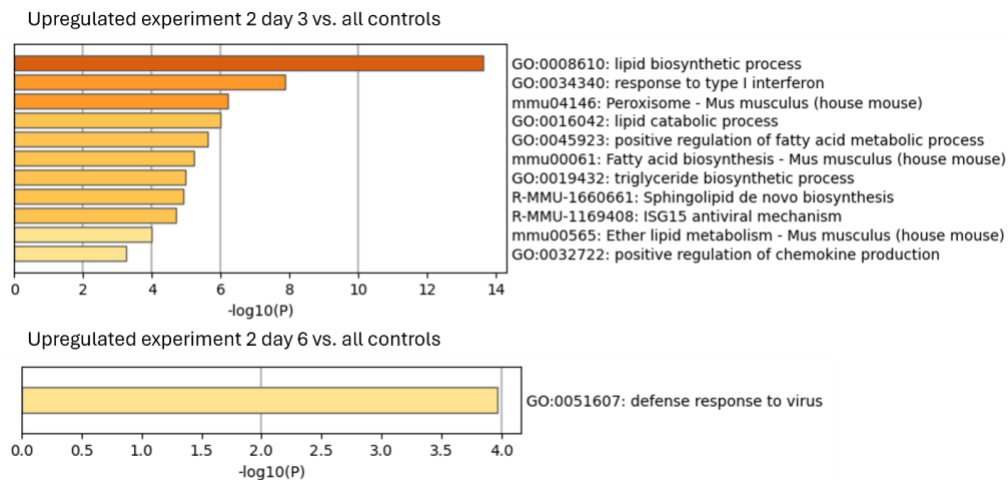
Figure 58 Gene ontology analysis in the brain

Gene ontology term analysis of *P. leucopus* brain with upregulated and downregulated genes on day 3 (2 upper panels) and day 6 (2 lower panels) compared to animals treated with media inhalation as controls and sacrificed on days 3 and 6 post-treatment. X-axes are -log10 P-value and bars, while bars represent a range of the -log10 p depicted by shades from yellow to dark brown.

Targeted RNA seq analysis in the brain

We use targeted gene expression to confirm bulk RNAseq data. Since gene expression correlates with

the viral copy number and viral reads, we are using both experiments to show that relationship, as

shown in figure 59. The analysis included the Interferon response genes Mx2, Ifi27Ia, Irf7, Isg15, Ifit1,

and Oas1a, all significantly upregulated in the brain.
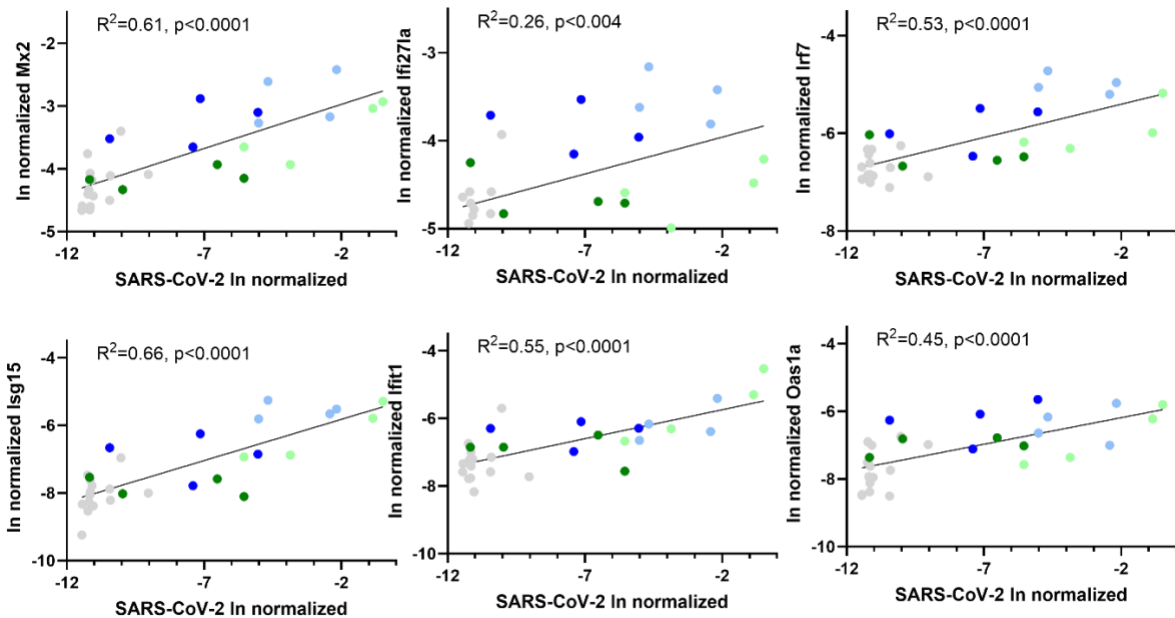
Figure 59 Targeted RNAseq analysis in the brain
Targeted and ln transformed and normalized to Gapdh gene expression targeted RNAseq of select genes correlated with SARS-CoV-2 reads. Green points represent day 3 cohort (light green is experiment 2 and dark green is experiment 1), blue points represent day 6 cohort (light blue is experiment 2 and dark blue is experiment 1).

Discussion

This study is the first to show viral detection, immunity, and respiratory pathology in response to

SARS-CoV-2 in *P. leucopus*. Our findings open new avenues for understanding the pathogenesis of

the virus and its effects on different organ systems, particularly the neuropathology of the brain at the transcriptional level—processes that have not been described before in *Peromyscus*.

P. leucopus and related *Peromyscus* species response to SARS-CoV-2

We show that neutralizing antibodies detected on day 6 correlate with SARS-Cov_2 spike protein (S) protein titer in *P. leucopus*, suggesting that by day 6, IgGs are successfully neutralizing the virus. Fagre et al. report that *P. sonoriensis* has detectable IgG to viral N protein in plasma on day 14 post-infection with SARS-CoV-2 while we detect IgG to N protein in experiment 2 on day 3 (108). Lewis et al. show neutralizing antibodies and anti-N IgG in *P. californicus* later in infection while reporting clinical signs, morbidity, and weight loss. Other *Peromyscus* species in the study (*P. sonoriensis*, *P. maniculatus,* and *P. polionotus*) did not show clinical signs of infection, but all were presenting with lung pathology, particularly on day 3, including lesions and neutrophil infiltration, similar to *P. leucopus* (107). Griffin et al. also showed pathology in the lungs of *P. maniculatus* along with abundant viral RNA in epithelial cells and overall higher viral load in lungs by RT-qPCR than in *P. leucopus*. This could be because this group used 10 to 100 x higher viral titer for inoculation, or because *P. leucopus* cleared infection sooner with neutralizing antibodies (109). We detected viral RNA in the feces, meaning there is a possibility of transmission even without close contact, although direct transmission in *Peromyscus* species has been shown previously (109).

Transcriptional analysis in the lungs

Transcriptional analysis has identified the upregulation of genes involved in viral responses. Notably, the lungs showed more transcriptional activity than the brain, with 315 DEGs in lungs having a false discovery rate below 0.05, compared to 20 DEGs in the brain using the same criteria. Genes associated with innate immunity were predominantly upregulated in the lungs on day 3, while genes linked to adaptive immunity were more prominent in the day 6 cohort. The day 3 cohort also showed significant upregulation of ISGs and specific isoforms of the Mx gene. A similar trend of gene

expression in response to viral infections is observed in bats. Bats have developed unique expansions and specializations in their interferon system, critical for their antiviral defense (421). This specialization includes constitutive and highly inducible expression of interferons, particularly type I interferons, which are the primary line of vertebrate immune defense against viruses (453). There is an expanded and divergent range of interferon-induced genes in bats, mainly Mx genes. Besides being constitutively expressed, Mx genes have expanded with several active isoforms and diverged in response to particular viruses representing a dynamic evolutionary arms race (29). We report this trend in *P. leucopus*, also – three isoforms of Mx2 were significantly differentially expressed (table 20). In homeostasis, ISGs are significantly upregulated compared to outbred *M. musculus*, as reported in Chapter 3 (figure 60).

antiviral effect: Mx Irf7 Isg15 Ifih1 Ifit1 Oas		
inflammation suppression: Arg1 Sod		
inflammation: Il-12 Nos2 ERV env ERV gag-pol		

Differentially expressed genes in blood of *P. leucopus* using *M. musculus* as a control		
log2 fold change	False Discovery Rate p-value	Gene
Upregulated in *P. leucopus* in homeostasis		
8.7	4.4E-18	Mx2
7.8	9.6E-20	Arg1
5.9	2.7E-06	Ifih1
5.3	3.3E-17	Oas1a
3.9	2.1E-09	Ifit1
3.6	4.2E-27	Ifi27l2a
2.9	5.4E-09	Sod2
2.1	3.1E-05	Irf7
1.9	1.6E-04	Isg15
Upregulated in *M. muscuslus* in homeostasis		
-1.0	1.0E+00	Il12a
-3.7	2.7E-10	Slpi

Figure 60 Constitutively upregulated genes in the blood of P. leucopus in comparison to M. musculus EdgeR analysis combined and confirmed by targeted analysis of RNAseq blood data presented in Chapter 3 show ISGs upregulation in *P. leucopus* at a base level. The inflammation suppressive gene Arg1 is also upregulated in *P. leucopus*. M. musculus is characterized by upregulation of genes characteristic of inflammation (IL-12 andNos2) and constitutively expressed ERV proteins.

When comparing RNAseq data from *P. leucopus* lungs to publicly available RNAseq data from hACE2 mouse lungs infected with SARS-CoV-2, we noted upregulation of the Mx2 gene in both infection and homeostasis scenarios using a targeted gene RNAseq approach (unpublished data) (454). In the

same analysis, *M. musculus* also demonstrated significant upregulation of Cxcl10 and Ifng. A limitation of this analysis is the control treatments: P. leucopus was exposed to Vero6 growing media, whereas *M. musculus* was not. Sequencing untreated P. *leucopus* lungs would help determine whether the media is responsible for the observed gene upregulation. However, we report upregulation of Cxcl10.

Fagre et al. report the same cytokine profile in the lungs on days 3 and 6 in *P. sonoriensis*: Isg15 significantly lowered by day 6, and Cxcl10 remains upregulated on both days.

We report that animals ID 25165 and ID 25255, presenting with periorbital edema and ruffled fur, respectively, had significantly higher expression of Arg1 and Slpi and lower expression of Il-12 in lungs (p=8.3 x 10^{-5}, p=0.03, p=0.03 respectively). We also report a nonsignificant change in IL-10 among all the other animals in the day 3

Transcriptional analysis in the brain

So far, there has been no report of any deermice transcriptional analysis of the brain in response to SARS-CoV-2. Significant DEGs in the brain are type I interferon genes like Mx2, Isg15, Irf7, and Cxcl10 on day 6. Besides these ISGs, the Ifi27l2a gene was found to be differentially expressed within the central nervous system upon interferon stimulation or viral infection elsewhere (455). DEGs in *P. leucopus* infected with SARS-CoV-2 correlate with infection with tick-borne *Flavivirus* - Powassan virus, causing life-threatening encephalitis in humans but lacking clinical signs of disease in *P. leucopus*. In early infection in the brain, DEGs in both SARS-CoV-2 and Powasan virus infection are Mx2, Irf7, and Ifitm3 (22).

Similarities between *Peromyscus* and hamster and K18-hACE2 mouse models of COVID-19

Clinical signs following infection of *P. leucopus* and lung pathology are also comparable to results in hamster lungs infected intranasally with similar viral loads (412), (105), (411). Antibody production in *Peromyscus* is comparable with hamster adaptive immune response to SARS-CoV-2 that peaked on

day 7 and returned to homeostasis on day 14 (456). None of the hamsters or P. leucopus died in these experiments. Mortality in humans infected with SARS-CoV-2 is less than 1-2% (457).

Lung histology shows pathology in line with hamster respiratory tract infection (105), (412), (411). Unlike K18-hACE2 mice infected with SARS-CoV-2, neither *Peromyscus* nor hamsters developed fatal diseases associated with lethal neurodissemination after intranasal inoculation of SARS-CoV-2. However, they both show changes predominantly localized in the olfactory epithelium. Hamster studies showed pathology in the olfactory bulb, and *P. leucopus* presented with GO terms pointing to changes in the epithelium of the olfactory bulb. This makes hamster and *Peromyscus* comparable to human COVID-19 disease (458), (413), (459), (460). With long COVID-19 on the rise, with over 65 million cases worldwide of debilitating illnesses affecting multiple organs, *P. leucopus* could be a valuable model for studying infection and infection tolerance (388).

Ecology

P. leucopus occupies a wide variety of ecological niches, some in close proximity to humans, which means they can be considered a reservoir for coronaviruses and other viruses. Moreover, there is usually one or more species of *Peromyscus* in all habitats in North America. The most abundant *Peromyscus* species*, P. maniculatus, P. leucopus, and P. sonoriensis,* warranted a separate genus based on similarities combining their nuclear and mitochondrial genomes (461), (462), (7). Like bats, they tolerate a variety of otherwise serious disease agents (3).

Since the onset of the COVID-19 pandemic, there has been a lingering question of where this zoonosis came from. There is still no definite answer. Viruses or antibodies to SARS-CoV-2 are detected in farm animals, pets, zoo animals, and animals in the wild, including *Peromyscus* species. A spillover from humans to other competent animal species is a viable scenario. To better understand a possible permanent reservoir of SARS-CoV-2 and its immunity to infection, we

analyzed *P. leucopus*, a widespread, genetically diverse and inflammation-tolerant small mammal found in heterogeneous habitats across Northern America.

Study limitations

This study was exploratory and not hypothesis-driven. We looked for differences within the population during the course of infection and compared our results to publicly available mouse and human data. Transcriptome analysis is confirmed using targeted gene analysis, but no reverse genetics and antibodies to specific proteins named in this study were used as there are no such resources for *P. leucopus*. This is why the opus of our work is constrained to sequencing. Animals used in the study are outbred and as diverse as some human populations. This is both a strength and a limitation since we detect a variety of responses, but this means that amplification of the genes can be silenced in the statistical analysis within the group, leading to type II error and false negative results, and overall increase of variances. We calculate the power of analysis based on previous studies, but there might be a higher uniformity of responses in the case of bacterial infection vs. viral inflammation. Low power for such a diverse group of animals could be the reason why we did not detect any differences between sexes or age-associated disease complications.

We analyzed data from the two experiments and saw no significant differences in viral copy number within both experiments in the lung and brain. The parameters of the experiments were consistent in viral titer, the route of infection, and durations. However, we used different viral stocks from two different sources for two experiments; even though they were both the same viral strains and titers, the effect on inflammation could be different. We did not collect the same tissues in two experiments, which lowered the power in some analyses; we only collected feces in experiment 1 and only olfactory bulbs in experiment 2.

Animals were observed daily for signs of distress. There was no significant weight loss except in one case when the animal lost 14% of its body mass. Since there was no other quantitative measure, we

described the severity of the disease qualitatively. Two animals we observed doing poorly were euthanized sooner than expected (the animals were initially randomly assigned to be euthanized on day 6). Moving sick animals to a different category could have artificially made the day 3 group more severe.

Media used for growing SARS-CoV-2 was used as a treatment for our control cohort to remove any effect of the media on the inflammation in the virus treatment group. We found lung pathology in media-treated animals with mild intraseptal inflammation and neutrophils, probably as a reaction to fetal bovine serum (FBS).

Although the animals were housed in the same facility and fed the same diet, we cannot exclude the cage effect before they were separated and used in the study. Deermice are social animals, and separating them causes additional stress. *P. leucopus* are coprophagous, and they share microbiomes within a cage. We showed in Chapter 2 that differences in microbiomes between animals and some microbial species can affect the immune response (463), (464).

Conclusion

P. leucopus is an appropriate model for SARS-CoV-2 and COVID-19. Some animals got sick in this experiment, but there is a spectrum of disease presentation. We propose this is because *P. leucopus* colony animals are genetically diverse. Inflammation in the lungs was more significant than in the brain based on the number of significantly upregulated genes. We detected SARS-CoV-2 RNA in the brain, predominantly in samples containing olfactory bulb.

Materials and Methods

Animals

All animal protocols described were carried out at the University of California, Irvine under Institutional Animal Care and Use Committee (IACUC) approved protocols AUP-18-020 and AUP-21-

007 and in accordance with the *Guide for the Care and Use of Laboratory Animals: Eighth Edition* of the National Academy of Sciences.

Peromyscus leucopus (white-footed deermouse) LL stock originated from 38 animals captured near Linville, NC, and maintained as a closed colony at The *Peromyscus* Genetic Stock Center at the University of South Carolina (246). Animals used for this study were bred without sib-sib matings at the University of California, Irvine, an AAALAC-approved facility, under the Institutional Animal Care and Use Committee (IACUC) approved breeding protocol AUP-21-045 used for all experiments.

Animals designated to get infected with the virus were brought to a biosafety level 3 animal room (ABSL-3) at least 24 hours in advance to acclimate. Infected animals were housed in Tecniplast USA's (Exton, PA) hermetic IsoCage, and negative pressure was maintained even when removed from the rack. All other experiments were performed in the biosafety level 2 (BSL-2) animal room or IsoCages on a rack with forced air to match the conditions in the biosafety level 3 animal room (ABSL-3). All animals were separated for at least 24 hours before being dosed with either virus or media and kept individually in cages with food and water ad libitum. The food supplied was 8604 Teklad Rodent chow (HarlanLaboratories), and we used soft bedding. Deermice were housed at a 12 hours light/dark schedule (lights turn on at 6 am and off at 6 pm) in temperature and humidity-controlled rooms. Animals were monitored daily for signs of disease or discomfort. No animals died before the termination point in both experiments described below. All animal procedures were performed under isoflurane anesthesia, and euthanasia was completed using carbon dioxide intracardiac exsanguination at the termination of the experiment in each animal's home cage. One animal was used untreated and processed the same way. All work with the virus was performed at the biosafety level 3 (BSL-3) conditions.

A priori sample size analysis between groups at α of 0.05 and effect size d of 1.2 computed that 40 animals would achieve a power of 0.95.

Virus

In experiment 1, we used SARS-SoV-2 isolate USA-WA1/2020 from Biodefense and Emerging Infections Research Resources Repository (BEI) (Catalog #NR-52281) obtained from Ilhem Messaoudi, UCI (isolate 1). Isolate received from BEI was passaged two times in Vero E6 cells (ATCC® CRL-1586™, female) before clarification by centrifugation (3000 rpm for 30 min) and storage at −80°C until we used it in the experiment. To determine viral titers by focus forming assay, we used Vero WHO cell line (ATCC® CCL-81™, female). Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma- D5796-500ML) containing 25mM glucose 1% HEPES (Sigma- H3537-100ML) and 2% FBS (Sigma- F0926) and at 37°C, 5% CO_2 (5% DMEM) according to a previously published protocol (465).

In experiment 2, we used the same isolate SARS-CoV-2 USA/WA/2020, originating from Microbiologics (Batch number: G2027B) provided by Lbachir BenMohamed, UCI (isolate 2). The stock was received from Microbiologics at 5.6×10^6 PFU/ml and further propagated to generate high-titer virus stocks. Vero E6 (ATCC® CRL-1586™, female) cells were used according to an earlier published protocol (466).

Experimental infection

We performed experiments in two parts but handled animals and experimental procedures the same way. We had two cohorts: SARS-CoV-2 infection and mock treatment.

In experiment 1, deermice were intranasally infected on day 0 with isolate 1 diluted in the same media used for propagating cells, to 2×10^4 particles in 20µl. During infection, the animals were anesthetized and held upright, and a drop of liquid was deposited in their nostrils. Animals were then observed for inhalation of the droplet in its entirety. They were weighed at day 0 and placed in a cage for daily observation. Animals were euthanized on either day 3 or day 6 (n=4 animals in each group). In the first experiment, we compared virus-infected animals to deermice dosed with DMEM with 2%

FBS as a mock treatment control group and euthanized them on day 3 or day 6 (n=3 animals in each group).

For experiment 2, we repeated the same procedure when infecting animals. We used 2×10^4 particles in 20 µl of isolate 2 for infection. Animals were euthanized on day 3 or day 6 post-infection (n=4 animals in each group). Virus-infected animals were compared to control animals dosed with 20 µl DMEM and 2% FBS (n=4 in each group). Deermice were euthanized with carbon dioxide either at 3 or 6 days post-infection.

In both experiments, each animal was exsanguinated by heart puncture, and blood was pelleted by centrifugation for 3min at 7000 rcf and separated from plasma. Plasma used for ELISA and neutralization assays were treated for 45 minutes at 56°C and collected in a 1.5 ml tube. Lungs, brains, and feces were collected. The right lungs were homogenized and kept in Trizol (TRI reagent, ZYMO Cat: R2050 -1-50), while the right brain and feces were kept in Trizol (Invitrogen TRIzol Cat: 10296-010 Lot). Trizol was shown to inactivate SARS-CoV-2 (467). Part of the right lung was frozen immediately on dry ice and kept at -80°C until TCID50 analysis. Control animals were treated the same way.

Histopathology

The left lung was submerged in 4% buffered paraformaldehyde (PFA) (Sigma HT501128) for 24h. Formalin-perfused tissue was then transferred to 70% ethanol. Formalin-perfused tissues were further paraffin-embedded and sliced to make slides. Slides used for pathology were stained with hematoxylin and eosin, and those not stained were used for RNAscope. We analyzed and scored slides according to the Histopathology scoring system presented in Table 22.

Table 22 Histopathology scoring system

Histopathology scoring system			
Scoring system	Score per field		
	0	1	2
A. neutrophils in the alveolar space	none	1 to 5	>5
B. neutrophils in the interstitial space	none	1 to 5	>5
C. hyaline membranes	none	1	>1
D. proteinaceous debris filling the airspaces	none	1	>1
E. alveolar septal thickening	<2x	2x-4x	>4x
score: [(20xA)+(14XB)+(7XC)+(7XD)+(2XE)]/(numbers of fieldsX100)			
at least 20 random high-power fields evaluated			

RNA isolation

Brain tissue was kept in Trizol at -80°C until RNA extraction. Before RNA extraction, the brain was homogenized, and 50mg was used for extraction. Tissue was vigorously homogenized on Tissue Lyser followed by chloroform precipitation of proteins and lipids. The upper aqueous phase with RNA was removed, and RNA was precipitated with ethanol and treated with DNase I. Further concentration and purification of RNA were performed using RNeasy Mini Kit (Qiagen, Cat: 74104). Lung RNA was isolated using Direct-zol Miniprep kit (Cat: R2051) according to instructions. Fecal pellets were stored in DNA/RNA Shield (Zymo) and frozen at -80°C until further processed. RNA was isolated from 250 mg of fecal pellets in BashingBead Lysis Tubes and S/F RNA Lysis Buffer (Zymo) followed by homogenization using a Bead Ruptor 24 (OMNI). We used a Quick-RNA Fecal/Soil Microbe Microprep Kit (Zymo) to perform the RNA extractions of the fecal pellets. RNA extracts were then stored in RNase-free water and quantified using Qubit fluorometer (ThermoFisher Scientific).

RNAscope *IN SITU* hybridization of SARS-CoV-2 spike RNA

Deermouse lung fixed sections were analyzed using RNAscope 2.5 HD Red assay kit (Advanced Cell Diagnostics, Cat: 322350) according to the manufacturer's instructions. Slides were deparaffinized and treated with H_2O_2 and Protease IV before hybridization. Probes targeting SARS-CoV-2 spike (Cat: 848561), positive-control Hs-PPIB (Cat: 313901) for housekeeping gene positive-control Hs-PPIB

(Homo sapiens control probe for peptidyl isomerase B) for the housekeeping gene, or negative control DapB bacterial protein Bacillus subtilis dihydrodipicolinate reductase (dapB) gene (Cat: 310043) were hybridized followed by proprietary assay signal amplification and detection. Tissues were counterstained with Gill's hematoxylin. As a negative control, we used a deermouse control animal and stained it, and as a positive control, we used mouse brain K18. Tissue staining was visualized using an Olympus BX60 microscope and imaged with a Nikon camera ay 40x magnification.

Focus forming assay

Frozen lungs were homogenized in DMEM media, cell debris was removed by centrifugation, and the supernatant was transferred onto Vero E6 cells (ATCC, C1008) seeded in a 96-well plate, followed by overlay using 1% methylcellulose (MilliporeSigma). After 24 hours, the medium was removed, and the plates were fixed. The number of infected foci was determined using anti–SARS-CoV-2 Nucleocapsid antibody (Novus Biologicals, NB100-56576) and HRP anti-rabbit IgG antibody (BioLegend). Plates were developed using True Blue HRP substrate (Sigma-Aldrich) and imaged on an ELISPOT reader (Autoimmun Diagnostika Gmbh). Each plate included viral stock used in the experiment as a positive control and negative control. We also tested the heat-treated viral sample used in the experiment using a focus-forming assay.

SARS-CoV-2 S protein expression and purification

The day before transfection, 8×10^6 human embryonic kidney (HEK) 293T cells were seeded into 143 cm^2 cell culture dishes (GenClone) and incubated overnight in growth medium (DMEM, 10% FBS,1X Pen/Strep, and 20 mM glutamine). The next day, cells were washed twice with DPBS (10 mL each) and further cultured in BalanCD HEK293 serum-free medium (Irvine Scientific) supplemented with 1X Pen/Strep and 20 mM glutamine. Cells were then transfected with pPP14 SARS CoV2 NFL 2P Foldon-His /polyethylenimine (PEI) complexes formulated in serum-free medium at a DNA/PEI ratio

of 1:3 (i.e., 10 µg DNA and 30 µg PEI per plate). Four to five days post-transfection, SARS-CoV-2 S protein-containing supernatants were collected, cleared (2,500 rpm for 5 min), filtered (0.45 µm), and concentrated with a 50-kDa Amicon Ultra-15 centrifugal filter device (Millipore) according to the manufacturer's recommendation. HIS-tagged SARS-CoV-2 S proteins were then purified over a gravity flow Ni-NTA column. The SARS-CoV-2 S protein expression plasmid pPP14 SARS CoV2 NFL 2P Foldon-His was provided by Dr. Rogier Sanders. According to the manufacturer's recommendations, eluted fractions containing the purified SARS-CoV-2 S protein were concentrated and buffer-exchanged with DPBS using Amicon Ultra 0.5 mL centrifugal filter (50 kDa) units (Millipore). Antibody concentrations were measured at an optical density of 280 nm (OD_{280}) with a NanoDrop spectrophotometer (Thermo Scientific). Purity was determined by SDS-PAGE followed by a simple blue safe stain procedure (Invitrogen) according to the manufacturer's suggestions.

