Mechanisms of *Lactobacillus*-mediated protection against Respiratory Syncytial Virus infection and pathogenesis

By

Britton Alexander Strickland

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbe-Host Interactions

December 17, 2022

Nashville, Tennessee

Approved:

James E. Cassat, M.D., Ph.D.

R. Stokes Peebles, M.D.

Fang Yan, M.D., Ph.D.

Seth Bordenstein, Ph.D.

Suman R. Das, Ph.D.

To Taylor and Wilma, my greatest loves and best friends.

To Mom and Dad, for always empowering and believing in me.

To Suman, Meghan, Raj, Hunter, and the Das Lab, who made all this work possible.

ARI = Acute Respiratory Infection

CFU: Colony-Forming Units

GO = gene ontology

FDA = Food and Drug Administration

IFN = interferon (IFNα, IFNβ, IFNγ)

Ig = immunoglobulin (IgA, IgE)

IL-# = interleukin (IL-4, IL-33)

KEGG = Kyoto Encyclopedia of Genes and Genomes

LBP = live biotherapeutic products

LDA: Linear Discriminant Analysis

LEfSe: Linear discriminant analysis [LDA] Effect Size

LactoX = probiotic cocktail containing *Lactobacillus paragasseri BRTN* and *Limosilactobacillus reuteri SMN*

LRTI = Lower Respiratory Tract Infection

OTU: Operational Taxonomic Unit

p40 = 40 kDa protein secreted by some but not all *Lactobacilli*

PBS = Phosphate Buffered Saline

PFU = Plaque-Forming Units

P.I. = Post-Infection

PRR = pattern recognition receptor

rRNA: ribosomal Ribonucleic Acid

RSV: Respiratory Syncytial Virus

RT-qPCR = Reverse Transcription quantitative Polymerase Chain Reaction

SCFA = short-chain fatty acid

TLR = Toll-Like Receptor

TMM = trimmed mean of M values, which is a weighted trimmed mean of the log expression ratios

WMS: Whole Metagenomic Sequencing, or, Wilma Moon Stri

CHAPTER 1

Introduction

A. THESIS OVERVIEW

Beneficial microbes have been implemented into human health practices for centuries; however, there are many unanswered questions on the role and function of these bacteria as a therapeutic. Researchers have only begun to understand that probiotic bacteria in the commensal microbiome can prime the immune system and protect against many infections during early life and adulthood, particularly those caused by respiratory viruses.

The exact mechanisms by which these symbiotic organisms convey protection to the host are not fully understood, limiting the clinical application of the microbiome in disease treatment and diagnostics. In this thesis work, I examine the effect of the probiotic bacterial genera *Lactobacillus* on the outcomes of acute viral infection by Respiratory Syncytial Virus (RSV) in cotton rats and explore potential mechanisms of action to bring therapeutic insight to commensal microbes. Chapter 1 of this thesis is focused on the significance and pathogenesis of RSV, as well as the history and role of the microbiome. Further, the microbiome is defined in the context of the gut and respiratory tract and how bacterial interactions can protect from disease. In Chapter 2, I described the first comprehensive characterization of the cotton rat microbiome and revealed how genetics contribute to host-microbiome structure. In Chapter 3, I generated and applied two *de novo* cotton rat transcriptome references to analyze gene changes during RSV infection. In Chapter 4, I utilized the cotton rat model to study the contribution of commensal *Lactobacillus* in the protection from severe RSV outcomes. In Chapter 5, I further investigated a mechanism of action for *Lactobacillus*-mediated protection through the secreted protein *p40*. Finally, in Chapter 6, I discussed how these findings present a model for understanding the relationship between host and microbe during respiratory viral infections and the future directions of these studies.

B. <u>RESPIRATORY SYNCYTIAL VIRUS (RSV)</u>

i. The significant health burden of RSV in humans

Respiratory syncytial virus (RSV) is the leading cause of lower respiratory tract infection (LRTI) and hospitalization in children below the age of two years. While RSV is a health burden in all age groups, it is a severe global threat to preterm infants, immunocompromised individuals, and the elderly (1, 2). More than 33 million episodes of lower respiratory tract infections, 3.6 million hospital admissions, and over 100,000 in-hospital deaths in children ≤ 5 years occur each year due to RSV (3), with mortality rates estimated to be higher due to underestimation in low-income countries (4).

An estimated 30-60% of children are infected with RSV in the first year of life, with nearly all children having been infected before the age of two (5, 6). Most infections cause mild to moderate acute respiratory infection (ARI) with cold-like symptoms (7). However, some patients may develop severe infection and airway inflammation, resulting in respiratory complications that persist into early adulthood (8, 9). RSV LRTIs are one of the most common causes of hospitalization of infants in the United States (10, 11). RSV-related hospitalization in early life is associated with 4-fold increased odds of developing recurrent wheezing or childhood asthma (12). There is also a significant association between severe RSV ARI in infancy and an increased risk of recurrent wheezing and childhood asthma (13). In addition, severe RSV infection in early life is also a risk factor

for developing sensitization to common allergens (14, 15). Although all children are infected with RSV, not all develop severe outcomes, suggesting intrinsic and extrinsic factors in disease risk that remain unclear.

Once the host airways become infected, RSV replicates in the upper respiratory tract for 2-8 days, with viral titers peaking around day 3-5 (16). In response to viral-induced damage signals, the airways are rapidly infiltrated with immune cells and develop edema in the interstitium (17). In severe cases, the immune cell infiltration may lead to uncontrolled inflammation called a "cytokine storm" that limits the host's ability to sequester infected tissue and limit viral-induced cell death. Subsequent necrosis of the airway tissue leads to epithelial sloughing and increased mucus production that obstructs the airways and spreads the infection to the lower respiratory tract. LRTIs are often associated with other complications, such as bacterial superinfection and pneumonia. Since the lung is among the last organs to mature after birth, there are more incidences of RSV LRTI in young infants (< 6 months) that often result in life-threatening bronchiolitis (inflammation of the lung's small airways) and pneumonia (infection of the lungs) (17).

The host immune response to RSV does not elicit long-lasting natural immunity (18), leaving individuals susceptible to multiple infections with RSV in their lifetime. Unfortunately, to date, there is no approved RSV vaccine and only one monoclonal antibody preventative strategy (Palivizumab), which is used only in high-risk children (19, 20). The failure of the formalin-inactivated RSV vaccine in the 1960s, which counteractively enhanced disease in vaccinees upon encounter with the virus, hampered the development of new RSV vaccines for decades (21-23). However, there is a recently renewed effort to develop RSV preventatives, with 14 vaccine candidates and alternative

anti-viral strategies against RSV (recombinant antibodies (24), nanobodies (25), and small molecule inhibitors analogs (26)) that are at various stages of development (reviewed in (27)). As 93% of RSV LRTI cases and 99% of RSV mortality occurs in developing countries, the need for effective vaccines and low-cost preventatives is critical (28). Additionally, previous vaccine complications highlight the need for an appropriate pre-clinical model for vaccine and drug development against RSV.

ii. Virus-host interactions during RSV infection and pathogenesis

RSV was initially discovered in chimpanzees suffering from colds and coryza (29) but was not associated with humans until 1957 when isolated from two children with respiratory disease (30). RSV is an enveloped virus of the *Pneumoviridae* family (Genus *Orthopneumovirus*). The viral structure can be spherical, filamentous, or asymmetric depending on its exposure and life cycle (31), and the capsid contains a negative-sense, single-stranded, non-segmented RNA genome. RSV infects the respiratory epithelia, specifically primary ciliated columnar cells, and pneumocytes (17). The viral envelope contains three transmembrane glycoproteins that mediate entry into the cell and delay programmed cell death. The G glycoprotein facilitates attachment via toll-like receptor 2 (*TLR2*) (32), CXC3 chemokine receptor 1 (*CX3CR1*) (33), and heparan sulfate proteoglycans (*HSPGs*; although only implicated *in vitro*) (34). These receptors are crucial in contributing to host protein mimicry (G protein contains a conserved CX3C motif that binds to CX3CR1 (35)) and immune evasion, such as suppressing type 1/3 interferon-

THE GLOBAL BURDEN OF
RESPIRATORY SYNCYTIAL VIRUS (RSV)

Paramyxoviridae
Pneumovirinae
Orthopneumovirus
Human RSV

No vaccine, antiviral, or long-lasting natural immunity.

Only 1 preventative used only in high-risk infants.

93% of cases and 99% of mortality occurs in developing countries.

Severe infection →→ hospitalization, recurrent wheezing, & childhood asthma.

Every child infected within two years of age

150,000 deaths in children ≤5 years

33 Million annual cases

3.3 million young children hospitalized

producing plasmacytoid dendritic cells and TNF-α-producing monocytes (36) and promoting Th2-polarized responses (37). Non-structural proteins (NS1, NS2) also promote Th2-polarized responses (37). Non-structural proteins (NS1, NS2) also interfere with the host transcription of type 1 interferons (38). The fusion (F) protein mediates viral-cell fusion via multiple host receptors, including epidermal growth factor receptor (*EGFR*; inducing micropinocytosis and airway mucus secretion) (39), nucleolin (viral internalization) (40), insulin-like growth factor 1 (*IGF1R*; translocates nucleolin from nucleus to cell membrane via protein kinase activity (41)), and intracellular adhesion molecule 1 (*ICAM-1*; promotes neutrophils and eosinophil adhesion to airway tissues (42)). F protein binding also modulates actin polymerization in the host membrane to allow better infected-cell motility, neighboring cell fusion, and cell-to-cell infection (43, 44). These clusters of fused, multi-nucleated cells are called "syncytia," which is the histological hallmark of this aptly named virus.

Recognition of the virus by the host initiates an innate immune response dominated by type 1 interferons (IFN-α, IFN-β) and interferon-stimulated genes. Adaptive immune responses embody a spectrum of cell-mediated responses that vary depending on which type of CD4+ T helper cells (Th1 or Th2) are activated and recruited (45). Th1 cells are activated by INF-γ and function to clear intracellular infections via production of type 1 cytokines (IL-2, IFN-γ, TNF-α) and high CD8+ cytotoxic T cells, which can effectively clear RSV from the tissue while causing minimal pathology. On the other hand, Th2 cells are activated by IL-4 and produce type 2 cytokines (e.g., IL-4, IL-5, IL-13) that induce IgE isotype switching, eosinophil and mast cell recruitment, and mucus production. As these cells mostly function to fight extracellular infections such as helminths, Th2

imbalances during RSV infection impairs viral clearance, increases inflammation, and significantly increases the risk wheezing outcomes and childhood asthma (46, 47). The balance between Th1- and Th2-dominant responses can be influenced by several host risk factors/comorbidities, including immunodeficiency (48), heart/lung disease (49), and SNPs in airway remodeling and innate immune genes (50). However, the most prominent risk factor is dependent on age and prematurity, as preterm and young infants have underdeveloped airways and immune system (51) that produce low levels of Th1-activating cytokines (52) and high levels of Th2-activating cytokines in nasal secretions (53). Still, researchers and clinicians have struggled to find effective prevention and intervention for at-risk individuals.
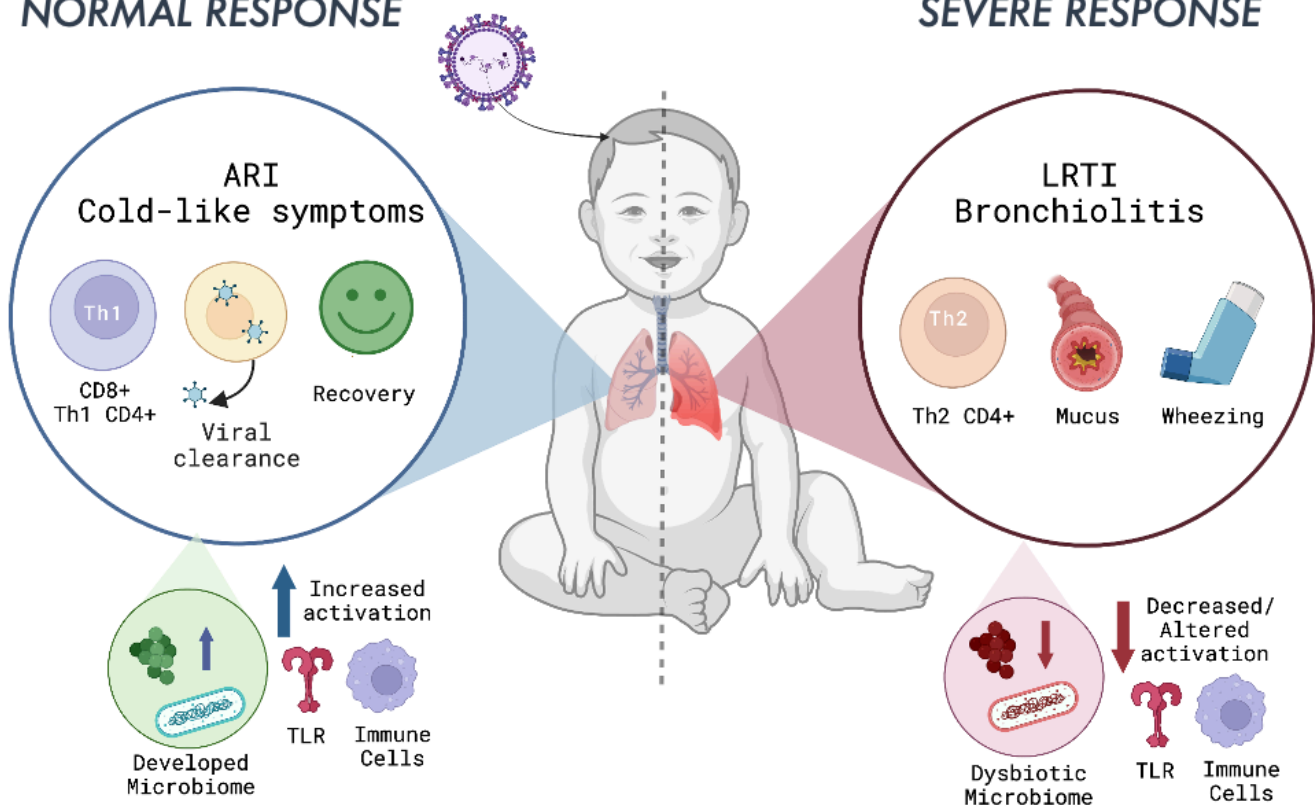
Recent literature on the gut and airway commensal bacteria has suggested a microbiome-mediated influence on immune homeostasis during RSV infection (54, 55). Furthermore, several studies have linked the microbiome's composition to RSV outcomes in animal and clinical models, which will be discussed in detail. A better understanding of these mechanisms can aid in developing targeted therapeutics and preventatives for RSV and other viruses.

i. A brief history of bacteria as a human therapeutic

The relationship between human-associated commensal bacteria and overall health has been studied since the microscopic observation of bacteria by Antonie van Leeuwenhoek in 1675. After bacteria were identified in the production and preservation of food products in the early 1700s (56), scholars toyed with the idea of utilizing bacteria to maintain or "improve" health. In the early 1900s, Nobel-Prize-winning Russian embryologist Elie Metchnikoff inferred that health and longevity could be promoted by repopulating the colonic flora, which he termed a "vestigial cesspool" of microbes harmful to the host, with host-friendly bacteria. This idea stemmed from his observation of impoverished but century-living residents of Eastern European communities and their diet

of bacteria-laden yogurt, which he isolated and implemented into his diet with self-reported success that was never scientifically-validated (57).

Microbial ecology studies by Stamen Grigorov and Ernst Moro later classified this beneficial bacterium under the genus *Lactobacillus* (58, 59) due to their production of beneficial lactic acid and their rod-like shape. Species of *Lactobacilli* (which have been recently classified into 23 different genera (60)) are among microorganisms that, when administered in adequate amounts, confer a health benefit on the host–formally known as "probiotics" (61). These species can coexist with several other probiotic genera, including *Bifidobacterium*, *Enterococcus*, *Streptococcus*, *Pediococcus,* and *Bacillus* species*.* Probiotic*s* are not only found in fermented foods and over-the-counter supplements but also in the oral, gut, respiratory, and vaginal microbiome of humans and animals (62). However, probiotic therapy has remained on the fringes of medical practice due to an unclear mechanistic understanding of exactly how these bacteria convey health benefits.

The microbiome was first showcased as a clinical therapeutic in 1958 when Benjamin Eiseman used fecal transplantation from healthy patients to cure those with severe diarrhea caused by *Clostridiodes difficile* enterocolitis (63). With the utilization of germ-free mice in 1963, scientists tested how individual members of the healthy microbiome interacted with the host in the absence of other microbes (64). Next-generation sequencing revolutionized the field, allowing the identification and study of thousands of unculturable bacteria on a community-wide scale. The early use of 16S rRNA gene sequencing described the contents and dynamics of the host microbiome, and studies began to connect unique microbiome community structures to different

10

populations and diseases. Concordantly, immunologists were working towards uncovering how the host distinguishes pathogens from self (65), danger from non-danger (66), and commensal bacteria from non-commensal bacteria (67). The overlap of these two fields rapidly uncovered evolutionarily conserved crosstalk between microbiota and the host immune system that redefined the study of immunology and infectious disease.

Today, the commensal microbiome is being extensively explored to the species and strain level in various populations and clinical studies. These bacteria have been implicated in many aspects of host health and disease, such as the homeostatic signaling and nutrient acquisition (68). These beneficial properties also significantly affect the protection from or exacerbation of pathogenic infections (69, 70). However, modulating the types and number of bacterial species present within the host depends on several additional factors.

ii. The interplay between commensal bacteria and the host

Several host factors affect the microbial dynamics and activity of bacteria. While environmental factors play an essential role in shaping the microbiome, host genetics play a significant role in establishing and maintaining the microbiome based on genetic polymorphisms and heritability (71, 72). Additionally, co-evolution of the host and microbe suggests that the microbiome is directly associated with the phylogenetics of the host–a term known as phylosymbiosis (73, 74). To this end, many co-speciating bacteria have been linked to neurological, metabolic, digestive, and circulatory diseases (75). In laboratory mice, the genetic background has a greater influence on the microbiome than

environmental stimuli (76). These relationships allow cohesive host-microbiome fidelity across time, but the question remains if these host-mediated patterns are correlative or causative.

Aside from genetics, environmental factors are major drivers of microbiome variability between hosts. Rothschild et al. studied 1,046 healthy individuals of different ancestry in a common environment. They concluded that over 20% of the variance in microbiome diversity could be attributed to environmental factors such as diet and lifestyle (77). Many elements of the environment can disrupt the commensal microbiome, including large-scale factors (chemical pollutants/toxins, socioeconomic setting) and individual-based factors (smoking, diet, antibiotic use).

From birth, the microbiome community structure is established within the first few years of life (78). Alterations in the gut and respiratory microbiota during these crucial years can predispose individuals to health disorders, including obesity (79), autoimmunity such as inflammatory bowel disease (80), and asthma (81). Several birth cohorts have significantly advanced the understanding of microbiome development, as young infants undergo rapid colonization by many different bacteria. While the origins of the human microbiome are poorly understood, mother-to-infant transmission of bacteria through delivery and breastfeeding is a major source of early-life microbes (82). Mode of delivery is thought to be the most critical factor in infant microbiome development due to vaginally-delivered neonates having a microbiome resembling the *Lactobacillus*-dominant birth canal compared to the skin-derived microbiome of cesarean section-delivered neonates (*Staphylococcus-, Corynebacterium-, Propionibacterium*-dominant) (83). In the first few months of life, breastfeeding is a source of probiotic *Bifidobacterium* species and

associated prebiotics (i.e., human milk oligosaccharides) with known metabolic functions that are beneficial to the host (84-86). Other environmental factors contributing to infant microbiome development include living in rural communities and the presence of siblings and pets in the home (87-89). Interestingly, deficiency of *Lactobacillus* and *Bifidobacterium* can predispose the host to developing health complications such as asthma (90) and allergy (91).

iii. The microbiome and host immunity

Studies aimed at understanding the gap between the microbiome and immunity have heavily focused on allergy, infection, and the early-life microbiome. While the initial idea that a "dirtier" early-life environment can protect from diseases has been redefined (92), much of our knowledge still follows the "hygiene hypothesis" (93), which associated early-life exposure to allergens and commensals with the protection from allergy, autoimmunity, and severe infection. This is primarily due to the evolution of the infant immune system driven by the recognition of environmental stimuli. Early exposure to commensal bacteria such as vaginal *Lactobacillus* during delivery is a large factor in protection from allergic disease and severe respiratory infections (94, 95). Other environmental factors that disrupt or modify the infant microbiome, such as the use of antibiotics and no siblings in the home environment, can significantly increase the odds of developing asthma (96). In addition to the environment, many host factors such as genetics and age correlate with both microbiome development and risk for allergy and infection (97, 98).

Indigenous commensals and probiotic bacteria can educate and enhance both innate and adaptive immune responses (99). These symbionts maintain tolerance and reduce inflammation by processing secondary metabolites and stimulating cytokines, in addition to other functional factors that are being rapidly discovered (100). Bacteria can interact directly with the host via pattern recognition receptors (PRRs) to maintain intestinal homeostasis and prevent inflammation, particularly during epithelial injury (67, 101, 102). For example, bacterial polysaccharide A (PSA) can activate host CD4+ T cells via toll-like receptor 2 (TLR-2) and class II major histocompatibility complex II (MHC-II), which stimulates INF-γ-dependent promotion of Th1 and suppression of Th2 cells (103, 104). Bacteria can also metabolize many host carbohydrates (e.g., dietary fiber) into short-chain fatty acids (SCFAs) such as acetate and butyrate that can reduce viral cytotoxicity (105) and promote antiviral interferon signaling (106). These mechanisms are most evident in the gut (where the bacterial load is the highest). Still, other sites such as the respiratory tract can utilize their distinct microbiome to combat common viruses and allergies (107-109).

iv. The gut-lung axis

The commonality of epithelial immune responses throughout the body allows the host to integrate metabolites and subsequent signaling and immunological responses from gut microbes at distal sites, particularly in the lung via the "gut-lung axis" (110). Microbiome communities differ based on their location in the human body, but evidence suggests a strong association between the respiratory tract and gut microbiome (111,

112). Clinically, there is a high rate (~50%) of pulmonary disorders in patients with inflammatory bowel disease (113), which also correlates to microbiome dysbiosis in these patients compared to healthy groups (114).

Proposed mechanisms of intercommunication include bacterial manufacturing or secretion of metabolites (i.e., lipopolysaccharide, SCFA from dietary fiber) that are transported through the bloodstream and modulate tissue sites across the body. For example, alveolar macrophages and alveolar type II cells can sense such metabolites through SCFA-specific receptors to modulate cytokines such as IL-1β (115). *Proteobacteria* in the gut can regulate select cytokines such as IL-33 to direct the migration of group 2 innate lymphoid cells (ILC2s) from the gut to the lung (116). During influenza infection, lung-derived CD4+ T cells can be recruited into the small intestine and increase inflammatory cytokines and injury. This leads to alterations of intestinal microbiota that stimulate epithelial production of IL-15 and promotes Th17 cell polarization in the small intestine (117). Other mechanisms are rapidly being uncovered, including other axes such as the gut-brain axis (118) and the gut-joint axis (119).

The Gut-Lung Axis

Soluble factor and immune cell transport via blood and lymphatic vessels

Lactobacillus spp.
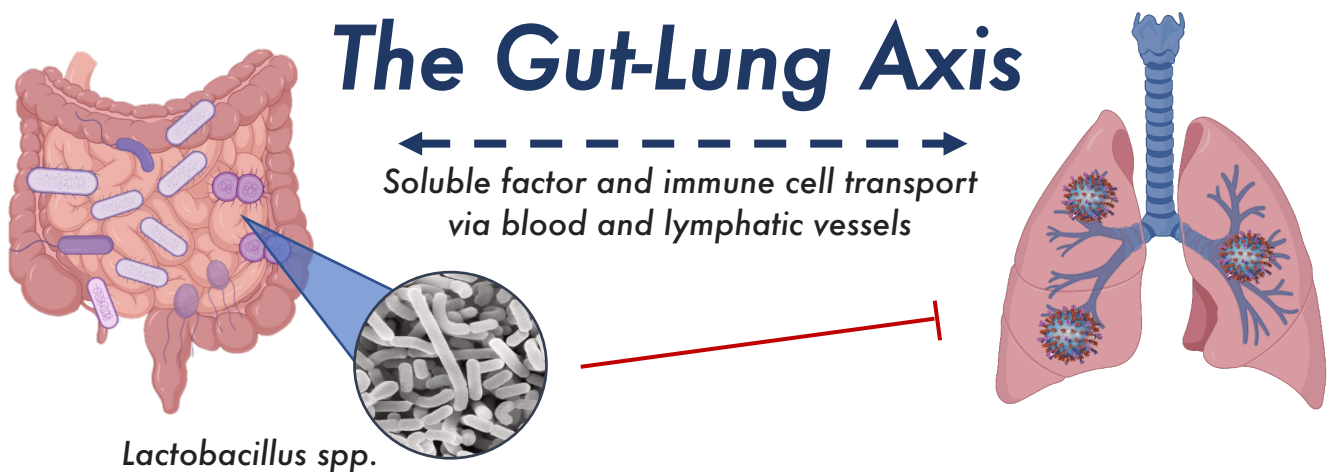
D. COMMENSAL BACTERIA ASSOCIATED WITH REDUCED RSV DISEASE SEVERITY

i. Protective properties of bacteria and probiotic supplementation

As with allergy and asthma, there have been many associational links between the composition of the gut and airway microbiome and the severity of acute respiratory infection in humans. Both large-scale differences (i.e., diversity of the microbiome, multiple taxa) and single-taxa differences (presence/absence of genus, species, or strain) have been attributed to the protection from or exacerbation of disease severity. Many of these differences are attributed to intrinsic (age, genetics) and extrinsic (environmental) factors that impede the early-life colonization of commensal bacteria. Consequently, these factors also have been shown to increase the risk for severe RSV bronchiolitis.

Two studies found breastfeeding to be associated with less severe RSV infection, as determined by the need for oxygen therapy and duration of hospital stay (120, 121). A national study in Denmark found that delivery by C-section was associated with increased RSV severity and hospitalization in the first two years of life (122). The use of perinatal antibiotics is also associated with severe RSV infections (123). Recent literature has longitudinally investigated the infant gut and airway microbiome to understand how microbiome composition affects RSV severity and outcomes. Additionally, others have used *in vivo* studies in mice to examine the effect of intranasal and oral probiotic supplementation on RSV disease severity.

Owing to the gut-lung axis, the role of the gastrointestinal tract microbiome in RSV disease severity has been previously explored in humans and animals. Using a clinical cohort of infants hospitalized with RSV and healthy controls, Harding et al. found significant differences in microbiome community composition between the two groups. The gut microbiome in RSV-infected infants had a higher abundance of *Muribaculacae, Clostridiales*, *Odoribacteraceae*, *Lactobacillaceae*, and *Actinomyces* compared to healthy controls. The abundance of *Muribaculacae* also correlated with increased RSV severity. However, it is unclear whether these trends contributed to RSV severity or were caused by the infection itself, as the samples were collected 72 hours post-hospitalization (124). In another case-controlled study of infants hospitalized with bronchitis (with 65% of cases caused by RSV), gut microbiome communities dominated by *Bacteroides* had higher instances of bronchitis compared to *Enterobacter/Veillonella*-dominant microbiomes (125).

Animal studies have greatly developed our knowledge of how the gut microbiome can contribute to RSV disease severity. In a series of studies led by Fonseca et al., oral supplementation of *Lactobacillus johnsonii* in adult mice is suggested to protect against RSV infection, measured by decreased airway pathology, Th2 cytokines, dendritic cell function, and increased T regulator cells (54). This effect was determined to be modulated by increased docosahexaenoic acid (DHA) and 2-hydroxyisobutyrate (a butyrate precursor) in the plasma of supplemented mice (54), which was validated *in vitro* in reducing Th2 cytokines, IL-12, IL-6, IL-1β, IFN-β, and TNF-α via PPAR-γ signaling in

dendritic cells (126). Additionally, maternal supplementation with *L. johnsonii* resulted in offspring with reduced airway mucus and Th2-dominant responses to RSV infection, signifying a role of the maternal microbiome in RSV immunity (127). Oral *L. reuteri* but not *L. salivarius* was also shown to protect against severe ovalbumin-induced allergic responses, signified by reduced airway eosinophils, TNF, MCP-1, IL-5, and IL-13 but not IL-10 in bronchoalveolar lavage fluid of antigen-challenged animals (128).

Ji et al. used oral supplementation of neonatal mice with *L. rhamnosus GG, Escherichia coli* Nissle 1917, and VSL#3 (a medical-grade 8-strain probiotic product) before RSV infection, which suppressed viral-induced lung pathology and increased antiviral IFN-β in alveolar macrophages (129). Supplementation also increased the abundance of SCFA-producing bacteria in the lung (*Corynebacterium*, *Lactobacillus*) and acetate levels in the serum.

Another study found that oral administration of *L. rhamnosus* CRL1505 to infant mice reduced RSV viral load and tissue injury by enhancing antiviral signaling, Th1 cells, and Th1-promoting dendritic cells (CD103[+] and CD11b[high] dendritic cells) (130). They further identified alveolar macrophages in producing type 1 interferon, IL-6, and IFN-γ to drive Th1 dominance, T-regulatory cell IL-10 production, and viral clearance (131). These results suggest a prominent effect of *Lactobacillus* in protecting against severe RSV that has not been tested in the clinic. However, in an Argentinian cohort, *L. rhamnosus* CRL1505 has also been administered to humans via probiotic yogurt, which successfully reduced the incidences of upper airway tract infections, pharyngitis/tonsilitis, and acute diarrhea (132).

Eguchi et al. also showed that oral administration of *L. gasseri* LG2055 in mice before RSV infection decreased viral titers, weight loss, proinflammatory cytokines, and interferon/interferon-stimulated genes (133). This effect was proposed to be exerted by the downregulation of SWI2/SNF2-related CREB-binding protein activator protein (SRCAP), a scaffolding protein to which viral non-structural proteins bind (133). Although different strains of bacteria produce variable transcriptomic results, this suggests several gut-lung axis mechanisms for treating and preventing RSV.

iii. Airway bacteria

The airway microbiome has also been directly implicated in RSV severity. In a clinical cohort of 132 patients less than 2 years of age, RSV infection and hospitalization were positively associated with nasal *Haemophilus influenzae* and *Streptococcus* but negatively associated with *Staphylococcus aureus* abundance, independent of age (134). Other studies have associated a higher relative abundance of *Haemophilus* in the nasopharynx with increased RSV viral load (135) and delayed viral clearance (136). Transcriptomic interactions of *Haemophilus* during RSV infection were examined by the Das Lab using the INSPIRE national infant microbiome cohort at Vanderbilt University Medical Center, where *Haemophilus* abundance was positively associated with many local immune mediators with antiviral, proinflammatory, chemotactic, and cell differentiation/proliferation activity (e.g., TNF-α, IFN-α, IL-1β, IL-6, IL-8, IL-12, IL-13, IL-33) but not IFN-γ, IL-4, IL-5, and IL-17A. This suggests that *Haemophilus* may increase airway inflammation during RSV infection but not CD4+ T cell imbalance. Another study in the INSPIRE cohort correlated severe RSV infection and increased asthma risk with

higher *Haemophilus*, *Moraxella*, and *Streptococcus* compared to higher *Staphylococcus* and *Corynebacterium* in healthy patients (137). To this end, in another RSV hospitalization cohort, the rate of ICU admittance in RSV-hospitalized infants was highest in *Haemophilus*-dominant patients and lowest in *Moraxella*-dominant patients (138). This data suggests that the presence and abundance of multiple taxa can exacerbate RSV disease severity.