SARS-CoV-2 binding enzyme-linked immunosorbent assay (ELISA)

Half-well area ELISA plates (Corning) were coated with either 100 ng SARS-CoV-2 S protein diluted in 1 X DPBS per well or 50 ng SARS-CoV-2 N protein diluted in 1 X DPBS per well and incubated overnight at 4ºC. Plates were washed three times with wash buffer (1 X DPBS containing 0.5% Tween-20) and incubated with blocking buffer (wash buffer supplemented with 5% non-fat dry milk). After 1 hour at 37ºC, plates were washed twice with wash buffer and further incubated for 1 hour at 37ºC with serial dilutions in blocking buffer of the deer mouse plasma samples. Unbound antibodies were removed by washing the plates three times with wash buffer, and bound antibodies were incubated with an HRP-conjugated goat anti-deer mouse serum IgG diluted in blocking buffer. After 1 hour at 37ºC, plates were washed four times with wash buffer, and bound antibodies were detected by adding 50 µL 1-step TMB solution (Thermo Scientific) per well. The reaction was stopped by adding 50 µL of 1 M H_2SO_4. Plates were read at 450 nm on an ELx808 Bio Tek plate reader. Endpoint binding

titers of deer mouse samples were determined by calculating the highest serum dilution, which gives a reading above the blank, including three standard deviations.

Pseudotyped virus production and neutralization assay

Pseudotyped HIV-1/SARS-CoV-2 S virions were generated by co-transfecting 1.5×10^7 HEK 293-T cells grown in 143 cm^2 cell culture dishes with a single round infectious HIV-1 NL4-3 Gag-iGFP ΔEnv plasmid as well as a SARS-CoV-2 spike protein-expressing plasmid (pcDNA 3.1 SARS CoV-2 S). HIV-1 NL4-3 Gag-iGFP ΔEnv (12455) was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. Benjamin Chen. Plasmids (1 μg plasmid per 1×10^6 cells) were mixed with PEI at a DNA/PEI ratio of 1:3 both formulated in serum-free medium and added dropwise to the cells cultivated in growth medium. After 3-4 days, cell supernatants were harvested and cleared from cells. Virus aliquots were stored at -80°C. 96-well tissue culture flat bottom plates (GenClone) were used to mix heat-inactivated (45 min at 56 °C) and serially diluted deer mouse plasma samples with pseudotyped HIV-1/SARS-CoV-2 S virions. After 1 hour at 37°C, 2.5×10^4 ACE2 expressing HEK 293T cells were added, and plates were further incubated for 2 days at 37°C. The human embryonic kidney cells (HEK-293T) expressing human angiotensin-converting enzyme 2, HEK-293T-hACE2 cell line (NR-52511) were obtained through BEI Resources, NIAID, NIH. The supernatant was removed, and cells were detached by adding 100 μL of Accutase (Innovative Cell Technologies) per well. The cells were transferred into a 96-well round bottom plate, which was washed once with DPBS and fixed with 4% PFA. After 30 min at 4 °C, ACE2 HEK 293T cells were analyzed by flow cytometry for green fluorescence protein expression. The reduction of the median fluorescence intensity compared to the virus control was used to determine the ID_{50} of the tested plasma samples.

SARS-CoV-2 quantitative PCR (qPCR)

For SARS-C0V-2 qPCR, genomic RNA from selected deer mouse tissues, including lung and brain tissue, was isolated and used as a template for amplification. Per reaction (25 μL total) 500 ng of

genomic RNA was used in combination with 0.1 µM of each SARS-CoV-2 N-specific primer (5'-GGGGAACTTCTCCTGCTAGAAT-3' and 5'-CAGACATTTTGCTCTCAAGCTG-3') as well as the reagents from the qPCRBIO SyGreen 1-step Go Hi-ROX kit (PCRBIOSYSTEMS). The qPCR was carried out with a Rotor-Gene 6000 series cycler (QIAGEN) for 1 cycle at 50ºC (30 min), 1 cycle at 95ºC (15 min) and 45 cycles at 94ºC (15 sec) and 60ºC (20 sec). SARS-CoV-2 copy numbers were quantified and reported based on a standard curve generated by serial dilutions of SARS-CoV-2 viral RNA using the Rotor-Gene 6000 series software. All samples were run and analyzed in triplicates.

RNAseq

Lung and brain isolated RNA was sequenced using Illumina technology. The brain RNA library was prepped stranded with rRNA depletion. The stranded mRNA library was normalized, multiplexed, and sequenced on the NovaSeq6000 instrument at the UC Irvine genomic High Throughput Facility. We used paired-end chemistry and 150 cycles to achieve 40 million reads per sample. Lung RNA was stranded, as well as PE150 sequenced on the same instrument, to obtain 40M reads. The quality of sequencing reads was analyzed using FastQC (Babraham Bioinformatics). The reads were trimmed of low-quality reads (Phred score of <15) and adapter sequences, and corrected for poor-quality bases using Trimmomatic (214). A number of reads for each sample is provided in Table 23, with an average of 7.3×10^7 reads for the brain, and 6.8×10^7 reads for the lungs.

Table 23 Numbers of reads in lungs and brain

Number or reads for each tissue per sample						
Animal ID	Experiment	Treatment	Day	Sex	Reads brain	Reads lung
25160	1	C	3	F	8.1E+07	3.4E+07
25169	1	C	3	M	8.6E+07	4.8E+07
25124	1	C	3	M	7.4E+07	3.9E+07
25159	1	C	6	M	8.8E+07	4.1E+07
25167	1	C	6	F	6.7E+07	3.9E+07
25018	1	C	6	F	8.6E+07	4.3E+07
25117	1	V	3	M	7.4E+07	4.6E+07
25130	1	V	3	M	6.7E+07	4.3E+07
25013	1	V	3	F	7.8E+07	4.5E+07
25025	1	V	3	F	6.2E+07	4.7E+07
25157	1	V	6	M	7.9E+07	4.4E+07
25161	1	V	6	M	7.2E+07	4.0E+07
25128	1	V	6	F	9.6E+07	4.3E+07
25158	1	V	6	F	6.6E+07	4.1E+07
25225	2	C	3	F	5.2E+07	9.7E+07
25226	2	C	3	F	8.3E+07	8.9E+07
25281	2	C	3	M	6.4E+07	8.6E+07
25152	2	C	3	M	7.8E+07	8.9E+07
25287	2	C	6	M	7.5E+07	9.1E+07
25219	2	C	6	F	7.5E+07	8.3E+07
25313	2	C	6	F	7.5E+07	8.8E+07
25280	2	C	6	M	5.3E+07	9.7E+07
25165	2	V	3	M	7.6E+07	8.5E+07
25255	2	V	3	M	7.6E+07	9.3E+07
25298	2	V	3	F	7.2E+07	9.0E+07
25304	2	V	3	F	8.9E+07	7.9E+07
25136	2	V	6	M	4.9E+07	9.4E+07
25306	2	V	6	F	8.3E+07	8.2E+07
25163	2	V	6	M	5.6E+07	9.5E+07
25228	2	V	6	F	5.2E+07	9.4E+07

For this experiment, reads were mapped to the genome transcript reference set of 54,475 based on the reference genome transcript on GenBank version GCF_004664715.2 on CLC Genomics workbench version 23 (Qiagen) (468). The settings for PE150 reads were as follows for both strands: length fraction of 0.35, similarity fraction of 0.9, and costs for mismatch, insertion, or deletion of 3.

Genome-wide differential gene expression

Differential expression between experimental conditions was assessed using RStudio 1.3.1093 and edgeR version 3.42.4 software package for the differential expression analysis of digital gene expression data, that is, of count data arising from DNA sequencing technologies. It is specially designed for differential expression analyses of RNA-Seq or SAGE data or differential marking analyses of ChIP-Seq data (319). Fold changes in TPM (transcripts per million) were log2 transformed. The False Discovery Rate (FDR) with corrected p-value was estimated by the method of Benjamini and Hochberg (227).

Gene ontology term analysis

Differentially expressed genes following EdgeR analysis and meeting the criteria of a FDR p-value < 0.05 and |log2 FC|>1.5 for lungs and p-value < 0.05 and |log2 FC|>1.5 for brain were used in each cohort. The analysis was implemented with the tools of Metascape (380). Functional enrichment analysis was carried out first with the hypergeometric test and FDR p-value correction by the method of Benjamini and Hochberg (227). Then, pairwise similarities between any two enriched terms were computed based on a Kappa-test score (469). Similarity matrices were then hierarchically clustered, and a 0.3 similarity threshold was applied to trim resultant trees into separate clusters. The lower the p-value, the less the likelihood that the observed enrichment is due to randomness (382). The lowest p-value term represented by each cluster is shown in the horizontal bar graph. Besides the terms beginning with 'GO' and referring to the Gene Ontology resource (http://geneontology.org), others 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.

Targeted RNAseq analysis

RNA-seq of the selected set of protein-coding sequences (CDS) listed below was carried out using CLC Genomics Workbench v. 23 (Qiagen). Paired-end reads were mapped with a length fraction of 0.35 for ~150 nt reads and 0.40 for ~100 nt reads, a similarity fraction of 0.9, and costs of 3 for mismatch, insertion, or deletion to the CDS gene set. Preliminary expression values were unique reads normalized for total reads across all the samples without adjustment for reference sequence length, as described in Chapter 3. For comparisons, we normalized reads based on overall reads and based on Gapgh. This is described in more detail in Results. Following the recommendation of Hedges et al. we used the natural logarithm (ln) of ratios (383).

The target CDS were as follows: *Acod1, Akt1, Akt2, Arg1, Bcl3, Camp, Ccl2, Ccl3, Ccl4, Cd14, Cd177, Cd3d, Cd4, Cd69, Cd8, Cfb, Cgas, Csf1, Csf1r, Csf2, Csf3, Csf3r, Cx3cr1, Cxcl1, Cxcl10, Cxcl2, Cxcl3, Dhx58, Fcer2, Fcgr2a, Fcgr2b, Fcgr3, Fgr, Fos, Fpr2, Gapdh, Gbp4, Glrx, Gzmb, Hif1a, Hk3, Hmox1, Ibsp, Icam, Ifih1, Ifit1, Ifng, Il10, Il12, Il18, Il1b, Il1rn, Il2ra, Il4ra, Il6, Il7r, Irf7, Isg15, Itgam, Jak1, Jak2, Jun, Lcn2, Lpo, Lrg, Lrrk2, Ltf, Mapk1, Mmp8, Mmp9, Mpo, MT-Co1, Mt2, Mtor, Mx2, Myc, Myd88, Ncf4, Nfkb1, Ngp, Nos2, Nox1, Nr3c1, Oas1, Olfm4, Padi4, Pbib, Pkm, Ptx, Ptprc, Retn, Rigi* (*Ddx58*), *S100a9, Saa3, Serpine1, Slc11a1, Slpi, Socs1, Socs3, Sod2, Stat1, Stat2, Stat4, Steap1, Sting, Tgfb, Thy1, Timp1, Tlr1, Tlr2, Tlr4, Tnf, Tnfrsf1a*, and *Tnfrsf9*. The sources for these coding sequences were the reference genome transcript sets for *P. leucopus* listed above. If there were two or more isoforms of the mRNAs and the amino acid sequences differed, the default selection for the coding sequence was the first listed isoform.

Statistical methods

Statistical analysis, namely student t-test, graphing, and ANOVA, was performed on GraphPad Prism version 10.2.2. Edwards-Venn diagram was generated using an online visualization tool (470). Volcano graph was generated using EnhancedVolcano version 1.18.0 (471).

Data availability

Sequencing data as fastq files of Illumina reads (SRA), along with descriptions of the samples (BioSamples) they are associated with, have been deposited with NCBI under BioProjects PRJNA1026327, PRJNA1026365. Data generated and analyzed for this study and included in the manuscript are deposited on Dryad database for lungs and brain separately. Data upload pending.

CHAPTER SIX: Conclusion and Future Directions

This study's main objective was to illustrate how *Peromyscus leucopus* has adapted to enhance tolerance and dampen inflammation. Our interest in *P. leucopus* was prompted by the species' status as a competent zoonosis reservoir suffering little morbidity from infection and living up to eight years. Throughout our broad and unbiased *P. leucopus* research, we investigated how *P. leucopus* avoids morbidity and mortality when infected with pathogens that are disabling or lethal for humans. Since the publication of *P. leucopus* genome sequencing and gene annotation in 2019, we have taken advantage of these tools to understand how a tolerant reservoir handles infection (5). Access to its annotated genome enabled us to proceed with transcriptome analysis and comparison between rodent species in Chapters 3 and 4 and to investigate immunity in viral infection in Chapter 5. In Chapter 2, we focused on another aspect of *P. lecusopus'* "omics," which describes the microbiota of this rodent. To better define and highlight the microbiome features of *P. leucopus* and contributors in reaction to pathogens, our experimental approach compared *P. leucopus* to rodents, widely used in research but not known for longevity or tolerance. By integrating genome sequencing, transcriptome analysis, and microbiota characterization, our study sheds light on the robustness of *P. leucopus* against infections, offering insights into its potential implications for understanding disease tolerance and management in wildlife reservoirs.

Microbiome and tolerance

Microbes are not the sole drivers of mortality and morbidity—the degree of the reaction to microbes determines the severity of the disease. We show that *P. leucopus* has a diverse microbiome and virome, harboring endogenous retroviruses. Most microbes do not cause disease in *P. leucopus* in any measurable or visible way, as the immune system controls them. Only a portion of microbes elicit a host response. In *P. leucopus*, this response is highly regulated and antagonized by an amplitude

of tolerance factors that dampen the antimicrobial response. We compare and discuss *P. leucopus* to other rodent specie and bats and highlight its features. We then examine the acquired knowledge from this study in the context of tolerance.

P. leucopus microbiome and other rodents

Rodents form the largest group of mammals in terms of diversity, and some synanthropic species have been involved in pandemics caused by zoonotic diseases, such as the 14th-century black plague (*Yersinia pestis*). However, despite their role in spreading the plague, rats were not tolerant to the *Y. pestis* bacterium and typically died from the disease within a week (472). Mice and rats do show some tolerance traits in laboratory conditions. Still, after embryonal transfer into wild animals, mice exhibit diminished tolerance since the microbiome of the gut and epithelial barrier changes (473), (474). Hence, the tolerance and gut microbiome are connected, and the microbiome in mice is influenced not only by environmental factors but also strongly by genetic factors (475), (476). The study provides a foundational understanding of the gastrointestinal microbiota of *P. leucopus*, highlighting its differences and similarities to more commonly studied rodents. We report that *P. leucopus* has higher diversity within individuals (alpha diversity) and between them (beta diversity) than *M. musculus*, particularly in the numbers of lactobacilli.

P. leucopus harbors a diverse gastrointestinal microbiota, including a notable abundance of *Lactobacillus* species. Four distinct *Lactobacillus* species were isolated and studied in our study, revealing differences in abundance and distribution between *P. leucopus* and *M. musculus* (figure 9). Notably, female *M. musculus* had 100-fold more reads mapping to lactobacilli than males, while in *P. leucopus*, the sex differences were not as pronounced (figure 15) as confirmed in the natural population in this study and previously by Baxter et al. (138). We report that *Lactobacillus johnsonii* is the predominant lactobacilli species in *M. musculus* (figure 15), confirming the findings of Buhnik-Rosenblau et al. Further, we report that *P. leucopus* has a higher alpha diversity of not only overall

microbial species but also lactobacilli (figure 16) (477). Among *P. leucopus* lactobacilli, the study identified host-specific microbial species such as *Lactobacillus peromysci sp, nov.* found in the colony and wild *P. leucopus* but not in *M. musculus* (figure 22). As is the case with hamsters' *Lactobacillus hamsteri, P. leucopus* has species-specific lactobacilli - *L. peromysci sp. nov.,* found in both colony animals or nature. *Peromyscus eremicus* colony animals and *Peromyscus boyliii* did not have this lactobacillus in the gut when cultivated and sequenced for 16S fragment (unpublished data). DNA of isolated colonies from these *Peromyscus* species could have novel species-specific lactobacilli. If there are indeed species-specific acclimated lactobacilli, there is the potential for developing targeted interventions for zoonotic disease control, like bait vaccines targeting specific *Peromyscus* species.

P. leucopus and bat microbiome and in the context of tolerance

After rodents, bats account for the second-highest number of mammalian species; they are the only mammals capable of flying, covering long distances, and occupying diverse niches (478). They are also considered a perfect host for emerging zoonotic pathogens since, like *Peromyscus,* they exhibit tolerance. Their microbiomes change with diet and thus differ between species and habitats (122), (478). Bats' stomachs are globular without specialization. Some bats lack caecum and glands in certain regions; overall, they have decreased intestinal tissue and reduced carried food loads with a higher density of villi supporting absorption (479), (480). This differentiates them from *Peromyscus*, which has a morphologically distinct caecum. However, like bats, *Peromyscus* have localized glandular tissue covering only a fraction of the stomach. Although these two mammals are known as zoonosis reservoirs with exceptional tolerance, they do not share the same microbial constituents in their gut. Their similarities lie in how both species allow for microbiome diversity and adapt to environmental changes. In the case of *Peromyscus*, there has long been a critical knowledge gap since deep microbiome analysis was unavailable. Further microbiome research is crucial for

developing strategies to manage the transmission of zoonotic diseases, particularly in areas where *P. leucopus* is prevalent.

P. leucopus microbiota and tolerance – tryptophan pathway

Determining the causal relationship between gut microbiota and mammal immunity was beyond the scope of this work, but since this relationship is usually studied using germ-free animal models, our study could contribute to future research directions. The impact of microbial colonization on immune cell repertoire and regulation could also be investigated by administering broad-spectrum antibiotics like vancomycin, neomycin, ampicillin, and metronidazole.

A more developed understanding of *Peromyscus* gut immunity and microbiota could fill a gap in knowledge of recently described intestinal innate lymphoid cells (ILCs) which play a critical role in microbiota and maintain homeostasis (481). Innate lymphoid cells sense changes in microbiota and diet indirectly through local dendritic and intestinal epithelial cells, while in return ILCs produce cytokines recruiting B and T cells (482). Dendritic cells of the intestines express Toll-Like Receptors (TLRs) and react to bacterial pathogen-associated molecular patterns (PAMPs) by producing IL-12 or type I interferon in stimulation with viruses. Although ILCs do not sense PAMPs, they have other environmental sensors, such as the aryl hydrocarbon receptor (Ahr). Ligands for Ahr originate from either diet or microbial activity and include metabolites in tryptophan catabolism indole and its derivatives, such as indole-3-aldehyde (IAld), indole-3-acetic-acid (IAA) and indole-3-propionic acid (IPA) (483). Indoles promote anti-inflammatory effects in tissues colonized by commensal microbiota like intestinal lumen, skin, lungs, and other mucosal environments (484), (483), (482), (485), (486), (487), (488), (489), (490). The binding of indoles to Ahr on T cells, dendritic cells, macrophages, or ILCs causes translocation of Ahr to the nucleus and acts as a transcription factor for IL-10, transforming growth factor (TGF)-α, TGF-β, and arginase (Arg1) (483), (491). Bacterial and mammalian metabolic pathways converge in the conversion of tryptophan to indole metabolites,

facilitating tolerance at the organismal level. Tryptophan is oxidized in mammals via kynurenine pathway by indoleamine-2,3-dioxygenase (Ido1) expressed in various organs (490), (487), (485). Bacteria initially metabolize tryptophan into indoles using the enzyme tryptophanase (TnaA). TnaA can be found in Gram-negative and Gram-positive bacterial species, including *Escherichia coli* and *Bacteroides sp*. (488), (483), (490). We detected TnaA in *P. leucopus'* isolated *E. coli* LL2 assembly GCF_005869145.1, as well as in the whole metagenome across both colony and wild animals. We mapped it predominantly in the *Bacteroidaceae* family, but also in *Clostridiales* and *Elusimicrobiaceae* when mapped against Kyoto Encyclopedia of Genes and Genomes (KEGG) function on MG-rast platform (125). Besides *Bacteroides,* Hezaveh et al. show that *M. musculus* C57BL gastrointestinal tract lactobacilli can produce Ahr indole ligands when tryptophan is introduced as an energy source (483), (492), (493). Hezaveh et al. indicate that *L. johnsonii, L. reuteri, L. intestinalis,* and *L. murinus* produce several indole derivatives. *P. leucopus* lactobacilli do not have annotated genes for tryptophan metabolism, but growing these bacilli with tryptophan as a carbon and energy source could cause phase variation. Notably, the lactobacilli we sequenced were not growing in their natural environment but in MRS media instead and could have already altered their genomes due to the change in environment. Further investigating this phenomenon could be a future research direction.

We report further convergence of tolerance, gut microbiome, and diet in *P. leucopus* in tryptophan metabolism and kynurenine pathway in Chapter 3. The kynurenine pathway is a highly regulated arm of the immune system promoting immunosuppression and tolerance. IDO1 is the initial rate-limiting inducible enzyme of the kynurenine pathway (Figure 1) (494), (495). During sepsis, Ido1 gene transcription is upregulated either directly by lipopolysaccharides (LPS) or indirectly by cytokines, mainly interferon γ (Ifn- γ) (496).
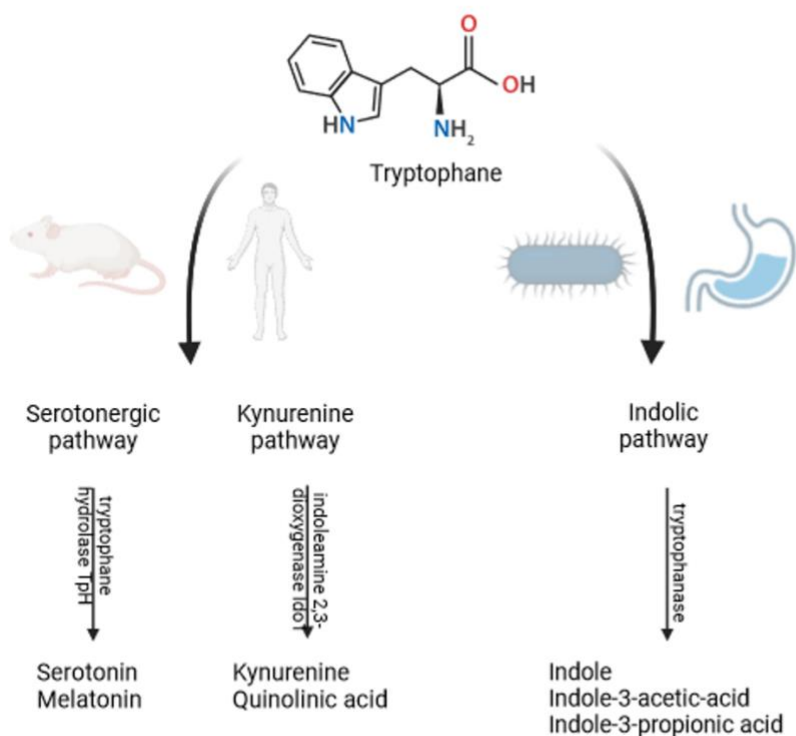
Figure 61 Tryptophan catabolism pathways and products in eukaryotes and procaryotes
The figure shows main catabolic pathways of tryptophan: serotonergic, kyneurine, and indolic. Created by BioRender.com

Under the same LPS-induced conditions, *P. leucopus* expresses Ido1 64, 15, 15, and 15 times more than controls in blood, spleen, liver, and adrenal glands, respectively (Dryad WGCNA, and adrenal RNAseq unpublished data). *M. musculus* does not show significant differential expression of Ido1 under these conditions. We report the same trend in Chapter 5, in which we describe how lungs infected with SARS-CoV-2 also differentially express Ido1 four times more than controls three days post-infection. We recapitulate the activity of Ido1 in Chapter 3 and report a significant depletion of Ido1 substrate-tryptophan levels in the plasma of *P. leucopus*. At the same time, in *M. musculus,* Ido1 did not change significantly (figure 26).

We further report in Chapter 3 that tryptophan metabolites IAA and IPA found in plasma and originating from microbial activity (497), (489), (498), (499) are significantly reduced in *P. leucopus*

treated with LPS compared to controls (student t-test p = 2.5 x 10^{-5}, p=2.8 x 10^{-7}, respectively) and did not significantly change in *M. musculus* under the same conditions. IAA attenuated inflammation and the effects of LPS by directly neutralizing free radicals or as through binding to Ahr (500), (489). IAA's effect on free radicals and nitric oxide (NO) production is direct. At the same time, it acts as a free radical scavenger, but it also suppresses inducible nitric oxide synthase (Nos2) expression induced by LPS (500). We conclude that IAA and IPA depletion in the plasma of *P. leucopus* is due to the higher efficacy of anti-oxidative and anti-inflammatory activity. At the same time, in the case of IPA, *P. leucopus* had a higher initial concentration in the plasma of control animals.