Additionally, the Das Lab found that the presence and increased abundance of *Lactobacillus* in the nasopharynx during infant RSV infection is associated with reduced inflammatory cytokines, viral titers, and risk of developing subsequent wheeze at 2 years of age (55). De Boeck et al. demonstrated that specific strains within the *Lactobacillus* genus (specifically *L. casei* AMBR2) occupy a niche in the nasopharynx due to bacterial fimbriae that enable strong adherence and colonization of upper respiratory tract epithelium in humans upon intranasal inoculation (139). *L. casei* AMBR2 was also shown *in vitro* to inhibit the growth and virulence of several respiratory pathogens (*S. aureus*, *M. catarrhalis*, and *H. influenzae*) (139), which are also taxa shown to be dominant in patients with severe RSV. However, it is unclear whether the protective effect of nasal *Lactobacillus* is due to modulation of the resident microbiota or directly involved in promoting protective outcomes.

Several animal studies have also explored the protective mechanisms of intranasal *Lactobacillus* and RSV outcomes. Intranasal administration of live *L. rhamnosus* CRL1505, heat-killed CRL1505, and its peptidoglycan improved the resistance to primary RSV infection (decreased viral titers, protection from weight loss) and secondary pneumococcal pneumonia through the promotion of Th1 cells (increased IFN-γ), T-

regulatory cells (IL-10), and alveolar macrophages (IFN-β) (140, 141). These effects mimic those of gut *L. rhamnosus* CRL1505 in RSV outcomes (130), and the success of heat-killed bacteria and isolated peptidoglycan from the bacterial membrane suggest the protective effect may be elicited through intrinsic factors and proteins within the structures of *Lactobacilli*. Similar studies have also found intranasal administration of various live *Lactobacillus* strains to induce protection during influenza infection (142).

As both gut and lung *Lactobacillus* have been implicated in RSV protection, the mechanism of direct or distal (gut-lung axis) action is still unclear. While these analyses of microbial effects are essential to understanding the pathogenesis and prevention of RSV and long-term sequelae, appropriate animal modeling has not been explored for mechanistic analysis of these microbial contributions. In each of these cases, a site-specific dominant *Lactobacillus* species was found to be associated with protection. Yet, it is poorly understood how these individual species and the entire genera could exert these protective effects.

iv. Probiotic Interventions in the Clinic

There is no published data on clinical trials to examine the effect of probiotic supplementation on RSV and wheezing outcomes in humans. However, other clinical trials have presented probiotic-induced protective phenotypes in infants and adults with influenza and rhinovirus, although with variable results. In summary, these studies have shown reduced incidence of infection (143-145), increased antiviral immune response (146), enhanced vaccine efficacy (147), decreased viral replication (148) (although not in

all studies (144, 149)), reduced clinical symptoms (150, 151), and shorter duration of upper respiratory symptoms (152, 153). The results of these studies have been previously reviewed in detail (154, 155). All these clinical trials used different species and strains of bacteria (mostly *Lactobacillus*, *Bifidobacterium*, *Lactococcus*, and *Bacillus* spp.) at varying concentrations and durations, but all via oral supplementation. Additionally, some of these studies are at an unclear risk for bias due to selection methods for randomness, lack of matching placebos, no sample size calculations, missing outcome data, and funding or employment by probiotic manufacturer (154). This data suggests the modulation of the gut microbiome as a promising approach to protecting against respiratory viral infection, but variable data highlights the need for an appropriate preclinical model to better test these mechanisms of microbe-host interaction.

While most probiotics are considered safe food supplements, patient and consumer options have been limited to over-the-counter applications rather than medical settings. Due to the lack of approved guidelines and variable clinical results, there have been many challenges in developing Food and Drug Administration (FDA)-approved probiotic therapies, also called live biotherapeutic products (LBPs). However, this has not hindered worldwide advances in using "bugs as drugs." In 2020, the first FDA-approved live biotherapeutic product from Seres Therapeutics (SER-109) hit the market, which is a stool-derived gut microbiota capsule for the treatment of recurrent *Clostridium difficile* (156). This product presented an excellent alternative to using complex but effective fecal microbiome transplants to treat *C. difficile* infections. This progress represents a shift in medicine towards a microbiome-focused approach to health and disease, opening a "wild-west"-like avenue for prebiotic, probiotic, and postbiotic therapy development.

LABORATORY SUPPLEMENTATION OF *LACTOBACILLUS* IN RSV

1.   2.   *RSV*
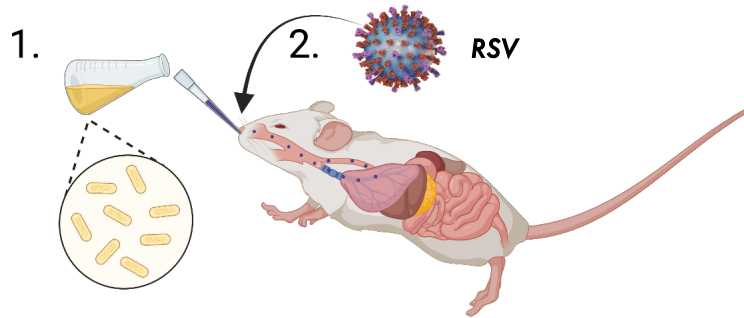
↓ decreased airway pathology[1,5]
↓ Type 2 cytokines[1,3,4]
↓ viral replication[2,4,6]

↑ antiviral response[2,4,5,7]
↑ regulatory T cells[1,3,7]
↑ Type 1 cytokines[2,3,7]
↑ weight retention[4,6]

[1]Fonseca et al, Mucosal Immunology, 2017 (gut)
[2]Chiba et al, Inter. Immunopharma, 2013 (gut)
[3]Garcia-Castillo et al, Front Immunol, 2020 (gut)
[4]Eguchi et al, Sci Reports, 2019 (gut)
[5]Ji et al, Acta Pharmacol Sin, 2021 (gut)
[6]Tomosada et al, BMC Immuno, 2013 (airway)
[7]Clua et al, Front Immunol, 2017 (airway)
*all studies performed in mice

CLINICAL SUPPLEMENTATION OF *LACTOBACILLUS* IN OTHER VIRUSES

1.   2.   *INFLUENZA*

RHINOVIRUS

RSV
No clinical trials on
RSV and Lactobacillus.
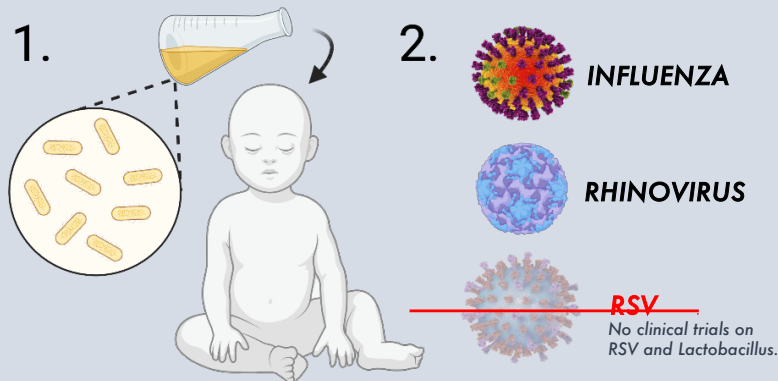
↓ incidence of infection[1,2,3]
↓ viral replication[4]
↓ clinical symptoms[5,6]

↑ antiviral immune response[7]
↑ vaccine efficacy[8]
↑ remission time[6]

[1]Lehtoranta et al, J Clin Virol, 2014 (gut)
[2]Luoto et al, JACI, 2014 (gut)
[3]Wang et al, J Am Geriatr Soc, 2018 (gut)
[4]Turner et al, Benef Microbes, 2017 (gut)
[5]Kumpu et al, Benef Microbes, 2015 (gut)
[6]Kumpu et al, J Med Virol, 2013 (gut)
[7]Sugimura et al, Br J Nutr, 2015 (gut)
[8]Davidson et al, EurJ Clin Nutr, 2011 (gut)

E. THE COTTON RAT AS A PRECLINICAL MODEL FOR RSV

Cotton rats (Genus *Sigmodon*) are considered to be the "gold standard" animal model for RSV infection because it manifests several characteristic features during human RSV infection, such as similar infection distribution, kinetics, and immune responses (157). Mice have been used extensively to study RSV immune responses, but several factors render the mouse impractical for understanding RSV pathology and viral kinetics:

- Low RSV replication: not translatable to humans; 100-fold higher in cotton rats (158)

- Resistance to upper respiratory infection: unlike humans and cotton rats, RSV does not infect the mouse nasal cavity (158-160)

- Divergent lung cell infection: RSV infects ciliated bronchial epithelial cells and alveolar cells in humans and cotton rats while only infecting pneumocytes in mice (161)

- Molecular function: cotton rats have functional *Mx* genes, a crucial component of the innate antiviral defense that is nonfunctional in laboratory mice (162, 163)

In addition to RSV, cotton rats are an important animal model for other viral diseases, including parainfluenza (164, 165), influenza (166, 167), measles (168), human metapneumovirus (169), and enterovirus (170) due to comparable human disease outcomes. Cotton rats have predicted the efficacy of several therapeutics and vaccines currently used in high-risk human populations (171-174). RSV infection persists in the animal's upper airway and causes pathology consistent with chronic histological changes

in human RSV-induced respiratory disease, recurrent wheeze, and asthma (175). However, the cotton rat is limited by the lack of immunological reagents (i.e., antibodies for cytometric analysis) and an available sequenced genome.

Two inbred species of cotton rats are commonly used in respiratory viral research: the Hispid cotton rat (*S. hispidus*) and the Tawny-bellied cotton rat (*S. fulviventer*). *S. hispidus* was the first and only rodent found susceptible to poliovirus at the National Institute of Health by Charles Armstrong in 1938 (176) and was inbred for laboratory research. *S. fulviventer* was later inbred by Virion Systems in the 1990s for work on RSV (177) and PIV-3 (178). The most recent phylogenetic analysis of *Sigmodon spp.* found that *S. hispidus* and *S. fulviventer* diverged 5.4 million years ago (179). In the wild, *S. hispidus* and *S. fulviventer* are sympatric species, with *S. fulviventer* being the more dominant animal (180).

Sigmodon hispidus　　　　*Sigmodon fulviventer*
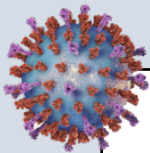
Most permissive animal model for human RSV infection

Develop upper respiratory infection

Infect lung epithelia similarly to humans
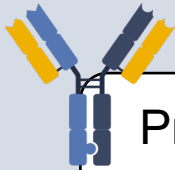
Predicted efficacy of current RSV prophylactic immunoglobulin
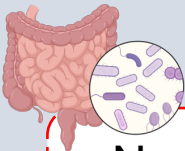
No established microbiome

CHAPTER 2

No genetic reference

CHAPTER 3

This dissertation argues that the cotton rat is an excellent preclinical model for studying microbiome mechanisms during respiratory infections. In Chapter 2, I explore the characteristics of the cotton rat microbiome for the first time. There, I found that two commonly used species of cotton rats have distinct microbiomes, which includes the differential abundance of *Lactobacillus* between species. The *S. fulviventer* gut had a significant lower abundance of *Lactobacillus* compared to *S. hispidus* based on 16S rRNA gene sequencing, whole metagenomic sequencing, and traditional culture methods (181). Additionally, *S. hispidus* gut was abundant in probiotic *Lactobacilli* (*L. reuteri, L. gasseri*) while *S. fulviventer* gut was, to a much lesser extent, populated with *Lactobacilli* with little probiotic value (*L. murinus, L. animalis*). This phenomenon presents an RSV-susceptible, *Lactobacillus*-deficient animal to better understand *Lactobacillus's* contribution to RSV disease severity. Genetic applications, such as next-generation sequencing, are limited due to the lack of a published genome for any species of cotton rat. To this end, in Chapter 3, I present reference transcriptomes for both *S. fulviventer* and *S. hispidus* and apply them to RSV transcriptomic analysis. Additionally, no microbiome modulation studies have been performed in cotton rats. In Chapter 4, I present the results on oral *Lactobacillus* and the protection from RSV severity in cotton rats.

F. DEFINING THE MECHANISM OF MICROBIOME-MEDIATED PROTECTION FROM SEVERE RSV AND OTHER VIRUSES

The anti-viral mechanisms of probiotic action are not fully understood, but overall mechanisms can be defined in three categories:

1. **Competition:** block pathogen colonization and production of toxins; compete for intestinal adherence by virus or other bacteria

2. **Immune modulation:** enhance innate immunity; modulate pathologic inflammation via pattern recognition receptor (PRR) signaling

3. **Homeostasis**: regulate non-immune epithelial homeostasis for proliferation, barrier function, and protective/anti-apoptotic signaling

Probiotics, unfortunately, have limits to their clinical application. Bioavailability is challenging because of potential degradation as the probiotic traverses the gastrointestinal tract and the lack of necessary nutrients for activation or replication. Live organisms also have differential abilities to colonize the host due to many intrinsic and extrinsic factors, which has presented challenges in effectively using probiotics in both mice and humans (182). However, properties that promote colonization, such as increased mucoadhesive properties, have been explored in manufacturing genetically modified probiotic strains (183). Safety is another concern, as the introduction of foreign strains has previously led to bacteremia in immunocompromised patients (184). *Prebiotics*–compounds in foods such as fibers that stimulate the growth of probiotic bacteria–are a viable option for promoting probiotic growth in the host, as indicated in the

use of dietary fiber to promote metabolite production by probiotics (106). *Postbiotics*–bioactive compounds produced by bacteria–have also been proposed as a safe alternative, as these convey the functional properties of probiotics without the need for live, active bacteria. These byproducts include metabolites (SCFA), extracellular polysaccharides (peptidoglycan, LPS, teichoic acid), cell lysates, pili-type structures, extracellular vesicles, bacteriocins, and other functional proteins (185).

Oral administration of postbiotics such as acetate, propionate, and butyrate are effective in protecting against respiratory viral infection. Antunes et al., 2019 showed that SCFA supplementation reduced RSV load, mediated IFN-β response via interferon-stimulated genes *OAS1* and *ISG15*, and increased virus recognition receptors MAVS and RIG-I in RSV-infected mice. RIG-I knockout mice did not exhibit protective phenotypes (106). Concordantly, quantification of SCFAs in stool from infants hospitalized with RSV bronchiolitis found that high levels of acetate were significantly associated with less severe bronchiolitis, higher oxygen saturation, and diminished lasting fever (105).

Other microbial-associated metabolites, such as desaminotyrosine (DAT), have also been shown to protect against influenza infection and inflammation following oral administration in mice. DAT is produced by *Clostridium orbiscindens* metabolism of food-derived flavonoids, and colonization of *C. orbiscindens* and flavonoid-enriched diet amplified type I interferon signaling and phagocytosis in the lung in response to influenza infection (186). The success of this study suggests that live biotherapeutic products (LBPs) may need to encompass a trivalent approach that employs active bacteria and pre- and post-biotics

Many other postbiotics have shown beneficial properties that may affect respiratory health but have not been directly studied in respiratory contexts. These include indole derivatives from dietary tryptophan metabolism (187), niacin (188), urolithin A (189), and pyruvic and lactic acids (190). One mechanism studied in a murine colitis model is the *L. rhamnosus GG-secreted* protein *p40*, which was identified to prevent TNFα-dependent epithelial cell death in a concentration-dependent manner (191, 192). *Lactobacillus* utilizes *p40* as a cell wall hydrolase that is later secreted in high abundance in the culture supernatant of several (but not all) species (193). As soluble *p40* traverses the epithelial membrane of the host, it interacts with metalloproteinase ADAM17 for cleavage of HB-EGF, which signals dimerization and phosphorylation of EGFR and PI3K to induce AKT, protecting against caspase-induced apoptosis (191, 192, 194). This protein has become of interest in RSV protection because of the strong cytopathic effect that the virus has on the host, often resulting from overstimulation of TNF-α and subsequent caspase-induced apoptosis (195-197). Based on these observations, I hypothesized that a similar mechanism is at work in the RSV disease pathway, as illustrated in **Figure 1.6**. This work is discussed in **Chapter 5.**
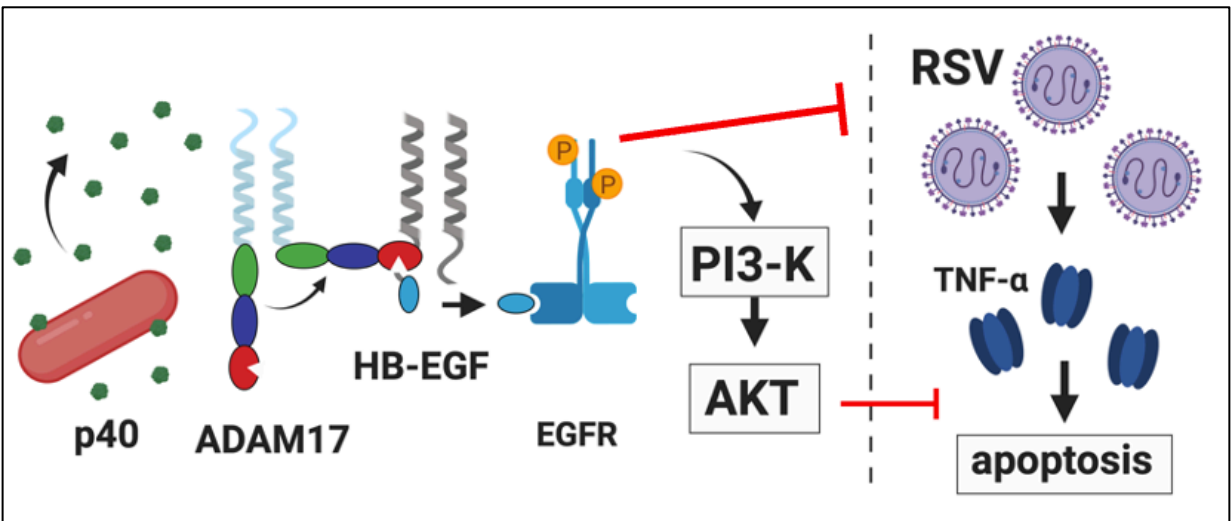
Figure 1.6. Proposed mechanism of *Lactobacillus*-secreted *p40* in RSV pathology. RSV infection mechanisms are well understood, and the *p40* mechanism of protection has been established. This work aims to link these two together in both the context of the gut and respiratory tract of cotton rats.

CHAPTER 2

Comprehensive characterization of the microbiome from two species of Sigmodon cotton rats and the influence of host genetics on microbial community structure

The commensal microbiome can dramatically influence host health and disease aspects, such as homeostatic signaling, nutrient acquisition, and protection from or exacerbation of infections (68, 69, 198). Many early studies established that environmental factors significantly shape and modulate the host microbiome (199-202). These include geographic regions, associated cultures, diet in humans (77), and vendor and housing facilities in animal model organisms (203, 204). In addition, recent studies have emphasized that host genetics plays a significant role in the co-evolution of the host and its associated microbiome (73, 75, 205-207). For example, the murine genetic background is a stronger determinant of microbiome composition and structure than environmental stimuli (76).

Similarly, genetic polymorphisms, heritability, and overall host genetics in humans can also shape how commensal bacteria evolve alongside the host (71, 72, 208). The microbiome has also been instrumental in predicting and protecting against severe viral disease outcomes (209, 210). However, this burgeoning field of bacteria-host-virus interactions has been limited by a lack of pre-clinical models to study mechanisms of virus-microbiome interaction.

Cotton rats (genus *Sigmodon*) are an important small animal model for studying various respiratory diseases, including Respiratory Syncytial Virus (RSV) (211), influenza

A Virus (166, 212), parainfluenza virus (164, 165), measles (168), human metapneumovirus (169), enterovirus (170), and human rhinovirus (213) due to comparable human disease outcomes. Cotton rats have also provided a valuable model for nasal colonization studies (especially with *Staphylococcus aureus*) due to their human-like nasal histology (214). Furthermore, cotton rats are a valuable tool for research since they harbor zoonotic viruses in the wild like Alphavirus (Equine Encephalitis virus), Hantavirus (Black Creek Canal virus, Bayou virus), Cardiovirus, Arenavirus (Tamiami virus), and Flavivirus (West Nile virus) (215-220).

While mice have been used extensively to study viral immune responses, several factors render the mouse impractical for understanding viral pathology and kinetics, such as those relating to RSV: low replication (158), resistance to upper respiratory infection [unlike humans and cotton rats, RSV does not infect the mouse nasal cavity (158-160), divergent lung cell infection [RSV infects ciliated bronchial epithelial cells and alveolar cells in humans and cotton rats, while only infecting pneumocytes in mice (161)], and histological outcomes inconsistent with those similarly seen in the upper and lower airway of both humans and cotton rats (158). However, mice do have advantages due to readily available reagents, antibodies, and genetic references. Yet, studies in cotton rats have accurately predicted the efficacy of several RSV therapeutics and vaccines currently used in high-risk human populations (171, 174, 221, 222). Considering all these factors, the cotton rat provides an excellent model for studying viral-bacterial interactions than mice.

Furthermore, understanding the cotton rat microbiome is instrumental for understanding microbial interactions with viral infections. Many associational effects of both nasal and gut microbiome composition and protection from severe viral outcomes

have been well described in mouse models (126, 223-225). The microbiome of humans (203), mice (204, 226), rats (227), and other animals have been comprehensively studied and characterized in publicly available databases used for answering questions relating to host microbiome and disease outcomes. However, there has not yet been a comprehensive analysis of the microbiome in healthy cotton rat species, making studies of viral-microbiota interactions in this animal model challenging. To date, only one study has examined the nasal microflora of healthy *S. hispidus* but was limited by the sample number and lack of longitudinal timepoints (228).

This study aimed to comprehensively characterize and establish the structure and composition of the cotton rat microbiome. Longitudinal samples were collected from four body sites of two commonly used inbred cotton rat species, *S. hispidus* and *S. fulviventer*, maintained under the same environment and dietary conditions. Microbiome characterization using both 16S rRNA gene and whole metagenomic sequencing (WMS) was used to comprehensively describe the microbiome community structure and composition of different body sites in cotton rats and showed distinct community structures based on the cotton rat species and sex. WMS also showed differential metabolic potential of the community between species. Overall, this study not only adds to the small but rapidly expanding literature on the influence of host genetics on the microbiome but also describes an appropriate animal model for studying microbiome influences on viral and bacterial diseases.

i. Characterization of cotton rat microbiome from multiple body sites.

To characterize the healthy cotton rat microbiome structure and composition, two groups of 10 young male cotton rats of S. fulviventer and S. hispidus were observed longitudinally for 111 days. A total of 140 samples were collected and processed for 16S rRNA gene sequencing: ear swabs (20 swabs/day 95), nasal brushes (20 swabs/day 95), skin swabs (20 swabs/day 95), and fecal samples (80 swabs/days 0, 4, 34, and 111) (**Figure 2.1A**). DNA was extracted, and the V4 region of the 16S rRNA gene was amplified and then sequenced on the Illumina MiSeq platform with 2x250 base pair reads, generating an average of 35,194 reads per sample. Microbiome data was processed by following the mothur SOP, and operational taxonomic units (OTUs) were clustered at 97% identity. For diversity testing with the *vegan* R package, a sample read cutoff of >10,000/sample was implemented, which utilized 92.9% of samples for analysis with the smallest library size of 11,820 reads. The remaining tests to assess differences in the abundance of specific taxa (i.e., DESeq2, stabsel, GeneSelector, and LEfSE) used samples passing a per sample read cutoff of >1000/sample, which utilized 96.4% of samples, with the smallest library size of 3,678 reads.

Phyla within *S. fulviventer* and *S. hispidus* external sites (aggregation of ear, nose,

and skin samples) were similarly dominated by *Proteobacteria*, *Actinobacteria*, and *Firmicutes*, with only *Tenericutes* being significantly more abundant in *S. fulviventer* (DESeq2 testing; $\log_2$ fold change=1.04, q=3.79E05). Fecal communities mainly consisted of two dominant phyla with opposite abundances between cotton rat species: higher *Bacteroidetes* abundance in *S. fulviventer* compared to *S. hispidus* (50.0% vs. 42.4% respectively, q=1.69E-06) and higher *Firmicutes* abundance in *S. hispidus* compared to S. *fulviventer* (36.2% vs. 31.8% respectively, q=1.18E-03) (**Table 2.1**). The distribution and differential abundance of the top 20 genera in each cotton rat species are shown in **Figure 2.1 B-C**.

Phyla	Body		Ear		Nose		Feces	
	S.hispidus	*S.fulviventer*	*S.hispidus*	*S.fulviventer*	*S.hispidus*	*S.fulviventer*	*S.hispidus*	*S.fulviventer*
Proteobacteria	61.1% ± 6.3	56.1% ± 6.0	60.0% ± 9.3	55.7% ± 3.9	63.0% ± 9.9	51.4% ± 8.8	2.4% ± 1.1	2.7% ± 1.0
Actinobacteria	21.6% ± 6.8	21.4% ± 5.1	22.5% ± 7.7	19.5% ± 2.8	16.0% ± 3.4	18.5% ± 4.5	0.8% ± 0.5	0.3% ± 0.3
Tenericutes	8.6% ± 1.1	11.2% ± 3.0	9.5% ± 3.0	13.9% ± 5.2	13.4% ± 9.5	18.6% ± 13.1	0.3% ± 0.5	2.1% ± 2.8
Firmicutes	6.6% ± 1.8	8.9% ± 1.2	6.4% ± 1.7	8.4% ± 1.2	5.8% ± 1.6	9.0% ± 1.8	42.4% ± 9.8	31.8% ± 6.8
Bacteroidetes	1.6% ± 0.5	1.9% ± 0.5	1.3% ± 0.4	2.0% ± 0.4	1.4% ± 0.5	1.7% ± 0.5	36.2% ± 9.1	50.0% ± 5.8
other	0.50%	0.50%	0.30%	0.60%	0.30%	0.70%	18.0%	13.2%

Table 2.1. Percentage abundance of phyla in both *S. hispidus/S.fulviventer* with standard deviations.
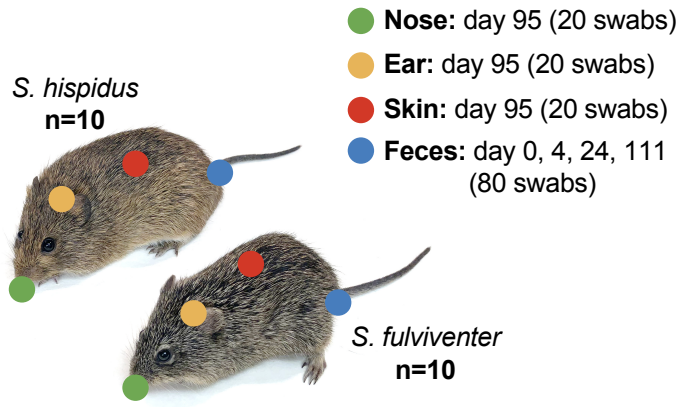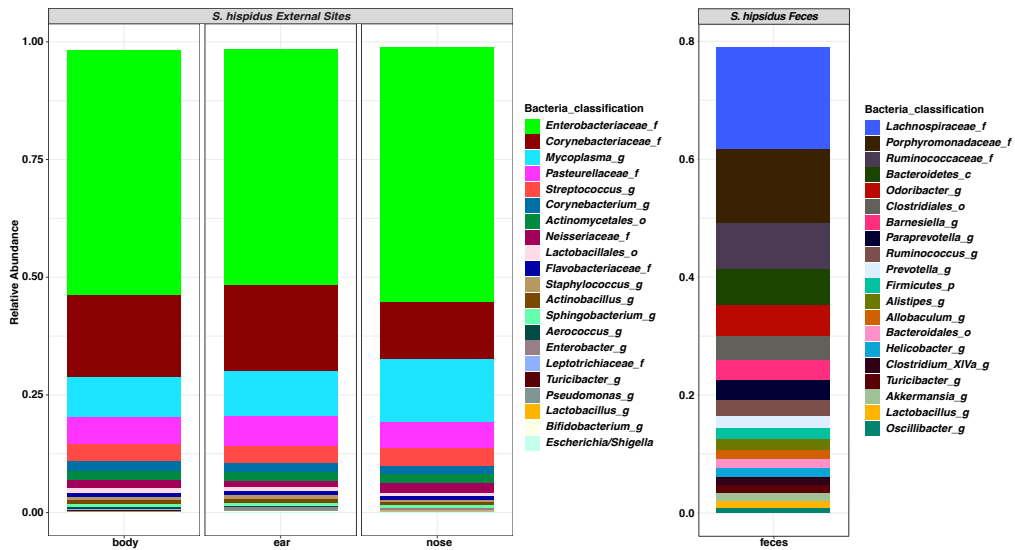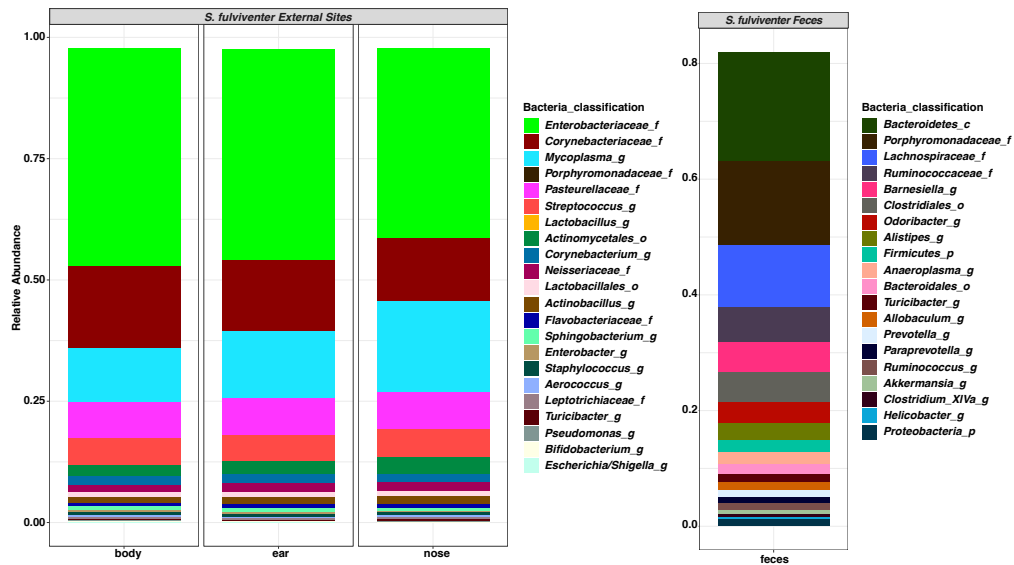
Figure 2.1. **The cotton rat microbiome using 16S rRNA sequencing.** (A) Study design includes ten cotton rats from both *S. fulviventer* and *S. hispidus.* Nose, ear, skin, and feces were taken across 111 days. (B-C) Top 20 most abundant bacteria genera at each body site of (B) *S. hispidus* and (C) *S. fulviventer.* External sites (body, ear, nose) share similar dominating genera.

To first compare community characteristics of cotton rats with other species, beta diversity (Bray-Curtis dissimilarity) was compared between cotton rat fecal samples from Day 0 and publicly available 16S rRNA sequencing data from multi-ethnic humans (203) and mice (226). Comparisons revealed significant differences in community composition between individual species (**Figure 2.2**).
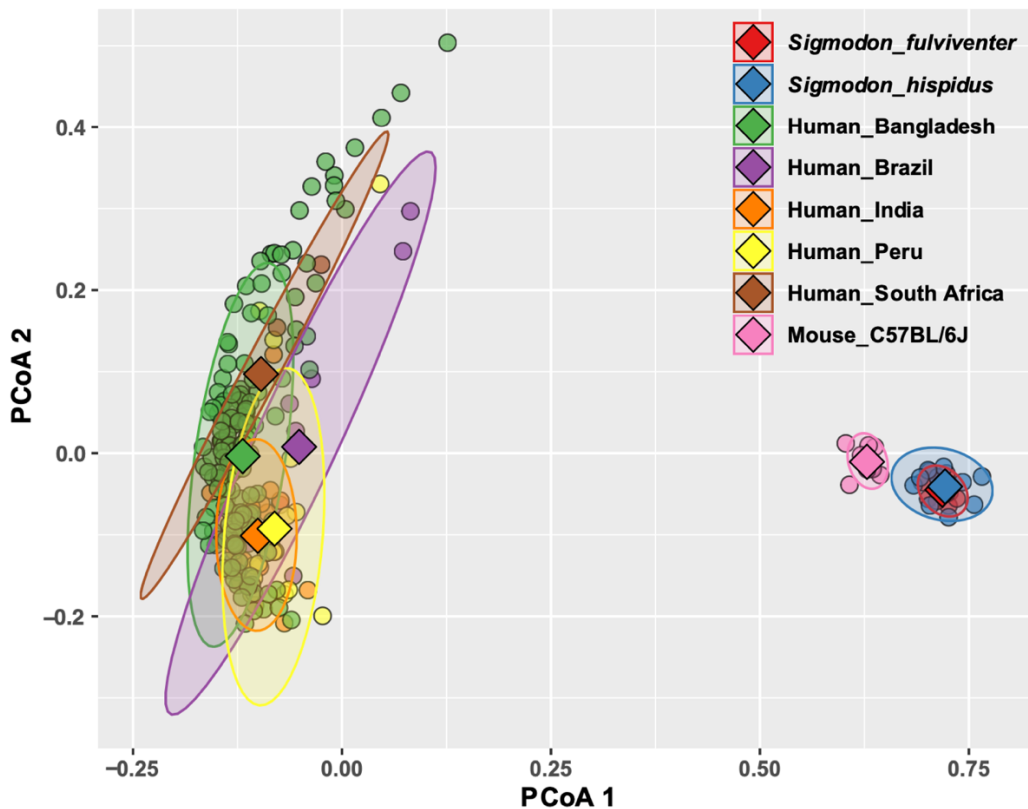
Figure 2.2. Ordination of human, mouse, and two *Sigmodon* cotton rat species (Bray-Curtis dissimilarities, OTU level) reveals that the cotton rat fecal microbiome is very distinct from that of humans and more similar to, but still distinct from, mice.