Although we detected significant differential expression of Ido1 in several tissues and decreased concentration of tryptophan and its catabolites in plasma, we cannot ascertain whether their origin in plasma is mammal or microbial. Methods and studies understanding the origins of tryptophan derivatives are limited (501), (502). Future directions determining tryptophan metabolism and their roles in infection tolerance can include analysis of species-specific compounds in the blood that are unique to either mammal or microbial enzyme products. Other options include isotope labeling, gene knockout, and silencing genes in tryptophan metabolism and antibiotic and gnotobiotic animal studies. So far, the literature review has not shown concrete evidence of an association between tryptophan metabolites and septic shock survival (496). Due to the direct causal relationship of indoles, Aht, and the production of genes involved in alternative macrophage polarization and tolerance, finding the ligand with the most significant effect could change outcomes in patients with septic shock. Although there is interest in tryptophan metabolites primarily due to its role in the gut-brain axis, serotonin production, and appetite suppression, there is still a gap in knowledge regarding its applications in inflammatory diseases (501), (489), (490). However, the opposite approach should be adopted for cancer treatment as it was shown that tryptophan and derivatives promote disease in cancer (483).

P. leucopus and protozoa

Besides bacteria, protozoa are an integral part of the mammalian microbiome, although not nearly as commonly studied. We report a high density of tritrichomonas flagellated protozoan in the cecal fluid of *P. leucopus* LL stock. Recent work implicates tritrichomonas in mucosal immunity (177).

This comprehensive analysis of *P. leucopus*'s gastrointestinal microbiota enhances our understanding of its unique microbial ecosystem, highlighting commensal microorganisms that co-evolved with *P. leucopus* toward mutualism and homeostasis. Our analysis sets the stage for developing targeted interventions to control zoonotic disease transmission in North American ecosystem. Subsequent parts of the dissertation focused on *P. leucopus'* response to infection and inflammation stimuli.

Countering the inflammatory effects of endotoxin

The study aimed to understand the mechanisms behind the infection tolerance observed in *P. leucopus*, a reservoir host for several zoonotic agents, including the agents of Lyme disease. By comparing the responses of *P. leucopus* and *M. musculus* to TLR4 ligand-LPS, we found that both species exhibited sickness behavior within the same time period and experienced similar increases in stress hormone and pro-inflammatory cytokine levels. However, the *P. leucopus* response was characterized by a dominance of neutrophil activity and alternatively activated macrophage response, indicative of a more regulated and less damaging inflammatory response.

Slpi and tolerance

Network analysis indicated that the *P. leucopus* response to LPS involved neutrophil activity as well as a low Nos2/Arg1 and high Il-10/Il-12 ratio, contrasting with the cytokine-associated response in *M. musculus*. One of the neutrophil activity drivers is the secretory leukocyte protease inhibitor (Slpi). Slpi has diverse biological roles, particularly in inflammation and innate immunity, by controlling excessive protease activity in inflammation. Slpi is a small, non-glycosylated protein with 131 amino

acids in *P. leucopus*. Its expression is not limited to leukocytes; it is also found in mucosal secretions synthesized by various epithelial cells, especially those lining the respiratory, reproductive, and gastrointestinal tracts. Possibly due to its small size, Slpi can cross membranes without a receptor (291). By contrasting the immune response patterns of *P. leucopus* with *M. musculus*, particularly in the context of LPS stimulation, the pivotal role of Slpi in moderating immune dynamics becomes increasingly evident.

Knockout mice (Slpi -/-) were used to understand the function of SLPI on a systemic level (292). When exposed to pathogens or inflammatory stimuli, these mice may exhibit enhanced inflammatory responses and tissue damage followed by aberrant innate and adaptive immune responses. This is due to the absence of SLPI's protease inhibitory and anti-inflammatory function. Increased protease activity can delay wound healing since increased elastase levels maximize collagen degradation (292).

The primary function of SLPI is to inhibit serine proteases, such as neutrophil elastase, which degrades extracellular matrix proteins. By inhibiting these proteases, SLPI protects tissues from damage during inflammatory responses. SLPI's anti-inflammatory properties reduce the production of pro-inflammatory cytokines and inhibit the activation of nuclear factor kappa B (NF-κB). Lentsch et al. showed a causal relationship between SLPI and NF-κB, effectively proving that SLPI regulates endogenous lung inflammation (503). In human asthma patients' lungs and a mouse model of severe asthma, interferon-gamma (Ifn-γ) expression is inversely correlated with Slpi. In a mouse model of severe asthma, overexpressed Slpi reduced signs of inflammation (504). Alterations in Slpi expression and function can have clinical implications. In Chapter 5, we also report that Slpi has four times higher expression in the lungs of *P. leucopus* infected with SARS-CoV-2 than controls, and that Slpi is significantly expressed in *P. leucopus* lungs exhibiting moribund state. Although in Chapter 3 we report 1,280 times ($p = 10^{-57}$) higher expression of Slpi in the blood of LPS-treated *P. leucopus*, Slpi

is typically not characterized as a systemic antiprotease, and it is primarily produced at local sites in response to inflammatory cytokines and bacterial products (505). Its prolonged presence in the systemic circulation could be a sign of chronic inflammation and a biomarker of diseases, as Sawicki et al. report in patients with heart failure (506). Besides immune suppression, SLPI possesses antimicrobial activity against bacteria (507), (508), fungi (509), and viruses (510), (511), contributing to the innate immune defense, especially in mucosal surfaces, facilitating tissue repair (292), (505). SLPI has an additional anti-inflammatory role: it inhibits neutrophil extracellular traps (NETs) formation since SLPI blocks cleavage of neutrophil elastase substrate histone H4 and blocks decondensation of chromatin necessary for expelling DNA and forming extracellular nets (290). SLPI is involved in tissue repair and regeneration, as it promotes epithelial cell growth and migration, facilitating the healing of mucosal surfaces (512). In that context, human recombinant Slpi controls pulmonary infection with *Pseudomonas aeruginosa* or systemic inflammation in *M. musculus* with LPS and lipoteichoic acid mimicking infection with Gram-negative and Gram-positive bacteria, respectively (513) Aerosolized SLPI was successfully used in patients with cystic fibrosis (514), (515). Ongoing research continues to unravel the multifaceted roles of Slpi in health and disease, making Slpi a potential target for therapeutic intervention, particularly sepsis.

SLPI garnered attention as a therapeutic agent due to its multifaceted roles in inflammation, immune response modulation, tissue protection, and short peptide structure. Human SLPI has 107 amino acids, and *P. leucopus* peptide has 131 amino acids. The N-terminal domain has modest antimicrobial activity, while the C-terminal domain is an inhibitor of neutrophil elastase, cathepsin G, trypsin, chymotrypsin, tryptase, and chymase (512). SLPI domains are 50 amino acids each, similar in size to insulin, which is widely used as a peptide therapeutic. The use of SLPI in modulating local and systemic inflammation should be considered in acute and short-term conditions, as the peptide is implicated in the progression and metastasis of cancer (505), (516). The balance between

immune surveillance of malignant progression and anti-inflammatory effect is crucial, like our reporting on other tolerance factors of innate (tryptophan metabolites) and adaptive immunity (Tigit-highlighted gene in Sars-CoV-2 infection). We reported Slpi upregulation in *P. leucopus* four hours post-LPS exposure in Chapters 3 and 4, twelve hours in Chapter 4, and three days in the case of viral infection. We propose that Slpi upregulation is transient as *P. leucopus* is known for longevity and shown to live up to eight years in captivity (7), (10), (23). Further, it shows reduced mitogenic activity in response to carcinogenic stimulation in comparison to other *Peromyscus* species (517), (518). Hence, *P. leucopus* has found a delicate balance between immune regulation of tolerance and surveillance of malignant cells. Though our understanding of immunity in this zoonotic reservoir continues to develop, there are still gaps in our knowledge and research tools related to *P. leucopus*.

Slpi recapitulation and future directions

Our study focused on transcriptome analysis and used sequencing tools and bioinformatics. We have not recapitulated Slpi expression (or other reported genes) using antibodies due to the resource constraint. Synthesizing antibodies and producing knockout animals would be a future direction.

We also report the upregulation of terms for gene ontology of neutrophils in *P. leucopus*; gradient-based cell separation and analysis confirm this. So far, we have not isolated neutrophils to verify that this is indeed the origin of differentially expressed genes driving gene ontology (GO) terms. Isolating neutrophils is challenging due to the same antibody constraint and uniqueness of the species.

Despite the unavailability of the antibodies, there is a technical advantage to working with *P. leucopus*, a genetically diverse animal model. Due to the genetic diversity between individuals, we can attempt multiplexing animals across conditions and analyze single-cell RNAseq in different tissues (unpublished data). Each animal would have low-pass short-read DNA sequenced and imputed genotypes to demultiplex single-cell RNAseq data (261). After dimensionality reduction, clustering, and annotation, analyzing cell clusters is still challenging since we do not have a database

of known markers for cell populations. *In vitro* studies would resolve this and, in combination with acquiring new antibodies, could establish a *Peromyscus* database of known marker genes. Gene markers overlapping between humans, mice, and primates is a place to start.

In addition, we focused our analysis on inter-species differences among rodents, and while we used orthologous genes, we do not have complete corresponding gene isoforms of *P. leucopus*. This could be achieved by using packages for identifying isoforms based on long-read transcriptome data sets like Talon and visualized with Swan (519), (520). Identifying isoforms is a future direction, and sequence data from different platforms are already available. While these are some of the suggested future directions, in Chapter 4, we recapitulate findings from Chapter 3 by following up transcriptome analysis with hematology studies and including additional animal models while proposing further ways *P. leucopus* tempers interferon responses.

Tolerance and endogenous retroviruses

The main findings presented in Chapter 4 are that in comparison to two other rodent species, *M. musculus,* and *Rattus norvegicus, P. leucopus* showed less transcription of interferon-gamma and lower activity in type 1 interferon-stimulated genes, suggesting a unique mechanism of infection tolerance in comparison to other two rodent species, *M. musculus,* and *Rattus norvegicus*. We confirmed that CD1 outbred *M. musculus* also presents with upregulated genes related to classically activated monocytes, including *M. musculus* BalbC in Chapter 3. We recapitulated findings from Chapter 3 that neutrophils' GO terms and network analysis point to neutrophil influx, as hematology analysis of *P. leucopus* blood shows a higher neutrophil to lymphocyte ratio. This ratio is lower in *M. musculus* and *R. norvegicus*. We report that exposure to LPS increases the transcription of endogenous retrovirus (ERV) sequences in *M. musculus* and *R. norvegicus* and, to a lesser extent, *P. leucopus*. This response involves surface TLR4 and cytoplasmic pattern recognition receptors (PRR) like RIG-I, linking ERV activation to innate immune responses. *P. leucopus*, however, shows subdued

ERV and interferon-stimulated gene responses, possibly due to evolved mechanisms that silence ERV expression. This suggests a unique infection tolerance strategy (377), (521).

ERVs are ancient viral remnants embedded within the genomes of all vertebrates, including humans, mice, and deermice (522), (359), (5). Over millions of years, these viral sequences have lost their infectious capabilities but have integrated into the germline while retaining the ability to influence the host genome and immune system (523), (377). Once presumed to be inactive, it has recently become apparent that some ERVs express intact open reading frames, which we report in *P. leucopus*. ERVs are typically silenced by host defense mechanisms; however, environmental factors, stress, or immune dysregulation can reactivate them, leading to the transcription of viral genes and the production of viral proteins. ERV might be a way to communicate between the immune system and environmental changes such as perturbations in the microbiota (346)

The reactivation of ERVs can have profound effects on the host's immune system. For instance, the production of viral proteins by ERVs can mimic an ongoing viral infection, triggering an immune response even without an actual virus (524). ERV transcription can lead to chronic inflammation as the immune system continuously reacts to these endogenous viral elements. Moreover, ERVs can influence the expression of nearby host genes, including those involved in immune regulation, thereby modulating the immune response. Understanding the precise mechanisms by which ERVs influence inflammation and the immune response is critical for developing new therapeutic strategies to treat ERV-associated diseases. This includes the potential for antiretroviral therapies, currently used in HIV treatment, to be repurposed for treating diseases associated with ERV reactivation and consequent inflammation (525), (526).

Inflammation-induced by ERV activation is a double-edged sword. On one hand, it represents an ancient defense mechanism, keeping the potentially harmful ERV activities in check. On the other hand, if uncontrolled, ERV expression can contribute to the pathogenesis of various chronic

inflammatory and autoimmune diseases. The molecular mechanisms often involve the activation of PRRs by ERV-derived nucleic acids, leading to the production of type I interferons and other pro-inflammatory cytokines. Kwon et al. report this phenomenon in LPS-treated mice leucocytes and we report the same in Chapter 4 in the blood of *M. musculus* and *R. norvegicus* (523). However, we do not see this trend in *P. leucopus*, although we have identified *gag-pol* and *env* sequences homologous to mammalian ERVs. Gozashti et al. analyzed the landscape of retroviruses *P. maniculatus* (359). This is another recorded tolerance factor utilized by *P. leucopus* (figure 62).

There is still a gap in knowledge about the origins of ERVs in *P. leucopus*. We have mapped mammalian ERVs to the *P. leucopus* genome and we report that metagenome analysis shows an abundance of bacteriophages as well as an abundance of retroviral sequences, identified as Beta- and Gamaretroviruses. However, there was no comprehensive analysis of mapped retroviruses to the genome or the position of ERVs on the genome in areas of transcription-active euchromatin or heterochromatin (527). Additionally, we do not know if ERVs are like in the case of *P. maniculatus*, organized in "hot spots" and enriched in Kruppel-associated box (KRAB) domains, like in the case of *P. maniculatus*. If so, this could be the reason for silencing in inflammation, unlike in *M. musculus*. Research into the role of ERVs in inflammation is still evolving, with studies suggesting that ERV-driven inflammation might be a more common contributor to disease than previously thought.
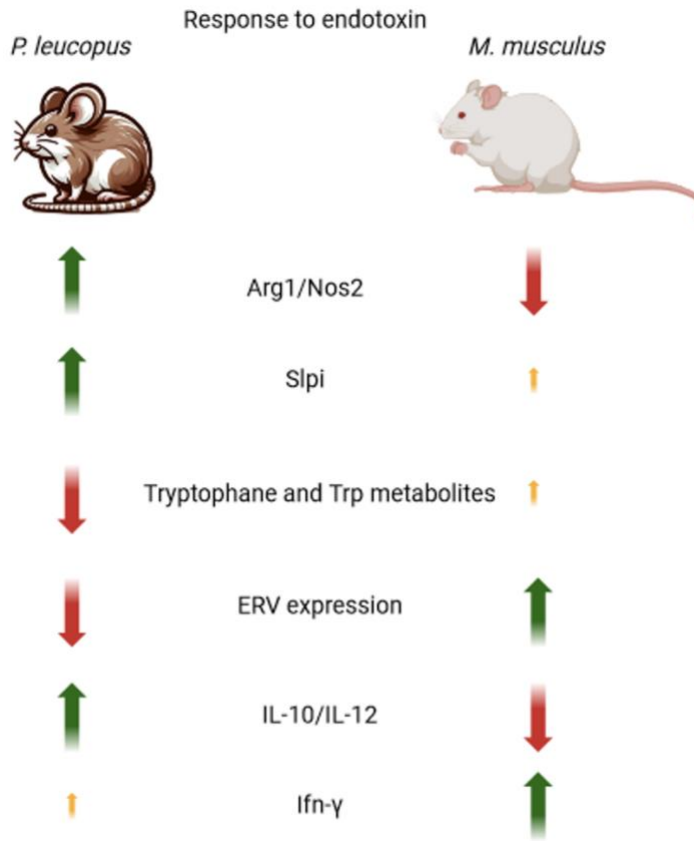
Figure 62 Factors supporting immune tolerance in P. leucopus and promoting inflammation in M. musculus
P. leucopus show tolerance in response to endotoxin in terms of Arg1, Nos2, Il10, Il12, Slpi, Ifng, ERV proteins and Tryptophan metabolism compared to *M. musculus*. Created by BioRender.com.

Implications of infection of *P. leucopus* with SARS-CoV-2

This study discusses variable responses of *P. leucopus* to coronavirus infection. We highlight that while *P. leucopus* is a tolerant reservoir of zoonoses; it can effectively clear an infection with minimal morbidity. Few studies have directly infected *Peromyscus* species with coronaviruses, and those have shown mild or no symptoms in the animals despite evidence of viral replication (108), (109), (107). However, none of those studies used *P. leucopus*; hence, we confirmed our initial hypothesis: that due to *P. leucopus* ACE2 receptor similarities with the human ACE2 protein, the coronavirus will be able to replicate in analyzed tissues while *P. leucopus* cleared infection within days. We propose that *P. leucopus* is a valuable animal model since it mimics the human response to the virus,

including the development of symptoms, transmission dynamics, and immune response. While no single animal model perfectly replicates human morbidity with SARS-CoV-2, several have been used extensively. Since *P. leucopus* are abundant mammals in many North American habitats, additional consideration of this animal model is warranted in terms of viral spillover into animal population.

P. leucopus are social animals living in groups of typically 30 animals per hectare, (2), (528) exhibiting longevity (10) and long-distance movements (up to 15 km) (529). *P. leucopus* exhibit minimal impact on their health while carrying high priority zoonoses: Lyme disease agent *Borrelia burgdorferi* detected in about 75% of the animals in the field, babesiosis, Powassan virus encephalitis, human granulocytic anaplasmosis as well as Sin Nombre-like hantavirus (3), (528), (530), (531). Their social behavior, ecological reasons, immune system, and tolerance make them ideal zoonoses hosts. Understanding the immunity of *P. leucopus* to SARS-CoV-2 could aid in managing the virus in nature in case of a spillover.

P. leucopus' immunity to viral infections is partially based on tolerance, or the ability to allow pathogens to persist without causing harm. In the case of SARS-CoV-2, we report that genetically diverse *P. leucopus*, after presenting with a spectrum of morbidities (primarily mild or none), initially clears the infection by upregulating type I interferon-stimulated genes; then, within six days of infection, it continues to clear the infection by producing neutralizing antibodies. Appropriate viral clearance of SARS-CoV-2 involves a series of coordinated immune responses that eliminate the virus from the body without causing excessive damage to the host's tissues.

Innate immunity in SARS-CoV-2 infection

Viral clearance starts immediately after infection at mucous barriers in the lungs and olfactory mucosa. Resident macrophages and dendritic cells recognize viral components using PRRs and respond by producing type I interferon. In *P. leucopus* we report increased expression of interferon-stimulated genes like Mx2, Ifr7, Isg15, and Ifi27l2a — mainly those involved in the reaction to viral

infection and cytokines like Cxcl10. This action helps contain the virus. We report upregulation of interferon-inducible Rtp4, which may bind to viral RNA and suppress viral replication (532), (423). Rtp4 is expressed in mice, humans, and bats, although the mouse ortholog is ineffective against SARS-CoV-2 and human is mildly effective, while bat Rtp4 ortholog successfully binds to viral RNA, inhibiting viral synthesis. Bat Rtp4 shares ~58% identity with the human ortholog, underscoring recent adaptations of Rtp4 and the virus in an ongoing host-virus arms race (423). To confirm that Rtp4 in *P. leucopus* indeed inhibits amplification of SARS-CoV-2, one future direction for research would be a screen of lung-originating fibroblasts infected with SARS-CoV-2 surviving infection and carry out RNAseq of surviving cells followed by extraction of RTP4 proteins and testing their antiviral activity.

In the case of viral infection with SARS-CoV-2, there was upregulation of cytoplasmic PRRs: RIG-I, Dhx58, Ifih1, and cGAS in the lungs of infected animals compared to control groups based on targeted analysis ($p=5.6 \times 10^{-5}$, $p=1.2 \times 10^{-5}$, $p=5.1 \times 10^{-5}$, and $p=4.5 \times 10^{-6}$, respectively). In this case, upregulation of PRRs leads to further response to the virus, unlike upregulation in these genes in response to endotoxin as a reaction to ERVs in *M. musculus* in homeostasis, causing inflammation. These sensors are integral to the intracellular innate immune response, allowing cells throughout the body to detect and respond to viral infections. Their wide expression profile across many cell types underscores the importance of a robust and ubiquitous antiviral defense mechanism. However, the upregulation of immune response genes is counterbalanced by the activation of immune suppressive genes in *P. leucopus.*

Anti-inflammatory genes strike a balance in the early infection

In concordance with our previous Chapters 3 and 4 results, we report upregulation of Slpi and Arg1 in animals experiencing morbidity following viral infection. We observed an upregulation of Sod2 on day 3, which normalized by day 6, potentially regulating oxidative stress during the early innate

process. As previously reported in this chapter, Ido1 is upregulated in the lungs three days post-infection. This leads to Trp degradation, shifting the balance towards alternative macrophage polarization and tolerance. If homeostasis post-infection is not reached and Trp is reduced by overexpressed Ido1, it could, in turn, deplete substrate for the serotonergic pathway, leading to serotonin reduction. Deficiency in serotonin can lead to reduced vagus activity, dysfunction of the hippocampus, and memory loss. Hence, serotonin depletion could manifest as post-acute sequelae of long COVID-19 (533). Early innate immune response is followed by activation of adaptive immunity to presented antigens.

Cellular immunity to SARS-CoV-2

After the initial upregulation of genes involved in innate immunity, in the day 6 cohort, there is an upregulation of genes related to cell-mediated immunity surveilling and targeting cells infected with a virus. We report the upregulation of genes pointing to the cell-mediated immune response by edgeR and confirmed by targeted analysis: Gzmb, Gzmk Perf1, Cxcr3, and CD8. Jonsson et al., report that CD8+ T cells expressing both GzmK and GzmB are a unique subset of immune cells with distinct characteristics. These CD8$^+$ GzmK$^+$ GzmB$^+$ T cells are primarily cytokine producers with lower cytotoxic potential compared to other T cell populations. Hence, their role might be more oriented towards regulating immune responses through signaling molecules rather than directly killing target cells (534). GzmK and Gzmb expression is upregulated in the bronchoalveolar lavage fluid from COVID-19 patients matching the pattern of expression of these genes in *P. leucopus* (Gzmk>Gzmb). We did not asses the expression of both genes in the T cells granules due to animal model restrictions, but expression of both Gzmk and Gzmb can be determined in ssRNAseq experiments.

Anti-inflammatory genes in cellular immunity

To combat the cytotoxic effect of T cells, we report upregulation of the immunosuppressive gene Tigit six days post-infection. Tigit is expressed in viral infection or as a response to LPS, shifting the

balance between homeostasis and inflammation in favor of alternative macrophage polarization and suppressing both T cell activation and prolonged tissue damage (435), (535), (436). The expression of Tigit is beneficial only if transient, but it can prevent tissue damage during infection, as is the case with other tolerance factors such as tryptophan metabolites and Slpi. The expression of Tigit on tumor-infiltrating T cells promotes cancer survival since its ligand is expressed and upregulated on cancer cells, inhibiting immune surveillance of the malignant phenotype (437).

Another upregulated anti-inflammatory gene is CD52. CD52 was initially detected in seminal fluid and is responsible for immunosuppression against complement-dependent sperm-immobilizing antibodies. Sperm-immobilizing antibodies have frequently been detected in women experiencing infertility (536). Rashidi et al., report that soluble CD52, suppresses T-cell activation, in modulating the innate immune response. *In vivo,* the same group report that CD52 mitigates symptoms of endotoxic shock induced by LPS, such as cytokine secretion and organ injury. (440) We report that expression of CD52 in SARS-CoV-2 infection may be another feature of immune tolerance. These findings suggest that soluble CD52 has broad immune-suppressive effects, making it a potential candidate for treating inflammatory and autoimmune disorders where modulation of innate immune responses is desirable.

Viral infection in the brain

We report that viral RNA was found in the brain of *P. leucopus*, although the virus may have been replicating in the olfactory bulb. Brain RNAseq showed upregulation of type I interferon stimulated genes. We propose that *P. leucopus* could be a valuable animal model for SARS-CoV-2 brain research and long COVID-19 since unlike hACE2 mouse, it can survive and clear infection.

Viral shedding and ecological surveillance

We also report the presence of the viral RNA in the feces as the opportunity for viral shedding and establishing a secondary reservoir of SARS-CoV-2. Although Earnest et al. sero-surveillance of SARS-

CoV-2 showed limited transmission potential between humans and *P. leucopus,* extensive human interactions and relevant ligands for viral S-protein in *P. leucopus* increase the chances for passing the virus to a previously naïve species (110). Our study fills the knowledge gap in the potential impact on long-term disease control and emerging zoonotic viruses.

Conclusion

This study aimed to present the adaptations of *P. leucopus* that enhance tolerance and mitigate inflammation, positioning it as a competent zoonosis reservoir with minimal morbidity from bacterial or viral infections. The investigation powered by the sequencing and annotation of the *P. leucopus* genome, delved into its transcriptome and microbiome milieu, contrasting these with common laboratory rodents. Supported by a diverse microbiome, *P. leucopus* exhibits a controlled immune response to microbial encounters, supported by a diverse microbiome, including with unique Lactobacillus species, as well as a well-regulated immune system marked by specific cytokine profiles and SLPI activity.