The cotton rat gut microbiome was stable in both cotton rat species over time.

Richness and alpha diversity did not significantly change over time, no taxa were significantly differentially abundant when the experimental day was set as the outcome variable, and beta diversity testing revealed no significant shifts in microbiome composition over time (data not shown). Subsequently, data were analyzed by computing mean counts for individual cotton rats across all four time points. Comparison of individual body site microbiomes of both *S. fulviventer* and *S. hispidus* across all time points showed community distinctions between species. All sites (ear, nose, skin, feces) from *S. fulviventer* consistently had higher richness (Observed OTUs, S.chao1 index) and alpha diversity (Shannon Index, Simpson's Index) when compared to *S. hispidus* (**Figure 2.3**; all values p<0.05). Beta diversity, computed by calculating Bray-Curtis dissimilarity between samples at the OTU level, showed unique composition between the fecal communities of *S. fulviventer* and *S. hispidus* (**Figure 2.4A**; PERMANOVA p=0.00025, PERMDISP p=0.00025; **Figure 2.4B**). However, comparison of beta diversity metrics of individual external sites from *S. fulviventer* and *S. hispidus* did not show significant differences (**Figure 2.4 C-E**).
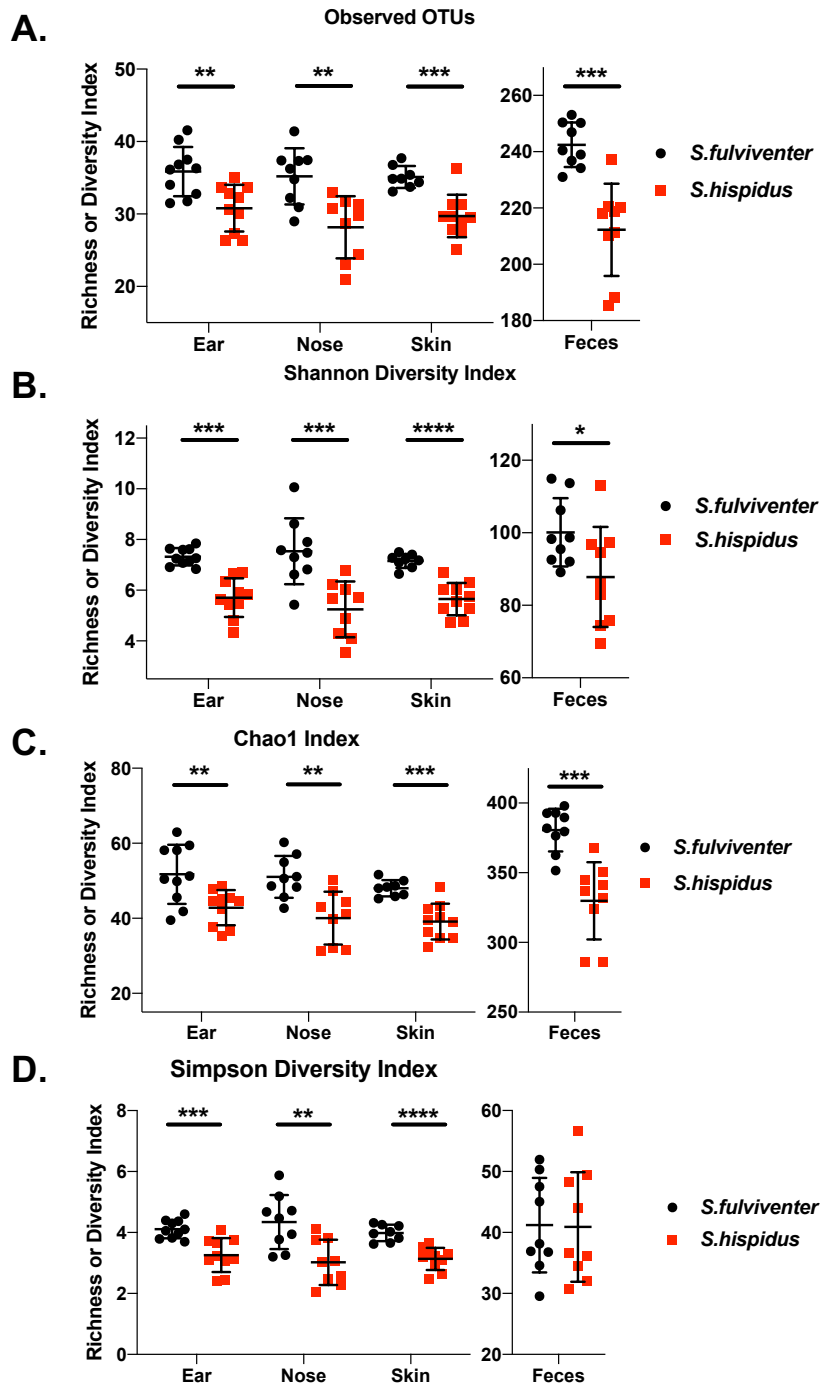
Figure 2.3. **Individual site alpha diversity.** (A) Site richness (Obs. OTUs) and alpha diversity measured by (B) Shannon index, (C) Chao1 Index, and (D) Simpson Diversity Index at the OTU level indicate that *S. fulviventer* is richer and more diverse across all sites. Statistical testing between each cotton rat species was performed using a Student's T-test. Statistical testing across body sites is not shown (not significant across external sites). Feces were plotted separately to account for discrepancies between Y axes.
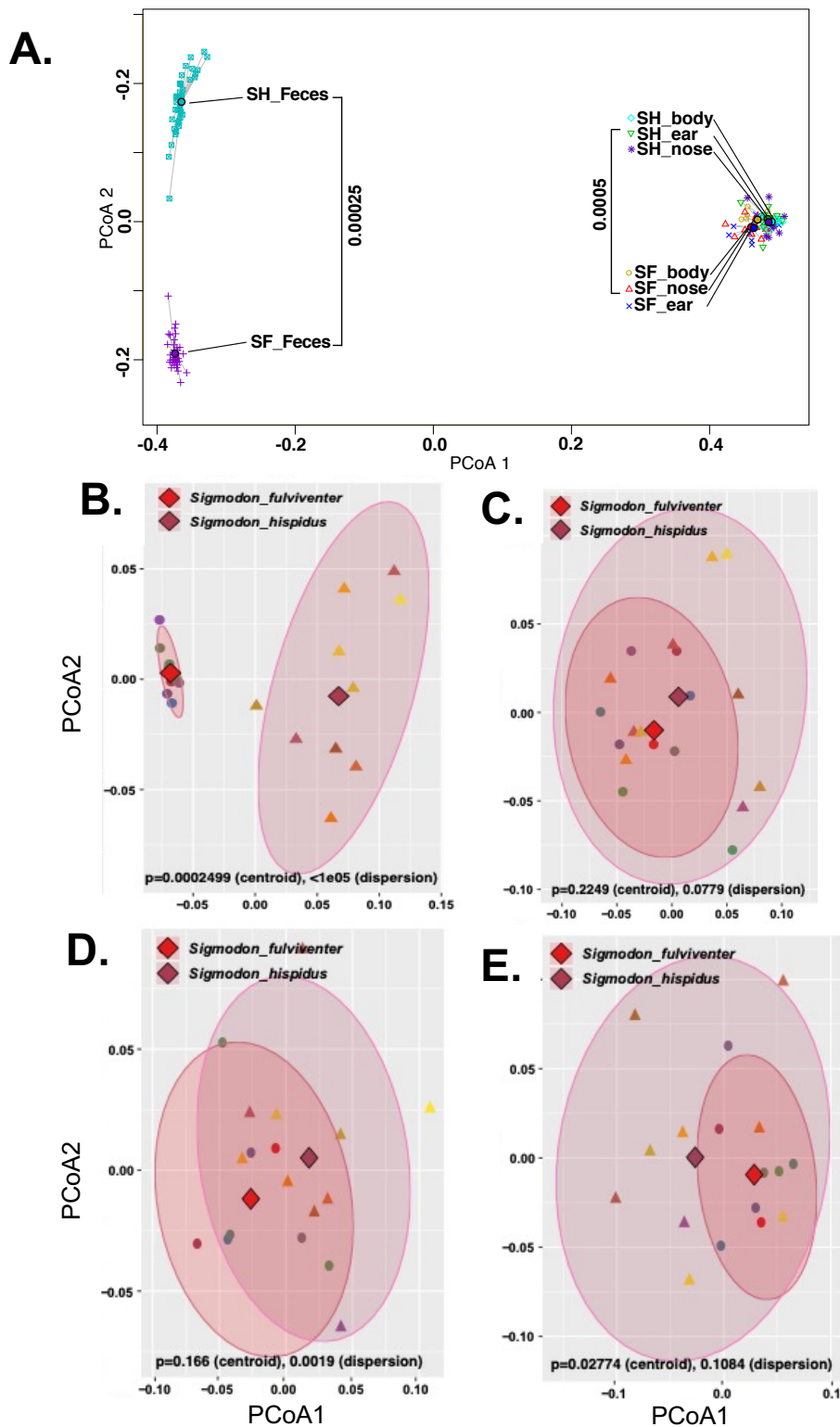
Figure 2.4. Individual site beta diversity. (A) Clustering of samples (Bray-Curtis dissimilarity) reveals distinct microbiota compositions at each site (feces and swabs from the body, ear, and nose), which are also affected by host species (*SF=S. fulviventer*, *SH=S. hispidus*). Statistical testing was performed using PERMANOVA between species. (B-C) Difference in body site beta diversity between *S. hispidus* and *S. fulviventer*. Clustering of samples (Bray-Curtis dissimilarity) shows separation by host species (B) Feces, (C) Ear, (D) Skin, and (E) Nose.

<u>Fecal microbiome:</u>

Analysis using the DESeq2 (229) package identified several bacterial genera that were differentially abundant in the two cotton rat species. These differences were most apparent in the gut (**Figure 2.5A**). *S. hispidus* had a higher abundance of 18 unique genera in the gut (q<0.05), including *Lactobacillus* ($\log_2$ fold change=3.12, q=3.27E-13), *Helicobacter* ($\log_2$ fold change=2.45, q=2.33E-33), *Anaerostipes* ($\log_2$ fold change=2.35, q=0.029), and *Bifidobacterium* ($\log_2$ fold change=1.99, q=1.03E-06). *Escherichia/Shigella* was more abundant in the *S. hispidus* gut ($\log_2$ fold change=7.30, q=3.79E-11) but had very low relative abundance compared to other genera. *S. fulviventer* had a higher abundance of 18 unique genera in the gut (q<0.05), including *Clostridium sensu stricto* ($\log_2$ fold change=-7.19, q=7.77E-12), *Elusimicrobium* ($\log_2$ fold change=-6.22, q=8.04E-18), and *Hespellia* ($\log_2$ fold change=-5.11, q=2.21E-04) (**Figure 2.5A**). To ensure that no observed differentially abundant taxa were false-positive observations due to low abundances, the conservative LEfSe test was utilized for differential taxa (230), which reported 38 genera and confirmed 35 of 36 DESeq2-calculated differentially abundant genera (except *Clostridia_unclassified*). A stability selection model showed *Lactobacillus* as one of the top genera (as well as those unclassified within the phyla *Bacteroidetes*) with a high probability of predicting whether a fecal sample was from *S. hispidus* or *S. fulviventer* (**Figure 2.5B**). While *Lactobacillus* was one of the top 20 most abundant bacteria of the skin, ear, and nose microbiomes of both *S. fulviventer* and *S. hispidus* (**Figure 2.1 B-C**), it was not significantly differentially abundant between the two species at any other sites except feces (**Figure 2.5C-F**).
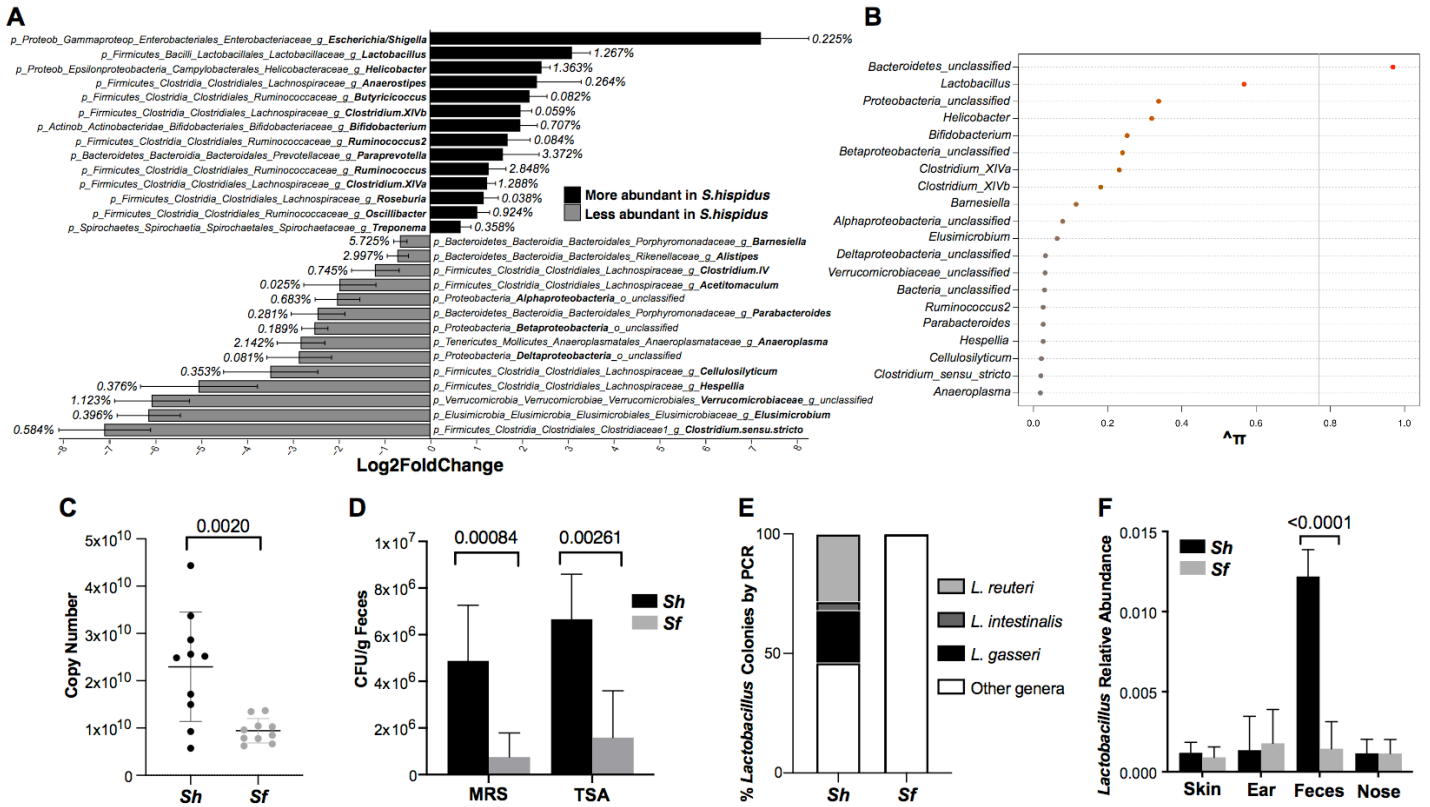
Figure 2.5. **Differential abundance of gut taxa using 16S rRNA and traditional microbiological techniques.** (A) Differential abundance of gut microbial taxa between *S. hispidus* vs. *S. fulviventer* that displayed significant differences (p < 0.05, q < 0.05, l2fc > ±0.65) between host species. Log2FoldChange is plotted along the x-axis, with genera ranked highest in *S. hispidus* (black) to highest in *S. fulviventer* (grey) on the y-axis. Error bars represent l2fc SE; relative abundances from either *S. hispidus* or *S. fulviventer are* denoted next to the corresponding bar. (B) Probability of a gut bacterial genus being selected into a stability selection model distinguishing cotton rat species. The probability is plotted along the x-axis, with the top 20 ranked genera along the y-axis. (C) Bacterial load depicted as copy number/uL of extracted DNA from normalized cotton rat stool. Data generated by qPCR; statistics performed using unpaired T-test. (D) Amount of aerobic colony forming units (CFU) per gram of feces on both *Lactobacillus*-selective (MRS) and general growth (TSA) media. *S. hispidus* displayed a higher amount of aerobic growth using both methods. Significance calculated by student's T-test. (E) Percentage of CFUs with positive detection of *Lactobacillus* amplicons determined by PCR with primers targeting *Lactobacillus* 16S rRNA region. Species-specific identity of colonies was confirmed by shotgun sequencing. "Other genera" include *Bacillus*, *Enterococcus*, and *Corynebacterium*. (F) Relative abundance of *Lactobacillus* between cotton rat body sites. Significance calculated by student's T-test. *Lactobacillus* was one of the higher-abundant taxa with differential abundance between cotton rat species.

Nasal microbiome:

DESeq2 testing revealed that the *S. hispidus* nose had a higher abundance of *Enterobacteriaceae* (log$_2$ fold change=1.13, q=1.39E-04) and *Corynebacteriaceae* (log$_2$ fold change=0.445, q=0.00589), while the *S. fulviventer* nose had a higher abundance of *Leuconostocaceae* (log$_2$ fold change=-1.56, q=0.064). LEfSE also revealed that *S. fulviventer* nose had higher abundance of *Facklamia* (LDA=3.343, p=0.0287), *Bifidobacterium* (LDA=3.259, p=0.0343), *Turcibacter* (LDA=3.498, p= 0.0161), *Streptococcus* (LDA=4.023, p=0.00389), and *Actinobacillus* (LDA=3.969, p=0.00285).

Ear Microbiome:

By DESeq2 testing, the *S. hispidus* ear had a higher abundance of *Enterobacteriaceae* (log$_2$ fold change=0.48, q=0.022, confirmed by LEfSe), *Corynebacteriaceae* (log$_2$ fold change=0.571, q=0.0470), and *Pseudomonas* (log$_2$ fold change=1.33, q=0.070, confirmed by LEfSe). DESeq2 testing showed that the *S. fulviventer* ear had a higher abundance of *Leptotrichiaceae* (log$_2$ fold change=0.571, q=0.0470), *Barnesiella* (log$_2$ fold change=-4.58, q=0.066), and *Porphyromonadaceae* (log$_2$ fold change=-2.11, q=0.063). LEfSE confirmed these 3 taxa, as well as higher abundance of *Sphingobacterium* (LDA=3.299, p= 0.00145), *Streptococcus* (LDA= 3.907, p= 0.00426), and *Actinobacillus* (LDA= 3.830, p= 0.00145) in *S. fulviventer*

Skin Microbiome:

Enterobacteriaceae was more abundant on the *S. hispidus* skin (DESeq2 log$_2$ fold change=0.47, q=0.0031, confirmed by LEfSe). LEfSe also found more abundant

Streptococcus (LDA= 3.956, p=0.00769), *Lactococcus* (LDA= 3.494, p=0.0456), *Actinobacillus* (LDA= 3.730, p= 0.0129), and *Mycoplasma* (LDA= 4.101, p= 0.0330) on the *S. fulviventer* skin.

iii. Confirmation of 16S rRNA gene sequencing data using traditional culture methods.

For quantitative comparison of bacterial load between cotton rat species, qPCR analysis of total bacterial DNA extracted from homogenized stool (equal weight/volume) found that the bacterial load was significantly higher in *S. hispidus* than *S. fulviventer* (**Figure 2.5C**). Variance of the bacterial load was different between the two species: while all *S. fulviventer* generally had a low bacterial load, the bacterial load had a large range in *S. hispidus*. An aliquot of normalized, homogenized stool was also plated on *Lactobacillus*-specific agar (De Man, Rogosa, and Sharpe agar); the number of colony-forming units (CFU) per gram in *S. hispidus* stool was significantly higher than in *S. fulviventer* stool (**Figure 2.5D**). 86% of colonies picked from *S. hispidus* stool were *Lactobacillus*-positive, compared to zero *Lactobacillus*-positive colonies grown from *S. fulviventer* stool (**Figure 2.5E**). Sequencing of colonies showed *Lactobacillus gasseri* and *Lactobacillus. reuteri* to be the two prominent bacterial species found in *S. hispidus* stool (**Figure 2.5E**). This significant trend supports the relative abundance of *Lactobacillus* as determined by 16S rRNA gene sequencing, where *Lactobacillus* was significantly more abundant in *S. hispidus* than *S. fulviventer* (**Figure 2.5F**). *Corynebacterium* and *Bacteroides* species were also identified in *S. hispidus* stool samples. Sanger sequencing of isolates from *S. fulviventer* stool identified the presence of *Enterococcus gallinarum*

and *E. casselifavus*.

iv. Differences in the microbiome community structure and composition based on sex

A secondary analysis was conducted to assess if there were any associations between host sex and microbiome community structure and composition. This cohort included 13 *S. fulviventer* cotton rats (10 males, 3 females) and 15 *S. hispidus* cotton rats (5 males, 9 females). Animals in both groups were 4-6 weeks old and were observed for 28 days, with fecal samples collected on days 0, 7, 13, 21, and 28 and nose, ear, and skin swabs collected on days 7 and 28. 16S rRNA gene sequencing was performed to examine any effect of host sex. Alpha diversity metrics indicated significant differences in richness (Observed OTUs, Chao1) and diversity (Shannon Index, Simpson Index) between male and female *S. fulviventer* at both the ear and fecal microbiomes (**Figure 2.6 A-D**). Still, there were no significant differences in richness and diversity of the microbiomes of male and female cotton rats in the nose and skin for both *S. fulviventer* and *S. hispidus* (except *S. hispidus* skin diversity, **Figure 2.6D**). Overall, differences between host sex were most pronounced in the gut compared to external sites (**Figure 2.6**) but only in *S. fulviventer*. Beta-diversity measurements of each species revealed that microbial composition of the gut was significantly dissimilar between male and female cotton rats for both *S. fulviventer* and *S. hispidus* (**Figure 2.7 A-B**; *S. hispidus* PERMANOVA p=0.00025, PERMDISP p=0.1116; *S. fulviventer* PERMANOVA p=0.00025, PERMDISP p=7E-04). There were also notable differences between males and females in *S. hispidus* skin and nose (**Figure 2.7 E, G**). Differential abundance

analysis using DESeq2 was conducted between males and females at each site (**Table 2.2**). While the fecal community structure differed, there were only 3 differentially abundant genera due to sex in the *S. hispidus* gut and no different genera in *S. fulviventer*. There were differential taxa between sexes at external sites of both *S. hispidus* (21 genera) and *S. fulviventer* (13 genera).
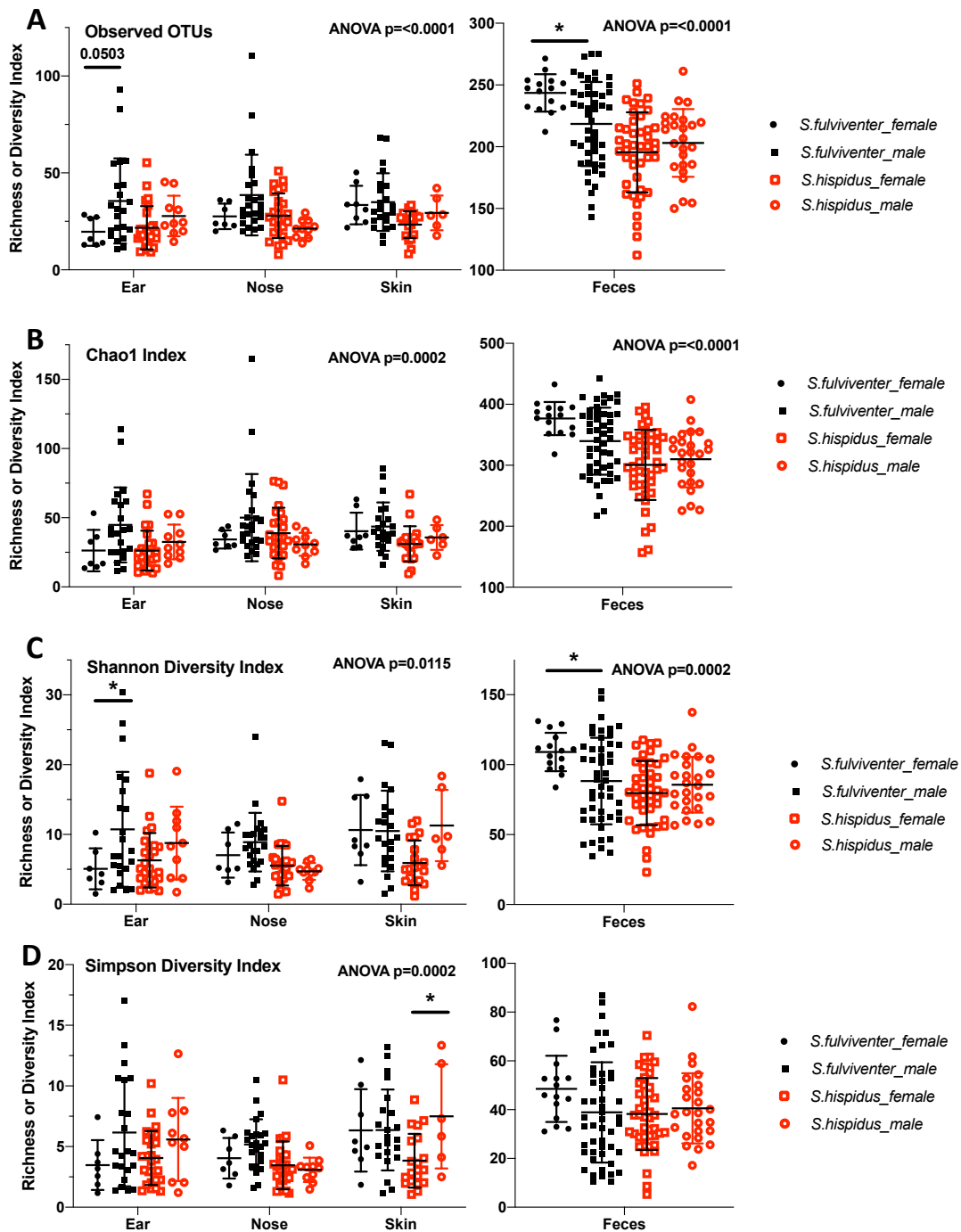
Figure 2.6. Alpha diversity metrics of ear, nose, skin, and feces between male and female _S. fulviventer_ and _S. hispidus_. Richness and diversity were determined using the following methods: (A) Observed OTUs, (B) Chao1 index, (C) Shannon Diversity Index, and (D) Simpson Diversity Index. Statistical testing between sex of each cotton rat species was performed using a one-way (feces) and two-way (external sites) ANOVA, followed by a Tukey's posthoc test. Statistical testing across species is not shown. Feces were plotted separately to account for discrepancies between Y axes.
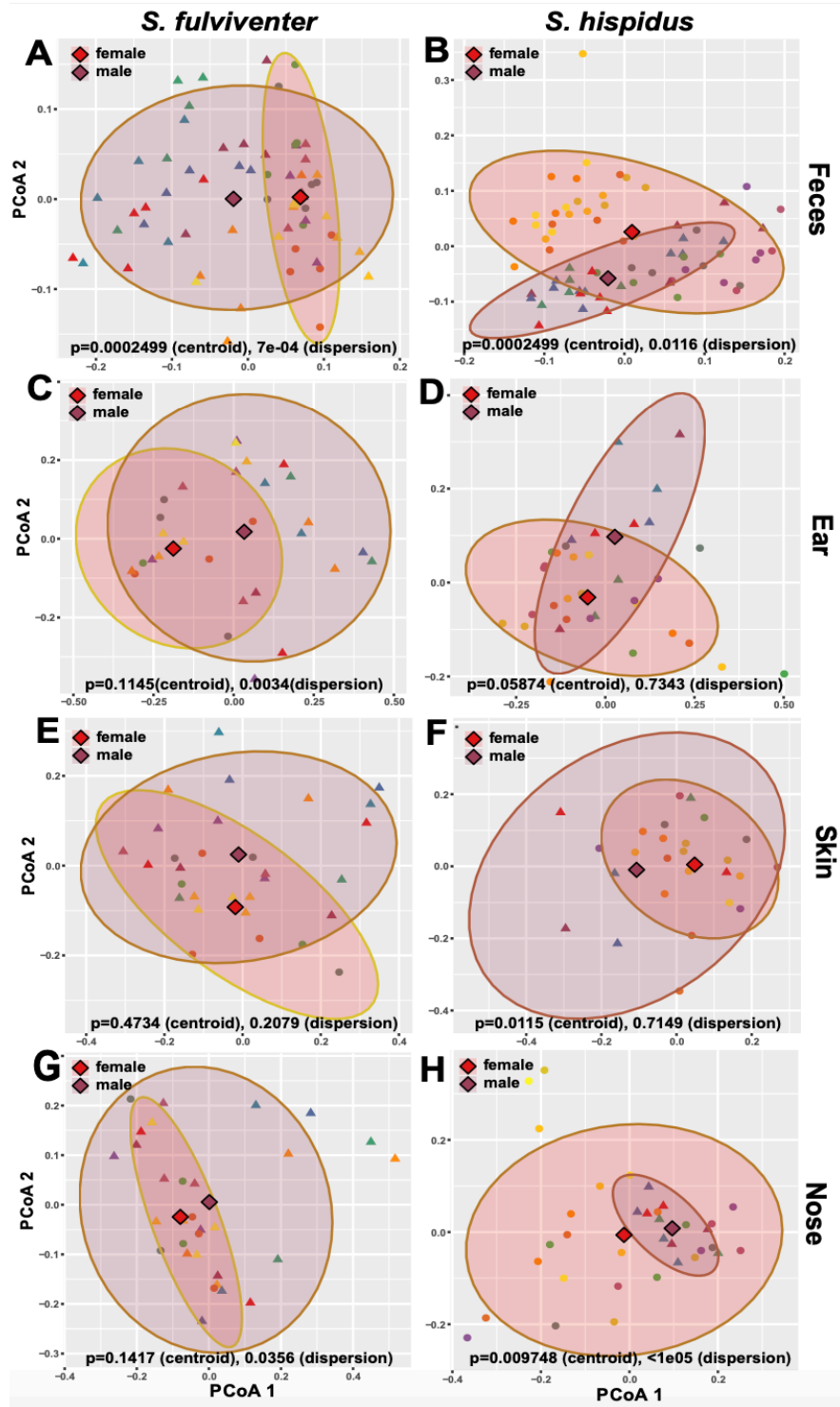
Figure 2.7. Clustering of site- and species-specific samples (Bray Curtis, OTU level) revealed host sex-dependent communities at most sites. Statistical testing was performed using PERMANOVA to compare geometric mean (or centroid); PERMDISP was used to calculate significant differences in dispersion or variance. (A-B) Gut communities showed significant differences between both *S. fulviventer* and *S. hispidus* males and females. (C-H) External sites (ear, skin, and nose) showed sex-based community trends based on both sample mean distances and dispersion. Longitudinal samples from the same cotton rat are represented by matching point colors.

Species/ SampleType	Bacteria (genus)	BaseMean	log2FoldChange (femaleVmale)	pvalue	qvalue
S. fulviventer Nose	*Anaerostipes*	27.559	19.968	3.53E-08	2.22E-06
	Kineococcus	17.698	19.351	9.19E-08	2.31E-06
	Comamonadaceae_unc	94.021	7.820	1.65E-03	1.73E-02
	Prevotellaceae_unc	87.015	7.711	6.92E-04	8.72E-03
	Ruminococcaceae_unc	828.485	5.317	5.00E-04	7.87E-03
	Lachnospiraceae_unc	553.467	3.228	3.90E-03	3.51E-02
	Staphylococcus	6130.222	-4.815	1.10E-07	2.31E-06
S. fulviventer Ear	*Rhizobium*	4.744	18.286	4.63E-07	2.50E-05
	Allobaculum	47.850	7.700	1.57E-04	4.23E-03
S. fulviventer Skin	*Alistipes*	74.458	22.537	1.19E-12	8.12E-11
	Hymenobacter	59.414	22.219	8.50E-10	2.89E-08
	Ruminococcaceae_unc	401.669	5.385	1.12E-03	1.90E-02
	Mucilaginibacter	50.502	-10.252	8.79E-04	1.90E-02
S. hispidus Feces	*Erysipelotrichaceae_unc*	142.661	2.677	5.71E-06	3.20E-04
	Allobaculum	5401.554	1.775	1.34E-03	2.50E-02
	Clostridium_IV	1171.401	-3.571	1.20E-03	2.50E-02
S. hispidus Nose	*Flavobacteriaceae_unc*	2853.705	2.635	4.61E-05	7.68E-04
	Betaproteobacteria_unc	63.140	2.510	4.99E-03	3.12E-02
	Neisseriaceae_unc	3774.233	2.246	1.18E-04	1.47E-03
	Mycoplasma	14945.290	1.952	2.95E-04	2.95E-03
	Streptococcus	8346.633	1.462	6.93E-03	3.85E-02
	Cellulosilyticum	72.399	-3.547	8.66E-03	4.33E-02
	Corynebacteriaceae_unc	1515.694	-4.873	3.70E-06	9.24E-05
	Alistipes	44.780	-7.300	1.78E-03	1.27E-02
	Odoribacter	39.825	-7.715	1.16E-03	9.68E-03
	Leuconostoc	34.683	-22.369	1.23E-17	6.13E-16
S. hispidus Ear	*Planococcaceae_unc*	46.965	24.491	3.42E-19	1.78E-17
	Burkholderia	76.137	6.968	1.74E-04	2.26E-03
	Ruminococcaceae_unc	36.895	-8.112	8.54E-05	1.48E-03
	Clostridium_sensu_stricto	131.715	-8.288	5.45E-06	1.42E-04
S. hispidus Skin	*Lactococcus*	73.115	5.542	1.71E-03	7.83E-03
	Turicibacter	180.005	-4.568	1.63E-03	7.83E-03
	Sphingobacterium	5341.196	-4.646	1.33E-04	1.07E-03
	Enterobacteriaceae_unc	10677.297	-5.124	5.72E-06	9.15E-05
	Flavobacteriaceae_unc	261.934	-5.649	3.78E-05	4.03E-04
	Bacteroidetes_unc	39.244	-6.347	1.18E-03	7.57E-03
	Clostridium_IV	91.701	-24.197	6.27E-15	2.01E-13

Table 2.2. Differential abundance analysis of taxa between individual body sites across female and male *S. hispidus* and *S. fulviventer*. Positive Log2FoldChange = higher in females; negative Log2FoldChange higher in males. There were no significant taxa in *S. fulviventer* feces.

Due to the dramatic differences between the *S. fulviventer* and *S. hispidus* gut microbiomes detected by 16S rRNA gene sequencing, further analysis was pursued to understand community differences at the species and strain level as well as differences in microbiome functional potential. DNA extracted from 10 male cotton rat stool samples (*S. hispidus*=5, *S. fulviventer*=5) on both days 34 and 111 (20 samples total) from the first experimental group were processed for shotgun metagenomic sequencing, which generated 1.11×10^9 reads (5.57×10^7 average reads per sample), comprising of 334,037 megabases (16,701 average megabases per sample) and 17.2% duplicate reads. Taxonomic classification was performed by MetaPhlAn2.

Whole metagenomic sequencing data showed differences at the species level that validated the 16S rRNA gene sequencing data. Abundances of several bacterial species were found to be statistically different ($q < 0.05$) between cotton rat species based on differential abundance analysis by both hierarchical clustering (based on Bray-Curtis dissimilarity) of the top 25 most abundant species and DESeq2 (**Figure 2.8**). One sample from each *S. fulviventer* and *S. hispidus* were removed from differential expression analysis due to reduced sequence quality and outlier testing by hierarchical clustering. *Lactobacillus reuteri*, *L. gasseri*, and the novel *L.* sp. ASF360 predominated the gut of *S. hipsidus* (**Figure 2.8A**), and many other *Lactobacillus* species were significantly more abundant in *S. hispidus* samples than *S. fulviventer* (**Figure 2.8B**). Total *Lactobacillus* within the *S. fulviventer* gut was significantly less abundant but included species unique to *S. fulviventer*, including *L. murinus*, *L.* BHWM-4, and *L. animalis*. *Akkermansia*

muciniphilia was significantly more abundant in *S. fulviventer* compared to *S. hispidus*. *Ruminococcus torques, Helicobacter cinaedi,* and *Oscillibacter* spp. were of higher abundance in the *S. hispidus* gut. *Parabacteroides* spp. (including *P. johnsonii*) and *Odoribacter* were more abundant in the *S. fulviventer* gut.
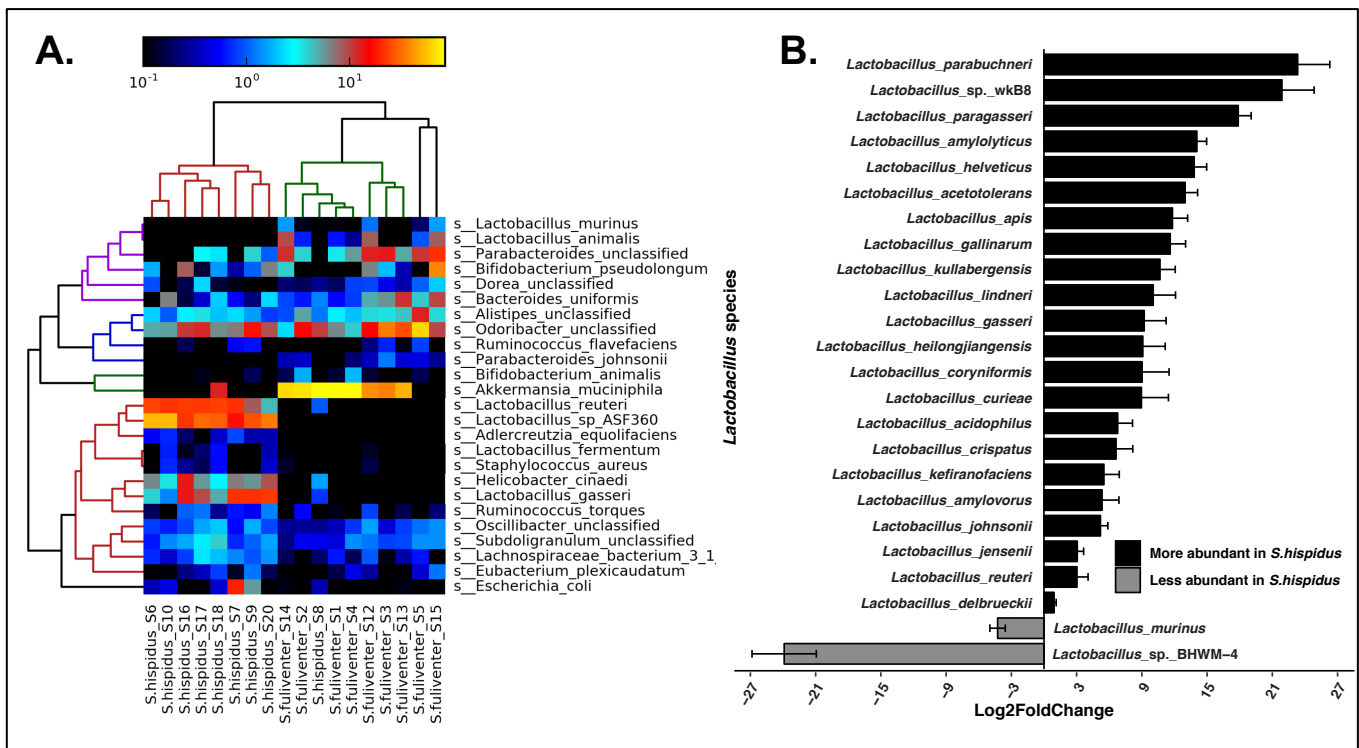
Figure 2.8. Differential abundance of cotton rat gut taxa and corresponding pathways using whole-genome sequencing. (A) Hierarchical clustering (Bray-Curtis) of the top 25 differentially abundant bacterial species between *S. fulviventer* vs. *S. hispidus*. *Lactobacillus reuteri* and *L. gasseri* were drastically higher in *S. hispidus* stool, while *Akkermansia muciniphila* was higher in *S. fulviventer* stool. (B) DESeq2 results representing the differential abundance of *Lactobacillus* species and strains between *S. hispidus* and *S. fulviventer*. Data generated from whole metagenome sequencing.

To understand the biological implications of these differences, HUMAnN2 (231) was used to map any functional differences [MetaCyc pathway database (232)] defined by identified gene families and bacterial profiles. 418 pathways (with nearly all associated with bacteria present in the sample) were identified in the two cotton rat species based on the MetaCyc database, with all 418 pathways represented in *S. hispidus* but only 334 pathways represented in *S. fulviventer* (**Figure 2.9A**). Most of these pathways included biosynthesis (39.60%) and degradation/utilization/assimilation (18.79%) pathways, as well as several overarching superpathways (27.18%) and energy/metabolite production pathways (10.57%). More specifically, these pathways were part of several instrumental superclass ontologies that metabolize (including *de novo* pathways) electron carriers, vitamins, fatty acids, lipids, amino acids, carbohydrates, secondary metabolites, and fermentation-derived energy (**Figure 2.9B**). Interestingly, several pathways were differentially abundant between cotton rat species. Each cotton rat species had unique pathways contributed to by their microbiomes (*S. fulviventer*=14 and *S. hispidus*=27, p<0.05), most of which involved biosynthesis pathways.

In relation to differentially abundant bacteria species, 44 pathways were solely driven by *Lactobacillus gasseri, L. reuteri*, and L. *ASF360* by matching reads from MetaPhlAn2 bacterial identifications with HUMANn2 predicted pathways. Several of these pathways were more highly expressed in *S. hispidus* (**Figure 2.10**). These included L-proline biosynthesis from arginine (catalyzed by bacterial enzymes), inosine-5'-phosphate biosynthesis (for *de novo* synthesis of purines), pyruvate fermentation to

acetate/lactate (for anaerobic energy production), adenosine deoxyribonuclease *de novo* biosynthesis (to promote ADP production), and D-galactose degradation (breakdown of D-galactose to a useable form in glycolysis). *Akkermansia municiphilia* was the driver of 25 other pathways, many of which were highly expressed in *S. fulviventer* compared to *S. hispidus* (**Figure 2.11**). These included L-isoleucine biosynthesis (for production of leucine and isoleucine), phosphopantothenate biosynthesis (to produce vitamin B5 *de novo*, of which animals cannot produce, and to feed production of coenzyme A and acyl carrier protein), glycolysis (particularly the degradation of starches for reductants and energy for anabolic pathways), and L-valine biosynthesis.
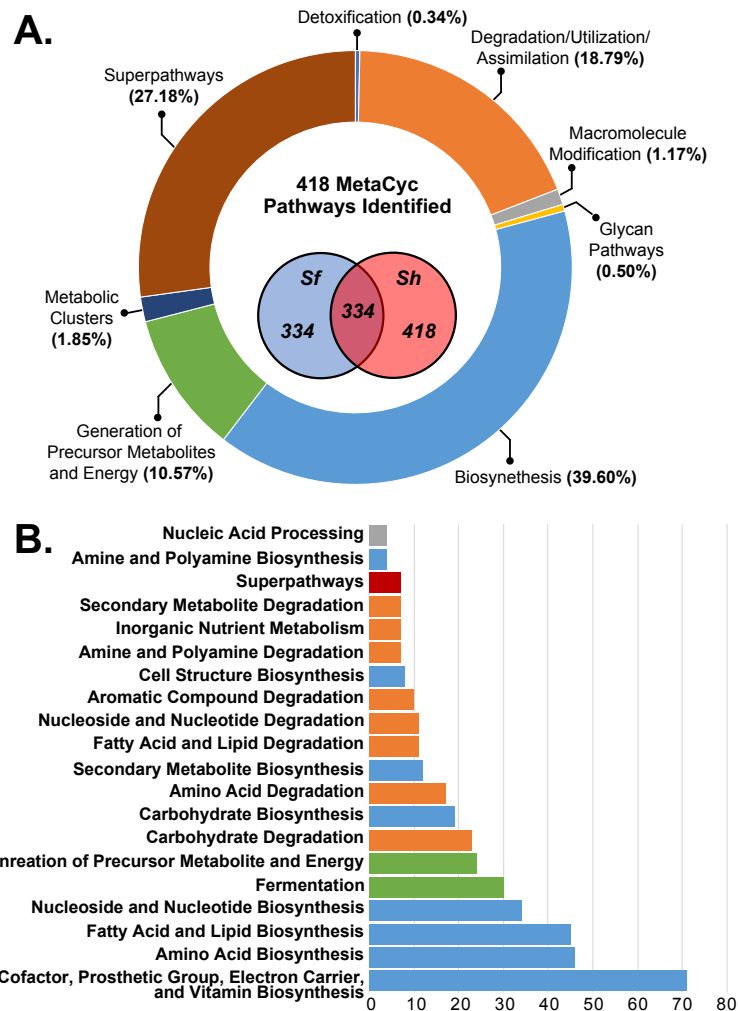
Figure 2.9. The functions of the cotton rat gut microbiome. (A) Distribution of MetaCyc metabolomic pathways predicted from bacterial sequences. 415 unique pathways were identified, and most pathways were classified within the Biosynthesis, Superpathway, and Degradation/Utilization/Assimilation pathway superclasses. (B) Distribution of pathway ontology. Of all the identified pathways, the largest group consisted of Cofactor, Prosthetic Group, Electron Carrier, and Vitamin Biosynthesis, which are often components of host biology sourced solely by commensal bacteria.

S. hispidus
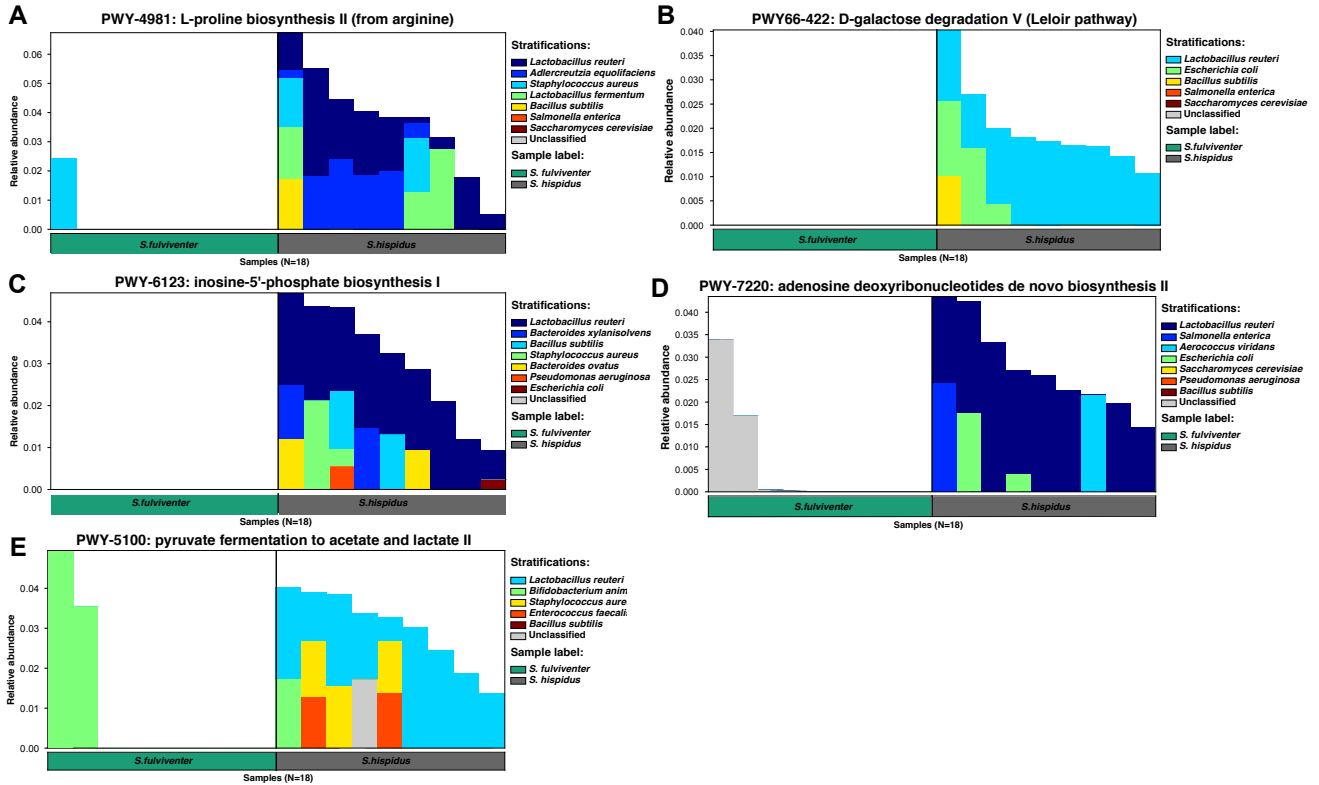
Figure 2.10. Several pathways are more active in *S. hispidus* gut and are contributed significantly to by *Lactobacillus.* (A) Catalysis of proline biosynthesis by bacterial enzymes (PWY-4981). (B) Catalysis of the conversion of D-galactose to D-glucopyranose 6-phosphate, the more metabolically versatile substrate that can feed directly into glycolysis, by enzymes of the Leloir pathway (PWY66-422). (C) *De novo* biosynthesis of purine nucleotide IMP (PWY-6123). (D) *De novo* biosynthesis of dATP for feeding of ATP generating pathways (PWY-7220). (E) Catalysis of the fermentation of the monosaccharide pyruvate to generate the energy source acetate and lactate (PWY-5100).

S. fulviventer
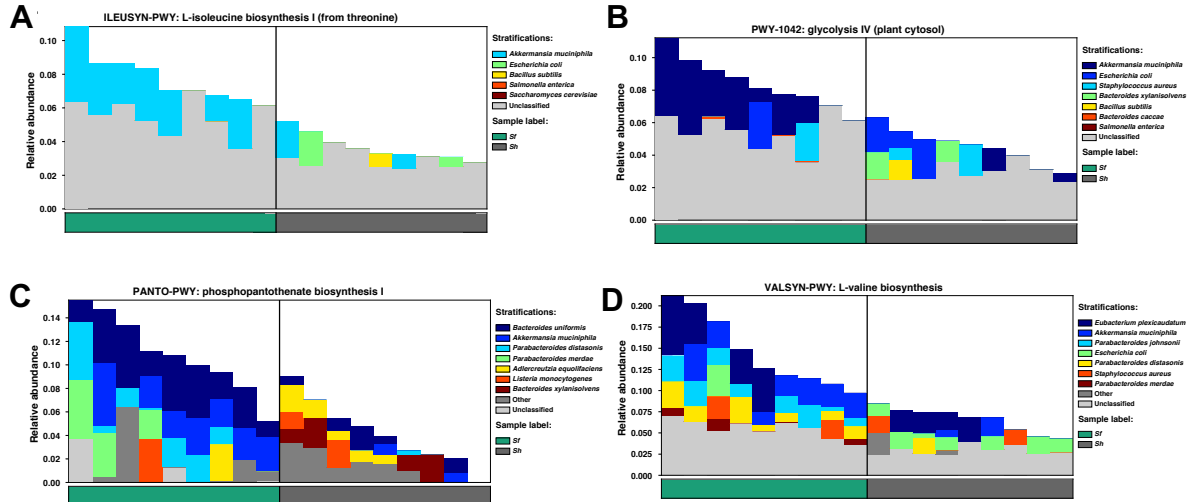
Figure 2.11. Several pathways that are more active in *S. fulviventer* than *S. hispidus* and are greatly contributed to by *Akkermansia* species. (A) A member of the superpathway of branched-chain amino acid biosynthesis that generates not only isoleucine but also leucine and valine (ILEUSYN-PWY). (B) Degradation of starch for the generation of carbon skeletons, reductants, and ATP for anabolic bacterial fatty acid pathway initiation via pyruvate decarboxylation to acetyl CoA (PWY-1042). (C) Biosynthesis of R-4'-phosphopantothenate, the universal precursor for synthesizing coenzyme A and acyl carrier protein. Only plants and microorganisms can synthesize pantothenate de novo; animals require a dietary supplement. Synonymous with Vitamin B5 synthesis. (PANTO-PWY). (D) A member of the superpathway of branched-chain amino acid biosynthesis that generates not only valine but also leucine and isoleucine.

and *Akkermansia* species.

This chapter presents the comprehensive characterization of the cotton rat microbiome and comparison of bacterial communities of two species: *S. hispidus* and *S. fulviventer*. These analyses uncovered species-specific differences in gut bacteria, even though the two species had the same diet over many generations and were housed in separate cages in the same facility and room. 16S analysis revealed that all external sites (skin, ear, nose) shared similar community structures within each species (**Figure 2.4**), but *S. fulviventer* had higher alpha diversity and richness than *S. hispidus* (**Figure 2.3**). There were no differentially abundant genera between species at external sites. *S. fulviventer* and *S. hispidus* had unique fecal microbiomes composed of several significantly differentially abundant genera (**Figure 2.5**). *Escherichia/Shigella, Lactobacillus*, *Helicobacter,* and *Anaerostipes* were higher in *S. hispidus*, whereas *Clostridium_sensu_stricto, Elusimicrobium, Verrucomicrobiaceae*_unclassified, and *Hespellia,* were higher in *S. fulviventer*. There were also significant differences in metabolic pathway functions due to differential host-microbiome contributions, including biosynthesis fermentation and degradation pathways. For sex-based comparisons, the microbiome composition of external sites showed significant differences between male and female animals, while fecal communities were largely similar, as measured by beta diversity. This data further supports that, while environmental factors play a vital role in

shaping microbiome structure and composition, underlying host genetics exerts homeostatic pressure for distinct microbiomes between populations. This study is also the first analysis of the cotton rat microbiome in two different inbred species.

The most recent phylogenetic analysis of the *Sigmodon sp.* found that *S. hispidus* and *S. fulviventer* diverged 5.4 million years ago (179). In the wild, *S. hispidus* and *S. fulviventer* are sympatric species, with *S. fulviventer* being the more dominant animal (180). Separate inbreeding of the two species has made them a useful small animal model in laboratory research (233, 234). This data shows that each species maintains a unique gut microbiome community structure and composition even upon inbreeding and adaptation to a controlled laboratory environment.

While other small animal models, including mice and rats, have been used to explore the relationship between genetic patterns and bacterial homeostasis (235-238), the cotton rat could be used to study the interplay between differing host microbiomes and host immune systems responses to viral infections. For example, *S. hispidus* had a significantly higher amount of probiotic gut bacteria genera (*Lactobacillus*, *Bifidobacterium*) that have been associated with protection against the severe outcomes from respiratory syncytial virus, influenza, and human rhinovirus (225, 239, 240). Other differentially abundant bacteria between the two species, such as *Escherichia coli* and *Bacillus cereus*, have also been shown to enhance both poliovirus and reovirus replication and pathogenesis (241). Host responses considering the microbiome could be elucidated with the presence/absence of these bacteria in this animal model. The cotton rat could also be an optimal model for supplementation studies of bacterial taxa in relation to these viruses.

This study also uncovered sex as a key factor impacting gut microbiome composition. In addition, metagenomic sequencing analysis revealed species level differences as well as the metabolic potential of the microbiome. Differentially abundant taxa were directly related to differentially abundant metabolic pathways between cotton rat species. Pathways such as biosynthesis and degradation pathways (cell structures, electron carriers, vitamins, fatty acids, lipids, amino acids, etc.) could be greatly implicated in mucosal reinforcement previously described in microbiome literature (242). This microbiome-metabolic characterization may provide an important resource in understanding mechanisms by which the microbiome protects against certain disease states.

This study has significant strengths compared to the lone study published so far (228), including large sample size, longitudinal sampling, and shotgun metagenomic sequencing to characterize the gut microbiome at the species level. However, we acknowledge several shortcomings:

1. Due to the lack of an assembled cotton rat genome, we could not examine host genes, genetic patterns, or polymorphisms that may be driving differences in microbial colonization.

2. Whole metagenomic sequencing was only performed on the gut microbiome of male cotton rats from the first cohort. While there is no cotton rat genome in the public databases, a *de novo* assembly of the cotton rat transcriptome could further integrate microbiome data and gene expression patterns to uncover more relevant information regarding differences in host and microbiome interactions.

3. Characterizing the microbiome of a unique animal species can be challenging due to host-adapted bacteria not matching the microbiome sequence databases. Many databases and tools that are commonly available are specifically designed for human microbiome analyses and can result in the misclassification of bacteria when used for new animal species. However, until better databases are generated, interpretation of data must be done with caution.

Further research is warranted to understand species-level microbiome differences and their impact on immune response in all small animal models to better interpret preclinical studies of vaccines and anti-microbiological agents. Despite some limitations, this study creates a stepping stone for future research into these pressing questions of host-microbiome interactions during infection.

Animals

Four- to six-week-old cotton rats (~100 g) were obtained from the inbred colony maintained at Sigmovir Biosystems, Inc. (SBI). Cotton rats in the colony were seronegative by ELISA to adventitious respiratory viruses (i.e., Pneumonia Virus of Mice, Rat parvovirus, Rat coronavirus, Sendai virus). Animals were individually housed in large polycarbonate cages and fed a diet of standard rodent chow and water *ad libitum*.

For rigor and reproducibility, two independent animal experiments were carried out to characterize and establish the healthy cotton rat microbiome structure and composition by comparing two species, *S. fulviventer* and S. *hispidus.* In the first experimental group, 20 young male cotton rats were examined: *S. fulviventer* (n=10) and *S. hispidus* (n=10). Each animal was observed for 111 days, with nose, ear, and skin swabs collected on day 95 and fecal samples collected on days 0, 4, 34, and 111. These samples were used for microbiome characterization. To analyze any sex bias in the microbiome, a second experimental group (at a later time) included 13 young *S. fulviventer* (10 males, 3 females) and 16 young *S. hispidus* (5 males, 9 females). Healthy animals were monitored for 28 days, with fecal samples collected on days 0, 7, 13, 21, and 28 and nose, ear, and skin swabs collected on days 7 and 28. To avoid fighting, all the animals were housed individually in large polycarbonate cages (with proper enrichment; nylon bone and glass jar). The cotton rat colony was maintained free of human and rodent viruses. All animal

procedures followed NIH and USDA guidelines and were approved by the Sigmovir Biosystems, Inc. IACUC.

Sample Collection

Feces collection: One day before feces collection, cage beddings were changed in the late afternoon for each animal. Samples were collected between 10 am and 1 pm with sterile forceps. On average, 10-15 feces pellets were collected from each animal. Immediately after collection, samples were frozen at −80°C.

Nose swab: Sterile saline (~100 µl) was pipetted into both nostrils of anesthetized cotton rats positioned face down; Fisherbrand Sterile Swabs (Calcium Alginate Fiber Tipped Ultrafine Aluminum Applicator Swab) were then immediately placed in the nostrils to absorb the saline. Swabs were broken into sterile DNase/RNase-free 1.5 ml tubes and stored at −80°C.

Ear swab: Sterile saline (~100 µl) was pipetted up and down into both ears of each anesthetized cotton rat while the animal was kept in an anesthesia chamber for 1-2 minutes, and residual liquid was absorbed from each ear with Beaver Visitec Ultracell PVA Eye Spears pack of 5 (intended for fluid absorption and tissue manipulation). Swabs were broken into sterile DNase/RNase-free 1.5 ml tubes and stored at −80°C.

Skin swab: Sterile saline (~200 µl) was put at the back of each anesthetized cotton rat (at approximately the same site for each animal) and rubbed vigorously using Fisherbrand Sterile Swabs (Calcium Alginate Fiber Tipped wood applicator swab). Swabs were broken in sterile DNase/RNase-free 1.5 ml tubes and stored at −80°C.

Microbiome DNA extraction and 16S rRNA gene sequencing

Genomic DNA was extracted from all samples at Vanderbilt University Medical Center using the Qiagen DNeasy PowerSoil HTP Kit (96-well plates) following the manufacturer's protocol, except the optional 4°C incubations were skipped. Stool samples were thawed on ice and added directly to the kit plate. Nose, ear, and skin swabs were vortexed in tubes with 800 uL Qiagen PowerBead solution for 5 mins; this PowerBead solution was then added to the kit plate. An extraction negative, which did not contain any template but was otherwise processed the same as the rest of the samples, was included on each extraction plate. To mechanically lyse the cells, plates were shaken at 20 Hz in a TissueLyser II system (Qiagen) for 20 minutes. Steps 16–33 of the kit manufacturer's protocol were performed on a QIAcube HT (Qiagen). One-step PCR targeting the V4 region of the 16S rRNA gene was performed using 515F/806R primers (243). MyTaq HS Mix (Bioline) was used to create amplicons with the following cycling conditions: 95°C for 2 min; 30 cycles of 95°C for 20 sec, 50°C for 15 sec, 72°C for 5 min; 72°C for 10 min; 4°C indefinitely. Positive PCR results were confirmed by the presence of a 400bp band in 1% agarose gel electrophoresis; all negative controls were verified at this step to not have a visible band. The PCR products were cleaned and normalized using the SequalPrep Normalization Kit (Invitrogen). Samples and complementary controls (extraction negative, PCR negative, and ZymoBIOMICS Microbial Community Standard) were pooled and cleaned using 1X AMPure XP beads. Sequencing was done on an Illumina MiSeq platform with 2x250bp reads at the Vanderbilt Technologies for Applied Genomics (VANTAGE) core facility.

16S rRNA Gene Data Processing and Statistical Analysis

After sequencing, reads were processed using the mothur SOP (https://mothur.org/wiki/miseq_sop/) (244). Operational taxonomic units (OTUs) were clustered at 97% sequence identity. Non-bacterial sequences, low-quality sequences (1.5% of total reads), and chimeras, as identified with UCHIME (245) were removed during data processing. Sequences were taxonomically assigned by the Ribosomal Database Project (RDP) database 14 (246) using the SILVA database release 128 (247). Samples with <10000 final reads (n=10) were removed before alpha and beta diversity analysis, and samples with <1000 final reads (n=5) were removed before the remaining analyses to examine specific differentially abundant taxa (i.e., DESeq2, GeneSelector, stability selection, and LEfSE). Statistical analyses were performed using MGSAT [https://bitbucket.org/andreyto/mgsat] (209, 240), which facilitates data analysis by wrapping the R packages described below.