The study highlighted the intricate balance between immune tolerance and pathogen defense in *P. leucopus*, uncovering potential avenues for zoonotic disease control and offering insights into the mammalian immune adaptation. The role of SLPI in mitigating inflammatory responses, along with the detailed examination of gut microbiota, tryptophan metabolism, ERV transcription, and type I interferon-regulated genes, illustrate the complex interplay between genetics, microbiome, and immune regulation. These findings pave the way for future research into targeted interventions for zoonotic disease management and deepen our understanding of the immune mechanisms underlying disease tolerance in wildlife reservoirs (figure 63).
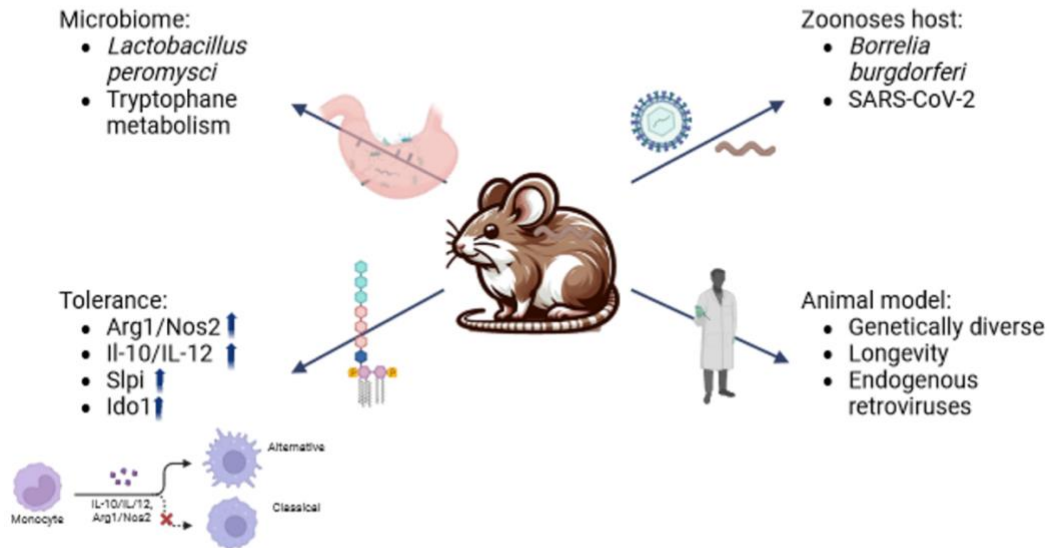
Figure 63 Graphical summary of P. leucopus as a model of tolerance and a zoonoses host

Created and adapted from "M2 differentiation" and "Lipopolysaccharide structure" by BioRender.com (2024).

References

1. M.J. Dewey WDD. 2001. Deer mice: the Drosophila of north american mammalogy. Genesis 29 (3) pp. 105-109.
2. Lackey JA, Huckaby DG, Ormiston BG. 1985. Mammalian Species. "Peromyscus leucopus" doi:10.2307/3503904. The American Society of Mammalogists.
3. Barbour AG. 2017. Infection resistance and tolerance in Peromyscus spp., natural reservoirs of microbes that are virulent for humans. Semin Cell Dev Biol 61:115-122.
4. Chen FC, Li WH. 2001. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am J Hum Genet 68:444-56.
5. Long AD, Baldwin-Brown J, Tao Y, Cook VJ, Balderrama-Gutierrez G, Corbett-Detig R, Mortazavi A, Barbour AG. 2019. The genome of Peromyscus leucopus, natural host for Lyme disease and other emerging infections. Sci Adv 5:eaaw6441.
6. Barbour AG, Shao H, Cook VJ, Baldwin-Brown J, Tsao JI, Long AD. 2019. Genomes, expression profiles, and diversity of mitochondria of the White-footed Deermouse Peromyscus leucopus, reservoir of Lyme disease and other zoonoses. Sci Rep 9:17618.
7. Bedford NL, Hoekstra HE. 2015. Peromyscus mice as a model for studying natural variation. Elife 4.
8. Havighorst A, Crossland J, Kiaris H. 2017. Peromyscus as a model of human disease. Semin Cell Dev Biol 61:150-155.
9. Shorter KR, Owen, A., Anderson, V., Hall-South, A.C., Hayford, S., Cakora, P., Crossland, J.P., Georgi, V.R., Perkins, A., Kelly, S.J. and Felder, M.R. 2014. Natural genetic variation underlying differences in Peromyscus repetitive and social/aggressive behaviors. Behavior genetics:.126-135.
10. Sacher GA, Hart, R W. 1978. 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; (United States) 14:1.
11. Netski D, Thran BH, St Jeor SC. 1999. Sin Nombre virus pathogenesis in Peromyscus maniculatus. J Virol 73:585-91.
12. Perlman RL. 2016. Mouse models of human disease: An evolutionary perspective. Evol Med Public Health 2016:170-6.
13. Lippens C, Estoup A, Hima MK, Loiseau A, Tatard C, Dalecky A, Ba K, Kane M, Diallo M, Sow A, Niang Y, Piry S, Berthier K, Leblois R, Duplantier JM, Brouat C. 2017. Genetic structure and invasion history of the house mouse (Mus musculus domesticus) in Senegal, West Africa: a legacy of colonial and contemporary times. Heredity (Edinb) 119:64-75.
14. Persons WE, Eason PK. 2019. White-footed mouse (Peromyscus leucopus) habitat selection and Amur honeysuckle (Lonicera maackii) canopy use in an urban forest. Urban Ecosystems 22:471-482.
15. Bunikis J, Tsao J, Luke CJ, Luna MG, Fish D, Barbour AG. 2004. Borrelia burgdorferi infection in a natural population of Peromyscus Leucopus mice: a longitudinal study in an area where Lyme Borreliosis is highly endemic. J Infect Dis 189:1515-23.
16. Schwan TG, Kime KK, Schrumpf ME, Coe JE, Simpson WJ. 1989. Antibody response in white-footed mice (Peromyscus leucopus) experimentally infected with the Lyme disease spirochete (Borrelia burgdorferi). Infect Immun 57:3445-51.
17. Tufts DM, Diuk-Wasser MA. 2018. Transplacental transmission of tick-borne Babesia microti in its natural host Peromyscus leucopus. Parasit Vectors 11:286.

18. Hofmeister EK, Ellis BA, Glass GE, Childs JE. 1999. Longitudinal study of infection with Borrelia burgdorferi in a population of Peromyscus leucopus at a Lyme disease-enzootic site in Maryland. Am J Trop Med Hyg 60:598-609.

19. Baum E, Hue F, Barbour AG. 2012. Experimental infections of the reservoir species Peromyscus leucopus with diverse strains of Borrelia burgdorferi, a Lyme disease agent. mBio 3:e00434-12.

20. Cook V, Barbour AG. 2015. Broad diversity of host responses of the white-footed mouse Peromyscus leucopus to Borrelia infection and antigens. Ticks Tick Borne Dis 6:549-58.

21. Schwanz LE, Voordouw MJ, Brisson D, Ostfeld RS. 2011. Borrelia burgdorferi has minimal impact on the Lyme disease reservoir host Peromyscus leucopus. Vector Borne Zoonotic Dis 11:117-24.

22. Mlera L, Meade-White K, Dahlstrom E, Baur R, Kanakabandi K, Virtaneva K, Porcella SF, Bloom ME. 2018. Peromyscus leucopus mouse brain transcriptome response to Powassan virus infection. J Neurovirol 24:75-87.

23. Schug MD, Vessey, S.H. and Korytko, A.I. 1991. Longevity and survival in a population of white-footed mice (Peromyscus leucopus). Journal of Mammalogy:360-366.

24. Shi Y, Pulliam DA, Liu Y, Hamilton RT, Jernigan AL, Bhattacharya A, Sloane LB, Qi W, Chaudhuri A, Buffenstein R, Ungvari Z, Austad SN, Van Remmen H. 2013. Reduced mitochondrial ROS, enhanced antioxidant defense, and distinct age-related changes in oxidative damage in muscles of long-lived Peromyscus leucopus. Am J Physiol Regul Integr Comp Physiol 304:R343-55.

25. Voordouw MJ, Lachish S, Dolan MC. 2015. The lyme disease pathogen has no effect on the survival of its rodent reservoir host. PLoS One 10:e0118265.

26. Ayres JS, Schneider DS. 2012. Tolerance of infections. Annu Rev Immunol 30:271-94.

27. Soares MP, Gozzelino R, Weis S. 2014. Tissue damage control in disease tolerance. Trends Immunol 35:483-94.

28. Randolph HE, Barreiro LB. 2018. Holy Immune Tolerance, Batman! Immunity 48:1074-1076.

29. Pavlovich SS, Lovett SP, Koroleva G, Guito JC, Arnold CE, Nagle ER, Kulcsar K, Lee A, Thibaud-Nissen F, Hume AJ, Muhlberger E, Uebelhoer LS, Towner JS, Rabadan R, Sanchez-Lockhart M, Kepler TB, Palacios G. 2018. The Egyptian Rousette Genome Reveals Unexpected Features of Bat Antiviral Immunity. Cell 173:1098-1110 e18.

30. Dantes RB, Kaur H, Bouwkamp BA, Haass KA, Patel P, Dudeck MA, Srinivasan A, Magill SS, Wilson WW, Whitaker M, Gladden NM, McLaughlin ES, Horowitz JK, Posa PJ, Prescott HC. 2023. Sepsis Program Activities in Acute Care Hospitals - National Healthcare Safety Network, United States, 2022. MMWR Morb Mortal Wkly Rep 72:907-911.

31. Reinhart K, Daniels R, Kissoon N, Machado FR, Schachter RD, Finfer S. 2017. Recognizing Sepsis as a Global Health Priority - A WHO Resolution. N Engl J Med 377:414-417.

32. Majno G. 1991. The Ancient Riddle of σῆψις (Sepsis). The Journal of Infectious Diseases 163:937-945.

33. Angus DC, van der Poll T. 2013. Severe sepsis and septic shock. N Engl J Med 369:840-51.

34. Ward PA. 2012. New approaches to the study of sepsis. EMBO Mol Med 4:1234-43.

35. Cerra FB. 1985. The Systemic Septic Response: Maltiple Systems Organ Failure. Critical Care Clinics 1:591-607.

36. Lewis AJ, Seymour CW, Rosengart MR. 2016. Current Murine Models of Sepsis. Surg Infect (Larchmt) 17:385-93.

37. Evans T. 2018. Diagnosis and management of sepsis. Clin Med (Lond) 18:146-149.

38. Dellinger RP, Carlet JM, Masur H, Gerlach H, Calandra T, Cohen J, Gea-Banacloche J, Keh D, Marshall JC, Parker MM, Ramsay G, Zimmerman JL, Vincent JL, Levy MM, Surviving Sepsis

Campaign Management Guidelines C. 2004. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. Crit Care Med 32:858-73.

39. Kylat RI, Ohlsson A. 2012. Recombinant human activated protein C for severe sepsis in neonates. Cochrane Database Syst Rev 2012:CD005385.

40. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. 2012. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. Cochrane Database Syst Rev 12:CD004388.

41. Angus DC. 2011. The search for effective therapy for sepsis: back to the drawing board? JAMA 306:2614-5.

42. Lai PS, Thompson BT. 2013. Why activated protein C was not successful in severe sepsis and septic shock: are we still tilting at windmills? Curr Infect Dis Rep 15:407-12.

43. Buras JA, Holzmann B, Sitkovsky M. 2005. Animal models of sepsis: setting the stage. Nat Rev Drug Discov 4:854-65.

44. Rautanen A, Mills TC, Gordon AC, Hutton P, Steffens M, Nuamah R, Chiche JD, Parks T, Chapman SJ, Davenport EE, Elliott KS, Bion J, Lichtner P, Meitinger T, Wienker TF, Caulfield MJ, Mein C, Bloos F, Bobek I, Cotogni P, Sramek V, Sarapuu S, Kobilay M, Ranieri VM, Rello J, Sirgo G, Weiss YG, Russwurm S, Schneider EM, Reinhart K, Holloway PA, Knight JC, Garrard CS, Russell JA, Walley KR, Stuber F, Hill AV, Hinds CJ, Investigators EEG. 2015. Genome-wide association study of survival from sepsis due to pneumonia: an observational cohort study. Lancet Respir Med 3:53-60.

45. Hernandez-Beeftink T, Guillen-Guio B, Lorenzo-Salazar JM, Corrales A, Suarez-Pajes E, Feng R, Rubio-Rodriguez LA, Paynton ML, Cruz R, Garcia-Laorden MI, Prieto-Gonzalez M, Rodriguez-Perez A, Carriedo D, Blanco J, Ambros A, Gonzalez-Higueras E, Espinosa E, Muriel A, Tamayo E, Martin MM, Lorente L, Dominguez D, de Lorenzo AG, Giannini HM, Reilly JP, Jones TK, Anon JM, Soro M, Carracedo A, Wain LV, Meyer NJ, Villar J, Flores C, Genetics of Sepsis N. 2022. A genome-wide association study of survival in patients with sepsis. Crit Care 26:341.

46. Scheiermann C, Frenette PS, Hidalgo A. 2015. Regulation of leucocyte homeostasis in the circulation. Cardiovasc Res 107:340-51.

47. Mestas J, Hughes CC. 2004. Of mice and not men: differences between mouse and human immunology. J Immunol 172:2731-8.

48. O'Connell KE, Mikkola AM, Stepanek AM, Vernet A, Hall CD, Sun CC, Yildirim E, Staropoli JF, Lee JT, Brown DE. 2015. Practical murine hematopathology: a comparative review and implications for research. Comp Med 65:96-113.

49. Doeing DC, Borowicz JL, Crockett ET. 2003. Gender dimorphism in differential peripheral blood leukocyte counts in mice using cardiac, tail, foot, and saphenous vein puncture methods. BMC Clin Pathol 3:3.

50. Smith C, Jarecki, A. 2011. Atlas of Comparative Diagnostic and Experimental Hematology

51. Janeway CA, Jr. 1989. Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol 54 Pt 1:1-13.

52. Rialdi A, Campisi L, Zhao N, Lagda AC, Pietzsch C, Ho JSY, Martinez-Gil L, Fenouil R, Chen X, Edwards M, Metreveli G, Jordan S, Peralta Z, Munoz-Fontela C, Bouvier N, Merad M, Jin J, Weirauch M, Heinz S, Benner C, van Bakel H, Basler C, Garcia-Sastre A, Bukreyev A, Marazzi I. 2016. Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation. Science 352:aad7993.

53. Krieg AM. 2002. CpG motifs in bacterial DNA and their immune effects. Annu Rev Immunol 20:709-60.

54. Metzemaekers M, Gouwy M, Proost P. 2020. Neutrophil chemoattractant receptors in health and disease: double-edged swords. Cell Mol Immunol 17:433-450.

55. Roberts AW, Kim C, Zhen L, Lowe JB, Kapur R, Petryniak B, Spaetti A, Pollock JD, Borneo JB, Bradford GB, Atkinson SJ, Dinauer MC, Williams DA. 1999. Deficiency of the hematopoietic cell-specific Rho family GTPase Rac2 is characterized by abnormalities in neutrophil function and host defense. Immunity 10:183-96.

56. Lacy P. 2006. Mechanisms of degranulation in neutrophils. Allergy Asthma Clin Immunol 2:98-108.

57. Kraus RF, Gruber MA. 2021. Neutrophils-From Bone Marrow to First-Line Defense of the Innate Immune System. Front Immunol 12:767175.

58. Nguyen GT, Green ER, Mecsas J. 2017. Neutrophils to the ROScue: Mechanisms of NADPH Oxidase Activation and Bacterial Resistance. Front Cell Infect Microbiol 7:373.

59. Zwarthoff SA, Berends ETM, Mol S, Ruyken M, Aerts PC, Jozsi M, de Haas CJC, Rooijakkers SHM, Gorham RD, Jr. 2018. Functional Characterization of Alternative and Classical Pathway C3/C5 Convertase Activity and Inhibition Using Purified Models. Front Immunol 9:1691.

60. Winterbourn CC, Kettle AJ, Hampton MB. 2016. Reactive Oxygen Species and Neutrophil Function. Annu Rev Biochem 85:765-92.

61. Koshland DE, Jr. 1992. The molecule of the year. Science 258:1861.

62. Lambden S. 2019. Bench to bedside review: therapeutic modulation of nitric oxide in sepsis-an update. Intensive Care Med Exp 7:64.

63. Vincent JL, Zhang H, Szabo C, Preiser JC. 2000. Effects of nitric oxide in septic shock. Am J Respir Crit Care Med 161:1781-5.

64. Tejero J, Hunt AP, Santolini J, Lehnert N, Stuehr DJ. 2019. Mechanism and regulation of ferrous heme-nitric oxide (NO) oxidation in NO synthases. J Biol Chem 294:7904-7916.

65. Bowdish DM, Davidson DJ, Hancock RE. 2006. Immunomodulatory properties of defensins and cathelicidins. Curr Top Microbiol Immunol 306:27-66.

66. Sorensen OE, Follin P, Johnsen AH, Calafat J, Tjabringa GS, Hiemstra PS, Borregaard N. 2001. Human cathelicidin, hCAP-18, is processed to the antimicrobial peptide LL-37 by extracellular cleavage with proteinase 3. Blood 97:3951-9.

67. Zeng W, Song Y, Wang R, He R, Wang T. 2023. Neutrophil elastase: From mechanisms to therapeutic potential. J Pharm Anal 13:355-366.

68. Voynow JA, Shinbashi M. 2021. Neutrophil Elastase and Chronic Lung Disease. Biomolecules 11.

69. Wang Y, Wang M, Zhang H, Wang Y, Du Y, Guo Z, Ma L, Zhou Y, Zhang H, Liu L. 2022. Sivelestat improves clinical outcomes and decreases ventilator-associated lung injury in children with acute respiratory distress syndrome: a retrospective cohort study. Transl Pediatr 11:1671-1681.

70. Pan T, Tuoerxun T, Chen X, Yang CJ, Jiang CY, Zhu YF, Li ZS, Jiang XY, Zhang HT, Zhang H, Wang YP, Chen W, Lu LC, Ge M, Cheng YQ, Wang DJ, Zhou Q. 2023. The neutrophil elastase inhibitor, sivelestat, attenuates acute lung injury in patients with cardiopulmonary bypass. Front Immunol 14:1082830.

71. Tagami T, Tosa R, Omura M, Fukushima H, Kaneko T, Endo T, Rinka H, Murai A, Yamaguchi J, Yoshikawa K, Saito N, Uzu H, Kase Y, Takatori M, Izumino H, Nakamura T, Seo R, Kitazawa Y, Sugita M, Takahashi H, Kuroki Y, Irahara T, Kanemura T, Yokota H, Kushimoto S. 2014. Effect of a selective neutrophil elastase inhibitor on mortality and ventilator-free days in patients with increased extravascular lung water: a post hoc analysis of the PiCCO Pulmonary Edema Study. J Intensive Care 2:67.

72. Chiang CC, Korinek M, Cheng WJ, Hwang TL. 2020. Targeting Neutrophils to Treat Acute Respiratory Distress Syndrome in Coronavirus Disease. Front Pharmacol 11:572009.

73. Takei H, Araki A, Watanabe H, Ichinose A, Sendo F. 1996. Rapid killing of human neutrophils by the potent activator phorbol 12-myristate 13-acetate (PMA) accompanied by changes different from typical apoptosis or necrosis. J Leukoc Biol 59:229-40.

74. Lu J, Liu J, Li A. 2022. Roles of neutrophil reactive oxygen species (ROS) generation in organ function impairment in sepsis. J Zhejiang Univ Sci B 23:437-450.

75. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. 2004. Neutrophil extracellular traps kill bacteria. Science 303:1532-5.

76. Cicco S, Cicco G, Racanelli V, Vacca A. 2020. Neutrophil Extracellular Traps (NETs) and Damage-Associated Molecular Patterns (DAMPs): Two Potential Targets for COVID-19 Treatment. Mediators Inflamm 2020:7527953.

77. Mosser DM, Hamidzadeh K, Goncalves R. 2021. Macrophages and the maintenance of homeostasis. Cell Mol Immunol 18:579-587.

78. Hamidzadeh K, Christensen SM, Dalby E, Chandrasekaran P, Mosser DM. 2017. Macrophages and the Recovery from Acute and Chronic Inflammation. Annu Rev Physiol 79:567-592.

79. Decano JL, Aikawa M. 2018. Dynamic Macrophages: Understanding Mechanisms of Activation as Guide to Therapy for Atherosclerotic Vascular Disease. Front Cardiovasc Med 5:97.

80. Gordon S, Martinez FO. 2010. Alternative activation of macrophages: mechanism and functions. Immunity 32:593-604.

81. Leopold Wager CM, Wormley FL, Jr. 2014. Classical versus alternative macrophage activation: the Ying and the Yang in host defense against pulmonary fungal infections. Mucosal Immunol 7:1023-35.

82. McNab F, Mayer-Barber K, Sher A, Wack A, O'Garra A. 2015. Type I interferons in infectious disease. Nat Rev Immunol 15:87-103.

83. Schoggins JW. 2019. Interferon-Stimulated Genes: What Do They All Do? Annu Rev Virol 6:567-584.

84. Platanias LC. 2005. Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat Rev Immunol 5:375-86.

85. Mazewski C, Perez RE, Fish EN, Platanias LC. 2020. Type I Interferon (IFN)-Regulated Activation of Canonical and Non-Canonical Signaling Pathways. Front Immunol 11:606456.

86. Zhang Q, Wang C, Ma F, Yao L, Gao H, Zhu L, Zheng L. 2020. Development and biological activity of long-acting recombinant human interferon-alpha2b. BMC Biotechnol 20:16.

87. Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, Rice CM. 2011. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature 472:481-5.

88. Chemudupati M, Kenney AD, Bonifati S, Zani A, McMichael TM, Wu L, Yount JS. 2019. From APOBEC to ZAP: Diverse mechanisms used by cellular restriction factors to inhibit virus infections. Biochim Biophys Acta Mol Cell Res 1866:382-394.

89. Barbour AG. 1984. Isolation and cultivation of Lyme disease spirochetes. Yale J Biol Med 57:521-5.

90. Burgdorfer W, Barbour AG, Hayes SF, Benach JL, Grunwaldt E, Davis JP. 1982. Lyme disease-a tick-borne spirochetosis? Science 216:1317-9.

91. Barbour AG, Zuckert WR. 1997. Genome sequencing. New tricks of tick-borne pathogen. Nature 390:553, 555.

92. Barbour AG. 2015. Lyme Disease. Johns Hopkins University Press.

93. Bourgeois JS, You SS, Clendenen LH, Shrestha M, Petnicki-Ocwieja T, Telford SR, Hu LT. 2024. doi:10.1101/2023.09.28.559638.

94. Gaber AM, Mandric I, Nitirahardjo C, Piontkivska H, Hillhouse AE, Threadgill DW, Zelikovsky A, Rogovskyy AS. 2023. Comparative transcriptome analysis of Peromyscus leucopus and C3H mice infected with the Lyme disease pathogen. Front Cell Infect Microbiol 13:1115350.

95. Crowder CD, Ghalyanchi Langeroudi A, Shojaee Estabragh A, Lewis ERG, Marcsisin RA, Barbour AG. 2016. Pathogen and Host Response Dynamics in a Mouse Model of Borrelia hermsii Relapsing Fever. Vet Sci 3.

96. Nieto NC, Teglas MB. 2014. Relapsing fever group Borrelia in Southern California rodents. J Med Entomol 51:1029-34.

97. Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S, Yuen KY. 2020. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg Microbes Infect 9:221-236.

98. Meekins DA, Gaudreault NN, Richt JA. 2021. Natural and Experimental SARS-CoV-2 Infection in Domestic and Wild Animals. Viruses 13.

99. Napoli MCMRAASCDRD. 2024. Evaluation, and Treatment of Coronavirus (COVID-19), StatPearls [Internet]. Treasure Island (FL): StatPearls.

100. V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19:155-170.

101. Pfaender S, Mar KB, Michailidis E, Kratzel A, Boys IN, V'Kovski P, Fan W, Kelly JN, Hirt D, Ebert N, Stalder H, Kleine-Weber H, Hoffmann M, Hoffmann HH, Saeed M, Dijkman R, Steinmann E, Wight-Carter M, McDougal MB, Hanners NW, Pohlmann S, Gallagher T, Todt D, Zimmer G, Rice CM, Schoggins JW, Thiel V. 2020. LY6E impairs coronavirus fusion and confers immune control of viral disease. Nat Microbiol 5:1330-1339.

102. Beitari S, Duque D, Bavananthasivam J, Hewitt M, Sandhu JK, Hadzisejdic I, Tran A. 2023. Cross protection to SARS-CoV-2 variants in hamsters with naturally-acquired immunity. Virol J 20:167.

103. Putcharoen O, Wacharapluesadee S, Chia WN, Paitoonpong L, Tan CW, Suwanpimolkul G, Jantarabenjakul W, Ruchisrisarod C, Wanthong P, Sophonphan J, Chariyavilaskul P, Wang LF, Hemachudha T. 2021. Early detection of neutralizing antibodies against SARS-CoV-2 in COVID-19 patients in Thailand. PLoS One 16:e0246864.

104. Wan Y, Shang J, Graham R, Baric RS, Li F. 2020. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. J Virol 94.

105. Chan JF, Zhang AJ, Yuan S, Poon VK, Chan CC, Lee AC, Chan WM, Fan Z, Tsoi HW, Wen L, Liang R, Cao J, Chen Y, Tang K, Luo C, Cai JP, Kok KH, Chu H, Chan KH, Sridhar S, Chen Z, Chen H, To KK, Yuen KY. 2020. Simulation of the Clinical and Pathological Manifestations of Coronavirus Disease 2019 (COVID-19) in a Golden Syrian Hamster Model: Implications for Disease Pathogenesis and Transmissibility. Clin Infect Dis 71:2428-2446.

106. Pach S, Nguyen TN, Trimpert J, Kunec D, Osterrieder N, Wolber G. 2021. ACE2-Variants Indicate Potential SARS-CoV-2-Susceptibility in Animals: A Molecular Dynamics Study. Mol Inform 40:e2100031.

107. Lewis J, Zhan S, Vilander AC, Fagre AC, Aboellail TA, Kiaris H, Schountz T. 2023. SARS-CoV-2 infects multiple species of North American deer mice and causes clinical disease in the California mouse. Frontiers in Virology 3.

108. Fagre A, Lewis J, Eckley M, Zhan S, Rocha SM, Sexton NR, Burke B, Geiss B, Peersen O, Bass T, Kading R, Rovnak J, Ebel GD, Tjalkens RB, Aboellail T, Schountz T. 2021. SARS-CoV-2

infection, neuropathogenesis and transmission among deer mice: Implications for spillback to New World rodents. PLoS Pathog 17:e1009585.

109. Griffin BD, Chan M, Tailor N, Mendoza EJ, Leung A, Warner BM, Duggan AT, Moffat E, He S, Garnett L, Tran KN, Banadyga L, Albietz A, Tierney K, Audet J, Bello A, Vendramelli R, Boese AS, Fernando L, Lindsay LR, Jardine CM, Wood H, Poliquin G, Strong JE, Drebot M, Safronetz D, Embury-Hyatt C, Kobasa D. 2021. SARS-CoV-2 infection and transmission in the North American deer mouse. Nat Commun 12:3612.

110. Earnest R, Hahn AM, Feriancek NM, Brandt M, Filler RB, Zhao Z, Breban MI, Vogels CBF, Chen NFG, Koch RT, Porzucek AJ, Sodeinde A, Garbiel A, Keanna C, Litwak H, Stuber HR, Cantoni JL, Pitzer VE, Olarte Castillo XA, Goodman LB, Wilen CB, Linske MA, Williams SC, Grubaugh ND. 2023. Survey of white-footed mice (Peromyscus leucopus) in Connecticut, USA reveals low SARS-CoV-2 seroprevalence and infection with divergent betacoronaviruses. npj Viruses 1.

111. Berg G, Rybakova D, Fischer D, Cernava T, Verges MC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH, Kazou M, Kinkel L, Lange L, Lima N, Loy A, Macklin JA, Maguin E, Mauchline T, McClure R, Mitter B, Ryan M, Sarand I, Smidt H, Schelkle B, Roume H, Kiran GS, Selvin J, Souza RSC, van Overbeek L, Singh BK, Wagner M, Walsh A, Sessitsch A, Schloter M. 2020. Microbiome definition re-visited: old concepts and new challenges. Microbiome 8:103.

112. Wang B, Yao M, Lv L, Ling Z, Li L. 2017. The Human Microbiota in Health and Disease. Engineering 3:71-82.

113. Tyler AD, Smith MI, Silverberg MS. 2014. Analyzing the human microbiome: a "how to" guide for physicians. Am J Gastroenterol 109:983-93.

114. Ursell LK, Metcalf JL, Parfrey LW, Knight R. 2012. Defining the human microbiome. Nutr Rev 70 Suppl 1:S38-44.

115. Biggs MB, Medlock GL, Moutinho TJ, Lees HJ, Swann JR, Kolling GL, Papin JA. 2017. Systems-level metabolism of the altered Schaedler flora, a complete gut microbiota. Isme j 11:426-438.

116. Martin AM, Sun EW, Rogers GB, Keating DJ. 2019. The Influence of the Gut Microbiome on Host Metabolism Through the Regulation of Gut Hormone Release. Front Physiol 10:428.

117. Marin IA, Goertz JE, Ren T, Rich SS, Onengut-Gumuscu S, Farber E, Wu M, Overall CC, Kipnis J, Gaultier A. 2017. Microbiota alteration is associated with the development of stress-induced despair behavior. Sci Rep 7:43859.

118. Round JL, Mazmanian SK. 2009. The gut microbiota shapes intestinal immune responses during health and disease. doi:10.1038/nri2515.

119. LeCureux JS, Dean GA. 2018. Lactobacillus Mucosal Vaccine Vectors: Immune Responses against Bacterial and Viral Antigens. 3.

120. Demian WL, Cormier O, Mossman K. 2024. Immunological features of bats: resistance and tolerance to emerging viruses. Trends Immunol 45:198-210.

121. Federici L, Masulli M, De Laurenzi V, Allocati N. 2022. An overview of bats microbiota and its implication in transmissible diseases. Front Microbiol 13:1012189.

122. Popov IV, Mazanko MS, Kulaeva ED, Golovin SN, Malinovkin AV, Aleshukina IS, Aleshukina AV, Prazdnova EV, Tverdokhlebova TI, Chikindas ML, Ermakov AM. 2021. Gut microbiota of bats: pro-mutagenic properties and possible frontiers in preventing emerging disease. Sci Rep 11:21075.

123. Kohl KD, Dearing MD, Bordenstein SR. 2018. Microbial communities exhibit host species distinguishability and phylosymbiosis along the length of the gastrointestinal tract. Mol Ecol 27:1874-1883.

124. Pollock J, Glendinning L, Wisedchanwet T, Watson M. 2018. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. Appl Environ Microbiol 84.

125. Keegan KP, Glass EM, Meyer F. 2016. MG-RAST, a Metagenomics Service for Analysis of Microbial Community Structure and Function. Methods Mol Biol 1399:207-33.

126. Dubos R. 1981. Autobiographical Manuscript. New York: Rockefeller University p. Folder 14.

127. Sangodeyi F. 2012. Rene Dubos and the Emerging Science of Human Microbial Ecology New York: Rockefeller University.

128. ER H. 1981. The Mammals of North America. econd ed New York: John Wiley and Sons; .

129. MUNSHI-SOUTH JASONaK, K. 2010. Rapid, pervasive genetic differentiation of urban white-footed mouse (Peromyscus leucopus) populations in New York City. Molecular ecology 19:4242-4254.

130. Crossland JP, Dewey MJ, Barlow SC, Vrana PB, Felder MR, Szalai GJ. 2014. Caring for Peromyscus spp. in research environments. Lab Anim (NY) 43:162-6.

131. Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. Mol Biol Evol 34:1812-1819.

132. Steppan SJ, Schenk JJ. 2017. Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. PLoS One 12:e0183070.

133. Schrago CG, Voloch CM. 2013. The precision of the hominid timescale estimated by relaxed clock methods. J Evol Biol 26:746-55.

134. Veres M, Duselis AR, Graft A, Pryor W, Crossland J, Vrana PB, Szalai G. 2012. The biology and methodology of assisted reproduction in deer mice (Peromyscus maniculatus). Theriogenology 77:311-9.

135. Gomes-Solecki MJ, Brisson DR, Dattwyler RJ. 2006. Oral vaccine that breaks the transmission cycle of the Lyme disease spirochete can be delivered via bait. Vaccine 24:4440-9.

136. Najjar DA, Normandin AM, Strait EA, Esvelt KM. 2017. Driving towards ecotechnologies. Pathog Glob Health 111:448-458.

137. Specter M. 2017. Rewriting the code of life. . The New Yorker 2.

138. Baxter NT, Wan JJ, Schubert AM, Jenior ML, Myers P, Schloss PD. 2015. Intra- and interindividual variations mask interspecies variation in the microbiota of sympatric peromyscus populations. Appl Environ Microbiol 81:396-404.

139. Bassam K, Milovic A, Barbour AG. 2019. Genome Sequences of Three Lactobacillus Species Strains of the Stomach of the White-Footed Deermouse (Peromyscus leucopus). 8.

140. Herlemann DP, Geissinger O, Ikeda-Ohtsubo W, Kunin V, Sun H, Lapidus A, Hugenholtz P, Brune A. 2009. Genomic analysis of "Elusimicrobium minutum," the first cultivated representative of the phylum "Elusimicrobia" (formerly termite group 1). Appl Environ Microbiol 75:2841-9.

141. Soo RM, Skennerton CT, Sekiguchi Y, Imelfort M, Paech SJ, Dennis PG, Steen JA, Parks DH, Tyson GW, Hugenholtz P. 2014. An expanded genomic representation of the phylum cyanobacteria. Genome Biol Evol 6:1031-45.

142. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, Tito RY, Chaffron S, Rymenans L, Verspecht C, De Sutter L, Lima-Mendez G, D'Hoe K, Jonckheere K, Homola D, Garcia R, Tigchelaar EF, Eeckhaudt L, Fu J, Henckaerts L, Zhernakova A, Wijmenga C, Raes J. 2016. Population-level analysis of gut microbiome variation. Science 352:560-4.

143. Walter J. 2008. Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. Appl Environ Microbiol 74:4985-96.

144. KC TGIM. 1990. The microecology of lactobacilli inhabiting the gastrointestinal tract. Advances in Microbial Ecology, Boston, MA: Springer; 11:147–71.

145. Gomes-Solecki M, Richer L. 2018. Recombinant E. coli Dualistic Role as an Antigen-adjuvant Delivery Vehicle for Oral Immunization. Methods Mol Biol 1690:347-357.

146. Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C, Colles FM, Wimalarathna H, Harrison OB, Sheppard SK, Cody AJ, Maiden MCJ. 2012. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology (Reading) 158:1005-1015.

147. Fujisawa T, Itoh K, Benno Y, Mitsuoka T. 1990. Lactobacillus intestinalis (ex Hemme 1974) sp. nov., nom. rev., isolated from the intestines of mice and rats. Int J Syst Bacteriol 40:302-4.

148. Kim D, Cho MJ, Cho S, Lee Y, Byun SJ, Lee S. 2018. Complete Genome Sequence of Lactobacillus reuteri Byun-re-01, Isolated from Mouse Small Intestine. Microbiol Resour Announc 7.

149. Sun Z, Harris HM, McCann A, Guo C, Argimon S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui Y, Zhang H, O'Toole PW. 2015. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nat Commun 6:8322.

150. Dent VE, Williams RAD. 1982. Lactobacillus animalis sp. nov., a new species of lactobacillus from the alimentary canal of animals. Zentralblatt für Bakteriologie Mikrobiologie und Hygiene: I Abt Originale C: Allgemeine, angewandte und ökologische Mikrobiologie 3:377-386.

151. Nam SH, Choi SH, Kang A, Kim DW, Kim RN, Kim A, Kim DS, Park HS. 2011. Genome sequence of Lactobacillus animalis KCTC 3501. J Bacteriol 193:1280-1.

152. Claesson MJ, Li Y, Leahy S, Canchaya C, van Pijkeren JP, Cerdeno-Tarraga AM, Parkhill J, Flynn S, O'Sullivan GC, Collins JK, Higgins D, Shanahan F, Fitzgerald GF, van Sinderen D, O'Toole PW. 2006. Multireplicon genome architecture of Lactobacillus salivarius. Proc Natl Acad Sci U S A 103:6718-23.

153. Frese SA, Benson AK, Tannock GW, Loach DM, Kim J, Zhang M, Oh PL, Heng NC, Patil PB, Juge N, Mackenzie DA, Pearson BM, Lapidus A, Dalin E, Tice H, Goltsman E, Land M, Hauser L, Ivanova N, Kyrpides NC, Walter J. 2011. The evolution of host specialization in the vertebrate gut symbiont Lactobacillus reuteri. PLoS Genet 7:e1001314.

154. Bandara M, Skehel JM, Kadioglu A, Collinson I, Nobbs AH, Blocker AJ, Jenkinson HF. 2017. The accessory Sec system (SecY2A2) in Streptococcus pneumoniae is involved in export of pneumolysin toxin, adhesion and biofilm formation. Microbes Infect 19:402-412.

155. Hidalgo-Cantabrana C, Goh YJ, Pan M, Sanozky-Dawes R, Barrangou R. 2019. Genome editing using the endogenous type I CRISPR-Cas system in Lactobacillus crispatus. Proc Natl Acad Sci U S A 116:15774-15783.

156. Wilson CM, Aggio RB, O'Toole PW, Villas-Boas S, Tannock GW. 2012. Transcriptional and metabolomic consequences of LuxS inactivation reveal a metabolic rather than quorum-sensing role for LuxS in Lactobacillus reuteri 100-23. J Bacteriol 194:1743-6.

157. Saulnier DM, Santos F, Roos S, Mistretta TA, Spinler JK, Molenaar D, Teusink B, Versalovic J. 2011. Exploring metabolic pathway reconstruction and genome-wide expression profiling in Lactobacillus reuteri to define functional probiotic features. PLoS One 6:e18783.

158. Moreno-Vivian C, Cabello P, Martinez-Luque M, Blasco R, Castillo F. 1999. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. J Bacteriol 181:6573-84.

159. Hwang H, Lee JH. 2018. Characterization of Arginine Catabolism by Lactic Acid Bacteria Isolated from Kimchi. Molecules 23.

160. Singer JR, Blosser EG, Zindl CL, Silberger DJ, Conlan S, Laufer VA, DiToro D, Deming C, Kumar R, Morrow CD, Segre JA, Gray MJ, Randolph DA, Weaver CT. 2019. Preventing dysbiosis of the neonatal mouse intestinal microbiome protects against late-onset sepsis. Nat Med 25:1772-1782.

161. Pena JA, Li SY, Wilson PH, Thibodeau SA, Szary AJ, Versalovic J. 2004. Genotypic and phenotypic studies of murine intestinal lactobacilli: species differences in mice with and without colitis. Appl Environ Microbiol 70:558-68.

162. Shen Z, Sheh A, Young SK, Abouelliel A, Ward DV, Earl AM, Fox JG. 2014. Draft genome sequences of six enterohepatic helicobacter species isolated from humans and one from rhesus macaques. Genome Announc 2.

163. Suerbaum S, Josenhans C, Sterzenbach T, Drescher B, Brandt P, Bell M, Droge M, Fartmann B, Fischer HP, Ge Z, Horster A, Holland R, Klein K, Konig J, Macko L, Mendz GL, Nyakatura G, Schauer DB, Shen Z, Weber J, Frosch M, Fox JG. 2003. The complete genome sequence of the carcinogenic bacterium Helicobacter hepaticus. Proc Natl Acad Sci U S A 100:7901-6.

164. Fox JG, Ge Z, Whary MT, Erdman SE, Horwitz BH. 2011. Helicobacter hepaticus infection in mice: models for understanding lower bowel inflammation and cancer. Mucosal Immunol 4:22-30.

165. Bracken TC CC, Ali Z, Truong H, Moore JM. . 2017. Helicobacter infection significantly alters pregnancy success in laboratory mice. . J Am Assoc Lab Animal Sci 56(3):322-9.

166. Nordhoff M, Taras D, Macha M, Tedin K, Busse HJ, Wieler LH. 2005. Treponema berlinense sp. nov. and Treponema porcinum sp. nov., novel spirochaetes isolated from porcine faeces. Int J Syst Evol Microbiol 55:1675-1680.

167. AG. B. 2018. Borreliaceae. In: Whitman WB, Rainey R, Kämpfe P, Trujillo M, Chun J, DeVos P. Bergey's Manual of Systematics of Archaea and Bacteria:p. 1–9.

168. Geissinger O, Herlemann DP, Morschel E, Maier UG, Brune A. 2009. The ultramicrobacterium "Elusimicrobium minutum" gen. nov., sp. nov., the first cultivated representative of the termite group 1 phylum. Appl Environ Microbiol 75:2831-40.

169. Lagkouvardos I, Lesker TR, Hitch TCA, Galvez EJC, Smit N, Neuhaus K, Wang J, Baines JF, Abt B, Stecher B, Overmann J, Strowig T, Clavel T. 2019. Sequence and cultivation study of Muribaculaceae reveals novel species, host preference, and functional potential of this yet undescribed family. Microbiome 7:28.

170. Doran DJ. 1954. A Catalogue of the Protozoa and Helminths of North American Rodents. I. Protozoa and Acanthocephala. American Midland Naturalist 52:118.

171. Ericsson AC, Gagliardi J, Bouhan D, Spollen WG, Givan SA, Franklin CL. 2018. The influence of caging, bedding, and diet on the composition of the microbiota in different regions of the mouse gut. Sci Rep 8:4065.

172. Steiner JM SS, Pantchev N, Balzer HJ, Vrhovec MG, Lesina M. 2016. Use of ronidazole and limited culling to eliminate Tritrichomonas muris from laboratory mice. . J Am Assoc Lab Animal Sci 55(4):480-3.

173. Koyama T, Endo T, Asahi H, Kuroki T. 1987. Life cycle of Tritrichomonas muris. Zentralbl Bakteriol Mikrobiol Hyg A 264:478-86.

174. Hackstein JH, Akhmanova A, Boxma B, Harhangi HR, Voncken FG. 1999. Hydrogenosomes: eukaryotic adaptations to anaerobic environments. Trends Microbiol 7:441-7.

175. Rae DO, Crews JE. 2006. Tritrichomonas foetus. Vet Clin North Am Food Anim Pract 22:595-611.

176. Yao C, Koster LS. 2015. Tritrichomonas foetus infection, a cause of chronic diarrhea in the domestic cat. Vet Res 46:35.

177. Chudnovskiy A, Mortha A, Kana V, Kennard A, Ramirez JD, Rahman A, Remark R, Mogno I, Ng R, Gnjatic S, Amir ED, Solovyov A, Greenbaum B, Clemente J, Faith J, Belkaid Y, Grigg ME, Merad M. 2016. Host-Protozoan Interactions Protect from Mucosal Infections through Activation of the Inflammasome. Cell 167:444-456 e14.

178. Howitt MR, Lavoie S, Michaud M, Blum AM, Tran SV, Weinstock JV, Gallini CA, Redding K, Margolskee RF, Osborne LC, Artis D, Garrett WS. 2016. Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. Science 351:1329-33.

179. Escalante NK, Lemire P, Cruz Tleugabulova M, Prescott D, Mortha A, Streutker CJ, Girardin SE, Philpott DJ, Mallevaey T. 2016. The common mouse protozoa Tritrichomonas muris alters mucosal T cell homeostasis and colitis susceptibility. J Exp Med 213:2841-2850.

180. Lange BM, Croteau R. 1999. Isopentenyl diphosphate biosynthesis via a mevalonate-independent pathway: isopentenyl monophosphate kinase catalyzes the terminal enzymatic step. Proc Natl Acad Sci U S A 96:13714-9.

181. Hammes WP VR. 1995. The genus Lactobacillus. In: Wood B, Holzapfel W, editors. The Genera of Lactic Acid Bacteria. Springer;:19-54.

182. Salzman NH, de Jong H, Paterson Y, Harmsen HJM, Welling GW, Bos NA. 2002. Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. Microbiology (Reading) 148:3651-3660.

183. MD C. 1973. A survey of gross stomach morphology in New World Cricetinae (Rodentia, Muroidea), with comments on functional interpretations. Miscellaneous Publications, Museum of Zoology, University of Michigan:1-43.

184. Savage DC, Dubos R, Schaedler RW. 1968. The gastrointestinal epithelium and its autochthonous bacterial flora. J Exp Med 127:67-76.

185. Wesney E TG. 1979. Association of rat, pig, and fowl biotypes of lactobacilli with the stomach of gnotobiotic mice. Microbial Ecol:35-42.

186. Dubos R, Schaedler RW, Costello R, Hoet P. 1965. Indigenous, Normal, and Autochthonous Flora of the Gastrointestinal Tract. J Exp Med 122:67-76.

187. Salas-Jara MJ, Ilabaca A, Vega M, Garcia A. 2016. Biofilm Forming Lactobacillus: New Challenges for the Development of Probiotics. Microorganisms 4.

188. Almiron M, Traglia G, Rubio A, Sanjuan N. 2013. Colonization of the mouse upper gastrointestinal tract by lactobacillus murinus: a histological, immunocytochemical, and ultrastructural study. Curr Microbiol 67:395-8.

189. Dertli E, Mayer MJ, Narbad A. 2015. Impact of the exopolysaccharide layer on biofilms, adhesion and resistance to stress in Lactobacillus johnsonii FI9785. BMC Microbiol 15:8.

190. Lin XB, Wang T, Stothard P, Corander J, Wang J, Baines JF, Knowles SCL, Baltrunaite L, Tasseva G, Schmaltz R, Tollenaar S, Cody LA, Grenier T, Wu W, Ramer-Tait AE, Walter J. 2018. The evolution of ecological facilitation within mixed-species biofilms in the mouse gastrointestinal tract. ISME J 12:2770-2784.

191. Huang CI, Kay SC, Davis S, Tufts DM, Gaffett K, Tefft B, Diuk-Wasser MA. 2019. High burdens of Ixodes scapularis larval ticks on white-tailed deer may limit Lyme disease risk in a low biodiversity setting. Ticks Tick Borne Dis 10:258-268.

192. Finch C, Al-Damluji MS, Krause PJ, Niccolai L, Steeves T, O'Keefe CF, Diuk-Wasser MA. 2014. Integrated assessment of behavioral and environmental risk factors for Lyme disease infection on Block Island, Rhode Island. PLoS One 9:e84758.

193. Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165-90.

194. Zheng J, Wittouck S, Salvetti E, Franz C, Harris HMB, Mattarelli P, O'Toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Ganzle MG, Lebeer S. 2020. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. Int J Syst Evol Microbiol 70:2782-2858.

195. Qiao N, Wittouck S, Mattarelli P, Zheng J, Lebeer S, Felis GE, Ganzle MG. 2022. After the storm-Perspectives on the taxonomy of Lactobacillaceae. JDS Commun 3:222-227.

196. Jiang J, Li K, Xiao Y, Zhong A, Tang J, Duan Y, Li Z. 2022. Limosilactobacillus reuteri Regulating Intestinal Function: A Review. Fermentation 9:19.

197. Hazards EPoB, Koutsoumanis K, Allende A, Alvarez-Ordonez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Cocconcelli PS, Fernandez Escamez PS, Prieto-Maradona M, Querol A, Sijtsma L, Evaristo Suarez J, Sundh I, Vlak J, Barizzone F, Hempen M, Herman L. 2022. Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 15: suitability of taxonomic units notified to EFSA until September 2021. EFSA J 20:e07045.

198. Salvetti E, O'Toole PW. 2017. The Genomic Basis of Lactobacilli as Health-Promoting Organisms. Microbiol Spectr 5.

199. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595.

200. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851-D860.

201. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Goker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acids Res 50:D801-D807.

202. Bohlin J, Eldholm V, Pettersson JH, Brynildsrud O, Snipen L. 2017. The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes. BMC Genomics 18:151.

203. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen CY, Graham M, Van Domselaar G, Stothard P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res 51:W484-W492.

204. Mitsuoka T, Fujisawa T. 1987. Lactobacillus-Hamsteri, a New Species from the Intestine of Hamsters. Proceedings of the Japan Academy Series B-Physical and Biological Sciences 63:269-272.

205. Reimer LC, Sarda Carbasse J, Koblitz J, Ebeling C, Podstawka A, Overmann J. 2022. BacDive in 2022: the knowledge base for standardized bacterial and archaeal data. Nucleic Acids Res 50:D741-D746.

206. Surve S, Shinde DB, Kulkarni R. 2022. Isolation, characterization and comparative genomics of potentially probiotic Lactiplantibacillus plantarum strains from Indian foods. Sci Rep 12:1940.

207. Dubos R, Schaedler RW. 1962. Some biological effects of the digestive flora. Am J Med Sci 244:265-71.

208. Dubos RJ, Schaedler RW. 1962. The effect of diet on the fecal bacterial flora of mice and on their resistance to infection. J Exp Med 115:1161-72.

209. Schaedler RW, Dubos RJ. 1962. The fecal flora of various strains of mice. Its bearing on their susceptibility to endotoxin. J Exp Med 115:1149-60.

210. Rosshart SP, Vassallo BG, Angeletti D, Hutchinson DS, Morgan AP, Takeda K, Hickman HD, McCulloch JA, Badger JH, Ajami NJ, Trinchieri G, Pardo-Manuel de Villena F, Yewdell JW, Rehermann B. 2017. Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance. Cell 171:1015-1028 e13.

211. del Rio B, Seegers JF, Gomes-Solecki M. 2010. Immune response to Lactobacillus plantarum expressing Borrelia burgdorferi OspA is modulated by the lipid modification of the antigen. PLoS One 5:e11199.

212. Carroll IM, Threadgill DW, Threadgill DS. 2009. The gastrointestinal microbiome: a malleable, third genome of mammals. Mamm Genome 20:395-403.

213. Center PGS. 2017. Peromyscus Genetic Stock Center, Columbia, SC: University of South Carolina;. http://stkctr.biol.sc.edu. Accessed

214. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-20.

215. de Jong A vHA. 2019. BAGEL4 Groningen, the Netherlands: University of Groningen;. http://bagel4.molgenrug.nl. Accessed

216. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res 46:W278-W281.

217. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9:386.

218. Gomez-Alvarez V, Teal TK, Schmidt TM. 2009. Systematic artifacts in metagenomes from complex microbial communities. ISME J 3:1314-7.

219. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11:485.

220. Hill MO. 1973. Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology 54:427-432.

221. Bray JR, Curtis JT. 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs 27:325-349.

222. Ward JH. 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58:236-244.

223. Chong J, Liu P, Zhou G, Xia J. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat Protoc 15:799-821.

224. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037-1039.

225. Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221-4.

226. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-7.

227. Benjamini Y, Hochberg Y. 2018. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289-300.

228. Paules CI, Marston HD, Bloom ME, Fauci AS. 2018. Tickborne Diseases - Confronting a Growing Threat. N Engl J Med 379:701-703.

229. ER. H. 1981. The mammals of North America,. John Wiley and Sons, New York, NY 2nd ed, vol 2.

230. Mather TN, Wilson ML, Moore SI, Ribeiro JM, Spielman A. 1989. Comparing the relative potential of rodents as reservoirs of the Lyme disease spirochete (Borrelia burgdorferi). Am J Epidemiol 130:143-50.

231. Lindsay LR, Barker IK, Surgeoner GA, McEwen SA, Campbell GD. 1997. Duration of Borrelia burgdorferi infectivity in white-footed mice for the tick vector Ixodes scapularis under laboratory and field conditions in Ontario. J Wildl Dis 33:766-75.

232. Barbour AG, Restrepo BI. 2000. Antigenic variation in vector-borne pathogens. Emerg Infect Dis 6:449-57.

233. Moody KD, Terwilliger GA, Hansen GM, Barthold SW. 1994. Experimental Borrelia burgdorferi infection in Peromyscus leucopus. J Wildl Dis 30:155-61.

234. Wright SD NS. 1990. Experimental infection of the white-footed mouse with Borrelia burgdorferi. Am J Vet Res 51:1980–1987.

235. Raberg L, Sim D, Read AF. 2007. Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. Science 318:812-4.

236. Schneider DS, Ayres JS. 2008. Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. Nat Rev Immunol 8:889-95.

237. 2014. N. 2021. Immune tolerance. . https://www.niaid.nih.gov/research/immune-tolerance. Accessed Accessed 7 January 2021. .

238. Jackson JA, Hall AJ, Friberg IM, Ralli C, Lowe A, Zawadzka M, Turner AK, Stewart A, Birtles RJ, Paterson S, Bradley JE, Begon M. 2014. An immunological marker of tolerance to infection in wild rodents. PLoS Biol 12:e1001901.