Alpha- and beta-diversity analyses were performed using the R package *vegan* (248). Before alpha- and beta-diversity analysis, counts were rarefied to the smallest library size, and richness, alpha-, and beta-estimates were calculated. This process was repeated 400 times, and the results were averaged. Richness was estimated with the observed OTUs and Chao1 indices; alpha diversity was assessed with the Shannon and Simpson indices, which were converted into their corresponding Hill numbers (249). Where applicable, statistical testing between site alpha diversity was calculated using Mann-Whitney U or Kruskal-Wallace/Dunn's Post Hoc test. Counts were normalized to simple proportions for beta-diversity analysis, and pairwise Bray-Curtis dissimilarities were estimated. The PermANOVA (permutation-based analysis of variance) test as

implemented in the *Adonis* function from the R package *vegan* was used to test for significance between overall microbial composition and groups of interest (i.e., *S. hispidus* compared to *S. fulviventer* and males compared to females) over 4000 permutations; results are indicated by "centroid" p-values. Homogeneity of variance within sample groups was tested using *betadisper* function; results are indicated by "dispersion" p-values. Comparisons between *Sigmodon* cotton rats, human, and mouse fecal microbiome communities were performed using the same methods, and data were downloaded from the NCBI Short Read Archive database (BioProject PRJNA368790, PRJEB27068, and PRJEB27068). All downloaded data were sampled from a single time point and did not represent longitudinal sampling.

Differential abundance of taxa in association with metadata categories was analyzed using DESeq2 (229) . Before DESeq2 analysis, we eliminated all taxa that had an average number of <10 reads, taxa with a minimum quantile mean fraction <0.25, and taxa with a minimum quantile incidence fraction <0.25; taxa with a normalized base mean (generated by DESeq2) <10 were removed. Reported adjusted P values (q) values are the result of a Wald test with the Benjamini and Hochberg correction to adjust for multiple comparisons. To build alternative rankings of taxa regarding their prevalence in one cotton rat species over the other, we also used stabsel and GeneSelector. The stabsel stability selection (250) approach aims to build the relative ranking of the predictor variables (taxa in this case) according to their importance for predicting the outcome. It does so by building multiple "base" models on random subsamples of the data. The elastic net model from the R package glmnet was used as the base feature selection method to be wrapped by the stability protocol. The ranking of taxa and their probability of being

selected into the model were reported, as well as the probability cutoff corresponding to the per-family error rate that is controlled by this method. The GeneSelector package (251) was used as a stability feature ranking method that is based on a nonparametric univariate test. In brief, the same ranking method (package function RankingWilcoxon) was applied to multiple random subsamples of the full set of observations (400 replicates, sampling 50% of observations without replacement). RankingWilcoxon ranks features in each replicate according to the test statistic from Wilcoxon rank-sum test with regard to the outcome group (e.g., *S. hispidus* vs. *S. fulviventer*). Consensus ranking between replicates was then found with a Monte Carlo procedure (package function AggregateMC) and the features were reported in the order of that consensus. The absolute abundance counts were normalized to simple proportions within each observation to account for different sequencing depths. For each feature, we also obtained several types of the effect size, such as common language effect size and rank biserial correlation. LEfSe (Linear discriminant analysis [LDA] Effect Size) was executed using the online Galaxy module (230) to determine taxa most likely to explain differences between classes (species, sex, etc) using feature ranking followed by Kruskal-Wallis and pairwise Wilcoxon tests. These 4 statistical analyses (DESeq2, stabsel, GeneSelector, and LEfSe) allowed for rigorous testing of each particular taxon of interest.

Metagenomic Library Preparation

A subset of fecal samples from 20 total male cotton rats (10 from each species), taken on days 34 and 111 within the first cohort of cotton rats, underwent whole-metagenomic shotgun sequencing. From the same stool samples, genomic DNA was

extracted using the Qiagen DNeasy PowerSoil Kit (Cat No./ID: 12888-100) by following the manufacturer's protocol (skipping the optional 4ºC incubations). In addition, a negative sample (which did not contain any template but was otherwise processed the same as the rest of the samples) and a positive control (ZymoBIOMICS Microbial Community Standard) were processed in parallel with samples and sequenced. Samples were normalized to 75 ng/ µL in 1X TE before library construction. Metagenomic libraries were prepared using the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina® following the manufacturer's protocol for inputs ≤ 100 ng. Samples were fragmented at 37ºC for 12 minutes, yielding a fragment size of 200-450 bp. NEBNext Multiplex Adaptors were diluted 10-fold. NEBNext Multiplex Oligos for Illumina (Set 1, NEB #E7335) were used for PCR enrichment of adaptor-ligated DNA, and 5 cycles of PCR were run. Library quality was assessed on an Agilent 2100 Bioanalyzer System using the Agilent High Sensitivity DNA Kit (5067-4626). Samples were sequenced via the NovaSeq 6000 2x150 platform for Illumina at the Vanderbilt Technologies for Advanced Genomics (VANTAGE) core, aiming for 40 million reads per sample.

Whole Metagenomic Sequence (WMS) Analysis

FastQC (252) followed by MultiQC (253) were used to examine data quality. Trimmomatic (254) was used to remove adaptors and trim low-quality reads using the parameters: TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75. An average of 85% of reads mapped to various host DNA databases, but reads were not filtered before functional classification. Microbial communities were then profiled using MetaPhlAn2 (255) . Differentially abundant bacteria were calculated using MetaPhlAn2's *hclust2.py*

function by hierarchical clustering (based on Bray-Curtis dissimilarity) of the top 25 most abundant species according to the 90th percentile of the abundance in each clade as well as DESeq2. Functional metabolic profiles were analyzed using HUMANn2, which aligns reads from UniRef (256) and clusters abundances to the ChocoPhlAn (231) database. This generates three outputs: UniRef IDs for gene families in reads per million, MetaCyc pathway coverage, and MetaCyc pathway abundances in copies per million. To identify differential pathways between sample groups, associations between cotton rat species were identified by the HUMANn2.associate script and statistical testing using the Kruskal-Wallis H-test. Data presented (generated by HUMANn2.barplot script) is from pathway abundances (normalized as relative abundance) within each sample with unmapped/unintegrated pathways removed and was found statistically significant ($p<0.05$ and $q<0.05$). Superclasses distribution of identified MetaCyc pathways was manually generated using the online MetaCyc database.

Enumeration of Lactobacillus

Two frozen stool pellets were taken from 20 male cotton rats (10 *S. hispidus*, 10 *S. fulviventer*), weighed, and diluted to 45 mg/mL in sterile 1x PBS. Samples were rocked for 20 min on ice and resuspended manually with pipette mixing. 10^{-1}-10^{-3} serial dilutions were plated on *Lactobacilli* MRS agar (BD 288210) and incubated at 37ºC for 48 h. Colonies on 10^{-2} were counted, and 95 colonies were randomly picked from each species and inoculated into 1.2 mL MRS broth (BD 288130) in a sterile 96-deep-well plate. The plate was incubated at 37ºC for 20 h with no shaking. Cultures were gently mixed by

pipetting, and 20% glycerol stocks were prepared for each culture. Colony PCR was performed on each isolate by boiling the culture at 95ºC for 10 minutes, then using 10 µL as the template with *Lactobacillus* species-specific primers (257) and MyTaq HS Red (Bioline®) with the following cycling conditions: 95°C for 2 min; 30 cycles of 95°C for 20 sec, 50°C for 15 sec, 72°C for 1 min; 72°C for 10 min; 4°C indefinitely. PCR reactions were spun at 3900g for 10 min to remove any bacterial debris from the boiled template and run on 1% agarose gel to verify *Lactobacillus*-positive colonies. A new PCR was then repeated using the universal primers Uni331F/Uni797R (258) (following cycling conditions listed above), and purified PCR products were sent for Sanger sequencing. Bacterial isolate identity was determined using NCBI BLAST database.

Determination of bacterial load by qPCR

DNA was extracted from an equal volume of normalized homogenates of cotton rat stool (described in Methods: Enumeration of *Lactobacillus*) using the DNeasy PowerSoil Kit (Qiagen). qPCR reactions were prepared in duplicate using BioRad iQ Supermix with Invitrogen SYBR Green, following the manufacturer's protocol. Universal eubacteria 16S rRNA primers (UniF340, UniR514) (259) equal volumes of extracted DNA, and targeted standards were used to determine copy number per gram of feces. Each qPCR plate included a corresponding extraction negative and a no-template negative control. As previously described, a serial dilution of standards containing known bacterial copy numbers specific to the primer pair was used as a standard curve. PCR reactions were run with a 15 sec 95ºC melting and 1 min 54ºC annealing step for 40 cycles. Cycle

threshold (CT) values were plotted against the standard curve to determine copy number, and figures and statistical testing (unpaired T-test) were generated using Prism v8.

Isolation and culture of Lactobacillus *strains from cotton rats' stool*

Glycerol stocks of identified *Lactobacillus* species were streaked on MRS agar plates, and a single colony was grown in culture using MRS broth. *Each species was incubated at 37ºC without shaking to determine growth parameters*, and growth efficiency was measured by turbidity. A growth curve was also estimated using a BioTek Synergy HTX plate reader at 37ºC for 24 h; OD_{600} was measured every 10 min following a brief 3-second shake to mix culture. CFU counts were also taken during the log phase by plating a 3-fold serial dilution on MRS agar plates.

E. <u>ACKNOWLEDGEMENTS AND AVAILABILITY OF DATA</u>

P30CA68485, P30EY08126, and G20RR030956) and Sigmovir Biosystems Inc's corporate funds.

The sequencing data are deposited at the NCBI Short Read Archive (SRA) under BioProject PRJNA721429. Many raw datasets can be found as supplemental files at https://doi.org/10.1186/s42523-021-00090-8. The full DESeq2 and GeneSelector data at both genus and family levels are shown in **Supplemental Files 1** and **2,** respectively. LEfSe data at the genus level is shown in **Supplemental File 3**. DESeq analysis of 16S rRNA gene sequencing data between male and female cotton rats (both *S. fulviventer* and *S. hispidus*) across all body sites at the genus and family level are shown in **Supplemental File 4**. DESeq2 analysis of whole metagenomic sequencing data between *S. fulviventer* and *S. hispidus* across all body sites at the species level is shown in **Supplemental File 5**. All relative abundance values at all phylogenetic levels are shown in **Supplemental File 6.** Raw counts from 16S rRNA sequencing data at all levels for both cotton rat cohorts are shown in **Supplemental File 7**. All annotated MetaCyc pathways identified by HUMANn2 from whole metagenomic sequencing data are shown in **Supplemental File 8**. Statistical MetaCyc pathway comparison between *S. fulviventer* and *S. hispidus* is shown in **Supplemental File 9**.

CHAPTER 3

*Generation of de novo transcriptome references for two species
of cotton rats and transcriptomic analysis during RSV infection*

Respiratory syncytial virus (RSV) is the leading cause of acute respiratory tract infection (ARI) in children below the age of two years, as well as in immunocompromised individuals and the elderly, resulting in 33 million ARIs, 3.2 million hospital admissions, and nearly 120,000 deaths worldwide each year (28). There is currently no approved RSV vaccine and only one preventative monoclonal antibody (Palivizumab), with use limited to high-risk children due to costs (260, 261). As 93% of RSV LRTI cases and 99% of RSV mortality occurs in developing countries, the need for an effective vaccine and low-cost preventatives is critical (28). The failure of the formalin-inactivated RSV vaccine in the 1960s, which induced enhanced disease in vaccinees upon encounter with the virus, hampered the development of new RSV vaccines for decades (21-23). However, there is a recently renewed effort to develop RSV preventatives, with 14 vaccine candidates and alternative anti-viral strategies against RSV (recombinant antibodies (24), nanobodies (25), small molecule inhibitors and analogs (26)) that are at various stages of development (27). This highlights the critical need for an appropriate pre-clinical model for vaccine and drug development against RSV.

The cotton rat (genus *Sigmodon*) is considered the "gold standard" animal model for RSV infection compared to mice and other animals because it is 100-fold more permissive than the majority of laboratory mice to RSV infection and RSV infects both its upper and lower respiratory tracts similar to humans (161, 262, 263). Cotton rats have

also accurately predicted the efficacy of the only two FDA-approved RSV therapeutics (RespiGam®, Palivizumab®) (171, 174, 221, 222). In addition to RSV, cotton rats have been used to study other human respiratory viruses of significance, i.e., influenza A virus (166, 212), parainfluenza virus (164, 165), measles (168), human metapneumovirus (169), enterovirus D68 (170), and human rhinovirus (213) due to the broad susceptibility and comparable human disease features. Unfortunately, studies comparing transcriptomic changes in cotton rats have been limited due to the lack of publicly available reference genomes for any cotton rat species. Previously, the lung transcriptome was assembled and analyzed upon RSV infection in *S. hispidus* (264). However, that study had only examined gene expression in one tissue type and was limited to only one cotton rat species (*S. hispidus*). As prior studies show, both *S. fulviventer*- and *S. hispidus*-specific differences exist in disease severity to viral pathogen (i.e., parainfluenza virus (178)) and microbiome community structure (181). The main objectives of this study were to develop comprehensive transcriptomes for both species and to compare gene expression changes upon RSV infection.

To this end, total RNA of multiple tissues (lung, spleen, heart, kidney, colon) from healthy animals were sequenced, assembled *de novo*, and functionally annotated to generate a comprehensive transcriptome for two species of cotton rats (*S. fulviventer* and *S. hispidus*). Each species of the cotton rat was then infected with RSV A/Long strain alongside uninfected controls, and the transcriptome references were used to determine differentially expressed genes after RSV infection.

i. RNA extraction and transcriptome sequencing

Total RNA was extracted from sections of spleen, heart, kidney, lung, and colon harvested from 4-6-week-old healthy *S. fulviventer* and *S. hispidus* for comprehensive transcriptome assembly. The RNA-seq libraries were sequenced to a depth of 50 million paired-end 2x150 per sample. After sequencing read QC, data were pooled from all healthy tissues (n=2 per species, lung n=4 per species) to generate 456 million paired-end reads for *S. fulviventer* (GC content 54.9%) and 465 million paired-end for *S. hispidus* (GC content 53.1%) **(Table 3.1)**. Transcriptome for each species was generated using *de novo* assembly, which was then used as a reference database to evaluate RSV-induced transcriptomic changes in the lungs of the infected animals compared to uninfected controls.

Species	Sample Type (n)	Raw Data (Mb)	Trimmed/QC PE Data (Mb)	GC%
S. fulviventer	*Lung (4 healthy, 5 infected)*	410.0	206.0	53.3
	Intestines (2)	138.0	72.3	52.5
	Heart (2)	132.8	70.5	56.0
	Spleen (2)	113.8	56.7	58.5
	Kidney (2)	112.9	50.5	54.0
		907.5	456	
S. hispidus	*Lung (4 healthy, 5 infected)*	421.8	211.9	52.0
	Intestines (2)	127.7	66.4	54.0
	Heart (2)	109.6	56.4	50.0
	Spleen (2)	135.9	66.4	54.5
	Kidney (2)	123.8	63.9	55.0
		918.8	465	

Species	Raw Assembly	Filter 200bp	CD-Hit 95%	Evidential Gene
S. fulviventer	1,399,089	1,326,282	1,323,179	620,569
S. hispidus	1,399,138	1,326,220	1,249,645	592,099

Table 3.1. (A) Sequencing statistics as raw data and Trimmomatic-processed paired-end reads separated by organ type. (B) Filtering statistics for final assembly. Contigs <200bp removed using seqkit. Contigs with taxonomic annotation (BlastX) as "virus", "bacteria", or "fungi" were removed from transcriptome assembly. Contigs with 95% similarity were removed with CD-HIT. Primary/non-redundant mRNAs were picked by EvidentialGene tr2aacds pipeline for final assembly file.

ii. De novo transcriptome assembly of S. fulviventer and S. hispidus reads

For each species, sequenced reads from all 5 tissue types were combined. Trinity (265) was used for *de novo* assembly of contigs (also referred to as 'transcripts') following *in silico* normalization to achieve an average coverage of 50x. The assembler generated over 1.3 million contigs per species, which were further filtered based on length (266) and redundancy (267, 268) (**Table 3.1**). The following annotation statistics are shown in **Table 3.2.** Each transcript was assigned a unique identifier that designated each 'gene' (*S. fulviventer*=587,619, *S. hispidus*=559,830) and their alternatively spliced 'isoforms' (*S. fulviventer*=620,569, *S. hispidus*=592,099). Assembled transcripts from both species had a GC content of 43% and contig N50>1600 (which exceeds N50 of other published transcriptome assemblies (269-271)). The Ex90N50, which is the N50 but limited to the top 90% of total normalized transcripts, was 2503 (*S. fulviventer*, #transcripts=108,995) and 2162 (*S. hispidus*, #transcripts=240,587) (**Table 3.2**).

To assess the quality of the assembly, quality-trimmed reads were mapped back to the full assembly using Bowtie2 (272) to find 81.9% and 87.9% of reads were utilized during assembly of the *S. fulviventer* and *S. hispidus* transcriptome respectively, for which >70% is indicative of good quality. Transcriptome completeness was also assessed by searching for evolutionarily-conserved BUSCO (273) groups from the vertebrata_odb10 lineage dataset (total n=3354) within the dataset. The *S. fulviventer* assembly had 92.7% complete BUSCOs (Complete:3109 [Complete&Single:785, Complete&Duplicated:2324], Fragmented:154, Missing:91), and *S. hispidus* had 90.2% complete BUSCOs (Complete:3025 [Complete&Single:2599,

Complete&Duplicated:426], Fragmented:194, Missing:135). All assembly quality statistics are also shown in **Table 3.2**.

Transcriptome Assembly	*S. fulviventer*	*S. hispidus*
Contigs/Isoforms	620,569	592,099
Genes	587,619	559,830
Genes (TPM>1)	270,451	474,882
Median/Mean Contig Length	580/1003.01	597/1032.27
GC%	43.28	43.32
Contig N50	1604	1655
Ex90-N50:#Genes	2503:108,995	2162:240,587
Reads Mapped Back to Assembly	81.9%	87.9%
BUSCO (vertebrata n=3354)	90.2%	90.4%
Annotation		
Identified Coding Sequences	118,159	113,999
Annotated transcripts (*BlastX* 1e-5)	118,060	117,153
Unique genes (*no duplicates*)	18,726	18,380
Annotated proteins (*BlastP* 1e-5)	31,855	30,953
Unique proteins (*no duplicates*)	17,154	16,887
KEGG Pathways	90,642	89,654
Pfam (protein domains)	5,042	6,298
TmHMM (transmembrane helices)	18,846	18,027
SignalP (signaling proteins)	7,915	7,778
EggNOG (orthology relationships)	2,902	2,205

Table 3.2. *De novo* assembly and annotations statistics.

The number of transcripts and their expression levels varied between tissue type, and several transcripts were tissue-specific, with the lung having the greatest number of unique transcripts (*S. fulviventer*=47,448, *S. hispidus*=38,270) (**Figure 3.1A-B**). Other transcripts were either found in all 5 tissues (*S. fulviventer*=216,307, *S. hispidus*=225,833) or a combination of two or more tissue types (**Figure 3.1A-B**). Many transcripts were differentially expressed between tissues (*S. fulviventer*=49,215, *S. hispidus*=55,415) as determined by DESeq2 and visualized by hierarchical heatmap clustering (274) (**Figure 3.1C-D**). The principal component analysis shows significant difference in gene expression patterns between species and different tissue samples, indicated by strong clustering based on species and tissue type (**Figure 3.1E**). Following assembly, transcripts were annotated using protein sequence database (SwissProt) to search for homology against other species. For both species, transcripts shared the highest homology with *Mus musculus* (54%), *Homo sapiens* (17%) and *Rattus rattus* (13%) (**Figure 3.1F**).
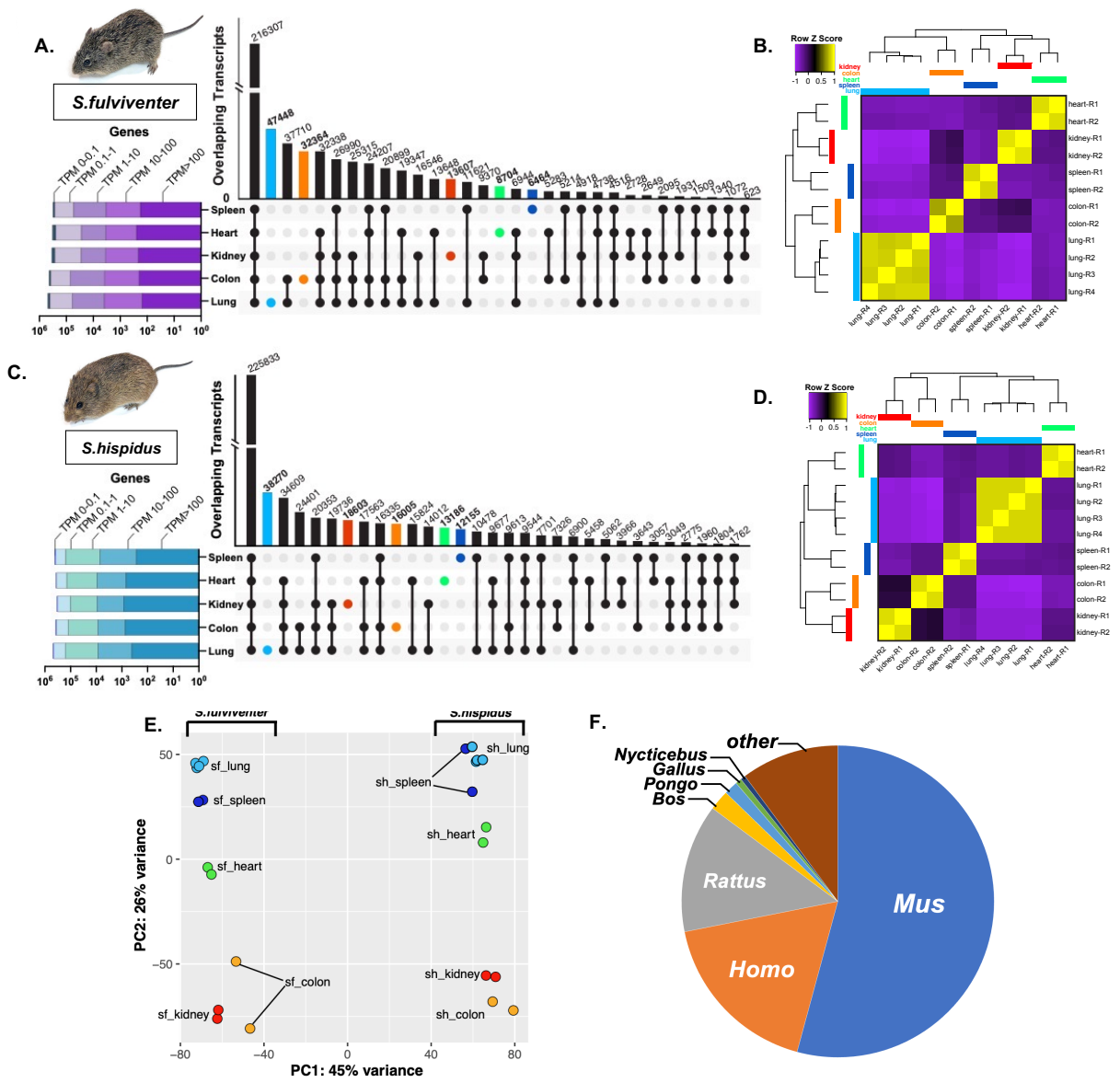
Figure 3.1. *Sigmodon* **transcriptome assembly using Trinity.** (A, C) UpSet plot of overlapping and tissue-specific transcripts expressed in the spleen (blue), heart (green), kidney (red), colon (orange), and lung (blue) of both (A) *S. fulviventer* (n=2) and (C) *S. hispidus* (n=2). Bars for overlapping sequences are shown in black, while bars for sequences unique to individual organs are shown in corresponding colors. Number of overlapping transcripts are shown above each bar, with individual organ comparison on the y-axis. Number of unique genes per tissue are represented by the horizontal barplot. (B, D) Hierarchical clustering and heatmap by Z-score of all differentially expressed transcripts (Q<0.001,L2fc>2, Euclidean distance) within individual tissue samples of both (B) *S. fulviventer* and (D) *S. hispidus.* (E) Principal component analysis (PCA) of expressed transcripts within healthy tissue of both species. (F) Taxonomic source of gene annotations determined by NCBI BlastX. Data represents annotations of *S. fulviventer* transcriptome; *S. hispidus* not shown but varies by ~1% in all categories except "other".

Transcripts were functionally annotated using the Trinotate pipeline (https://trinotate.github.io/). Transcriptome references were generated for each species independently. First, cotton rat transcripts were aligned to a non-redundant protein annotation database (UniProt (275))–which contains sequences from 14,132 different taxa (http://web.expasy.org/docs/relnotes/relstat.html)–based on sequence homology via BLASTx (276) (cutoff e-value <1e-05). 118,060 transcripts were annotated, comprising 18,726 unique genes for *S. fulviventer* and 117,153 transcripts comprising of 18,380 unique genes for *S. hispidus* out of the ~600,000 transcripts assembled for each species **(Table 3.2)**. Coding regions were then identified based on open reading frames (*S. fulviventer*=118,159, *S. hispidus*=113,999). Protein coding transcripts were annotated using BLASTp (276) (cutoff e-value <1e05) against the UniProt database (275) to identify >30,000 annotated proteins per species (~17,000 proteins after filtering duplicate annotations). Proteins were further annotated based on protein domains (>5,000), transmembrane helices (>18,000), and signaling proteins (>7700).

In addition to gene name annotations, BLAST alignment to UniProt also assigned several functional terms to describe genes, including Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways (277), Gene Ontology (GO) (278), and EggNOG orthology (279). Gene Ontology (GO) terms describe the functional properties of genes and their relationship to one another. One or more GO Terms were assigned to 98.35% *S. fulviventer* annotated transcripts (n=116,115) and 98.45% *S. hispidus* annotated transcripts (n=115,332). GO terms were grouped into 3 categories: Biological Processes,

Cellular Components, and Molecular Function (**Figure 3.2 A-C**). Genes from both *S. fulviventer* and *S. hispidus* had a similar ratio of associated GOs. The top Biological Processes GOs included "cellular nitrogen compound metabolic processes" (~25%), "DNA metabolic processes" (~17%), "transport" (~16%), and "anatomical structure development" (~15%) (**Figure 3.2A**). Cellular Component GOs consisted mostly of "organelles" (~38%), "cytoplasm" (~24%), "cytosol" (~18%), and "nucleus" (16%) (**Figure 3.2B**). The most common Molecular Function GOs include "nucleotidyltransferase" (~15%), "nuclease" (~15%), "enzyme binding" (~9%), and "nucleic acid binding" (RNA=8%, DNA=8%) (**Figure 3.2C**).

The majority of KEGG pathways identified belonged to metabolism and signaling, specifically secondary metabolite biosynthesis, ligand/cytokine/metal receptor interactions, and a variety of signaling pathways (**Figure 3.2D**). Immune-specific KEGG pathways were examined to showcase the large number of relevant genes identified in the *de novo* transcriptome, including chemokine signaling, innate and adaptive cell receptor signaling, complement cascades, and lymphocyte differentiation (**Figure 3.2E**). The top signaling and membrane transport KEGG pathways are shown in **Figure 3.2F.** KEGG also provides gene mapping against many infection pathways, including infectious agents successfully modeled in cotton rats i.e., influenza (166), HIV-1 (280), and measles (168). (**Figure 3.2G**). All annotated genes of both *S. fulviventer* and *S. hispidus* were mapped to the host genes associated with influenza infection KEGG pathway, and both species' references included annotations for 94.3% of the genes required during influenza infection (**Figure 3.3**).
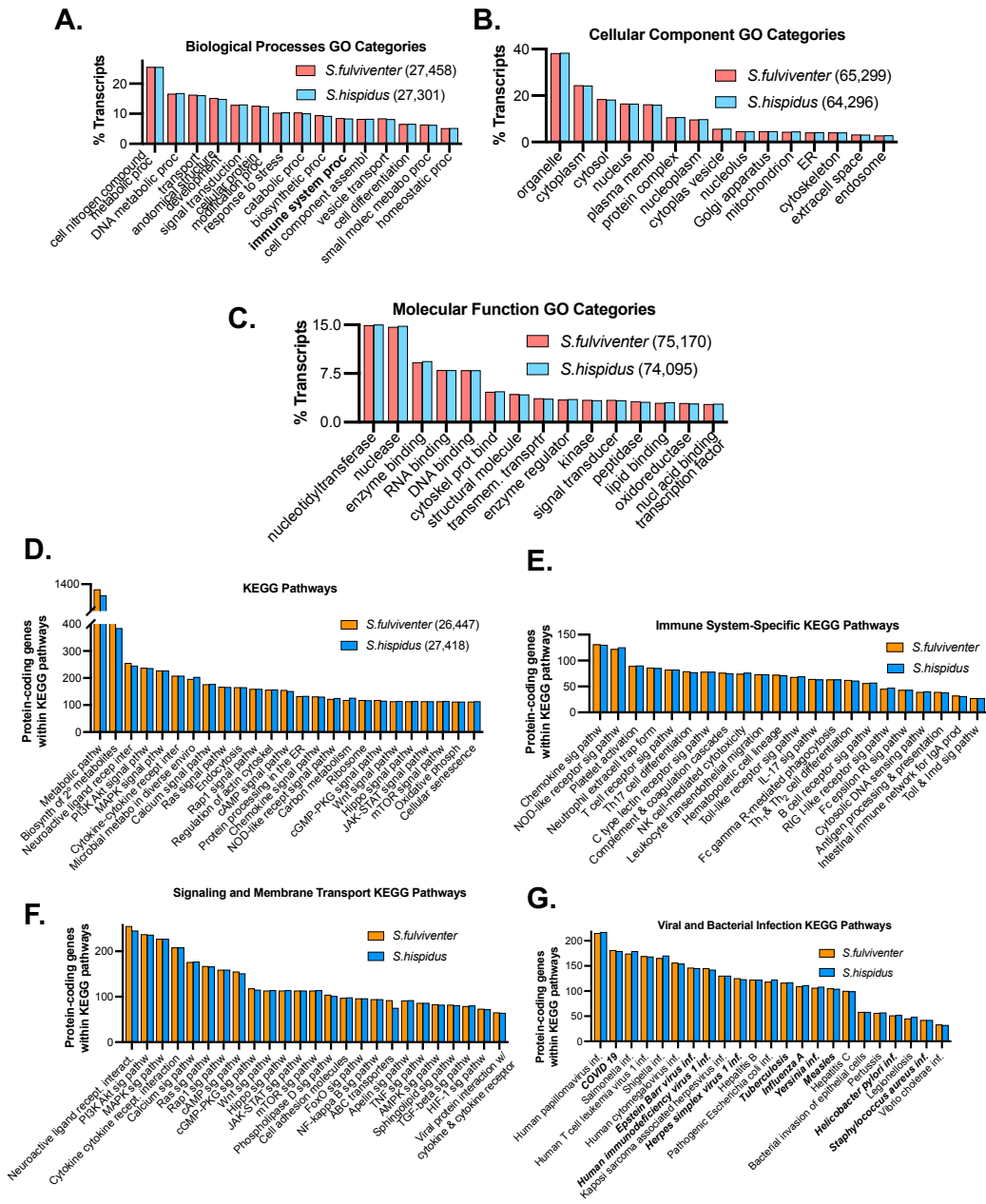
Figure 3.2. *Sigmodon* transcriptome annotation. (A) Biological processes, (B) cellular components, and (C) molecular function SwissProt Gene Ontology (GO) terms on the x-axis (total GO's within respective categories including overlaps represented in figure legends) with number of GO's (percentage of total ~107k annotated transcripts) on the y-axis; GO's assigned against TrEMBL/SwissProt database using Trinotate v.3.2.2. (D) Top 25 KEGG, (E) immune system-specific pathways, (F) Top Signaling and Membrane Transport, and (G) Viral and Bacterial Infection Pathways on the x-axis with total number of protein-coding genes on the y-axis. Pathways assigned using TransDecoder-determined CDS followed by GhostKOALA. (G) Infections in which there is published literature using cotton rats as a model are *bold italics*.
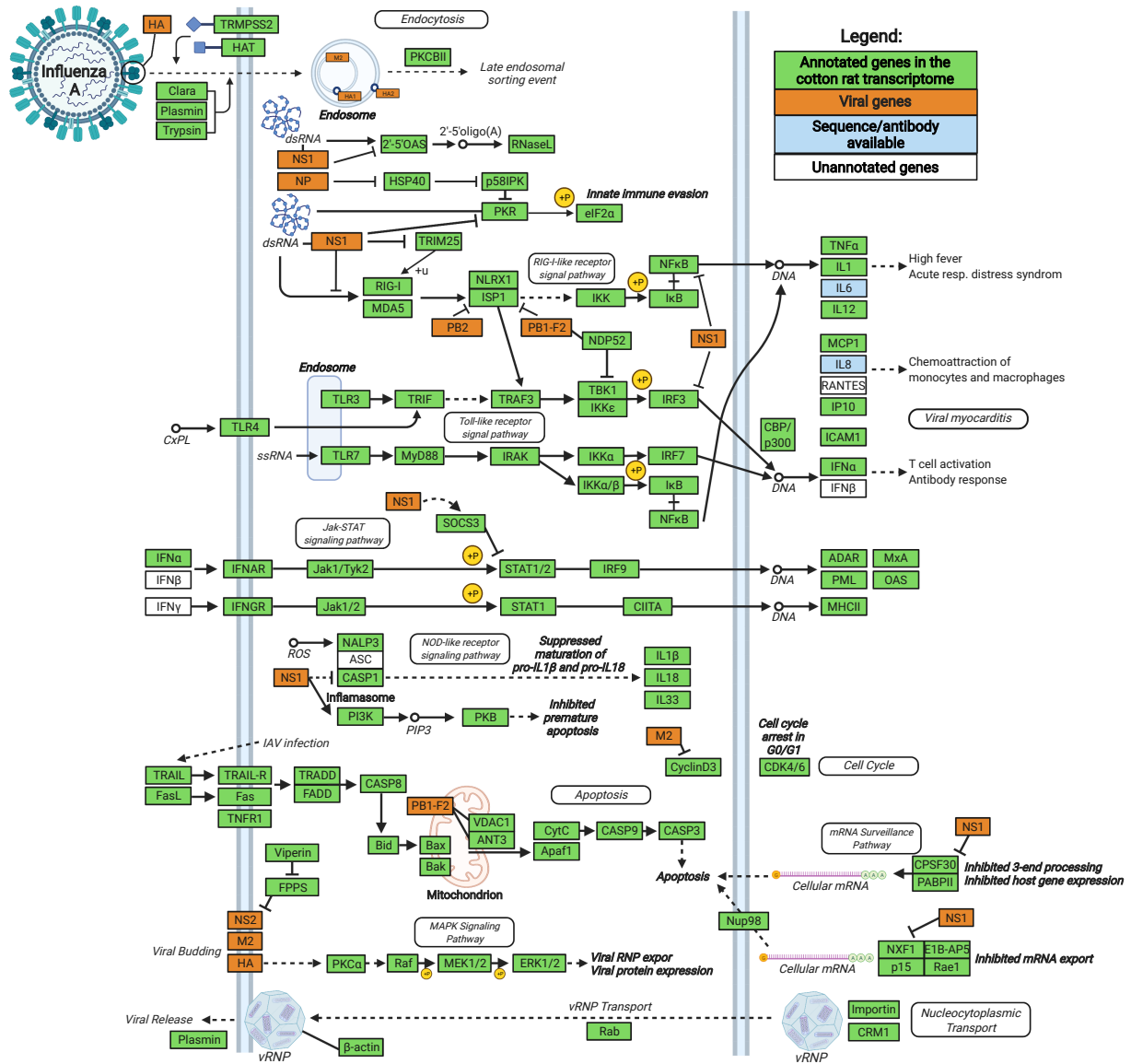
Figure 3.3. The Influenza KEGG Disease Pathway (adapted with permission from Kyoto Encyclopedia of Genes and Genomes [KEGG]). 99 of the 105 essential genes of influenza pathogenesis were successfully assembled and annotated in our *de novo* transcriptome of *S. fulviventer* and *S. hispidus*. Created with BioRender.