239. McCarville JL, Ayres JS. 2018. Disease tolerance: concept and mechanisms. Curr Opin Immunol 50:88-93.

240. Botten J, Mirowsky K, Kusewitt D, Bharadwaj M, Yee J, Ricci R, Feddersen RM, Hjelle B. 2000. Experimental infection model for Sin Nombre hantavirus in the deer mouse (Peromyscus maniculatus). Proc Natl Acad Sci U S A 97:10578-83.

241. Schountz T, Prescott J, Cogswell AC, Oko L, Mirowsky-Garcia K, Galvez AP, Hjelle B. 2007. Regulatory T cell-like responses in deer mice persistently infected with Sin Nombre virus. Proc Natl Acad Sci U S A 104:15496-501.

242. Mandl JN, Schneider C, Schneider DS, Baker ML. 2018. Going to Bat(s) for Studies of Disease Tolerance. Front Immunol 9:2112.

243. Wang LF, Anderson DE. 2019. Viruses in bats and potential spillover to animals and humans. Curr Opin Virol 34:79-89.

244. Klemme I, Karvonen A. 2017. Vertebrate defense against parasites: Interactions between avoidance, resistance, and tolerance. Ecol Evol 7:561-571.

245. Mandl JN, Ahmed R, Barreiro LB, Daszak P, Epstein JH, Virgin HW, Feinberg MB. 2015. Reservoir host immune responses to emerging zoonotic viruses. Cell 160:20-35.

246. Joyner CP, Myrick LC, Crossland JP, Dawson WD. 1998. Deer Mice As Laboratory Animals. ILAR J 39:322-330.

247. Barthold SW CD, Phillip MT. 2010. Animal models of borreliosis, Borrelia: molecular biology, host interaction, and pathogenesis. . Caister Academic Press, Norfolk, United Kingdom.

248. Galanos C. 1998. Endotoxin (Lipopolysaccharide (LPS)). doi:10.1006/rwei.1999.0213:806-809.

249. Beeson PB, Technical Assistance of Elizabeth R. 1947. Tolerance to Bacterial Pyrogens : I. Factors Influencing Its Development. J Exp Med 86:29-38.

250. Foster SL, Hargreaves DC, Medzhitov R. 2007. Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature 447:972-8.

251. Medvedev AE, Kopydlowski KM, Vogel SN. 2000. Inhibition of lipopolysaccharide-induced signal transduction in endotoxin-tolerized mouse macrophages: dysregulation of cytokine, chemokine, and toll-like receptor 2 and 4 gene expression. J Immunol 164:5564-74.

252. Fitzgerald KA, Palsson-McDermott EM, Bowie AG, Jefferies CA, Mansell AS, Brady G, Brint E, Dunne A, Gray P, Harte MT, McMurray D, Smith DE, Sims JE, Bird TA, O'Neill LA. 2001. Mal (MyD88-adapter-like) is required for Toll-like receptor-4 signal transduction. Nature 413:78-83.

253. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science 282:2085-8.

254. Mahieu T, Park JM, Revets H, Pasche B, Lengeling A, Staelens J, Wullaert A, Vanlaere I, Hochepied T, van Roy F, Karin M, Libert C. 2006. The wild-derived inbred mouse strain SPRET/Ei is resistant to LPS and defective in IFN-beta production. Proc Natl Acad Sci U S A 103:2292-7.

255. McCuskey RS, McCuskey PA, Urbaschek R, Urbaschek B. 1984. Species differences in Kupffer cells and endotoxin sensitivity. Infect Immun 45:278-80.

256. Raduolovic K, Mak'Anyengo R, Kaya B, Steinert A, Niess JH. 2018. Injections of Lipopolysaccharide into Mice to Mimic Entrance of Microbial-derived Products After Intestinal Barrier Breach. J Vis Exp doi:10.3791/57610.

257. Reynolds K, Novosad B, Hoffhines A, Gipson J, Johnson J, Peters J, Gonzalez F, Gimble J, Hill M. 2002. Pretreatment with troglitazone decreases lethality during endotoxemia in mice. J Endotoxin Res 8:307-14.

258. Schaedler RW, Dubos RJ. 1961. The susceptibility of mice to bacterial endotoxins. J Exp Med 113:559-70.

259. Wei XQ, Charles IG, Smith A, Ure J, Feng GJ, Huang FP, Xu D, Muller W, Moncada S, Liew FY. 1995. Altered immune responses in mice lacking inducible nitric oxide synthase. Nature 375:408-11.

260. Langeroudi AG, Hirsch CM, Estabragh AS, Meinardi S, Blake DR, Barbour AG. 2014. Elevated carbon monoxide to carbon dioxide ratio in the exhaled breath of mice treated with a single dose of lipopolysaccharide. Open Forum Infect Dis 1:ofu085.

261. Long PN, Cook VJ, Majumder A, Barbour AG, Long AD. 2022. The utility of a closed breeding colony of Peromyscus leucopus for dissecting complex traits. Genetics 221.

262. Ravishankar B, Liu H, Shinde R, Chaudhary K, Xiao W, Bradley J, Koritzinsky M, Madaio MP, McGaha TL. 2015. The amino acid sensor GCN2 inhibits inflammatory responses to apoptotic cells promoting tolerance and suppressing systemic autoimmunity. Proc Natl Acad Sci U S A 112:10774-9.

263. Sorgdrager FJH, Naude PJW, Kema IP, Nollen EA, Deyn PP. 2019. Tryptophan Metabolism in Inflammaging: From Biomarker to Therapeutic Target. Front Immunol 10:2565.

264. Murray PJ, Allen JE, Biswas SK, Fisher EA, Gilroy DW, Goerdt S, Gordon S, Hamilton JA, Ivashkiv LB, Lawrence T, Locati M, Mantovani A, Martinez FO, Mege JL, Mosser DM, Natoli G, Saeij JP, Schultze JL, Shirey KA, Sica A, Suttles J, Udalova I, van Ginderachter JA, Vogel SN, Wynn TA. 2014. Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity 41:14-20.

265. Lee JT, Bartolomei MS. 2013. X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell 152:1308-23.

266. Inoue H, Ogonuki N, Hirose M, Hatanaka Y, Matoba S, Chuma S, Kobayashi K, Wakana S, Noguchi J, Inoue K, Tanemura K, Ogura A. 2016. Mouse D1Pas1, a DEAD-box RNA helicase,

is required for the completion of first meiotic prophase in male germ cells. Biochem Biophys Res Commun 478:592-8.

267. Itoh Y, Golden LC, Itoh N, Matsukawa MA, Ren E, Tse V, Arnold AP, Voskuhl RR. 2019. The X-linked histone demethylase Kdm6a in CD4+ T lymphocytes modulates autoimmunity. J Clin Invest 129:3852-3863.

268. Ouchi N, Parker JL, Lugus JJ, Walsh K. 2011. Adipokines in inflammation and metabolic disease. Nat Rev Immunol 11:85-97.

269. van Stijn CM, Kim J, Lusis AJ, Barish GD, Tangirala RK. 2015. Macrophage polarization phenotype regulates adiponectin receptor expression and adiponectin anti-inflammatory response. FASEB J 29:636-49.

270. Zhong C, Zheng M, Cui K, Martins AJ, Hu G, Li D, Tessarollo L, Kozlov S, Keller JR, Tsang JS, Zhao K, Zhu J. 2020. Differential Expression of the Transcription Factor GATA3 Specifies Lineage and Functions of Innate Lymphoid Cells. Immunity 52:83-95 e4.

271. Boivin GA, Pothlichet J, Skamene E, Brown EG, Loredo-Osti JC, Sladek R, Vidal SM. 2012. Mapping of clinical and expression quantitative trait loci in a sex-dependent effect of host susceptibility to mouse-adapted influenza H3N2/HK/1/68. J Immunol 188:3949-60.

272. Han L, Diehl A, Nguyen NK, Korangath P, Teo W, Cho S, Kominsky S, Huso DL, Feigenbaum L, Rein A, Argani P, Landberg G, Gessler M, Sukumar S. 2014. The Notch pathway inhibits TGFbeta signaling in breast cancer through HEYL-mediated crosstalk. Cancer Res 74:6509-18.

273. H. F. 1988. Human mucus proteinase inhibitor (human MPI). Human seminal inhibitor I (HUSI-I), antileukoprotease (ALP), secretory leukocyte protease inhibitor (SLPI). Biol Chem Hoppe Seyler 369:79–82.

274. Levi M, van der Poll T. 2017. Coagulation and sepsis. Thromb Res 149:38-44.

275. Zhang Y, Zhang J, Sheng H, Li H, Wang R. 2019. Acute phase reactant serum amyloid A in inflammation and other diseases. Adv Clin Chem 90:25-80.

276. Mosser DM, Edwards JP. 2008. Exploring the full spectrum of macrophage activation. Nat Rev Immunol 8:958-69.

277. Luan HH, Wang A, Hilliard BK, Carvalho F, Rosen CE, Ahasic AM, Herzog EL, Kang I, Pisani MA, Yu S, Zhang C, Ring AM, Young LH, Medzhitov R. 2019. GDF15 Is an Inflammation-Induced Central Mediator of Tissue Tolerance. Cell 178:1231-1244 e11.

278. Johnson TL, Fischer RJ, Raffel SJ, Schwan TG. 2016. Host associations and genomic diversity of Borrelia hermsii in an endemic focus of tick-borne relapsing fever in western North America. Parasit Vectors 9:575.

279. Takayama K, Rothenberg RJ, Barbour AG. 1987. Absence of lipopolysaccharide in the Lyme disease spirochete, Borrelia burgdorferi. Infect Immun 55:2311-3.

280. Bergström S ZW. 2010. Borrelia: molecular biology, host interaction and pathogenesis. . In Radolf JD, Samuels DS (ed), , Caister Academic Press, Norwich, United Kingdom.

281. Li Q, Li L, Fei X, Zhang Y, Qi C, Hua S, Gong F, Fang M. 2018. Inhibition of autophagy with 3-methyladenine is protective in a lethal model of murine endotoxemia and polymicrobial sepsis. Innate Immun 24:231-239.

282. Bhowmick R, Clark S, Bonventre JV, Leong JM, McCormick BA. 2017. Cytosolic Phospholipase A(2)alpha Promotes Pulmonary Inflammation and Systemic Disease during Streptococcus pneumoniae Infection. Infect Immun 85.

283. McDermott AM. 2013. Antimicrobial compounds in tears. Exp Eye Res 117:53-61.

284. Aslam S GV. 2021. Carbonic anhydrase inhibitors. StatPearls Publishing,, Treasure Island, FL.

285. Hart BL. 1988. Biological basis of the behavior of sick animals. Neurosci Biobehav Rev 12:123-37.
286. Wiedmeyer CE CJ, Veres M, Dewey MJ, Felder MR, Barlow SC, Vrana PB, Szalai G. . 2014. Hematologic and serum biochemical values of 4 species of Peromyscus mice and their hybrids. , vol 53. J Am Assoc Lab Anim Sci
287. River. C. 2012. BALB/C mouse hematology. North American colonies. , *on* Charles River Laboratories International, . https://www.criver.com/sites/default/files/resources/BALBcMouseClinicalPathologyData.pdf. . Accessed Accessed 6 February 2021.
288. Wu PJ GE, Hansen LG, Segre M. . 1999. Hematology values from clinically healthy Peromyscus leucopus. , vol 30. J Zoo Wildl Med.
289. Majchrzak-Gorecka M, Majewski P, Grygier B, Murzyn K, Cichy J. 2016. Secretory leukocyte protease inhibitor (SLPI), a multifunctional protein in the host defense response. Cytokine Growth Factor Rev 28:79-93.
290. Zabieglo K, Majewski P, Majchrzak-Gorecka M, Wlodarczyk A, Grygier B, Zegar A, Kapinska-Mrowiecka M, Naskalska A, Pyrc K, Dubin A, Wahl SM, Cichy J. 2015. The inhibitory effect of secretory leukocyte protease inhibitor (SLPI) on formation of neutrophil extracellular traps. J Leukoc Biol 98:99-106.
291. Taggart CC, Greene CM, McElvaney NG, O'Neill S. 2002. Secretory leucoprotease inhibitor prevents lipopolysaccharide-induced IkappaBalpha degradation without affecting phosphorylation or ubiquitination. J Biol Chem 277:33648-53.
292. Ashcroft GS, Lei K, Jin W, Longenecker G, Kulkarni AB, Greenwell-Wild T, Hale-Donze H, McGrady G, Song XY, Wahl SM. 2000. Secretory leukocyte protease inhibitor mediates non-redundant functions necessary for normal wound healing. Nat Med 6:1147-53.
293. McCartney-Francis N, Jin W, Belkaid Y, McGrady G, Wahl SM. 2014. Aberrant host defense against Leishmania major in the absence of SLPI. J Leukoc Biol 96:917-29.
294. Nakamura A, Mori Y, Hagiwara K, Suzuki T, Sakakibara T, Kikuchi T, Igarashi T, Ebina M, Abe T, Miyazaki J, Takai T, Nukiwa T. 2003. Increased susceptibility to LPS-induced endotoxin shock in secretory leukoprotease inhibitor (SLPI)-deficient mice. J Exp Med 197:669-74.
295. Sugimoto MA, Vago JP, Teixeira MM, Sousa LP. 2016. Annexin A1 and the Resolution of Inflammation: Modulation of Neutrophil Recruitment, Apoptosis, and Clearance. J Immunol Res 2016:8239258.
296. Kong F, Saldarriaga OA, Spratt H, Osorio EY, Travi BL, Luxon BA, Melby PC. 2017. Transcriptional Profiling in Experimental Visceral Leishmaniasis Reveals a Broad Splenic Inflammatory Environment that Conditions Macrophages toward a Disease-Promoting Phenotype. PLoS Pathog 13:e1006165.
297. Osorio EY, Zhao W, Espitia C, Saldarriaga O, Hawel L, Byus CV, Travi BL, Melby PC. 2012. Progressive visceral leishmaniasis is driven by dominant parasite-induced STAT6 activation and STAT6-dependent host arginase 1 expression. PLoS Pathog 8:e1002417.
298. MacMicking JD, Nathan C, Hom G, Chartrain N, Fletcher DS, Trumbauer M, Stevens K, Xie QW, Sokol K, Hutchinson N, et al. 1995. Altered responses to bacterial infection and endotoxic shock in mice lacking inducible nitric oxide synthase. Cell 81:641-50.
299. Massung RF, Priestley RA, Levin ML. 2004. Transmission route efficacy and kinetics of Anaplasma phagocytophilum infection in white-footed mouse, Peromyscus leucopus. Vector Borne Zoonotic Dis 4:310-8.
300. Mlera L, Meade-White K, Saturday G, Scott D, Bloom ME. 2017. Modeling Powassan virus infection in Peromyscus leucopus, a natural host. PLoS Negl Trop Dis 11:e0005346.

301.   Kruger TE, Miller AH, Godwin AK, Wang J. 2014. Bone sialoprotein and osteopontin in bone metastasis of osteotropic cancers. Crit Rev Oncol Hematol 89:330-41.

302.   Castello LM, Baldrighi M, Molinari L, Salmi L, Cantaluppi V, Vaschetto R, Zunino G, Quaglia M, Bellan M, Gavelli F, Navalesi P, Avanzi GC, Chiocchetti A. 2019. The Role of Osteopontin as a Diagnostic and Prognostic Biomarker in Sepsis and Septic Shock. Cells 8.

303.   Guo H, Wai PY, Mi Z, Gao C, Zhang J, Kuo PC. 2008. Osteopontin mediates Stat1 degradation to inhibit iNOS transcription in a cecal ligation and puncture model of sepsis. Surgery 144:182-8.

304.   Belkacem N, Serafini N, Wheeler R, Derrien M, Boucinha L, Couesnon A, Cerf-Bensussan N, Gomperts Boneca I, Di Santo JP, Taha MK, Bourdet-Sicard R. 2017. Lactobacillus paracasei feeding improves immune control of influenza infection in mice. PLoS One 12:e0184976.

305.   Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK, Le A, Cowan TM, Nolan GP, Fischbach MA, Sonnenburg JL. 2017. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. Nature 551:648-652.

306.   Negatu DA, Gengenbacher M, Dartois V, Dick T. 2020. Indole Propionic Acid, an Unusual Antibiotic Produced by the Gut Microbiota, With Anti-inflammatory and Antioxidant Properties. Front Microbiol 11:575586.

307.   Beeson PB, Technical Assistance of Elizabeth R. 1947. Tolerance to Bacterial Pyrogens : Ii. Role of the Reticulo-Endothelial System. J Exp Med 86:39-44.

308.   Pena OM, Pistolic J, Raj D, Fjell CD, Hancock RE. 2011. Endotoxin tolerance represents a distinctive state of alternative polarization (M2) in human mononuclear cells. J Immunol 186:7243-54.

309.   Committee for the Update of the Guide for the Care and Use of Laboratory Animals IfLAR, National Research Council. 2011. Guide for the care and use of laboratory animals. National Institutes of Health, Bethesda, MD.

310.   Barbour AG, Bunikis J, Travinsky B, Hoen AG, Diuk-Wasser MA, Fish D, Tsao JI. 2009. Niche partitioning of Borrelia burgdorferi and Borrelia miyamotoi in the same tick vector and mammalian reservoir species. Am J Trop Med Hyg 81:1120-31.

311.   Zhang Y, Lucius MD, Altomare D, Havighorst A, Farmaki E, Chatzistamou I, Shtutman M, Kiaris H. 2019. Coordination Analysis of Gene Expression Points to the Relative Impact of Different Regulators During Endoplasmic Reticulum Stress. DNA Cell Biol 38:969-981.

312.   Fitzgerald BL, Graham B, Delorey MJ, Pegalajar-Jurado A, Islam MN, Wormser GP, Aucott JN, Rebman AW, Soloski MJ, Belisle JT, Molins CR. 2021. Metabolic Response in Patients With Post-treatment Lyme Disease Symptoms/Syndrome. Clin Infect Dis 73:e2342-e2349.

313.   Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak MY, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P. 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30:918-20.

314.   Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. 2006. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 78:779-87.

315.   Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH. 2003. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. Anal Chem 75:4818-26.

316. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, Jones DP, Pulendran B. 2013. Predicting network activity from high throughput metabolomics. PLoS Comput Biol 9:e1003123.

317. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966-8.

318. Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34:525-7.

319. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-40.

320. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14:128.

321. Team. RC. 2103. A language and environment for statistical computing, *on* Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org. Accessed

322. Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

323. Foroushani A, Agrahari R, Docking R, Chang L, Duns G, Hudoba M, Karsan A, Zare H. 2017. Large-scale gene network analysis reveals the significance of extracellular matrix pathway and homeobox genes in acute myeloid leukemia: an introduction to the Pigengene package and its applications. BMC Med Genomics 10:16.

324. Labinskyy N, Mukhopadhyay P, Toth J, Szalai G, Veres M, Losonczy G, Pinto JT, Pacher P, Ballabh P, Podlutsky A, Austad SN, Csiszar A, Ungvari Z. 2009. Longevity is associated with increased vascular resistance to high glucose-induced oxidative stress and inflammatory gene expression in Peromyscus leucopus. Am J Physiol Heart Circ Physiol 296:H946-56.

325. Hall E. 1983. l98l. The mammals of North America. John Wiley and Sons, New York 1:1-600.

326. Moscarella RA, Hoffman SMG, Myers P, Yahnke CJ, Lundrigan BL. 2019. Genetic and demographic analysis of invasive Peromyscus leucopus in the northern Great Lakes region. Journal of Mammalogy 100:345-353.

327. Biser JA, Vogel LA, Berger J, Hjelle B, Loew SS. 2004. Effects of heavy metals on immunocompetence of white-footed mice (Peromyscus leucopus). J Wildl Dis 40:173-84.

328. Levengood JM, Heske EJ. 2008. Heavy metal exposure, reproductive activity, and demographic patterns in white-footed mice (Peromyscus leucopus) inhabiting a contaminated floodplain wetland. Sci Total Environ 389:320-8.

329. Munshi-South J, Kharchenko K. 2010. Rapid, pervasive genetic differentiation of urban white-footed mouse (Peromyscus leucopus) populations in New York City. Mol Ecol 19:4242-54.

330. Bradley RD, Ammerman LK, Baker RJ, Bradley LC, Cook JA, Dowler RC, Jones C, Schmidly DJ, Stangl FB, Van Den Bussche RA. 2014. Revised checklist of North American mammals north of Mexico, 2014. Museum of Texas Tech University Lubbock, TX, USA.

331. Xiao W, Mindrinos MN, Seok J, Cuschieri J, Cuenca AG, Gao H, Hayden DL, Hennessy L, Moore EE, Minei JP, Bankey PE, Johnson JL, Sperry J, Nathens AB, Billiar TR, West MA, Brownstein BH, Mason PH, Baker HV, Finnerty CC, Jeschke MG, Lopez MC, Klein MB, Gamelli RL, Gibran NS, Arnoldo B, Xu W, Zhang Y, Calvano SE, McDonald-Smith GP, Schoenfeld DA, Storey JD, Cobb JP, Warren HS, Moldawer LL, Herndon DN, Lowry SF, Maier RV, Davis RW, Tompkins RG, Inflammation, Host Response to Injury Large-Scale Collaborative Research P. 2011. A genomic storm in critically injured humans. J Exp Med 208:2581-90.

332.   Bashirova AA, Apps R, Vince N, Mochalova Y, Yu XG, Carrington M. 2014. Diversity of the human LILRB3/A6 locus encoding a myeloid inhibitory and activating receptor pair. Immunogenetics 66:1-8.

333.   Schnizlein-Bick CT, Mandy FF, O'Gorman MR, Paxton H, Nicholson JK, Hultin LE, Gelman RS, Wilkening CL, Livnat D. 2002. Use of CD45 gating in three and four-color flow cytometric immunophenotyping: guideline from the National Institute of Allergy and Infectious Diseases, Division of AIDS. Cytometry 50:46-52.

334.   Murray PJ. 2017. Macrophage Polarization. Annu Rev Physiol 79:541-566.

335.   Arranz A, Doxaki C, Vergadi E, Martinez de la Torre Y, Vaporidi K, Lagoudaki ED, Ieronymaki E, Androulidaki A, Venihaki M, Margioris AN, Stathopoulos EN, Tsichlis PN, Tsatsanis C. 2012. Akt1 and Akt2 protein kinases differentially contribute to macrophage polarization. Proc Natl Acad Sci U S A 109:9517-22.

336.   Vergadi E, Ieronymaki E, Lyroni K, Vaporidi K, Tsatsanis C. 2017. Akt Signaling Pathway in Macrophage Activation and M1/M2 Polarization. J Immunol 198:1006-1014.

337.   Narasimhan PB, Marcovecchio P, Hamers AAJ, Hedrick CC. 2019. Nonclassical Monocytes in Health and Disease. Annu Rev Immunol 37:439-456.

338.   Italiani P, Boraschi D. 2014. From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. Front Immunol 5:514.

339.   Bothwell A, Pace PE, LeClair KP. 1988. Isolation and expression of an IFN-responsive Ly-6C chromosomal gene. Journal of immunology (Baltimore, Md: 1950) 140:2815-2820.

340.   Lowenstein CJ, Alley EW, Raval P, Snowman AM, Snyder SH, Russell SW, Murphy WJ. 1993. Macrophage nitric oxide synthase gene: two upstream regions mediate induction by interferon gamma and lipopolysaccharide. Proc Natl Acad Sci U S A 90:9730-4.

341.   Salkowski CA, Detore G, McNally R, van Rooijen N, Vogel SN. 1997. Regulation of inducible nitric oxide synthase messenger RNA expression and nitric oxide production by lipopolysaccharide in vivo: the roles of macrophages, endogenous IFN-gamma, and TNF receptor-1-mediated signaling. The Journal of Immunology 158:905-912.

342.   Quatrini L, Wieduwild E, Guia S, Bernat C, Glaichenhaus N, Vivier E, Ugolini S. 2017. Host resistance to endotoxic shock requires the neuroendocrine regulation of group 1 innate lymphoid cells. J Exp Med 214:3531-3541.

343.   Heinzelmann M, Polk HC, Jr., Chernobelsky A, Stites TP, Gordon LE. 2000. Endotoxin and muramyl dipeptide modulate surface receptor expression on human mononuclear cells. Immunopharmacology 48:117-28.

344.   Mazgaeen L, Gurung P. 2020. Recent Advances in Lipopolysaccharide Recognition Systems. Int J Mol Sci 21.

345.   Hurst TP, Magiorkinis G. 2015. Activation of the innate immune response by endogenous retroviruses. J Gen Virol 96:1207-1218.

346.   Lima-Junior DS, Krishnamurthy SR, Bouladoux N, Collins N, Han SJ, Chen EY, Constantinides MG, Link VM, Lim AI, Enamorado M, Cataisson C, Gil L, Rao I, Farley TK, Koroleva G, Attig J, Yuspa SH, Fischbach MA, Kassiotis G, Belkaid Y. 2021. Endogenous retroviruses promote homeostatic and inflammatory responses to the microbiota. Cell 184:3794-3811 e19.

347.   Rangel SC, da Silva MD, da Silva AL, Dos Santos JMB, Neves LM, Pedrosa A, Rodrigues FM, Trettel CDS, Furtado GE, de Barros MP, Bachi ALL, Romano CM, Nali L. 2022. Human endogenous retroviruses and the inflammatory response: A vicious circle associated with health and illness. Front Immunol 13:1057791.

348.   Barbour AG. 2023. Borreliaceae.

349. Salazar JC, Duhnam-Ems S, La Vake C, Cruz AR, Moore MW, Caimano MJ, Velez-Climent L, Shupe J, Krueger W, Radolf JD. 2009. Activation of human monocytes by live Borrelia burgdorferi generates TLR2-dependent and -independent responses which include induction of IFN-beta. PLoS Pathog 5:e1000444.