S. fulviventer and *S. hispidus* were intranasally infected with 10^5 PFU of RSV
A/Long or PBS (as mock infection control), and lungs were harvested 5 days post infection
for analysis. Infection was confirmed by qRT-PCR targeting the RSV G and F protein,
where significantly higher RNA copy was found in infected animals, whereas uninfected
controls did not have amplification of viral RNAs (**Figure 3.4A**). There was no significant
difference in viral RNA levels between *S. fulviventer* and *S. hispidus* following 5-day
infection (p=0.3491, Tukey's multiple comparisons test). Lung tissue was evaluated for
histopathology (**Figure 3.4B**) and blindly scored for 4 inflammatory parameters as
described in the methods: peribronchiolitis, perivasculitis, interstitial pneumonia, and
alveolitis (**Figure 3.4C**). Each infected animal had a significantly larger cumulative
pathology score compared to uninfected controls, indicating successful infection (*Sf*
p=0.0020, *Sh* p=<0.0001, Tukey's multiple comparisons test) (**Figure 3.4C**). RNA was
then sequenced as described, and reads were aligned to the previously annotated
reference transcriptome for each species. Following expression normalization, gene
expression analysis showed RSV infection significantly altered expression patterns. RSV
infection in the lungs of *S. hispidus* resulted in greater alterations in gene expression than
that of *S. fulviventer* (**Figure 3.4 F-H**).

Gene differential expression analyses (RSV A/Long vs. mock-infected Controls)
identified several genes that were differentially up- and down-regulated in infected lungs,
including genes unique and common to each cotton rat species (**Figure 3.4D**). RSV
infection in *S. fulviventer* resulted in 325 unique upregulated genes (98 annotated) and

140 unique downregulated genes (20 annotated). RSV infection in *S. hispidus* resulted in 750 upregulated genes (44 annotated) and 245 downregulated genes (8 annotated) (**Figure 3.4D-E**). By combining infection data from both species, 276 genes (12 annotated) were upregulated in both species while 9 genes (0 annotated) were downregulated in both species. Only annotated genes for each species were selected for further analysis. Twenty of the top-most differentially expressed annotated genes (selected based on function and Gene Ontology) are listed in **Table 3.3** (*S. fulviventer*) and **Table 3.4** (*S. hispidus*).
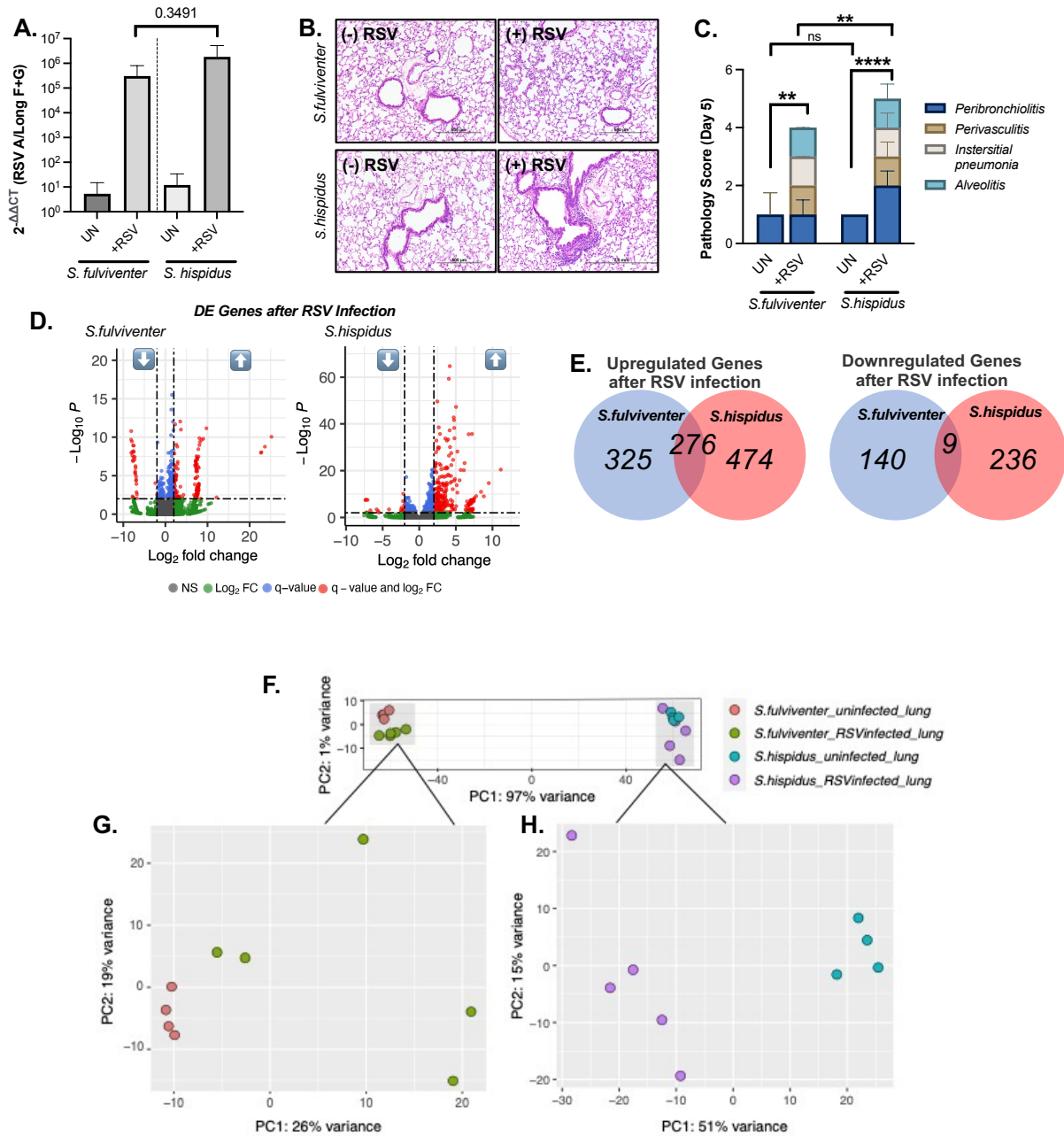
Figure 3.4. RSV-induced changes to the lung environment. (A) Viral titers (RSV G and F proteins) in uninfected and infected lungs determined by qRT-PCR normalized to β-actin; fold change (calculated by $2^{-\Delta\Delta CT}$) on the y-axis. (B) Histopathological imaging and (C) blindly generated pathology scoring for both healthy (n=4/species) and RSV-infected (n=5/species) lung tissue from *S. fulviventer* and *S. hispidus*. (D-E) DESeq2 analysis of assembled genes from lung tissues (same tissue imaged from Figures A-B) revealed several genes that are differentially expressed (DE) between healthy and infected tissues (p<0.05, q<0.05, l2fc>|1.0|) of *S. fulviventer* (blue) and *S. hispidus* (red), of which several genes were successfully annotated using Trinotate (as indicated in boxes). (F-H) Principal component analysis of RSV-infected lungs vs uninfected controls. *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

In *S. fulviventer*, several genes were downregulated upon infection related to immunoglobulin structure (*HVM44*), transcription factors (*CREB1*, *CRY1*), epithelial structure integrity (*K2C8*), formation and maintenance of tight junctions (*OCLD*), membrane trafficking regulation (*CPNE3*), targeted RNA degradation (*MOV10*), and tissue remodeling and homeostasis (*MFGM*). *GLYC*, which is the RSV G protein, was marked as upregulated in RSV lungs due to the absence of reads in the uninfected group (indicated by DESeq2 basemean=0). Upregulated host genes were related to the complement system (*CFAB*), immune receptor signaling (*UN93B*), vesicle trafficking (*EXOC5*), interferon-induced antiviral activity (*MX2, CXCL10, I27L2, OAS1A, IIGP1, ISG15*), prostaglandin metabolism (*PGDH*), and electron transport (*COX1*). These differentially expressed *S. fulviventer* genes following RSV infection are listed in **Table 3.3**.

In *S. hispidus*, only 1 annotated gene was downregulated: *LMTD1*, which is involved in cell proliferation. Other downregulated genes had annotations indicating inclusion of signaling peptides (via SignalP) and transmembrane regions (via TmHMM) but with unknown function. Several upregulated immune genes in *S. hispidus* were also upregulated in *S. fulviventer* (*I27L2, CXCL10, MX2*). Most upregulated genes were related to cytokine stimulation of antiviral activity (*OAS3, CXCL11, IRF9, IFIT1, NLRC5*), proteasomal degradation (*PSB9*), cytoskeletal reorganization in response to stress (*SYWC*, *K2C6A*), T cell activation (*HSH2D, LY6E*), post-transcriptional regulation (*ABEC1*), viral protein degradation (*RSAD2*/*Viperin*), complement pathway (*CO4A,*

C4BPA), chemical metabolism (*GBP4*), and cellular defense-related signal transduction

(*LG3BP*). These differentially expressed *S. hispidus* genes following RSV infection are

presented in **Table 3.4**.

Gene	Description	Transcript ID	Log2FC	p value	q value
HVM44	Ig heavy chain V region PJ14	*Sfulv_DN86481_c0_g1*	↓ -3.51	1.7E-04	3.4E-02
CREB1	Cyclic AMP-responsive element-binding protein 1	*Sfulv_DN150233_c0_g1*	↓ -1.94	1.5E-05	6.3E-03
CRY2	Cryptochrome-2	*Sfulv_DN8103_c29_g1*	↓ -1.69	4.8E-08	5.8E-05
K2C8	Keratin, type II cytoskeletal 8	*Sfulv_DN18044_c9_g1*	↓ -1.54	5.4E-07	4.7E-04
OCLN	Occludin	*Sfulv_DN14349_c3_g1*	↓ -1.49	2.7E-05	9.6E-03
CPNE3	Copine-3	*Sfulv_DN28516_c1_g1*	↓ -1.27	1.6E-04	3.4E-02
MOV10	Putative helicase MOV-10	*Sfulv_DN357483_c0_g1*	↓ -1.14	1.8E-05	7.1E-03
MFGM	Lactadherin	*Sfulv_DN100738_c3_g1*	↓ -1.07	3.4E-05	1.1E-02
ISG15	Interferon-induced ubiquitin-like protein, 15kDa	*Sfulv_DN2597_c2_g1*	↑ 1.02	1.3E-04	2.9E-02
IIGP1	Interferon-inducible GTPase 1	*Sfulv_DN1249_c0_g1*	↑ 1.33	3.0E-05	1.0E-02
OAS1A	2'-5'-oligoadenylate synthase 1A	*Sfulv_DN24798_c0_g1*	↑ 1.58	7.8E-05	2.0E-02
I27L2	Interferon alpha-inducible protein 27-like protein 2	*Sfulv_DN13614_c0_g1*	↑ 1.96	8.0E-13	5.2E-09
COX1	Cytochrome c oxidase assembly protein COX15 homolog	*Sfulv_DN104408_c3_g1*	↑ 2.04	4.6E-08	5.7E-05
CXCL10	C-X-C motif chemokine 10	*Sfulv_DN18689_c0_g1*	↑ 2.60	6.1E-09	1.0E-05
PGDH	15-hydroxyprostaglandin dehydrogenase [NAD(+)]	*Sfulv_DN56601_c1_g1*	↑ 2.95	5.8E-10	1.3E-06
MX2	Interferon-induced GTP-binding protein Mx2	*Sfulv_DN17914_c1_g1*	↑ 3.48	1.1E-05	4.8E-03
EXOC5	Exocyst complex component 5	*Sfulv_DN7267_c6_g1*	↑ 4.02	4.1E-05	1.3E-02
UN93B	Protein unc-93 homolog B1	*Sfulv_DN114523_c0_g1*	↑ 4.04	3.2E-05	1.1E-02
CFAB	Complement factor B	*Sfulv_DN212381_c0_g2*	↑ 6.99	1.3E-07	1.4E-04
GLYC	Major surface glycoprotein G_*Orthopneumovirus*	*Sfulv_DN7526_c1_g1*	↑ 9.74	2.5E-16	7.1E-12

Table 3.3. Twenty notable differentially expressed genes in *S. fulviventer* lung following RSV infection. Negative Log2FC = downregulated, positive Log2FC = upregulated. Gene names/descriptions determined by BlastX against UniRef/SwissProt. Transcript IDs correspond to contig names (Sfulv_xxx) in assembly FASTA file. Log2FC=log2FoldChange, q=FalseDiscoveryRate.

Gene	Description	Transcript ID	Log2FC	p value	q value
LMTD1	Lamin tail domain-containing protein 1	*Shisp_DN10656_c2_g1*	↓ -1.04	3.7E-04	3.5E-02
NLRC5	NLR family CARD domain containing 5	*Shisp_DN1039_c0_g1*	↑ 1.74	1.5E-06	3.3E-04
GBP4	Guanylate binding protein 4	*Shisp_DN1171_c0_g1*	↑ 1.88	2.8E-13	2.3E-10
IFIT1	Interferon induced protein with tetratricopeptide repeats 1	*Shisp_DN1078_c15_g1*	↑ 2.01	2.6E-06	5.3E-04
LY6E	Lymphocyte antigen 6 family member E	*Shisp_DN101026_c0_g1*	↑ 2.20	2.0E-12	1.4E-09
IRF9	Interferon regulatory factor 9	*Shisp_DN12125_c1_g1*	↑ 2.33	3.2E-07	8.4E-05
C4BPA	C4b-binding protein alpha chain	*Shisp_DN107540_c0_g1*	↑ 2.42	5.6E-13	4.3E-10
CO4A	Collagen alpha-1(IV) chain	*Shisp_DN122397_c0_g1*	↑ 2.43	3.1E-05	4.6E-03
LG3BP	Galectin-3-binding protein	*Shisp_DN11955_c0_g1*	↑ 2.45	1.1E-12	8.1E-10
PSB9	Proteasome subunit beta type-9	*Shisp_DN13011_c0_g1*	↑ 2.52	5.2E-10	2.3E-07
K2C6A	Keratin, type II cytoskeletal 6A	*Shisp_DN12075_c0_g1*	↑ 2.70	1.5E-05	2.4E-03
HSH2D	Hematopoietic SH2 domain-containing protein	*Shisp_DN0_c15_g8*	↑ 2.80	3.0E-12	2.1E-09
ABEC1	C->U-editing enzyme APOBEC-1	*Shisp_DN11754_c2_g1*	↑ 3.45	3.9E-25	1.0E-21
RSAD2	Radical S-adenosyl methionine domain-containing prot. 2	*Shisp_DN112437_c0_g1*	↑ 3.51	3.5E-12	2.3E-09
CXCL11	C-X-C motif chemokine 11	*Shisp_DN149434_c0_g1*	↑ 3.56	8.3E-15	8.1E-12
MX2	Interferon-induced GTP-binding protein Mx2	*Shisp_DN140604_c0_g1*	↑ 4.24	1.0E-18	1.5E-15
CXCL10	C-X-C motif chemokine 10	*Shisp_DN14446_c0_g1*	↑ 4.62	6.3E-29	2.5E-25
OAS3	2'-5'-oligoadenylate synthase 3	*Shisp_DN1184_c1_g1*	↑ 5.12	1.3E-17	1.6E-14
I27L2	Interferon alpha-inducible protein 27-like protein 2	*Shisp_DN12103_c7_g1*	↑ 6.20	1.6E-27	6.0E-24
SYWC	Tryptophan--tRNA ligase, cytoplasmic	*Shisp_DN144721_c0_g1*	↑ 8.44	1.1E-14	1.1E-11

Table 3.4. Twenty notable differentially expressed genes in *S. hispidus* lung following RSV infection. Negative Log2FC = downregulated, positive Log2FC = upregulated. Gene names/descriptions determined by BlastX against UniRef/SwissProt. Transcript IDs correspond to contig names (Shisp_xxx) in assembly FASTA file. Log2FC=log2FoldChange, q=FalseDiscoveryRate.

Of the differentially expressed genes (DEs) following RSV infection, 12 DEs were either up- or down-regulated in both *S. fulviventer* and *S. hispidus*. The chemokine *CXCL10*, *MX2*, and *I27L2* were among the top 10 upregulated genes in both species. Other DEs within the significance threshold include *GBP2, GBP4, IRF9, NLRC5, OAS3, PAR14, RN213, SYWC,* and *IIGP1.*

Differentially expressed Gene Ontology (GO) terms were also determined using the GoSeq R package(281). The top GOs enriched in infected lungs of both *S. fulviventer* (**Figure 3.5A**), and *S. hispidus* (**Figure 3.5B**) were "biological processes" such as response to cytokine and interferon-beta, immune system process, cell surface receptor signaling pathway, and viral process; "molecular functions" such as ion binding, catalytic activity, hydrolase activity, and "cellular components" such as intracellular membrane-bounded organelles and nucleus. Other enriched GOs in infected lungs were related to enhancing the components and function of cells, such as components of the membrane and vesicle, response to stress, regulation of apoptotic processes, regulation of viral genome replication, leukocyte activation, and binding of carbohydrates and other organic compounds. Only *S. hispidus* had enriched GOs in healthy tissues, including the cellular component involved in the guanyl-nucleotide exchange factor complex.
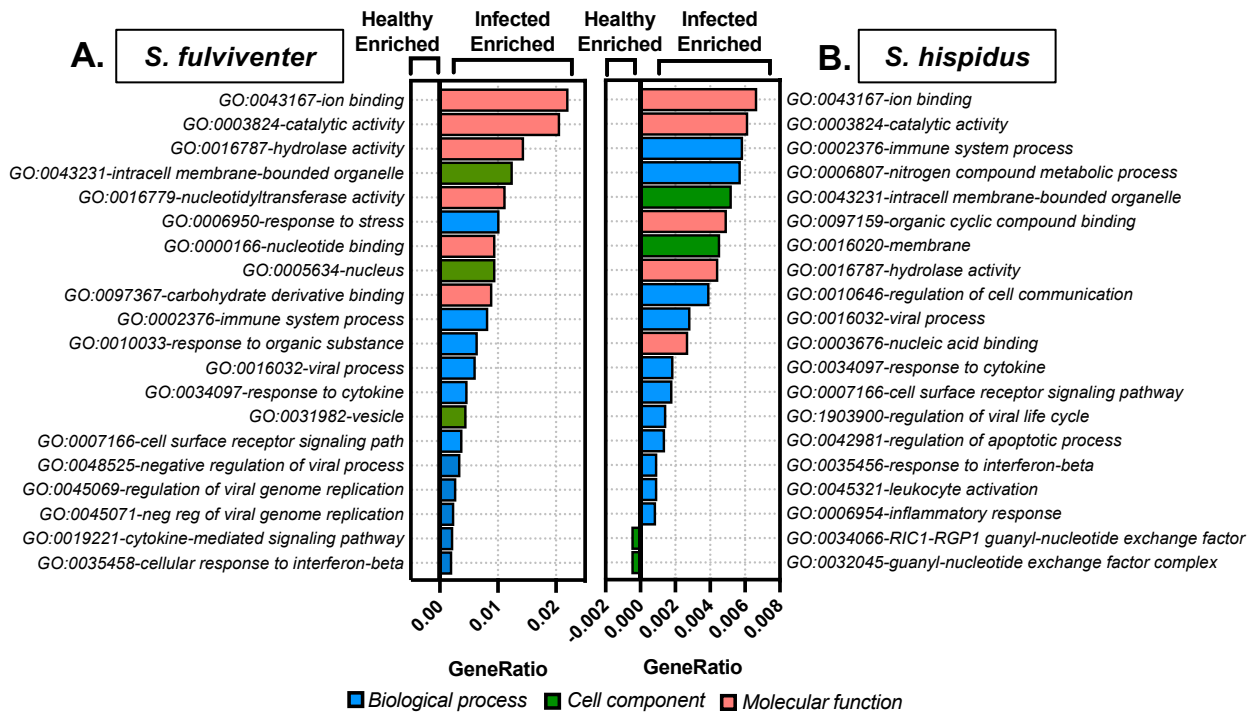
Figure 3.5. Categorical changes in gene expression after RSV infection in both (A) *S. fulviventer* and (B) *S. hispidus* using Gene Ontology (GO) annotations and GoSeq. SwissProt GO terms on the y-axis (color coordinated based on biological process [blue], cell component [green], or molecular function [red]) with relative expression of GOs on the x-axis (expressed as GeneRatio, which is the proportion of differentially expressed genes in each GO category to the total number of differentially expressed genes in all significant GOs). Positive GeneRatio = enriched in RSV-infected lungs; negative GeneRatio = enriched in healthy lungs (no significant GOs in healthy *S. fulviventer*). All GOs were statistically significant (p<0.05, q<0.05).

vi. Confirmation of differentially expressed genes

To confirm differentially expressed genes (DEs), three *de novo* assembled genes were selected based on their upregulation on day 5 post-infection (*Shisp_DN132151_c0_g1* [*IIGP1*], *Shisp_DN12103_c7_g1* [*I27L2*], *Sfulv_DN158_c1_g1* [*IIGP1*]). qRT-PCR with SYBR green was performed with primers designed to target each gene, normalized to the β-actin housekeeping gene. The three genes included IIGP1 in both *S. fulviventer* and *S. hispidus* and *I27L2* in *S. hispidus*. qRT-PCR confirmed gene upregulation post-RSV infection (i.e., positive fold changes) that correlated with the RNA-Seq data (**Table 3.5**).

GeneID	DESeq2 FC	qRT-PCR FC	Forward Primer	Reverse Primer
IIGP1_S.hispidus	1.59	2.28	ACAGCTGGTCCGGATTTGAG	CCTTGCCAAAGCCATCAGC
I27L2_S.hispidus	6.20	2.19	ACTGTGTGCTAGCCAACCTC	CACTGACGCCAGAGAACACT
IIGP1_S.fulviventer	1.33	2.45	AGTCCTCAGCCAGACTCTGT	TGCCAAAGCCATCAGCATGA
B-actin (282)	–	–	GGCCAACCGTGAAAAGATGACTC	GTCCGCCTAGAAGCATTTGCG

Table 3.5. RT-qPCR validation of differentially expressed genes and primer sequences.

Additionally, to further confirm other significant DEs, the results of this study were compared to the previously published DE analysis by Rajagopala et al. (264), where *S. hispidus* was infected with RSV A/Long in the same facility with the same strain and concentration of virus. The previously published study included two time points (Day 4

and 6 post-infection) compared to the one Day 5-time point but determined DEs using the same DESeq2 significance parameters (p<0.05, q<0.05, l2fc>|1.0|). Compared to the Rajagopala dataset, 7 of the 12 total DEs on Day 4 post-infection were viral genes (*FUS, GLYC, L, M2-2, MATRX, NCAP, SH*); however, this current study confirmed 4 of the 5 Eukaryotic DEs (*IFIT1, MX2, OASL2, RSAD2*). Compared to the DEs found by Rajagopala et al. on day 6, this current study confirmed 29 of the 81 DEs following infection by RSV (including *CFAB, GBP2/4/5/6, IFIT1, IIGP1, K2C6A, MX2, OASL2, RSAD2*).

Cotton rats are excellent pre-clinical animal models for studying infections with RSV and other viruses. However, without a published genome for either species, genetic or transcriptomic analysis in cotton rats is limited. In this study, comprehensive transcriptome references were generated from multiple organs of two inbred species of cotton rats: *S. fulviventer* and *S. hispidus*. Transcriptomes were assembled from each species' lung, spleen, kidney, heart, and intestines with a contig N50>1600. Annotation of contigs generated nearly 120,000 gene annotations for each species. Comparing these two transcriptomic references revealed large differences in transcript sequence homology and baseline gene expression levels between each species, although the total number of annotated genes and functional categories (gene ontology, KEGG terms) were similar (**Table 3.2**).

The transcriptomes of *S. fulviventer* and *S. hispidus* were then used to assess the immune response to RSV infection. 238 unique genes that are significantly differentially expressed during RSV infection were identified, including several genes implicated in RSV infection (e.g., *Mx2, I27L2*, *LY6E*, *Viperin*, *Keratin 6A, ISG15, CXCL10, CXCL11, IRF9*) as well as novel genes that have not previously described in RSV research (*LG3BP, SYWC, ABEC1, IIGP1, CREB1*). This study presents two comprehensive transcriptome references as resources for future gene expression analysis studies in the cotton rat model and provides gene sequences for mechanistic characterization of molecular

pathways. Overall, these results offer generalizable insights into the effect of host genetics on host-virus interactions, as well as identify new host therapeutic targets for RSV treatment and prevention.

These transcriptome assemblies and annotations surpass the quality and completeness standards of previously published transcriptome references in other animals (269, 270, 283). The quality and depth of the transcriptome assembly also improved upon the previous lung transcriptome of *S. hispidus* lung tissues (264) due to the inclusion of multiple tissues. This resulted in nearly 3 times the number of annotated transcripts (117,153 vs. 38,736) and an additional 6,169 unique gene annotations in *S. hispidus*. Furthermore, the application of the transcriptome references to infer RSV-induced gene expression alterations exhibits these references' utility in the study and treatment of many infections, including RSV, influenza, and other respiratory viruses.

This analysis confirmed many genes previously implicated in RSV infection and immunity. *RSAD2* (or Viperin) is a significantly upregulated gene upon RSV infection in *S. hispidus* that inhibits RSV filament formation and cell-cell viral transmission without inhibiting viral protein expression (284). Viperin is also the most upregulated gene in human nasal epithelium following intranasal challenge with Rhinovirus (285). Another RSV-related gene is *I27L2*, the top upregulated gene in *S. hispidus* and *S. fulviventer* that is also the top differentially expressed gene in preterm infants with severe RSV infection (286). Other upregulated genes also shown to be involved in RSV infection include chemokines *CXCL10* (287), *CXCL11* (288), *LY6E* (289), *MX2* (162), *OAS1A* (290), *ISG15* (291)*, IRF9* (292)*, NLRC5* (293) and *K2C6A* (294). Genes involved in the complement system (*CFAB, CO4A, C4BPA*) and prostaglandin metabolism (*PGDH*) were also

identified during infection, both of which have been previously described in the host response to RSV (295, 296). The results from this study support the importance of these genes in host-mediated protection against RSV and the cotton rat's translational value in viral research.

CRY2 (downregulated in *S. fulviventer*) has also been indirectly attributed to RSV disease severity. *CRY2* is a transcription factor that modulates circadian rhythms by suppressing *BMAL1*, a putative regulator of cellular factors essential for viral replication (297). Analysis of *BMAL1-/-* primary cells revealed low-*BMAL1* expression enhances susceptibility to influenza (298), parainfluenza virus 3, and RSV (299). *BMAL1* also shows seasonal variation in humans, with the lowest levels in winter months during peak respiratory virus season (297). While this gene will need further investigation in its role in RSV pathogenesis, this is the first association of modulated expression during RSV infection.

The cotton rat mimics the respiratory infections in humans due to the presence of human homologous genes absent in other laboratory rodents (e.g., *Mus musculus*), i.e., interferon-stimulated *MX* genes (300). This analysis is consistent with the upregulation of *MX2* in both *S. fulviventer* and *S. hispidus* (as well as *MX1*, *MX1B*, and *MX3* in *S. fulviventer*) upon RSV infection as previously described (301), highlighting the importance of this model in capturing the interferon-induced immune response to RSV and other viruses. Additionally, analysis captures *IIGP1* upregulation in *S. fulviventer* and *S. hispidus*. *IIGP1* is another human interferon-stimulated gene with equal importance and function of *Mx* genes, but it is only present in mice as a less-effective paralog *IGPT* (302). A study on IGPT-deficient mice found no increased susceptibility or interferon-induced

cytokine production towards intracellular pathogens such as *Listeria* and cytomegalovirus, which suggests a different upstream mechanism in mice (303). Contrarily, analysis resulted in the successful annotation of *IIGP1* in cotton rats and its significant modulation during interferon-dominant responses to intracellular infection (i.e., with RSV). These results further argue the translational utility of the cotton rat as a model for RSV and other respiratory viruses.

One interesting, upregulated gene in both *S. fulviventer* and *S. hispidus*, *SYWC,* is an interferon-induced activator of ERK, Akt, and eNOS pathways and cytoskeletal reorganization in response to stress and is a serum marker for pulmonary tuberculosis (304). While this study is the first association of *SYWC* with RSV infection, further study may uncover it as a marker for severe RSV infection in tissue and serum.