350. Chiaranunt P, Burrows K, Ngai L, Cao EY, Liang H, Tai SL, Streutker CJ, Girardin SE, Mortha A. 2022. NLRP1B and NLRP3 Control the Host Response following Colonization with the Commensal Protist Tritrichomonas musculis. J Immunol 208:1782-1789.

351. Hara I, Izui S, McConahey PJ, Elder JH, Jensen FC, Dixon FJ. 1981. Induction of high serum levels of retroviral env gene products (gp70) in mice by bacterial lipopolysaccharide. Proc Natl Acad Sci U S A 78:4397-401.

352. Jongstra J, Moroni C. 1981. Lipopolysaccharide induces retroviral antigen expression in 129/J mouse lymphocytes: evidence for assembly of a defective viral particle. J Virol 37:1044-50.

353. Stoye JP, Moroni C. 1983. Endogenous retrovirus expression in stimulated murine lymphocytes. Identification of a new locus controlling mitogen induction of a defective virus. J Exp Med 157:1660-74.

354. Kong L, Sun L, Zhang H, Liu Q, Liu Y, Qin L, Shi G, Hu JH, Xu A, Sun YP, Li D, Shi YF, Zang JW, Zhu J, Chen Z, Wang ZG, Ge BX. 2009. An essential role for RIG-I in toll-like receptor-stimulated phagocytosis. Cell Host Microbe 6:150-61.

355. De Cecco M, Ito T, Petrashen AP, Elias AE, Skvir NJ, Criscione SW, Caligiana A, Brocculi G, Adney EM, Boeke JD, Le O, Beausejour C, Ambati J, Ambati K, Simon M, Seluanov A, Gorbunova V, Slagboom PE, Helfand SL, Neretti N, Sedivy JM. 2019. Author Correction: L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature 572:E5.

356. Mommert M, Tabone O, Guichard A, Oriol G, Cerrato E, Denizot M, Cheynet V, Pachot A, Lepape A, Monneret G, Venet F, Brengel-Pesce K, Textoris J, Mallet F, Group MIS, Group RS. 2020. Dynamic LTR retrotransposon transcriptome landscape in septic shock patients. Crit Care 24:96.

357. Pisano MP, Tabone O, Bodinier M, Grandi N, Textoris J, Mallet F, Tramontano E. 2020. RNA-Seq Transcriptome Analysis Reveals Long Terminal Repeat Retrotransposon Modulation in Human Peripheral Blood Mononuclear Cells after In Vivo Lipopolysaccharide Injection. J Virol 94.

358. Norgard MV, Arndt LL, Akins DR, Curetty LL, Harrich DA, Radolf JD. 1996. Activation of human monocytic cells by Treponema pallidum and Borrelia burgdorferi lipoproteins and synthetic lipopeptides proceeds via a pathway distinct from that of lipopolysaccharide but involves the transcriptional activator NF-kappa B. Infect Immun 64:3845-52.

359. Gozashti L, Feschotte C, Hoekstra HE. 2023. Transposable Element Interactions Shape the Ecology of the Deer Mouse Genome. Mol Biol Evol 40.

360. Yang P, Wang Y, Macfarlan TS. 2017. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. Trends Genet 33:871-881.

361. Wolf G, Yang P, Fuchtbauer AC, Fuchtbauer EM, Silva AM, Park C, Wu W, Nielsen AL, Pedersen FS, Macfarlan TS. 2015. The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes Dev 29:538-54.

362. Treger RS, Pope SD, Kong Y, Tokuyama M, Taura M, Iwasaki A. 2019. The Lupus Susceptibility Locus Sgp3 Encodes the Suppressor of Endogenous Retrovirus Expression SNERV. Immunity 50:334-347 e9.

363. Hoen AG, Margos G, Bent SJ, Diuk-Wasser MA, Barbour A, Kurtenbach K, Fish D. 2009. Phylogeography of Borrelia burgdorferi in the eastern United States reflects multiple independent Lyme disease emergence events. Proc Natl Acad Sci U S A 106:15013-8.

364. Loria-Cervera EN, Sosa-Bibiano EI, Van Wynsberghe NR, Andrade-Narvaez FJ. 2018. Finding a model for the study of Leishmania (Leishmania) mexicana infection: The Yucatan Deer mouse (Peromyscus yucatanicus) as a suitable option. Acta Trop 187:158-164.

365. Childs JE, Ksiazek TG, Spiropoulou CF, Krebs JW, Morzunov S, Maupin GO, Gage KL, Rollin PE, Sarisky J, Enscore RE, et al. 1994. Serologic and genetic identification of Peromyscus maniculatus as the primary rodent reservoir for a new hantavirus in the southwestern United States. J Infect Dis 169:1271-80.

366. Deschambault Y, Klassen L, Soule G, Tierney K, Azaransky K, Sloan A, Safronetz D. 2023. Experimental Infection of North American Deer Mice with Clade I and II Monkeypox Virus Isolates. Emerg Infect Dis 29:858-860.

367. Oka K, Yamakawa M, Kawamura Y, Kutsukake N, Miura K. 2023. The Naked Mole-Rat as a Model for Healthy Aging. Annu Rev Anim Biosci 11:207-226.

368. Gorshkova EA, Gubernatorova EO, Dvorianinova EM, Yurakova TR, Marey MV, Averina OA, Holtze S, Hildebrandt TB, Dmitriev AA, Drutskaya MS, Vyssokikh MY, Nedospasov SA. 2023. Macrophages from naked mole-rat possess distinct immunometabolic signatures upon polarization. Front Immunol 14:1172467.

369. Hilton HG, Rubinstein ND, Janki P, Ireland AT, Bernstein N, Fong NL, Wright KM, Smith M, Finkle D, Martin-McNulty B, Roy M, Imai DM, Jojic V, Buffenstein R. 2019. Single-cell transcriptomics of the naked mole-rat reveals unexpected features of mammalian immunity. PLoS Biol 17:e3000528.

370. Emmrich S, Trapp A, Tolibzoda Zakusilo F, Straight ME, Ying AK, Tyshkovskiy A, Mariotti M, Gray S, Zhang Z, Drage MG, Takasugi M, Klusmann JH, Gladyshev VN, Seluanov A, Gorbunova V. 2022. Characterization of naked mole-rat hematopoiesis reveals unique stem and progenitor cell patterns and neotenic traits. EMBO J 41:e109694.

371. Diuk-Wasser MA, Hoen AG, Cislo P, Brinkerhoff R, Hamer SA, Rowland M, Cortinas R, Vourc'h G, Melton F, Hickling GJ, Tsao JI, Bunikis J, Barbour AG, Kitron U, Piesman J, Fish D. 2012. Human risk of infection with Borrelia burgdorferi, the Lyme disease agent, in eastern United States. Am J Trop Med Hyg 86:320-7.

372. Steere AC, Bartenhagen NH, Craft JE, Hutchinson GJ, Newman JH, Pachner AR, Rahn DW, Sigal LH, Taylor E, Malawista SE. 1986. Clinical manifestations of Lyme disease. Zentralbl Bakteriol Mikrobiol Hyg A 263:201-5.

373. Steere AC, Sikand VK, Meurice F, Parenti DL, Fikrig E, Schoen RT, Nowakowski J, Schmid CH, Laukamp S, Buscarino C, Krause DS. 1998. Vaccination against Lyme disease with recombinant Borrelia burgdorferi outer-surface lipoprotein A with adjuvant. Lyme Disease Vaccine Study Group. N Engl J Med 339:209-15.

374. Coburn J, Garcia B, Hu LT, Jewett MW, Kraiczy P, Norris SJ, Skare J. 2021. Lyme Disease Pathogenesis. Curr Issues Mol Biol 42:473-518.

375. Nathan C. 2022. Nonresolving inflammation redux. Immunity 55:592-605.

376. McClure M, Kaye S. 2010. Can detection of xenotropic murine leukemia virus-related virus be linked to chronic fatigue syndrome? Expert Rev Mol Diagn 10:537-9.

377. Russ E, Iordanskiy S. 2023. Endogenous Retroviruses as Modulators of Innate Immunity. Pathogens 12.

378. Hart SN, Therneau TM, Zhang Y, Poland GA, Kocher JP. 2013. Calculating sample size estimates for RNA sequencing data. J Comput Biol 20:970-8.

379. McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 40:4288-97.

380. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10:1523.

381. Cohen J, . 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20:37–46.

382. Zar JH. 1999. Biostatistical analysis. Pearson Education India.

383. Hedges L.V., Gurevitch J. CPS. 1999. THE meta-analysis of response ratios in experimental ecology Ecology 80:1150–1156.

384. dashboard WC-. 2024. Number of COVID-19 cases reported to WHO (cumulative total). https://data.who.int/dashboards/covid19/cases?n=c. Accessed 03/05/2024.

385. dashboard WC-. 2024. Number of COVID-19 deaths reported to WHO (cumulative total). https://data.who.int/dashboards/covid19/deaths?n=c. Accessed 03/05/2024.

386. Yang X, Shi F, Zhang J, Gao H, Chen S, Olatosi B, Weissman S, Li X. 2023. Disease severity of COVID-19 in different phases of the pandemic: Do healthcare workers have better outcomes? Vaccine X 15:100377.

387. Tang D, Comish P, Kang R. 2020. The hallmarks of COVID-19 disease. PLoS Pathog 16:e1008536.

388. Davis HE, McCorkell L, Vogel JM, Topol EJ. 2023. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 21:133-146.

389. Wu YC, Chen CS, Chan YJ. 2020. The outbreak of COVID-19: An overview. J Chin Med Assoc 83:217-220.

390. Paules CI, Marston HD, Fauci AS. 2020. Coronavirus Infections-More Than Just the Common Cold. JAMA 323:707-708.

391. Fan E, Beitler JR, Brochard L, Calfee CS, Ferguson ND, Slutsky AS, Brodie D. 2020. COVID-19-associated acute respiratory distress syndrome: is a different approach to management warranted? Lancet Respir Med 8:816-821.

392. Li X, Ma X. 2020. Acute respiratory failure in COVID-19: is it "typical" ARDS? Crit Care 24:198.

393. Tang Y, Liu J, Zhang D, Xu Z, Ji J, Wen C. 2020. Cytokine Storm in COVID-19: The Current Evidence and Treatment Strategies. Front Immunol 11:1708.

394. Zanza C, Romenskaya T, Manetti AC, Franceschi F, La Russa R, Bertozzi G, Maiese A, Savioli G, Volonnino G, Longhitano Y. 2022. Cytokine Storm in COVID-19: Immunopathogenesis and Therapy. Medicina (Kaunas) 58.

395. Voiriot G, Dorgham K, Bachelot G, Fajac A, Morand-Joubert L, Parizot C, Gerotziafas G, Farabos D, Trugnan G, Eguether T, Blayau C, Djibre M, Elabbadi A, Gibelin A, Labbe V, Parrot A, Turpin M, Cadranel J, Gorochov G, Fartoukh M, Lamaziere A. 2022. Identification of bronchoalveolar and blood immune-inflammatory biomarker signature associated with poor 28-day outcome in critically ill COVID-19 patients. Sci Rep 12:9502.

396. Azkur AK, Akdis M, Azkur D, Sokolowska M, van de Veen W, Bruggen MC, O'Mahony L, Gao Y, Nadeau K, Akdis CA. 2020. Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. Allergy 75:1564-1581.

397. Ceglarek L, Boyman O. 2024. Immune dysregulation in long COVID. Nat Immunol 25:587-589.

398. Lopez Lloreda C. 2024. COVID's toll on the brain: new clues emerge. Nature doi:10.1038/d41586-024-00828-9.

399. Greene C, Connolly R, Brennan D, Laffan A, O'Keeffe E, Zaporojan L, O'Callaghan J, Thomson B, Connolly E, Argue R, Martin-Loeches I, Long A, Cheallaigh CN, Conlon N, Doherty CP, Campbell M. 2024. Blood-brain barrier disruption and sustained systemic

inflammation in individuals with long COVID-associated cognitive impairment. Nat Neurosci doi:10.1038/s41593-024-01576-9.

400. Yin K, Peluso MJ, Luo X, Thomas R, Shin MG, Neidleman J, Andrew A, Young KC, Ma T, Hoh R, Anglin K, Huang B, Argueta U, Lopez M, Valdivieso D, Asare K, Deveau TM, Munter SE, Ibrahim R, Standker L, Lu S, Goldberg SA, Lee SA, Lynch KL, Kelly JD, Martin JN, Munch J, Deeks SG, Henrich TJ, Roan NR. 2024. Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. Nat Immunol 25:218-225.

401. Chen R, Wang K, Yu J, Howard D, French L, Chen Z, Wen C, Xu Z. 2020. The Spatial and Cell-Type Distribution of SARS-CoV-2 Receptor ACE2 in the Human and Mouse Brains. Front Neurol 11:573095.

402. Mavrikaki M, Lee JD, Solomon IH, Slack FJ. 2022. Severe COVID-19 is associated with molecular signatures of aging in the human brain. Nat Aging 2:1130-1137.

403. Iadecola C, Anrather J, Kamel H. 2020. Effects of COVID-19 on the Nervous System. Cell 183:16-27 e1.

404. Koyuncu OO, Hogue IB, Enquist LW. 2013. Virus infections in the nervous system. Cell Host Microbe 13:379-93.

405. Yachou Y, El Idrissi A, Belapasov V, Ait Benali S. 2020. Neuroinvasion, neurotropic, and neuroinflammatory events of SARS-CoV-2: understanding the neurological manifestations in COVID-19 patients. Neurol Sci 41:2657-2669.

406. Fan C, Wu Y, Rui X, Yang Y, Ling C, Liu S, Liu S, Wang Y. 2022. Animal models for COVID-19: advances, gaps and perspectives. Signal Transduct Target Ther 7:220.

407. Munoz-Fontela C, Dowling WE, Funnell SGP, Gsell PS, Riveros-Balta AX, Albrecht RA, Andersen H, Baric RS, Carroll MW, Cavaleri M, Qin C, Crozier I, Dallmeier K, de Waal L, de Wit E, Delang L, Dohm E, Duprex WP, Falzarano D, Finch CL, Frieman MB, Graham BS, Gralinski LE, Guilfoyle K, Haagmans BL, Hamilton GA, Hartman AL, Herfst S, Kaptein SJF, Klimstra WB, Knezevic I, Krause PR, Kuhn JH, Le Grand R, Lewis MG, Liu WC, Maisonnasse P, McElroy AK, Munster V, Oreshkova N, Rasmussen AL, Rocha-Pereira J, Rockx B, Rodriguez E, Rogers TF, Salguero FJ, Schotsaert M, Stittelaar KJ, Thibaut HJ, Tseng CT, et al. 2020. Animal models for COVID-19. Nature 586:509-515.

408. Munoz-Fontela C, Widerspick L, Albrecht RA, Beer M, Carroll MW, de Wit E, Diamond MS, Dowling WE, Funnell SGP, Garcia-Sastre A, Gerhards NM, de Jong R, Munster VJ, Neyts J, Perlman S, Reed DS, Richt JA, Riveros-Balta X, Roy CJ, Salguero FJ, Schotsaert M, Schwartz LM, Seder RA, Segales J, Vasan SS, Henao-Restrepo AM, Barouch DH. 2022. Advances and gaps in SARS-CoV-2 infection models. PLoS Pathog 18:e1010161.

409. Dong W, Mead H, Tian L, Park JG, Garcia JI, Jaramillo S, Barr T, Kollath DS, Coyne VK, Stone NE, Jones A, Zhang J, Li A, Wang LS, Milanes-Yearsley M, Torrelles JB, Martinez-Sobrido L, Keim PS, Barker BM, Caligiuri MA, Yu J. 2022. The K18-Human ACE2 Transgenic Mouse Model Recapitulates Non-severe and Severe COVID-19 in Response to an Infectious Dose of the SARS-CoV-2 Virus. J Virol 96:e0096421.

410. Takayama K. 2020. In Vitro and Animal Models for SARS-CoV-2 research. Trends Pharmacol Sci 41:513-517.

411. Imai M, Iwatsuki-Horimoto K, Hatta M, Loeber S, Halfmann PJ, Nakajima N, Watanabe T, Ujie M, Takahashi K, Ito M, Yamada S, Fan S, Chiba S, Kuroda M, Guan L, Takada K, Armbrust T, Balogh A, Furusawa Y, Okuda M, Ueki H, Yasuhara A, Sakai-Tagawa Y, Lopes TJS, Kiso M, Yamayoshi S, Kinoshita N, Ohmagari N, Hattori SI, Takeda M, Mitsuya H, Krammer F, Suzuki T, Kawaoka Y. 2020. Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. Proc Natl Acad Sci U S A 117:16587-16595.

412. Sia SF, Yan LM, Chin AWH, Fung K, Choy KT, Wong AYL, Kaewpreedee P, Perera R, Poon LLM, Nicholls JM, Peiris M, Yen HL. 2020. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. Nature 583:834-838.

413. Frere JJ, Serafini RA, Pryce KD, Zazhytska M, Oishi K, Golynker I, Panis M, Zimering J, Horiuchi S, Hoagland DA, Moller R, Ruiz A, Kodra A, Overdevest JB, Canoll PD, Borczuk AC, Chandar V, Bram Y, Schwartz R, Lomvardas S, Zachariou V, tenOever BR. 2022. SARS-CoV-2 infection in hamsters and humans results in lasting and unique systemic perturbations after recovery. Sci Transl Med 14:eabq3059.

414. Dowall S, Salguero FJ, Wiblin N, Fotheringham S, Hatch G, Parks S, Gowan K, Harris D, Carnell O, Fell R, Watson R, Graham V, Gooch K, Hall Y, Mizen S, Hewson R. 2021. Development of a Hamster Natural Transmission Model of SARS-CoV-2 Infection. Viruses 13.

415. Wang Z, Cormier RT. 2022. Golden Syrian Hamster Models for Cancer Research. Cells 11.

416. Matute-Bello G, Downey G, Moore BB, Groshong SD, Matthay MA, Slutsky AS, Kuebler WM, Acute Lung Injury in Animals Study G. 2011. An official American Thoracic Society workshop report: features and measurements of experimental acute lung injury in animals. Am J Respir Cell Mol Biol 44:725-38.

417. Olivarria GM, Cheng Y, Furman S, Pachow C, Hohsfield LA, Smith-Geater C, Miramontes R, Wu J, Burns MS, Tsourmas KI, Stocksdale J, Manlapaz C, Yong WH, Teijaro J, Edwards R, Green KN, Thompson LM, Lane TE. 2022. Microglia Do Not Restrict SARS-CoV-2 Replication following Infection of the Central Nervous System of K18-Human ACE2 Transgenic Mice. J Virol 96:e0196921.

418. Khalil BA, Elemam NM, Maghazachi AA. 2021. Chemokines and chemokine receptors during COVID-19 infection. Comput Struct Biotechnol J 19:976-988.

419. Kim K, Calabrese P, Wang S, Qin C, Rao Y, Feng P, Chen XS. 2022. The Roles of APOBEC-mediated RNA Editing in SARS-CoV-2 Mutations, Replication and Fitness. Res Sq doi:10.21203/rs.3.rs-1524060/v1.

420. Honarmand Ebrahimi K. 2018. A unifying view of the broad-spectrum antiviral activity of RSAD2 (viperin) based on its radical-SAM chemistry. Metallomics 10:539-552.

421. Fuchs J, Holzer M, Schilling M, Patzina C, Schoen A, Hoenen T, Zimmer G, Marz M, Weber F, Muller MA, Kochs G. 2017. Evolution and Antiviral Specificities of Interferon-Induced Mx Proteins of Bats against Ebola, Influenza, and Other RNA Viruses. J Virol 91.

422. Kuznetsov D, Tegenfeldt F, Manni M, Seppey M, Berkeley M, Kriventseva EV, Zdobnov EM. 2023. OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. Nucleic Acids Res 51:D445-D451.

423. Boys IN, Xu E, Mar KB, De La Cruz-Rivera PC, Eitson JL, Moon B, Schoggins JW. 2020. RTP4 Is a Potent IFN-Inducible Anti-flavivirus Effector Engaged in a Host-Virus Arms Race in Bats and Other Mammals. Cell Host Microbe 28:712-723 e9.

424. Xu F, Wang G, Zhao F, Huang Y, Fan Z, Mei S, Xie Y, Wei L, Hu Y, Wang C, Cen S, Liang C, Ren L, Guo F, Wang J. 2022. IFITM3 Inhibits SARS-CoV-2 Infection and Is Associated with COVID-19 Susceptibility. Viruses 14.

425. Hachim MY, Al Heialy S, Hachim IY, Halwani R, Senok AC, Maghazachi AA, Hamid Q. 2020. Interferon-Induced Transmembrane Protein (IFITM3) Is Upregulated Explicitly in SARS-CoV-2 Infected Lung Epithelial Cells. Front Immunol 11:1372.

426. Prelli Bozzo C, Nchioua R, Volcic M, Koepke L, Kruger J, Schutz D, Heller S, Sturzel CM, Kmiec D, Conzelmann C, Muller J, Zech F, Braun E, Gross R, Wettstein L, Weil T, Weiss J, Diofano F, Rodriguez Alfonso AA, Wiese S, Sauter D, Munch J, Goffinet C, Catanese A, Schon M, Boeckers TM, Stenger S, Sato K, Just S, Kleger A, Sparrer KMJ, Kirchhoff F. 2021. IFITM

proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro. Nat Commun 12:4584.

427. Shojaei M, Shamshirian A, Monkman J, Grice L, Tran M, Tan CW, Teo SM, Rodrigues Rossi G, McCulloch TR, Nalos M, Raei M, Razavi A, Ghasemian R, Gheibi M, Roozbeh F, Sly PD, Spann KM, Chew KY, Zhu Y, Xia Y, Wells TJ, Senegaglia AC, Kuniyoshi CL, Franck CL, Dos Santos AFR, de Noronha L, Motamen S, Valadan R, Amjadi O, Gogna R, Madan E, Alizadeh-Navaei R, Lamperti L, Zuniga F, Nova-Lamperti E, Labarca G, Knippenberg B, Herwanto V, Wang Y, Phu A, Chew T, Kwan T, Kim K, Teoh S, Pelaia TM, Kuan WS, Jee Y, Iredell J, O'Byrne K, Fraser JF, et al. 2022. IFI27 transcription is an early predictor for COVID-19 outcomes, a multi-cohort observational study. Front Immunol 13:1060438.

428. Sarkar L, Liu G, Gack MU. 2023. ISG15: its roles in SARS-CoV-2 and other viral infections. Trends Microbiol 31:1262-1275.

429. Ning S, Pagano JS, Barber GN. 2011. IRF7: activation, regulation, modification and function. Genes Immun 12:399-414.

430. Jefferies CA. 2019. Regulating IRFs in IFN Driven Disease. Front Immunol 10:325.

431. Jena KK, Mehto S, Nath P, Chauhan NR, Sahu R, Dhar K, Das SK, Kolapalli SP, Murmu KC, Jain A, Krishna S, Sahoo BS, Chattopadhyay S, Rusten TE, Prasad P, Chauhan S, Chauhan S. 2020. Autoimmunity gene IRGM suppresses cGAS-STING and RIG-I-MAVS signaling to control interferon response. EMBO Rep 21:e50051.

432. Bustamante S, Yau Y, Boys V, Chang J, Paramsothy S, Pudipeddi A, Leong RW, Wasinger VC. 2022. Tryptophan Metabolism 'Hub' Gene Expression Associates with Increased Inflammation and Severe Disease Outcomes in COVID-19 Infection and Inflammatory Bowel Disease. Int J Mol Sci 23.

433. Shin JH, Bozadjieva-Kramer N, Shao Y, Lyons-Abbott S, Rupp AC, Sandoval DA, Seeley RJ. 2022. The gut peptide Reg3g links the small intestine microbiome to the regulation of energy balance, glucose levels, and gut function. Cell Metab 34:1765-1778 e6.

434. Shin JH, Bozadjieva-Kramer N, Seeley RJ. 2023. Reg3gamma: current understanding and future therapeutic opportunities in metabolic disease. Exp Mol Med 55:1672-1677.

435. Wang J, Hou H, Mao L, Wang F, Yu J, Luo Y, Lin Q, Sun Z. 2021. TIGIT Signaling Pathway Regulates Natural Killer Cell Function in Chronic Hepatitis B Virus Infection. Front Med (Lausanne) 8:816474.

436. Yu X, Harden K, Gonzalez LC, Francesco M, Chiang E, Irving B, Tom I, Ivelja S, Refino CJ, Clark H, Eaton D, Grogan JL. 2009. The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. Nat Immunol 10:48-57.

437. Ge Z, Peppelenbosch MP, Sprengers D, Kwekkeboom J. 2021. TIGIT, the Next Step Towards Successful Combination Immune Checkpoint Therapy in Cancer. Front Immunol 12:699895.

438. Orecchioni M, Ghosheh Y, Pramod AB, Ley K. 2019. Macrophage Polarization: Different Gene Signatures in M1(LPS+) vs. Classically and M2(LPS-) vs. Alternatively Activated Macrophages. Front Immunol 10:1084.

439. Bandala-Sanchez E, Zhang Y, Reinwald S, Dromey JA, Lee BH, Qian J, Bohmer RM, Harrison LC. 2013. T cell regulation mediated by interaction of soluble CD52 with the inhibitory receptor Siglec-10. Nat Immunol 14:741-8.

440. Rashidi M, Bandala-Sanchez E, Lawlor KE, Zhang Y, Neale AM, Vijayaraj SL, O'Donoghue R, Wentworth JM, Adams TE, Vince JE, Harrison LC. 2018. CD52 inhibits Toll-like receptor activation of NF-kappaB and triggers apoptosis to suppress inflammation. Cell Death Differ 25:392-405.

441. Elemam NM, Talaat IM, Maghazachi AA. 2022. CXCL10 Chemokine: A Critical Player in RNA and DNA Viral Infections. Viruses 14.

442. Pulit-Penaloza JA, Scherbik SV, Brinton MA. 2012. Activation of Oas1a gene expression by type I IFN requires both STAT1 and STAT2 while only STAT2 is required for Oas1b activation. Virology 425:71-81.

443. Kuntova B, Stopkova R, Stopka P. 2018. Transcriptomic and Proteomic Profiling Revealed High Proportions of Odorant Binding and Antimicrobial Defense Proteins in Olfactory Tissues of the House Mouse. Front Genet 9:26.

444. Diamond MS. 2014. IFIT1: A dual sensor and effector molecule that detects non-2'-O methylated viral RNA and inhibits its translation. Cytokine Growth Factor Rev 25:543-50.

445. Singh DK, Aladyeva E, Das S, Singh B, Esaulova E, Swain A, Ahmed M, Cole J, Moodley C, Mehra S, Schlesinger LS, Artyomov MN, Khader SA, Kaushal D. 2022. Myeloid cell interferon responses correlate with clearance of SARS-CoV-2. Nat Commun 13:679.

446. Singer JM, Shew TM, Ferguson D, Renkemeyer MK, Pietka TA, Hall AM, Finck BN, Lutkewitte AJ. 2022. doi:10.1101/2022.02.14.480414.

447. Drokhlyansky E, Goz Ayturk D, Soh TK, Chrenek R, O'Loughlin E, Madore C, Butovsky O, Cepko CL. 2017. The brain parenchyma has a type I interferon response that can limit virus spread. Proc Natl Acad Sci U S A 114:E95-E104.

448. Boribong BP, LaSalle TJ, Bartsch YC, Ellett F, Loiselle ME, Davis JP, Gonye ALK, Hajizadeh S, Kreuzer J, Pillai S, Haas W, Edlow A, Fasano A, Alter G, Irimia D, Sade-Feldman M, Yonker LM. 2021. Neutrophil Profiles of Pediatric COVID-19 and Multisystem Inflammatory Syndrome in Children. bioRxiv doi:10.1101/2021.12.18.473308.

449. Tsang H-W, Bhatia I, Chan K-W, Chan GC-F, Ip P, Cheung P-T. 2022. Transmembrane 29 (Tmem29), a Newly Identified Molecule Showed Downregulation in Hypoxic-Ischemic Brain Damage. NeuroSci 3:41-51.

450. Grin'kina NM, Karnabi EE, Damania D, Wadgaonkar S, Muslimov IA, Wadgaonkar R. 2012. Sphingosine kinase 1 deficiency exacerbates LPS-induced neuroinflammation. PLoS One 7:e36475.

451. Bryan L, Kordula T, Spiegel S, Milstien S. 2008. Regulation and functions of sphingosine kinases in the brain. Biochim Biophys Acta 1781:459-66.

452. Al-Kuraishy HM, Al-Gareeb AI, Qusti S, Alshammari EM, Atanu FO, Batiha GE. 2021. Arginine vasopressin and pathophysiology of COVID-19: An innovative perspective. Biomed Pharmacother 143:112193.

453. Crespi B. 2020. Evolutionary medical insights into the SARS-CoV-2 pandemic. Evol Med Public Health 2020:314-322.

454. Winkler ES, Bailey AL, Kafai NM, Nair S, McCune BT, Yu J, Fox JM, Chen RE, Earnest JT, Keeler SP, Ritter JH, Kang LI, Dort S, Robichaud A, Head R, Holtzman MJ, Diamond MS. 2020. SARS-CoV-2 infection of human ACE2-transgenic mice causes severe lung inflammation and impaired function. Nat Immunol 21:1327-1335.

455. Lucas TM, Richner JM, Diamond MS. 2015. The Interferon-Stimulated Gene Ifi27l2a Restricts West Nile Virus Infection and Pathogenesis in a Cell-Type- and Region-Specific Manner. J Virol 90:2600-15.

456. Horiuchi S, Oishi K, Carrau L, Frere J, Moller R, Panis M, tenOever BR. 2021. Immune memory from SARS-CoV-2 infection in hamsters provides variant-independent protection but still allows virus transmission. Sci Immunol 6:eabm3131.

457. Hopkins J. 2023. MORTALITY ANALYSES. https://coronavirus.jhu.edu/data/mortality. Accessed

458. Carpenter KC, Yang J, Xu JJ. 2023. Animal Models for the Study of Neurologic Manifestations Of COVID-19. Comp Med 73:91-103.

459. Ferreira J. 2023. New insights into SARS-CoV-2 brain infection and olfactory impairment in the hamster model. Lab Anim (NY) 52:201.

460. de Melo GD, Perraud V, Alvarez F, Vieites-Prado A, Kim S, Kergoat L, Coleon A, Trueb BS, Tichit M, Piazza A, Thierry A, Hardy D, Wolff N, Munier S, Koszul R, Simon-Loriere E, Thiel V, Lecuit M, Lledo PM, Renier N, Larrous F, Bourhy H. 2023. Neuroinvasion and anosmia are independent phenomena upon infection with SARS-CoV-2 and its variants. Nat Commun 14:4485.

461. Platt RN, 2nd, Amman BR, Keith MS, Thompson CW, Bradley RD. 2015. What Is Peromyscus? Evidence from nuclear and mitochondrial DNA sequences suggests the need for a new classification. J Mammal 96:708-719.

462. Greenbaum IF, Honeycutt, Rodney L., Chirhart, Scott E. 2019. Taxonomy And Phylogenetics Of The Peromyscus Maniculatus Species Group. Special Publications of the Museum of Texas Tech University 71:559-575.

463. Kim S, Lee S, Kim TY, Lee SH, Seo SU, Kweon MN. 2023. Newly isolated Lactobacillus paracasei strain modulates lung immunity and improves the capacity to cope with influenza virus infection. Microbiome 11:260.

464. Nguyen QV, Chong LC, Hor YY, Lew LC, Rather IA, Choi SB. 2022. Role of Probiotics in the Management of COVID-19: A Computational Perspective. Nutrients 14.

465. Case JB, Bailey AL, Kim AS, Chen RE, Diamond MS. 2020. Growth, detection, quantification, and inactivation of SARS-CoV-2. Virology 548:39-48.

466. Prakash S, Dhanushkodi NR, Zayou L, Ibraim IC, Quadiri A, Coulon PG, Tifrea DF, Suzer B, Shaik AM, Chilukuri A, Edwards RA, Singer M, Vahed H, Nesburn AB, Kuppermann BD, Ulmer JB, Gil D, Jones TM, BenMohamed L. 2024. Cross-protection induced by highly conserved human B, CD4(+), and CD8(+) T-cell epitopes-based vaccine against severe infection, disease, and death caused by multiple SARS-CoV-2 variants of concern. Front Immunol 15:1328905.

467. Patterson EI, Prince T, Anderson ER, Casas-Sanchez A, Smith SL, Cansado-Utrilla C, Solomon T, Griffiths MJ, Acosta-Serrano A, Turtle L, Hughes GL. 2020. Methods of Inactivation of SARS-CoV-2 for Downstream Biological Assays. J Infect Dis 222:1462-1467.

468. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11:R25.

469. J. C. 2016. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20:37-46.

470. Chen T, Zhang H, Liu Y, Liu YX, Huang L. 2021. EVenn: Easy to create repeatable and editable Venn diagrams and Venn networks online. J Genet Genomics 48:863-866.

471. Blighe K, S Rana, and M Lewis. 2018. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. https://github.com/kevinblighe/EnhancedVolcano. Accessed

472. Benedictow OJ. 2005. The Black Death: The Greatest Catastrophe Ever, *on* History Today, Volume 55 Issue 3. https://www.historytoday.com/archive/black-death-greatest-catastrophe-ever. Accessed 04/01/2024.

473. Pereira PDC, Diniz DG, da Costa ER, Magalhaes NGM, da Silva AJF, Leite JGS, Almeida NIP, Cunha KN, de Melo MAD, Vasconcelos P, Diniz JAP, Brites D, Anthony DC, Diniz CWP, Guerreiro-Diniz C. 2023. Genes, inflammatory response, tolerance, and resistance to virus infections in migratory birds, bats, and rodents. Front Immunol 14:1239572.

474. Gravinatti ML, Barbosa CM, Soares RM, Gregori F. 2020. Synanthropic rodents as virus reservoirs and transmitters. Rev Soc Bras Med Trop 53:e20190486.

475. Hufeldt MR ND, Vogensen FK, Midtvedt T, Hansen AK. . 2010. Variation in the Gut Microbiota of Laboratory Mice Is Related to Both Genetic and Environmental Factors. Comp Med Oct;60(5):336-47.

476. Rosshart SP, Herz J, Vassallo BG, Hunter A, Wall MK, Badger JH, McCulloch JA, Anastasakis DG, Sarshad AA, Leonardi I, Collins N, Blatter JA, Han SJ, Tamoutounour S, Potapova S, Foster St Claire MB, Yuan W, Sen SK, Dreier MS, Hild B, Hafner M, Wang D, Iliev ID, Belkaid Y, Trinchieri G, Rehermann B. 2019. Laboratory mice born to wild mice have natural microbiota and model human immune responses. Science 365.

477. Buhnik-Rosenblau K, Danin-Poleg Y, Kashi Y. 2011. Predominant effect of host genetics on levels of Lactobacillus johnsonii bacteria in the mouse gut. Appl Environ Microbiol 77:6531-8.

478. Dhivahar J, Parthasarathy A, Krishnan K, Kovi BS, Pandian GN. 2023. Bat-associated microbes: Opportunities and perils, an overview. Heliyon 9:e22351.

479. Strobel S, Encarnacao JA, Becker NI, Trenczek TE. 2015. Histological and histochemical analysis of the gastrointestinal tract of the common pipistrelle bat (Pipistrellus pipistrellus). Eur J Histochem 59:2477.

480. Toshiro Kamiya PP. 1975. Comparative Gastric Morphology of Old World Bats. Journal of the Mammalogical Society of Japan Volume 6 Issue 4:145-154.

481. Gury-BenAri M, Thaiss CA, Serafini N, Winter DR, Giladi A, Lara-Astiaso D, Levy M, Salame TM, Weiner A, David E, Shapiro H, Dori-Bachash M, Pevsner-Fischer M, Lorenzo-Vivas E, Keren-Shaul H, Paul F, Harmelin A, Eberl G, Itzkovitz S, Tanay A, Di Santo JP, Elinav E, Amit I. 2016. The Spectrum and Regulatory Landscape of Intestinal Innate Lymphoid Cells Are Shaped by the Microbiome. Cell 166:1231-1246 e13.

482. Wang R, Cui W, Yang H. 2023. The interplay between innate lymphoid cells and microbiota. mBio 14:e0039923.

483. Hezaveh K, Shinde RS, Klotgen A, Halaby MJ, Lamorte S, Ciudad MT, Quevedo R, Neufeld L, Liu ZQ, Jin R, Grunwald BT, Foerster EG, Chaharlangi D, Guo M, Makhijani P, Zhang X, Pugh TJ, Pinto DM, Co IL, McGuigan AP, Jang GH, Khokha R, Ohashi PS, O'Kane GM, Gallinger S, Navarre WW, Maughan H, Philpott DJ, Brooks DG, McGaha TL. 2022. Tryptophan-derived microbial metabolites activate the aryl hydrocarbon receptor in tumor-associated macrophages to suppress anti-tumor immunity. Immunity 55:324-340 e8.

484. Zheng D, Liwinski T, Elinav E. 2020. Interaction between microbiota and immunity in health and disease. Cell Res 30:492-506.

485. Seo SK, Kwon B. 2023. Immune regulation through tryptophan metabolism. Exp Mol Med 55:1371-1379.

486. Yang W, Cong Y. 2021. Gut microbiota-derived metabolites in the regulation of host immune responses and immune-related inflammatory diseases. Cell Mol Immunol 18:866-877.

487. Badawy AA. 2022. Tryptophan metabolism and disposition in cancer biology and immunotherapy. Biosci Rep 42.

488. Kurnasov O, Goral V, Colabroy K, Gerdes S, Anantha S, Osterman A, Begley TP. 2003. NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria. Chem Biol 10:1195-204.

489. Roager HM, Licht TR. 2018. Microbial tryptophan catabolites in health and disease. Nat Commun 9:3294.

490. Gao K, Mu CL, Farzi A, Zhu WY. 2020. Tryptophan Metabolism: A Link Between the Gut Microbiota and Brain. Adv Nutr 11:709-723.

491. Franchini AM, Myers JR, Jin GB, Shepherd DM, Lawrence BP. 2019. Genome-Wide Transcriptional Analysis Reveals Novel AhR Targets That Regulate Dendritic Cell Function during Influenza A Virus Infection. Immunohorizons 3:219-235.

492. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, Zecchi R, D'Angelo C, Massi-Benedetti C, Fallarino F, Carvalho A, Puccetti P, Romani L. 2013. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. Immunity 39:372-85.

493. Wilck N, Matus MG, Kearney SM, Olesen SW, Forslund K, Bartolomaeus H, Haase S, Mahler A, Balogh A, Marko L, Vvedenskaya O, Kleiner FH, Tsvetkov D, Klug L, Costea PI, Sunagawa S, Maier L, Rakova N, Schatz V, Neubert P, Fratzer C, Krannich A, Gollasch M, Grohme DA, Corte-Real BF, Gerlach RG, Basic M, Typas A, Wu C, Titze JM, Jantsch J, Boschmann M, Dechend R, Kleinewietfeld M, Kempa S, Bork P, Linker RA, Alm EJ, Muller DN. 2017. Salt-responsive gut commensal modulates T(H)17 axis and disease. Nature 551:585-589.

494. Krupa A, Kowalska I. 2021. The Kynurenine Pathway-New Linkage between Innate and Adaptive Immunity in Autoimmune Endocrinopathies. Int J Mol Sci 22.

495. Pallotta MT, Rossini S, Suvieri C, Coletti A, Orabona C, Macchiarulo A, Volpi C, Grohmann U. 2022. Indoleamine 2,3-dioxygenase 1 (IDO1): an up-to-date overview of an eclectic immunoregulatory enzyme. FEBS J 289:6099-6118.

496. Troche G, Henry-Lagarrigue M, Soppelsa F, Legriel S, Yehia A, Bruneel F, Bedos JP, Spreux-Varoquaux O. 2020. Tryptophan pathway catabolites (serotonin, 5-hydroxyindolacetic acid, kynurenine) and enzymes (monoamine oxidase and indole amine 2,3 dioxygenase) in patients with septic shock: A prospective observational study versus healthy controls. Medicine (Baltimore) 99:e19906.

497. Sehgal R, de Mello VD, Männistö V, Lindström J, Tuomilehto J, Pihlajamäki J, Uusitupa M. 2022. Indolepropionic Acid, a Gut Bacteria-Produced Tryptophan Metabolite and the Risk of Type 2 Diabetes and Non-Alcoholic Fatty Liver Disease. Nutrients 14:4695.

498. Jiang H, Chen C, Gao J. 2022. Extensive Summary of the Important Roles of Indole Propionic Acid, a Gut Microbial Metabolite in Host Health and Disease. Nutrients 15:151.

499. Konopelski P, Mogilnicka I. 2022. Biological Effects of Indole-3-Propionic Acid, a Gut Microbiota-Derived Metabolite, and Its Precursor Tryptophan in Mammals' Health and Disease. Int J Mol Sci 23.

500. Ji Y, Yin W, Liang Y, Sun L, Yin Y, Zhang W. 2020. Anti-Inflammatory and Anti-Oxidative Activity of Indole-3-Acetic Acid Involves Induction of HO-1 and Neutralization of Free Radicals in RAW264.7 Cells. International Journal of Molecular Sciences 21:1579.

501. Shaw C, Hess M, Weimer BC. 2023. Microbial-Derived Tryptophan Metabolites and Their Role in Neurological Disease: Anthranilic Acid and Anthranilic Acid Derivatives. Microorganisms 11.

502. Pavlova T, Vidova V, Bienertova-Vasku J, Janku P, Almasi M, Klanova J, Spacil Z. 2017. Urinary intermediates of tryptophan as indicators of the gut microbial metabolism. Anal Chim Acta 987:72-80.

503. Lentsch AB, Jordan JA, Czermak BJ, Diehl KM, Younkin EM, Sarma V, Ward PA. 1999. Inhibition of NF-kappaB activation and augmentation of IkappaBbeta by secretory leukocyte protease inhibitor during lung inflammation. Am J Pathol 154:239-47.

504. Raundhal M, Morse C, Khare A, Oriss TB, Milosevic J, Trudeau J, Huff R, Pilewski J, Holguin F, Kolls J, Wenzel S, Ray P, Ray A. 2015. High IFN-gamma and low SLPI mark severe asthma in mice and humans. J Clin Invest 125:3037-50.

505. Nugteren S, Samsom JN. 2021. Secretory Leukocyte Protease Inhibitor (SLPI) in mucosal tissues: Protects against inflammation, but promotes cancer. Cytokine Growth Factor Rev 59:22-35.

506. Sawicki KT, Nannini DR, Bielinski SJ, Larson NB, Lloyd-Jones DM, Psaty B, Taylor KD, Shah SJ, Rasmussen-Torvik LJ, Wilkins JT, McNally EM, Patel RB. 2023. Secretory leukocyte protease inhibitor and risk of heart failure in the Multi-Ethnic Study of Atherosclerosis. Sci Rep 13:604.

507. Hiemstra PS, Maassen RJ, Stolk J, Heinzel-Wieland R, Steffens GJ, Dijkman JH. 1996. Antibacterial activity of antileukoprotease. Infect Immun 64:4520-4.

508. Miller KW, Evans RJ, Eisenberg SP, Thompson RC. 1989. Secretory leukocyte protease inhibitor binding to mRNA and DNA as a possible cause of toxicity to Escherichia coli. J Bacteriol 171:2166-72.

509. Tomee JF, Hiemstra PS, Heinzel-Wieland R, Kauffman HF. 1997. Antileukoprotease: an endogenous protein in the innate mucosal defense against fungi. J Infect Dis 176:740-7.

510. Chattopadhyay A, Gray LR, Patton LL, Caplan DJ, Slade GD, Tien HC, Shugars DC. 2004. Salivary secretory leukocyte protease inhibitor and oral candidiasis in human immunodeficiency virus type 1-infected persons. Infect Immun 72:1956-63.

511. Shugars DC. 1999. Endogenous mucosal antiviral factors of the oral cavity. J Infect Dis 179 Suppl 3:S431-5.

512. Doumas S, Kolokotronis A, Stefanopoulos P. 2005. Anti-inflammatory and antimicrobial roles of secretory leukocyte protease inhibitor. Infect Immun 73:1271-4.

513. Osbourn M, Rodgers AM, Dubois AV, Small DM, Humphries F, Delagic N, Moynagh PN, Weldon S, Taggart CC, Ingram RJ. 2022. Secretory Leucoprotease Inhibitor (SLPI) Promotes Survival during Acute Pseudomonas aeruginosa Infection by Suppression of Inflammation Rather Than Microbial Killing. Biomolecules 12.

514. Greene CM, McElvaney NG. 2009. Proteases and antiproteases in chronic neutrophilic lung disease - relevance to drug discovery. Br J Pharmacol 158:1048-58.

515. McElvaney NG, Nakamura H, Birrer P, Hebert CA, Wong WL, Alphonso M, Baker JB, Catalano MA, Crystal RG. 1992. Modulation of airway inflammation in cystic fibrosis. In vivo suppression of interleukin-8 levels on the respiratory epithelial surface by aerosolization of recombinant secretory leukoprotease inhibitor. J Clin Invest 90:1296-301.

516. Kumar A, Das SK, Emdad L, Fisher PB. 2023. Applications of tissue-specific and cancer-selective gene promoters for cancer diagnosis and therapy. Adv Cancer Res 160:253-315.

517. Kaza V, Farmaki E, Havighorst A, Crossland J, Chatzistamou I, Kiaris H. 2018. Growth of human breast cancers in Peromyscus. Dis Model Mech 11.

518. Gross L, Gluckman EC, Kershaw BB, Posselt AE. 1953. Resistance of the white-footed field mouse(peromyscus leucopus noveboracensis) to the carcinogenic action of urethane. Cancer 6:1241-1243.

519. Wyman D, Balderrama-Gutierrez G, Reese F, Jiang S, Rahmanian S, Forner S, Matheos D, Zeng W, Williams B, Trout D, England W, Chu S-H, Spitale RC, Tenner AJ, Wold BJ, Mortazavi A. 2020. doi:10.1101/672931.

520. Reese F, Mortazavi A. 2021. Swan: a library for the analysis and visualization of long-read transcriptomes. Bioinformatics 37:1322-1323.

521. Rauch E, Amendt T, Lopez Krol A, Lang FB, Linse V, Hohmann M, Keim AC, Kreutzer S, Kawengian K, Buchholz M, Duschner P, Grauer S, Schnierle B, Ruhl A, Burtscher I, Dehnert S, Kuria C, Kupke A, Paul S, Liehr T, Lechner M, Schnare M, Kaufmann A, Huber M, Winkler TH, Bauer S, Yu P. 2024. T-bet(+) B cells are activated by and control endogenous retroviruses through TLR-dependent mechanisms. Nat Commun 15:1229.

522.	Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.
523.	Kwon DN, Lee YK, Greenhalgh DG, Cho K. 2011. Lipopolysaccharide stress induces cell-type specific production of murine leukemia virus type-endogenous retroviral virions in primary lymphoid cells. J Gen Virol 92:292-300.
524.	Greenig M. 2019. HERVs, immunity, and autoimmunity: understanding the connection. PeerJ 7:e6711.
525.	Dhillon P, Mulholland KA, Hu H, Park J, Sheng X, Abedini A, Liu H, Vassalotti A, Wu J, Susztak K. 2023. Increased levels of endogenous retroviruses trigger fibroinflammation and play a role in kidney disease development. Nat Commun 14:559.
526.	Srinivasachar Badarinarayan S, Shcherbakova I, Langer S, Koepke L, Preising A, Hotter D, Kirchhoff F, Sparrer KMJ, Schotta G, Sauter D. 2020. HIV-1 infection activates endogenous retroviral promoters regulating antiviral gene expression. Nucleic Acids Res 48:10890-10908.
527.	Torre D, Fstkchyan YS, Ho JSY, Cheon Y, Patel RS, Degrace EJ, Mzoughi S, Schwarz M, Mohammed K, Seo JS, Romero-Bueno R, Demircioglu D, Hasson D, Tang W, Mahajani SU, Campisi L, Zheng S, Song WS, Wang YC, Shah H, Francoeur N, Soto J, Salfati Z, Weirauch MT, Warburton P, Beaumont K, Smith ML, Mulder L, Villalta SA, Kessenbrock K, Jang C, Lee D, De Rubeis S, Cobos I, Tam O, Hammell MG, Seldin M, Shi Y, Basu U, Sebastiano V, Byun M, Sebra R, Rosenberg BR, Benner C, Guccione E, Marazzi I. 2023. Nuclear RNA catabolism controls endogenous retroviruses, gene expression asymmetry, and dedifferentiation. Mol Cell 83:4255-4271 e9.
528.	Swanson KI, Norris DE. 2008. Presence of multiple variants of Borrelia burgdorferi in the natural reservoir Peromyscus leucopus throughout a transmission season. Vector Borne Zoonotic Dis 8:397-405.
529.	Maier TJ. 2002. Long-distance Movements by Female White-footed Mice, Peromyscus leucopus, in Extensive Mixed-wood Forest Canadian Field-Naturalist 108-111.
530.	Morzunov SP, Rowe JE, Ksiazek TG, Peters CJ, St Jeor SC, Nichol ST. 1998. Genetic analysis of the diversity and origin of hantaviruses in Peromyscus leucopus mice in North America. J Virol 72:57-64.
531.	Ebel GD, Campbell EN, Goethert HK, Spielman A, Telford SR, 3rd. 2000. Enzootic transmission of deer tick virus in New England and Wisconsin sites. Am J Trop Med Hyg 63:36-42.
532.	Kaur R, Tada T, Landau NR. 2023. Restriction of SARS-CoV-2 replication by receptor transporter protein 4 (RTP4). mBio 14:e0109023.
533.	Wong AC, Devason AS, Umana IC, Cox TO, Dohnalova L, Litichevskiy L, Perla J, Lundgren P, Etwebi Z, Izzo LT, Kim J, Tetlak M, Descamps HC, Park SL, Wisser S, McKnight AD, Pardy RD, Kim J, Blank N, Patel S, Thum K, Mason S, Beltra JC, Michieletto MF, Ngiow SF, Miller BM, Liou MJ, Madhu B, Dmitrieva-Posocco O, Huber AS, Hewins P, Petucci C, Chu CP, Baraniecki-Zwil G, Giron LB, Baxter AE, Greenplate AR, Kearns C, Montone K, Litzky LA, Feldman M, Henao-Mejia J, Striepen B, Ramage H, Jurado KA, Wellen KE, O'Doherty U, Abdel-Mohsen M, Landay AL, Keshavarzian A, et al. 2023. Serotonin reduction in post-acute sequelae of viral infection. Cell 186:4851-4867 e20.

534. Jonsson AH, Zhang F, Dunlap G, Gomez-Rivas E, Watts GFM, Faust HJ, Rupani KV, Mears JR, Meednu N, Wang R, Keras G, Coblyn JS, Massarotti EM, Todd DJ, Anolik JH, McDavid A, Accelerating Medicines Partnership RASLEN, Wei K, Rao DA, Raychaudhuri S, Brenner MB. 2022. Granzyme K(+) CD8 T cells form a core population in inflamed human tissue. Sci Transl Med 14:eabo0686.

535. Chen X, Lu PH, Liu L, Fang ZM, Duan W, Liu ZL, Wang CY, Zhou P, Yu XF, He WT. 2016. TIGIT negatively regulates inflammation by altering macrophage phenotype. Immunobiology 221:48-55.

536. Hardiyanto L, Hasegawa A, Komori S. 2012. The N-linked carbohydrate moiety of male reproductive tract CD52 (mrt-CD52) interferes with the complement system via binding to C1q. J Reprod Immunol 94:142-50.