There are some limitations to this study. As the annotated transcriptome references were only generated using tissues from 2 batched samples from male animals, additional male and female animals across multiple age groups may give further insight into the healthy transcriptome of cotton rats. Additionally, this study does not explore RSV-induced transcriptome changes at other age points other than 5–7-week-old animals. This study also only utilized RSV A/Long in the infection experiments. Previously published data from Pletneva et al. has shown that different RSV strains and isolates have differential induction of interferon-activated genes in cotton rats (301)*.* While the study by Pletneva et al. only includes *S. hispidus*, it is expected that this may be true for *S. fulviventer* due to similar susceptibility to RSV A/Long. To this end, this study's conclusions are limited to RSV A/Long, and transcriptomic comparison of RSV variants should be considered for future studies. Further, multiple genes were identified that have

not been previously implicated in RSV infection, such as (*LG3BP, SYWC, ABEC1, IIGP1, CREB1*). While this analysis suggests potential unknown mechanisms of pathogenesis, the association of these genes with RSV infection is severely limited in the absence of experimental data. Such efforts are beyond the scope of this study on the reference cotton rat transcriptome but will be considered for future studies.

Animals, infection, and tissue collection

Four- to six-week-old *S. hispidus* (n=8 male, 1 female) and *S. fulviventer* (n=8 male, 1 female) cotton rats (~100 g) were obtained from the inbred colony maintained at Sigmovir Biosystems, Inc. Cotton rats in the colony were seronegative by ELISA to adventitious respiratory viruses (*i.e.,* Pneumonia Virus of Mice, Rat parvovirus, Rat coronavirus, Sendai virus). Each species was randomly split into 2 groups: RSV-infected (n=5) and Uninfected controls (n=4). Transcriptome assembly and annotation were performed on all tissues (whole lung, large intestine, heart, spleen, and kidney) from 2 males within each uninfected group. Female animals were only included in RSV-infected groups. To avoid fighting, all animals were individually housed in large polycarbonate cages and fed a diet of standard rodent chow and water *ad libitum*. All animal procedures followed NIH and USDA guidelines approved by the Sigmovir Biosystems, Inc. IACUC. Study design, analysis, and reporting of methods and results are presented in accordance with the ARRIVE guidelines.

The Long strain of RSV A/Long (ATCC Cat. # VR-26) was propagated in HEp-2 cells, and viral titer was determined using plaque assay. Cotton rats were intranasally inoculated under isoflurane anesthesia with 10^5 plaque-forming units in 100 μL of either RSV suspension or PBS vehicle. Animals were sacrificed by carbon dioxide inhalation on day 5 of infection. Whole lung, large intestine, heart, spleen, and kidney were dissected

from all animals and frozen in RNA-later (Invitrogen AM7021) for processing and sequencing at VUMC.

Lung histopathology and viral titers

Lungs (right lobe) were dissected and inflated with 10% neutral buffered formalin and immersed in formalin for fixation. Lungs were embedded in paraffin blocks, sectioned, and stained with hematoxylin and eosin (H&E). Pathologists were blinded to the study group, and slides were examined for 4 parameters of pulmonary immune cell infiltration and inflammation as previously described (305): peribronchiolitis (bronchioles), perivasculitis (small blood vessels), interstitial pneumonia (alveolar walls), and alveolitis (alveolar space) (listed from least to greatest severity). Each condition was assigned a score from 0 to 4, where 0 is no pathology and 4 is max pathology. Scores correspond to the percentage of pathology present in the field of view pictured in Figure 4A (0=0%, 1=5%, 2=25%, 3=75%, and 4=100%). Cumulative pathology scores were calculated by summing the median score for each condition.

Viral titers were determined using Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) from RNA extracted from lung (lingular lobe) homogenates (see next section *RNA extraction and cDNA library preparation* for extraction methods). Following extraction, cDNA was generated using 1ug of total RNA and the SuperScript™ III Reverse Transcriptase (Invitrogen™) kit according to the manufacturer's instructions. cDNA was diluted 1:5, and 3uL was added to qPCR reactions using iQ™ SYBR® Green Supermix (10uL total). Reactions for each target were performed in duplicate for each sample using primers (IDT, 250nm final concentration)

targeting the G and F proteins of RSV and β-actin, which have been previously published and validated for use in both species of cotton rats (306). No-template-controls and a positive control (RNA extracted from RSV A/Long viral stock used for infection) were run on each plate. CT values for both G and F were averaged and normalized to β-actin. Results were calculated using the $2^{-\Delta\Delta CT}$ method (307). Figures and statistical analysis were performed using ANOVA/Tukey's multiple comparisons test in Prism 8.

RNA extraction and cDNA library preparation

RNA was prepared as previously described (264). In summary, small sections of cotton rat lung (lingular lobe), large intestines (~20mm colon, flushed with sterile PBS), spleen, kidney, and heart were homogenized using a NextAdvance Bullet Blender® with RED RINO™ tubes containing 3.2 mm stainless steel beads (SSB32) and 2x volume of QIAzol® Lysis Regent (Qiagen, cat. no. 79306) at max speed for 3 minutes. RNA was extracted from homogenates using RNeasy® Plus Universal Mini kit (Qiagen, cat. No. 73404) via additional QIAzol® (total volume 900uL), gDNA eliminator, and chloroform with column DNase digestion as recommended by the manufacturer's protocol. RNA quality was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA). Host rRNA was removed using the NEBNext rRNA Depletion Kit v2 (Cat no: NEB E7400X). The cDNA libraries were prepared using 1µg total RNA from each sample using the NEBNext Ultra II RNA Library Prep Kit for Illumina (Cat no: NEB E7805L) following the manufacturer's protocol for intact RNA (RIN >7), AMPure XP Beads for cleanup steps (Beckman, cat. No. A63881), and NEBNext Multiplex Oligos for Illumina (Set 1, cat no: E7600S) for pooling samples. Sequencing was performed using Illumina NovaSeq6000

2x150 bp sequencing at the Vanderbilt Technologies for Applied Genomics (VANTAGE) core facility.

Transcriptome assembly

Data was then parsed into individual samples by barcode. Adaptor sequences, low quality (minimum Phred 33), and short (<75bp) reads were removed using Trimmomatic (version 0.36, "ILLUMINACLIP: NEB_multiplexoligos.fa:2:30:10 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75") (308). Only paired-end reads that passed the quality threshold described in (309) were retained.

About 921 million paired-end reads from healthy tissues were utilized for *de novo* transcriptome assembly using the Trinity v2.13.1 software package (265), with default parameters of 50x coverage. An additional 314.5 million paired-end reads from viral-infected lungs were used for experimental analysis. Contigs were assembled for *S. fulviventer* and *S. hispidus* by first clustering individual contigs/transcripts into 'genes' followed by construction of de Bruijn graphs to report full-length, alternatively spliced isoforms. All transcripts were then filtered by transcript length cutoff of <200bp using SeqKit (266), contaminant annotation (Blast annotation of "virus," "bacteria," or "fungi"), sequence similarity of 95% or greater using CD-HIT (267), and removal of redundant mRNAs picked by the EvidentialGene tr2aacds pipeline (268). Assembly statistics of final transcripts, such as mean number of transcripts, transcript length, mean transcript length, and N50, are listed in **Table 3.1**. Raw reads and assembly data will be deposited in SRA under BioProject PRJNA816878 upon acceptance of the manuscript for publication.

The RSEM package was used for quantifying and normalizing gene and isoform abundances from paired-end RNA-Seq data (310). RSEM enables accurate transcript quantification per sample and sample type without a reference genome. TPM (transcript per million mapped reads) was calculated using the RSEM package with Bowtie2 aligner, and a cutoff of TPM > 1 was used to filter the low-quality assembled transcripts to use for differential expression analysis. The Ex90N50 was calculated using the Trinity *contig_ExN50_statistic.pl* script. This pipeline is depicted in **Figure 3.6.**

Transcriptome characterization

The Trinotate v3.2.2 annotation pipeline was used to annotate transcripts from both *S. hispidus* and *S. fulviventer*. BlastX and BlastP (e-value cutoff of 1e-05) (276) were used to find sequence homology of individual contigs and protein-coding regions (determined by Transdecoder, https://github.com/TransDecoder/TransDecoder) against the UniProt/SwissProt database (275). KEGG terms (277), Gene Ontology terms (278), and EggNOG terms (279) from the Swissprot (275) database alignment with BLAST (276). Other tools within the pipeline annotated transcripts based on protein domains via Pfam (311), transmembrane helices via TmHMM (312), and signaling proteins (313).

Transcriptome assembly pipeline

SEQUENCING

Sample Collection (lungs, large intestines)

↓

Extract total RNA from tissues
- rRNA Depletion (NEB)
- Bioagilent quality assessment

↓

cDNA Library Prep

↓

Paired end sequencing at VANTAGE Core at VUMC

- NovaSeq6000 2x150
- Total bases sequenced and processed: 2,596,512,335

ASSEMBLY/ ANNOTATION

FastQC for quality assessment

↓

Trimmomatic for removing sequencing adapters

↓

Trinity for assembly
A well-established tool for assembly of transcripts into unique contigs

↓

CD-HIT & EvidentialGene to remove similar transcripts and redundancy

↓

TransDecoder to identify coding regions

↓

Annotation with Trinotate

- BLASTp/x (protein/nt sequence)
- HMMER/PFAM (protein domain)
- signalP (signal peptides)
- tmHMM (transmembrane)
- eggNOG/GO/Kegg (gene function)

DIFFERENTIAL EXPRESSION ANALYSIS

Compare experimental conditions against transcriptome reference

↓

DESeq2 for differential expression analysis

↓

goseq to assign gene ontology terms

↓

RT-qPCR verification

Figure 3.6. Transcriptome Assembly Pipeline

Differential gene expression and GO terms

Lungs isolated from RSV-infected and Uninfected groups were processed and sequenced as described above. The Trinotate pipeline was used to annotate all reads from experimental groups against our reference. Normalized expression levels of transcripts across tissues were determined using Salmon (314) and the TPM (transcripts per million) metric; only transcripts with a TPM>1 were used for downstream differential expression analysis (*S. fulviventer*=270,451, *S. hispidus*=474,882). DESeq2 package (274) was used within the pipeline by comparing experimental group (*RSV-infected lungs*) containing biological replicates with the corresponding control group (*uninfected lungs*). Genes with a P<0.05, adjusted P < 0.05 ("q"/false discovery rate/Benjamini-Hochberg), and a $\log_2$ fold change > |1| were treated as differentially expressed. The goseq package in Bioconductor was used to detect differentially abundant GO terms (278). GOs with a P<0.05 and adjusted P < 0.05 ("q"/false discovery rate/Benjamini-Hochberg) were treated as differentially expressed.

Validation of differentially expressed genes

RNA extracted from the lung tissue was used for qRT-PCR assays. qRT-PCR was performed in duplicate using the SuperScript™ III Reverse Transcriptase (Invitrogen™) kit and iQ™ SYBR® Green Supermix as described above. qRT-PCR primers were designed for 3 randomly selected up-regulated genes based on *de novo* assembly and differential expression analysis (*Shisp_DN132151_c0_g1*, *Shisp_DN12103_c7_g1*, *Sfulv_DN158_c1_g1*). Primers were designed using Primer3Plus (315) and are reported in **Table 3.5**. No-template-controls were run on each plate. Technical replicate CT values

were averaged for each gene and normalized to β-actin housekeeping gene (primers acquired from (282)). Results were calculated as fold change induction over uninfected lungs using the $2^{-\Delta\Delta CT}$ method (307). Additionally, to further confirm differential expression of genes, results from this study were compared to the previous annotation of the *S. hispidus* lung transcriptome upon RSV infection (264) using the same DESeq2 significance parameters (p<0.05, q<0.05, l2fc>|1.0|).

E. ACKNOWLEDGMENTS AND AVAILABILITY OF DATA

The sequencing data are deposited at the NCBI Short Read Archive (SRA) under BioProject PRJNA816878. Other raw datasets can be found in the supplementary files

online at https://doi.org/10.1038/s41598-022-19810-4. **Supplementary Files 1** and **2** contain the reference transcriptomes (fasta) for *S. hispidus* and *S. fulviventer*. **Supplementary File 3** contains the full annotation file for both *Sigmodon hispidus* and *S. fulviventer* (separated by xlsx sheet). **Supplementary File 4** contains KEGG annotations, Gene Ontologies, and taxonomic sources of transcript annotation. All KEGG pathways can be reconstructed using data from the "KEGG_reconstruct" tab. **Supplementary File 5** contains differential expression analysis data for both *S. hispidus* and *S. fulviventer*, including analysis by DESeq2 (including comparisons to Rajagopala et al., 2018) and GOSeq. Results are reported in separate tabs for each species. **Supplementary File 6** includes TransDecoder output identifying coding regions in each transcript for both *S. hispidus* and *S. fulviventer*. **Supplementary File 7** contains raw expression count data (generated via Salmon; used for DESeq2 analysis) from tissues of *S. hispidus* and *S. fulviventer*.

CHAPTER 4

Transcriptomic and histopathological analysis of Lactobacillus-mediated protection against severe Respiratory Syncytial Virus infection in cotton rats

The microbiome community structure and composition have been instrumental in predicting viral-associated respiratory disease outcomes and have been found to exacerbate or protect against severe sequelae (108, 209). Numerous animal and clinical studies have orally and intranasally supplemented commensal bacteria (e.g., *Lactobacillus*, *Bifidobacterium*, *Lactococcus*, and *Bacillus* species) in the prevention and treatment of respiratory infection with rhinovirus and influenza (143-155). However, the use of different bacterial strains and concentrations of bacteria have produced variable outcomes. Clinical cohort studies have suggested the commensal microbiome in protecting from severe respiratory syncytial virus (RSV) infection and long-term outcomes like wheezing and asthma (55, 134-138). Only one recently completed (July 2022) randomized controlled clinical study in Vietnamese children determined that nasal inoculation of *Bacillus* spores (LiveSpo® Navax) could effectively reduce symptoms of RSV-induced acute respiratory infections, particularly in reducing viral load and inflammation (316). Additionally, there are four clinical studies (either active or recruiting) exploring the role of vaginal seeding (which will confer *Lactobacillus* and other commensal microbes) in allergy and infection outcomes in the first years of life in healthy children born via Cesarean section. One of these trails is being performed at Vanderbilt University Medical Center and will investigate the role of seeding in early-life viral infection

outcomes (MOTHER SEED, ClinicalTral NCT05505110). There are currently no clinical trials examining probiotic-administration for the protection and prevention of RSV (www.clinicaltrials.gov, accessed October 13, 2022), but multiple murine studies have shown *Lactobacillus* to decrease airway pathology, inflammatory cytokines, and Th2 bias during RSV (54, 126, 129-131, 133, 140, 141).

Current *in vivo* data on probiotic-mediated effects during RSV infection are limited due to the use of different strains and concentrations of *Lactobacillus* that were isolated from non-host environments, which may lead to challenges in bacterial survival and colonization in a foreign environment (317). These interactions must be thoroughly tested before clinical applications. Additionally, in the context of RSV, all published studies have been performed in mice, which have been shown to have less translational value than the cotton rat. The cotton rat (genus *Sigmodon*) is considered the gold standard for RSV and a variety of other human respiratory viruses (318) due to comparable human disease outcomes (300). This model has been instrumental in predicting the efficacy of several therapeutics and vaccines currently used in high-risk human populations with RSV (171-174), but bacteria-host-virus interactions have yet to be studied in this context. Furthermore, the mechanisms of how *Lactobacillus* conveys protection are unclear, highlighting the need for an appropriate pre-clinical model to develop and test microbial interventions during RSV.

This chapter discusses the application of host-derived *Lactobacillus* species in preventing severe RSV infection in cotton rats. An oral cocktail of *Lactobacillus paragasseri* BRTN and *Limosilactobacillus reuteri* SMN that had been previously isolated from cotton rat stool in **Chapter 2** was administered to *Lactobacillus*-deficient *S.*

fulviventer for 7 days (along with PBS vehicle controls) before being intranasally infected with RSV for 5 days (along with uninfected controls). 16S rRNA gene sequencing revealed no major alterations to the gut microbiome structure by *Lactobacillus* supplementation. However, significant gut alpha and beta diversity changes were observed upon RSV infection that was absent in *Lactobacillus*-gavaged animals. Further, *Lactobacillus*-gavaged animals were found to be protected from RSV-induced histopathology, of which transcriptomic analysis revealed several host gene changes in both the gut and lung tissue upon oral *Lactobacillus* supplementation.

i. The effect of oral Lactobacillus on the cotton rat gut microbiome

Both *S. hispidus* and *S. fulviventer* received a 7-day oral gavage with a cocktail of *Lactobacillus paragasseri* BRTN and *Limosilactobacillus reuteri* SMN at a 50:50 ratio (herby referred to as *LactoX*, see **Appendix A** for isolation and characterization methods) previously isolated from *S. hispidus* stool, along with PBS-gavaged controls. Stool samples were examined by plating assay on *Lactobacillus*-selective agar and 16S rRNA sequencing to determine longitudinal changes in the gut microbiome environment.

After 7-day gavage of *LactoX* in *S. fulviventer*, *Lactobacillus* abundance significantly increased in the gut compared to PBS controls (**Figure 4.1 A**). *LactoX* inoculation did not significantly alter the gut microbiome as measured by alpha diversity (**Figure 4.1 B**, all p>0.05) and beta diversity (**Figure 4.1C**, p=0.1147, R2=0.1278) between days 0 and 7. Longitudinal DESeq2 testing of differential abundant taxa found that *Turcibacter* (l2fc=1.53, q=0.00155) and *Proteobacteria_unclassified* (l2fc=1.45, q=0.0174) increased upon *LactoX* gavage while *Lachnospiraceae_unclassified* (l2fc=-1.22, q=0.00169) and *Helicobacter* (l2fc=-1.04, q=0.117) decreased upon *LactoX* gavage (**Figure 4.1 D**).

After 7-day gavage in *S. hispidus*, gut *Lactobacillus* abundance did not significantly change compared to PBS controls (**Figure 4.1 E**). *LactoX* inoculation significantly altered

gut richness (but not by Shannon/Simpson alpha diversity indices, **Figure 4.1 F**). Beta-diversity did not significantly differ between days 0 and 7 (**Figure 4.1G**, p=0.4741, R2=0.1095). Longitudinal DESeq2 testing of differential abundant taxa found that *Blautia* (l2fc=2.82, q=0.000111) and *Bacteroides* (l2fc=1.13, q=0.00759) increased upon *LactoX* gavage while *Akkermansia* (l2fc=-2.07, q=0.0759) decreased upon *LactoX* gavage (**Figure 4.1 H**).
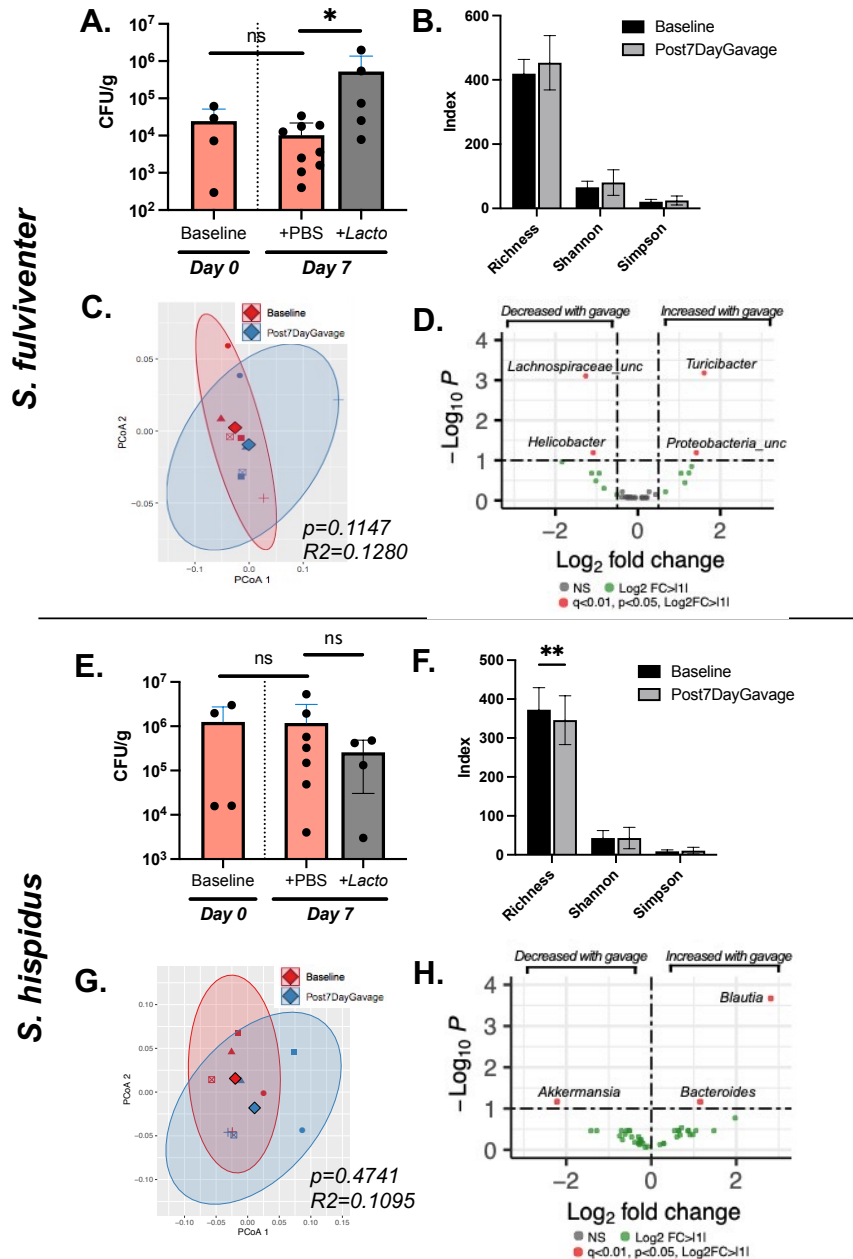
Figure 4.1. Microbiome changes upon 7-day oral supplementation of Lactobacillus in cotton rats (A-D. S. fulviventer, E-H. S. hispidus). (**A/E**) Abundance of Lactobacillus (represented by colony forming units on Lactobacillus-selective MRS agar) following Lactobacillus or PBS oral gavage. (**B/F**) Longitudinal alpha diversity metrics (richness, Shannon, Simpson) of Lactobacillus-supplemented cotton rats at baseline and following a 7-day gavage regimen. (**C/G**) Beta diversity (Bray-Curtis dissimilarity) representing community structure following a 7-day gavage regimen; statistical testing performed by PERMANOVA of geometric means. (**D/H**) Differentially abundant taxa (DESeq2) at day 7 of gavage regimen compared to baseline. All statistical testing performed using t-test unless otherwise noted.

119

ii. Oral Lactobacillus alters gene expression in healthy S. fulviventer gut and lung

As only *S. fulviventer could* be populated with *Lactobacillus* and no changes to the gut microbiome occurred following 7-day *LactoX* gavage, another experiment was conducted using only *S. fulviventer* species as outlined in **Figure 4.2** (+ *LactoX* or PBS, –RSV). Two groups of uninfected animals were administered either *LactoX* or PBS for 12 days with RSV or mock challenge on day 7. Lung and large intestine RNA was then sequenced for each animal to assess the effect of *Lactobacillus* on gut and lung gene expression.

	A +LactoX +RSV	B +LactoX -RSV	C PBS +RSV	D PBS -RSV
N	9 (lung) 9 (intestine)	5 (lung) 5 (intestine)	10 (lung) 10 (intestine)	5 (lung) 5 (intestine)
N (postQC)	9 (lung) 8 (intestine)	5 (lung) 5 (intestine)	9 (lung) 10 (intestine)	5 (lung) 5 (intestine)
M-reads	46.1 (lung) 42.7 (intest)	37.5 (lung) 32.9 (intest)	30.7 (lung) 38.3 (intest)	32.6 (lung) 41.4 (intest)
Av %GC	53% (lung) 56% (intest)	52% (lung) 56% (intest)	52% (lung) 55% (intest)	51% (lung) 54% (intest)
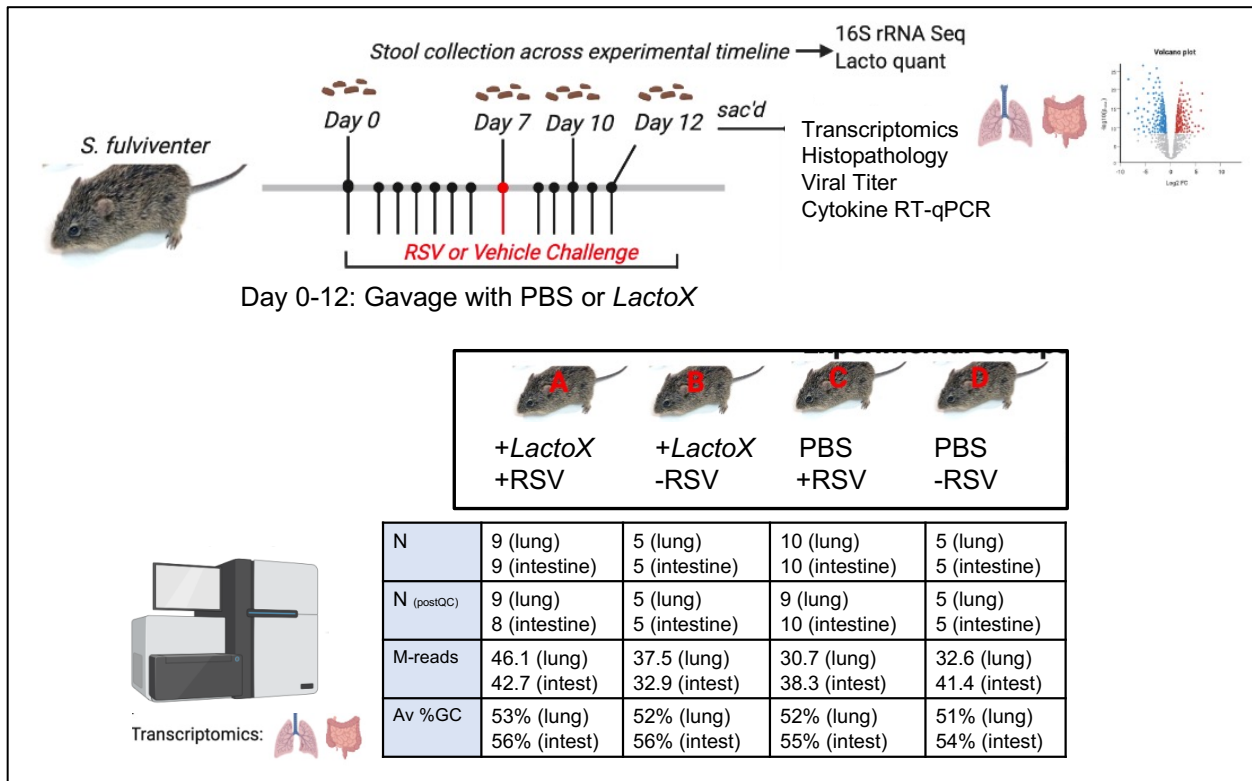
Figure 4.2. Oral *LactoX* experimental design and sequencing statistics

The presence and increased abundance of *Lactobacillus* upon *LactoX* oral gavage in *S. fulviventer* were determined by qPCR targeting all *Lactobacillus* species normalized to total bacterial load. *Lactobacillus* abundance was not significantly different between PBS and *LactoX* animals on Day 0 (p=0.8135) but increased in abundance in *LactoX* groups on Day 7 (p=<0.0001), Day 10 (p=0.0048), and Day 12 (p=0.0027) compared to PBS controls (**Figure 4.3**).
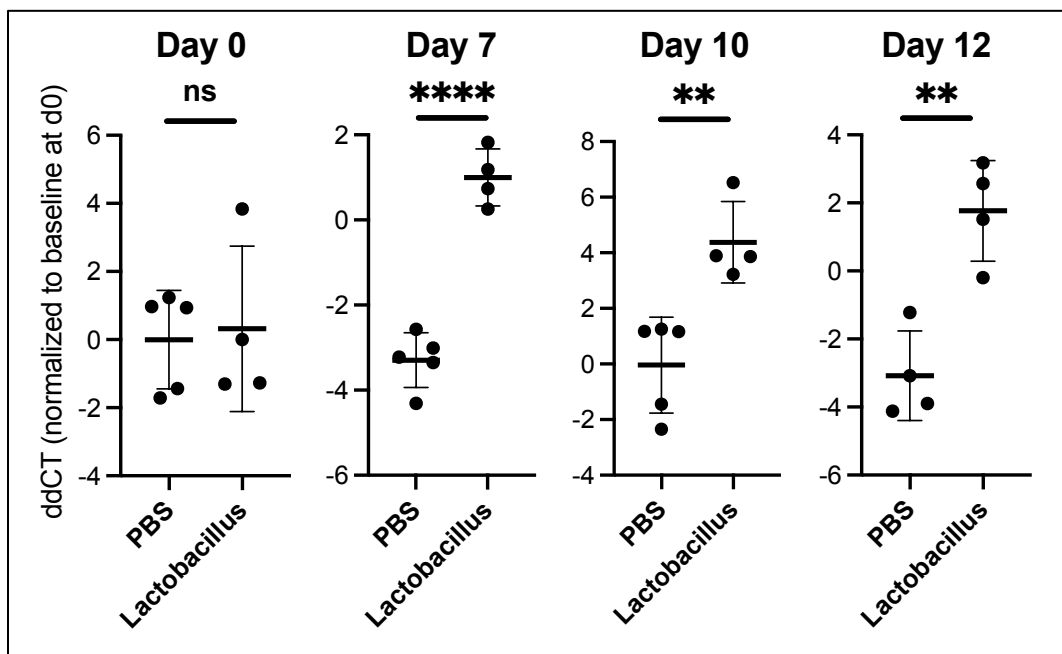
Figure 4.3. Relative abundance of *Lactobacillus* (qPCR; normalized to total bacteria) in *S. fulviventer* upon PBS Vehicle vs. *LactoX* gavage at Day 0, Day 7, Day 10, and Day 12. Delta-delta-CT values normalized to baseline abundance at Day 0 and total bacterial load on the x-axis. Statistics=students T-test; *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

In the large intestine of uninfected animals, 12-day *LactoX* gavage modulated 142 genes (+17 bacterial genes) and downregulated 9 genes (**Figure 4.4A**). Upregulated genes include Ig heavy chain V region IR2 (*HVRO1*), Low-density lipoprotein receptor-

related protein 2 (*LRP2*), Immunoglobulin gamma-1 heavy chain (*IGG1*), Phospholipase A2 membrane-associated (*PA2GA*), TRAF-interacting protein with FHA domain-containing protein A (*TIF1A*), Marginal zone B- and B1-cell-specific protein (*MZB1*), 2'-5'-oligoadenylate synthase-like protein 2 (*OASL2*), Mucin-16 (*MUC16*), Retinoic acid receptor responder protein 1 (*TIG1*), Heparin-binding EGF-like growth factor (*HB-EGF*), and Armadillo-like helical domain-containing protein 4 (*ARMD4*) as determined by DESeq2 (**Figure 4.4B,** blue cells). Downregulated genes included Mucin-6 (*MUC6*), Antithrombin-III (*ANT3*), Spermatogenesis-associated protein 46 (*SPT46*), Cilia- and flagella-associated protein 74 (*CFA74*), Cysteine and glycine-rich protein 2 (*CSRP2*), Carboxylesterase 1D (*EST1D*), Vomeronasal type-2 receptor 116 (*V2116*), E3 ubiquitin-protein ligase (*NEDD4*), and Collagen alpha-2(XI) chain (*COBA2*) (**Figure 4.4B**, red cells). To deduce function, all differentially expressed genes were classified and grouped based on their PANTHER protein classes. Upregulated genes were related to defense/immunity, metabolite interconversion, transmembrane signaling, intercellular signaling, and scaffolding/adapters/structural proteins (**Figure 4.4C**, blue bars). Downregulated genes were related to the extracellular matrix, cytoskeleton, transmembrane signaling, and protein-binding activity modulators (**Figure 4.4C**, red bars).

In the lung of uninfected animals, 12-day oral *LactoX* gavage resulted in 88 upregulated genes and 44 downregulated genes (**Figure 4.4D**). Upregulated genes include Mitochondrial phosphate carrier protein (*MPCP*), Proteoglycan 4 (*PRG4*), Mucin-16 (*MUC16*), R-spondin-1 (*RSPO1*), Espin-like protein (*ESPNL*), Testin (*TES*), Ras and Espin (*ESPN*), insulin-like growth factor binding protein 1 (*IGFBP1*), Armadillo-like helical

domain-containing protein 4 (*ARMD4*), and Retinoic acid receptor responder protein 1 (*TIG1*) as determined by DESeq2 (**Figure 4.4E**, blue cells). Downregulated genes included Adiponectin (*ADIPO*), Carbonic anhydrase 3 (*CAH3*), Myosin regulatory light chain 2 atrial isoform (*MLRA*), Myosin-6 (*MYH6*), Telethonin (*TELT*), Titin (*TITIN*), E3 ubiquitin-protein ligase NEDD4 (*NEDD4*), Thrombospondin-2 (*TSP2*), Vascular endothelial growth factor receptor 2 (*VGFR2*), and Transforming growth factor beta-1 proprotein (*TGFB1*) (**Figure 4.4E**, red cells). All upregulated genes were related to metabolite interconversion, transcriptional regulation, protein binding/modifying activity, and cell adhesion (**Figure 4.4 F**, blue bars). Downregulated genes were related to transporters, extracellular matrix, cytoskeleton, and protein binding/modifying activity (**Figure 4.4F**, red bars).

Of all differentially expressed genes, 5 genes were differentially expressed in both the lung and intestine upon *LactoX* gavage. These genes included *NEDD4* (downregulated in both lung and intestine), *COBA2* (upregulated in lung, downregulated in intestine), *ARMD4* (upregulated in both lung and intestine), *TIG1* (upregulated in both lung and intestine), and *MUC16* (upregulated in both lung and intestine).
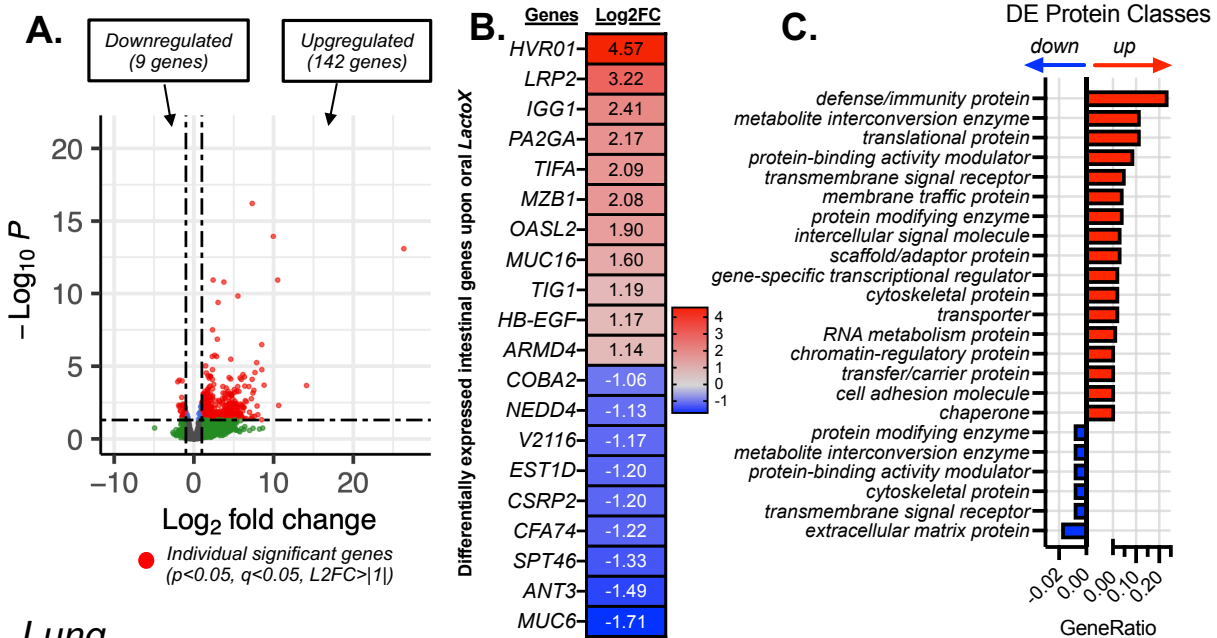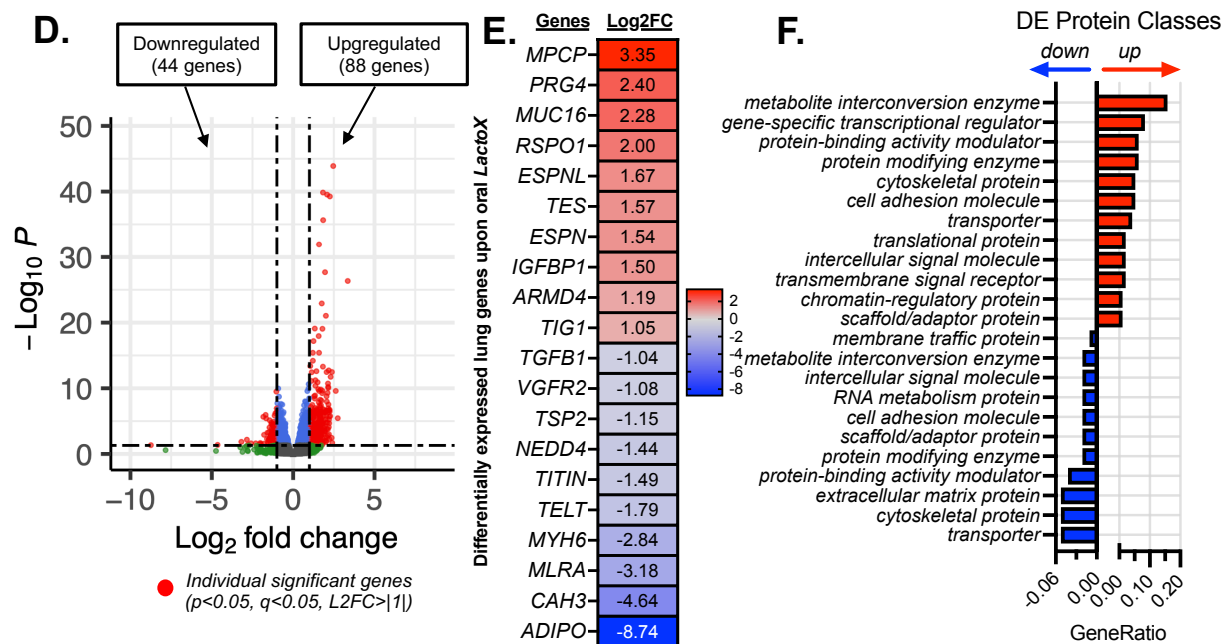
Figure 4.4. Modulation of gene expression in the uninfected (A-C) large intestine and (D-F) lung upon 12-day *LactoX* oral gavage regimen. (A, C) Differentially expressed genes (DEs) by oral *LactoX*. (B, E) Top 20 DEs upregulated (red) and downregulated (blue) by oral *LactoX*; all genes q<0.05. (C, F) Differentially expressed (DE) PANTHER Protein Classes (based on differentially expressed genes with p<0.05, q<0.05) on the y-axis with relative expression of protein classes on the x-axis (expressed as GeneRatio, which is the proportion of differentially expressed genes in each category to the total number of differentially expressed genes). Positive GeneRatio (red) = increased expression in *Lactobacillus*-gavaged animals; negative GeneRatio (blue) = decreased expression in *Lactobacillus*-gavaged animals.

Another group of *S. fulviventer* was orally administered either PBS or *LactoX* for 12 days and intranasally infected with 10^5 PFU of RSV A/Long or PBS (as mock infection control) on Day 7. Stool was collected throughout the experiment, which concluded on Day 12 (Day 5 post-infection) as outlined in **Figure 4.2** (+*LactoX* or PBS, +RSV). 16S rRNA sequencing was performed on DNA extracted from stool to assess RSV-induced changes to the gut microbiome.

Upon RSV infection in PBS-gavaged animals, gut microbiome richness was significantly reduced at Day 3 post-infection compared to baseline (Observed OTUs p=0.0108, **Figure 4.5A**), although not by alpha diversity metrics (Shannon/Simpson Index p=ns, **Figure 4.5 B-C**). In *LactoX* animals, no drop in richness occurred upon RSV infection compared to baseline (Observed OTUs, p=0.8874, **Figure 4.5D-F**). Changes in microbiome community structure were analyzed according to beta diversity (Bray-Curtis dissimilarity), which confirmed a significant change in PBS+RSV animals upon infection (PERMANOVA p=0.004, R^2=0.129) but not in *LactoX*+RSV animals (PERMANOVA p=0.062, R^2=0.111) (**Figure 4.6A-B**). These RSV-induced microbiome changes were apparent on Day 3 post-infection (adonis.pair adjusted p=0.027) and Day 5 post-infection (adonis.pair adjusted p=0.024) compared to baseline/Day 0 (**Figure 4.6A**). There were no significant RSV-induced changes to microbiome community structure in infected *LactoX* animals (Day 3 adjusted p=0.13, Day 5 adjusted p=0.13) (**Figure 4.6B**).
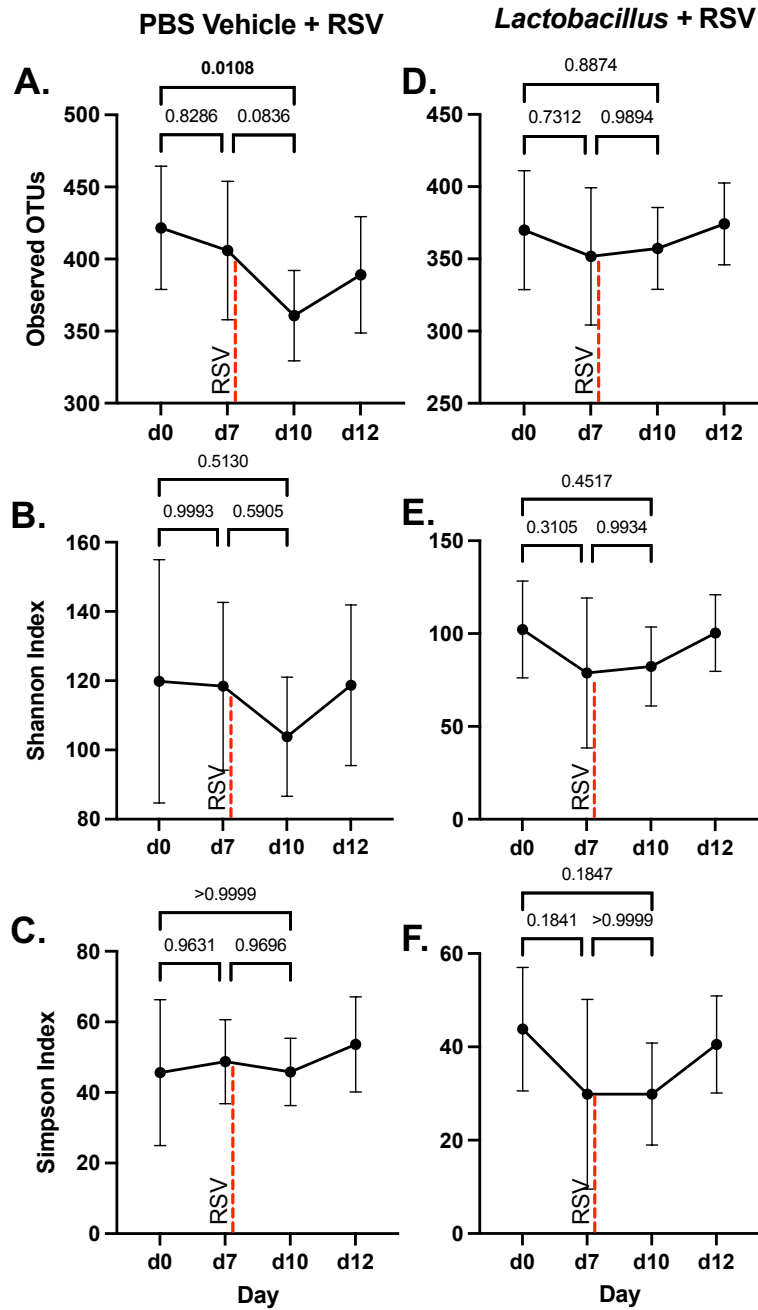
Figure 4.5. Alpha diversity changes in PBS Vehicle (A-C) or *LactoX* (D-F) gavaged animals pre-infection (Day0-Day7) and post-infection (Day7-Day10; indicated by red line) with RSV. Three alpha diversity metrics are presented on the x-axis: Richness (Observed OTUs), Shannon Index, and Simpson Index. Significant drop in richness occurs from Day 0 – Day 10 in PBS animals but not in *Lactobacillus*-gavaged animals. Statistics performed with ANOVA followed up Tukey posthoc test; figures generated in Prism 8.
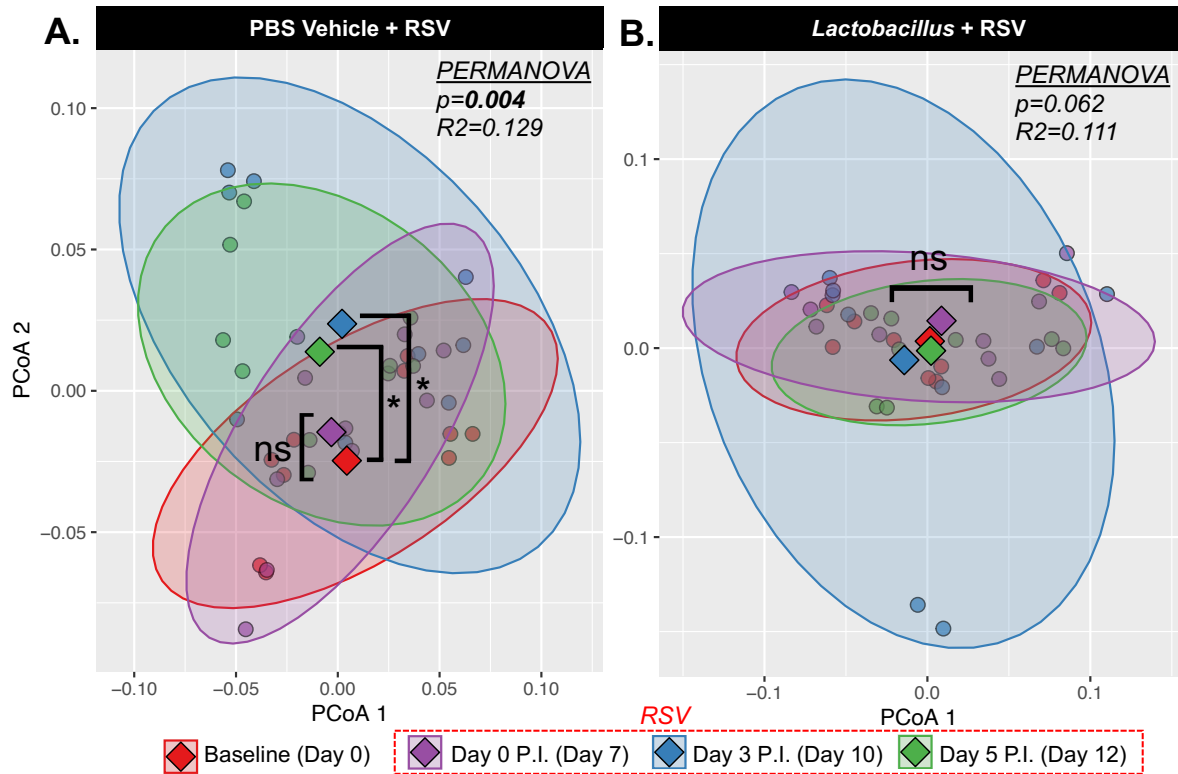
A. PBS Vehicle + RSV

PERMANOVA
p=0.004
R2=0.129

B. *Lactobacillus* + RSV

PERMANOVA
p=0.062
R2=0.111

ns

Day

Day

Day

ns

RSV

Baseline (Day 0)   Day 0 P.I. (Day 7)   Day 3 P.I. (Day 10)   Day 5 P.I. (Day 12)

PCoA 1

...a diversity (Bray-Curtis
...of fecal collection reveals
...-gavaged animals on Day
...s seen in *Lactobacillus*-
...RMANOVA between day
...ay comparisons, *=P≤0.05, **=P≤0.01,
...=P≤0.0001.

iv. Oral Lactobacillus reduces RSV histopathology but not viral titers

Upon completion of the 12-day gavage and 5-day infection, lungs from all animals were harvested for histopathology and blindly scored for 4 inflammatory parameters as described in the methods: peribronchiolitis, perivasculitis, interstitial pneumonia, and alveolitis (**Figure 4.7 A-B**). Infected animals (+RSV+PBS vehicle) had a significantly larger cumulative pathology score compared to uninfected controls, indicating successful infection (Tukey's multiple comparisons test p=0.0041) (**Figure 4.7A**). However, infected animals receiving *LactoX* (+RSV+*LactoX*) had a cumulative pathology score more like uninfected animals (p=0.7348) compared to RSV-infected animals receiving PBS vehicle (p=0.0093). These trends can also be seen in the reduction of inflammatory cell infiltration in **Figure 4.7B**. Viral titers were assessed by RT-qPCR and plaque assay in the nose (**Figure 4.7 C-D**) and the lung (**Figure 4.7 E-F**). Both viral RNA and plaques were significantly higher in all infected groups compared to uninfected groups, but *LactoX* treatment did not significantly alter virus levels compared to PBS-gavaged infected controls.
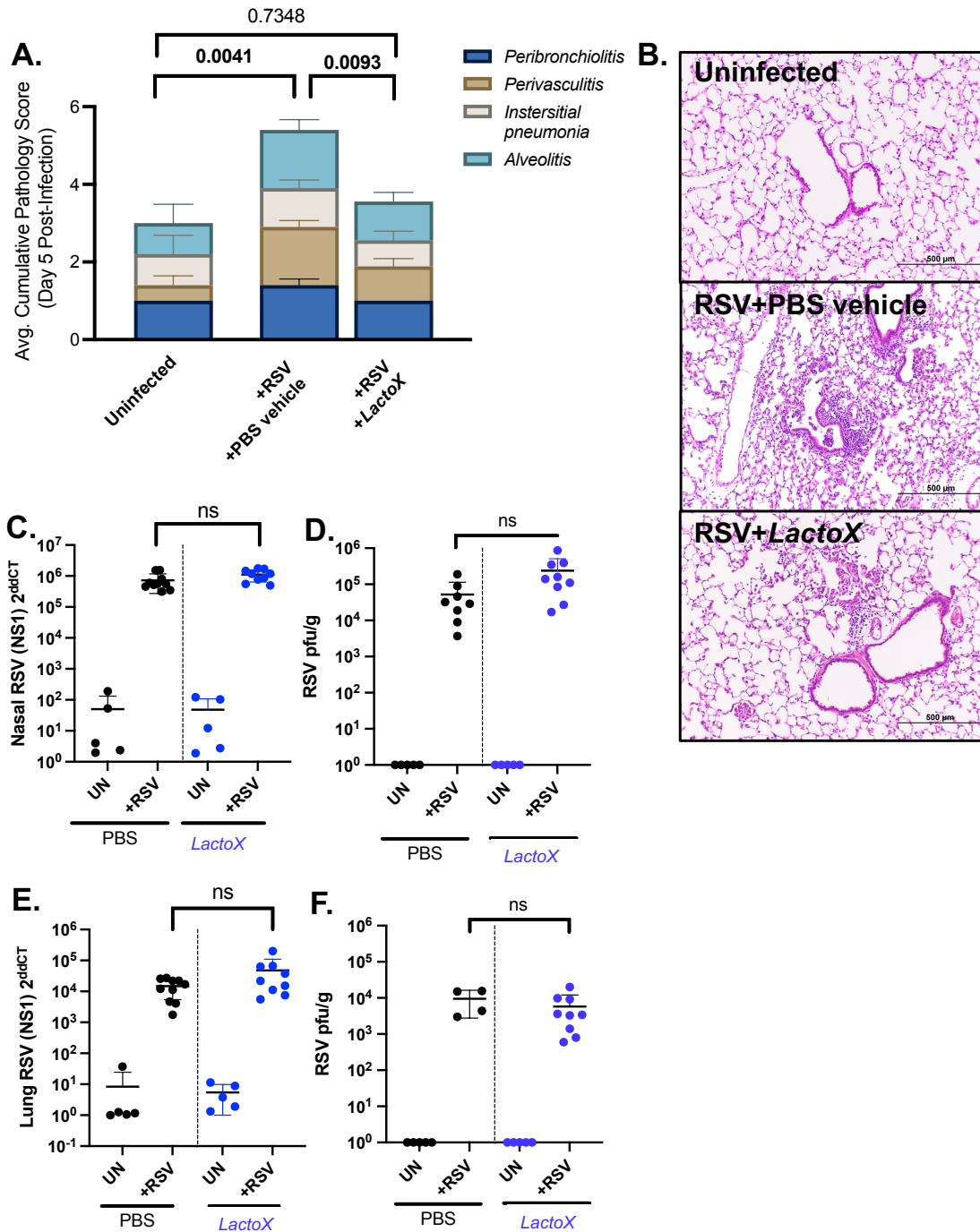
Figure 4.7. *Lactobacillus*-induced protection from severe lung histology but not viral load. (A) Blindly generated pathology scores were cumulated to compare healthy and RSV-infected lung tissue treated with PBS vehicle or *LactoX*. (B) Representative images for histological imaging. (C-D) Nasal RSV titers measured by (C) RT-qPCR and (D) plaque assay. (E-F) Lung viral titers measured by (E) RT-qPCR and (D) plaque assay. RT-qPCR performed with primers targeting the RSV NS1 protein and normalized to β-actin; fold change (calculated by $2^{-\Delta\Delta CT}$) on the y-axis. Statistics performed using ANOVA and Tukey posthoc in Prism 8; *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

v. Oral Lactobacillus significantly alters lung gene expression in RSV-infected animals

RNA was extracted from all harvested lungs and sequenced to assess transcriptomic changes by RSV and *Lactobacillus*. Clustering of samples by transcript counts (Euclidean distance) indicated differences between uninfected (PBS) and infected (PBS_RSV) groups; *LactoX*-gavaged groups infected with RSV clustered in between these groups, suggesting more similarity to uninfected samples (**Figure 4.8A**).

To determine genes modulated in the lung by gut *Lactobacillus* during RSV infection, comparisons were made between significantly modulated genes in RSV vs. Uninfected groups (**Figure 4.8B**) and *LactoX*+RSV vs. Uninfected groups (**Figure 4.8C**). In infected animals receiving *LactoX*, 648 unique transcripts were upregulated and 384 downregulated. In infected animals receiving PBS vehicle, 663 unique transcripts were upregulated and 113 downregulated. 454 transcripts were upregulated upon infection in both *LactoX* and PBS groups, and 33 were downregulated (**Figure 4.8D**). Transcripts were aligned to the *S. fulviventer* transcriptome reference for functional annotation (**Chapter 3**).
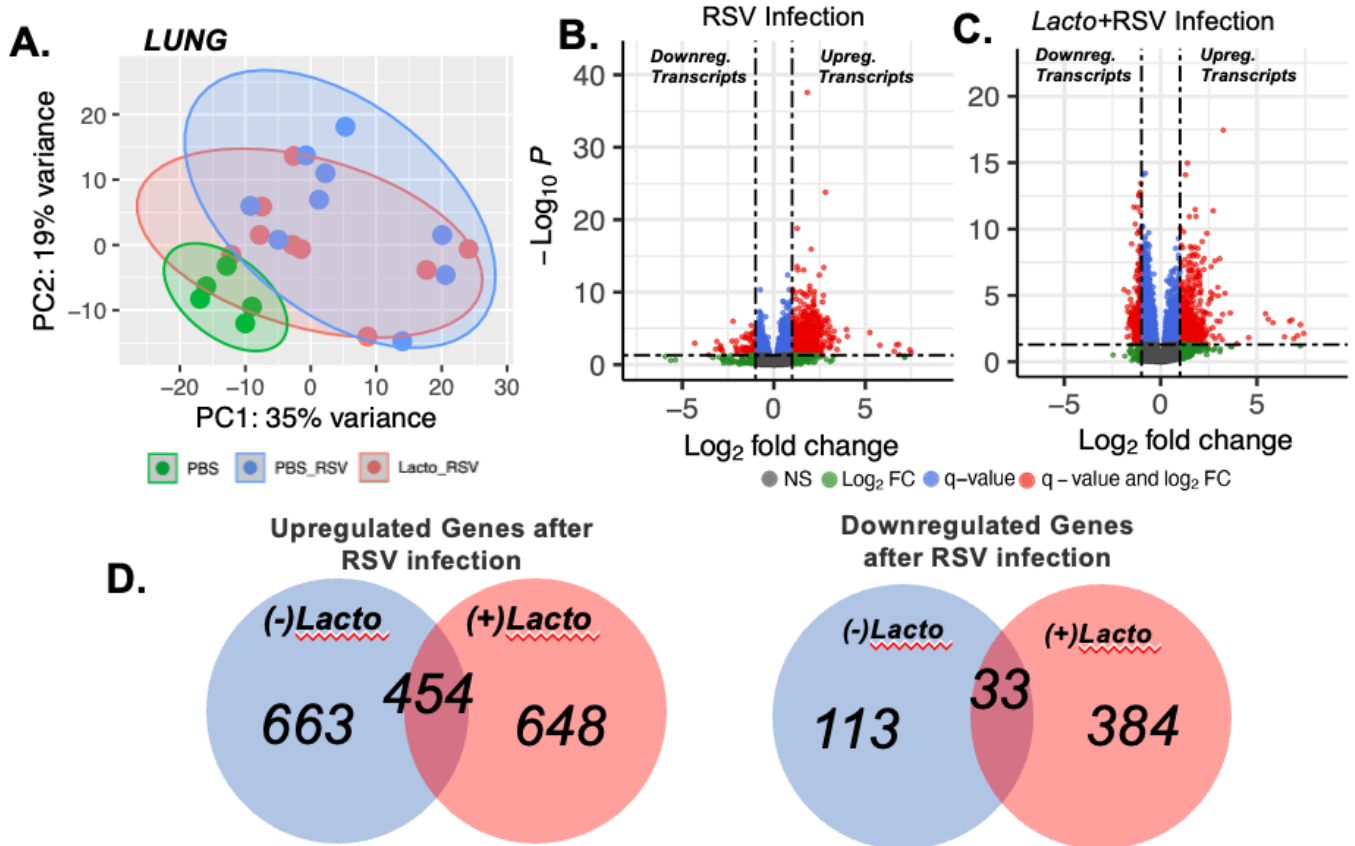
Figure 4.8. *Lactobacillus*-induced transcriptome changes during RSV infection. (A) Clustering of transcript counts by Euclidean distance using Principal Component Analysis. (B) DE transcripts in PBS+RSV and (C) *LactoX*+RSV animals. Red dots indicate transcripts that met significance threshold (Log2FC>|1|, p<0.05, q<0.05). Number of unique and overlapping DE transcripts during RSV infection with or without *LactoX*.

Many genes were differentially expressed between PBS+RSV and *LactoX*+RSV animal groups (**Figure 4.9**). For comparisons, genes that were differentially expressed in one group but not the other were investigated. Lung genes downregulated by *LactoX* during RSV infection include Neuropeptide Y Receptor Y4 (*NPY4R*)*,* Zinc transporter 2 (*ZNT2*), Cadherin-3 (*CADH3*)*,* Adhesion G protein-coupled receptor B1 (*AGRB1*), Claudin-15 (*CLD15*), Armadillo-like helical domain-containing protein 4 (*ARMD4*)*,* Vasoactive intestinal polypeptide receptor 2 (*VIPR2*)*,* Interleukin-17 receptor E (*I27L2*), Carboxypeptidase X, M14 Family Member 2 (*CPXM2*), Fibulin 2 (*FBLN2*), Pappalysin 2 (*PAPP2)*, and Complement C1q TNF-Related Protein 6 (*C1QT6*) (**Figure 4.9A**). Three of these genes were also found to be downregulated by *Lactobacillus* in the lung of uninfected animals: *CPXM2*, *FBLN2,* and *PAPP2* (**Figure 4.4E**). These downregulated genes are related to G-protein coupled receptors, transmembrane signaling receptors, scaffolding/adaptor proteins, and protein-binding activity monitors as determined by PANTHER protein functions.

Lung genes upregulated by *LactoX* during RSV infection include CXC motif chemokine ligand 9 (*CXCL9*), C-X-C Motif Chemokine Ligand 10 (*CXCL10)* Radical S-Adenosyl Methionine Domain Containing 2 (*RSAD2*), Interferon Induced Protein 44 (*IFI44*), NLR Family CARD Domain Containing 5 (*NLRC5*), Forkhead Box K2 (*FOXK2*), Retinoic acid receptor responder protein 1 (*TIG1*), Insulin-like growth factor binding protein 1 (*IGFBP1*), Titin (*TITIN*), Titin-cap (*TELT*)*,* and Carbonic Anhydrase 3 (*CAH3*) (**Figure 4.9B**). Many of these genes are related to chemokines, actin filament binding, immunoglobulin structure, and cell structural proteins. Th2-inducing cytokines IL-4, IL-13, and IL-33 were also downregulated in *LactoX*-gavaged groups (**Figure 4.9C-E**).

A. Downregulated Genes

Gene	PBS+RSV	LactoX+RSV	Protein Function
NPY4R	4.0	●	G-protein coupled receptor
EXTL1	2.3	●	G-protein coupled receptor
AGRB1	2.2	●	G-protein coupled receptor
ZNT2	3.1	●	Transporter
CADH3	1.9	●	Transmembrane signaling receptor
VIPR2	1.2	●	Transmembrane signaling receptor
I17RE	1.2	●	Transmembrane signaling receptor
GDF2	2.3	●	Growth factor
CLD15	1.8	●	Tight junction
ARMD4	1.8	●	Membrane protein
ESPN	1.6	●	Non-motor actin binding prot
IRGM1	1.0	●	IFN-inducible GTPase
I27L2	1.1	●	Apoptosis signaling protein
CPNE1	1.1	●	Calcium binding protein
HBB	●	-1.2	Globulin
CPXM2	●	-1.7	Protease
NMRL1	●	-1.4	Nitrogen metabolic protein
FBLN2	●	-1.4	ECM structural protein
SMTN	●	-1.2	Scaffolding/adapter protein
C1QRF	●	-1.0	Scaffolding/adapter protein
C1QT6	●	-1.0	Scaffolding/adapter protein
MK	●	-1.1	Carbohydrate kinase
PAPP2	●	-1.1	Protein-binding activity monitor
PDCD5	●	-1.1	Protein-binding activity monitor
FLNA	●	-1.1	Integrin signaling protein

FoldChange from Uninfected Controls

B. Upregulated Genes

Gene	PBS+RSV	LactoX+RSV	Protein Function
CXCL9	●	4.6	Chemokine
CXL10	●	2.7	Chemokine
RSAD2	●	3.1	Iron-sulfur binding protein
IFI44	●	2.8	IFNα/β-inducible protein
CYC	●	2.0	Mitochondrial electron transport
NLRC5	●	1.6	Scaffolding/adapter protein
TES	●	1.6	Scaffolding/adapter protein
TKN4	●	1.4	Receptor binding
TLN2	●	1.4	Integrin binding
FOXK2	●	1.4	Transcription factor
TIG1	●	1.2	Transcription factor
IGFBP1	●	1.3	Protease inhibitor
GBP5	●	1.3	G-protein
KVM5	-1.4	●	Immunoglobulin subunit
TELT	-1.7	●	Actin filament binding
NRAP	-2.1	●	Actin filament binding
MY18B	-2.2	●	Actin filament binding
MYH6	-3.2	●	Actin filament binding
EF1A2	-2.9	●	Translation factor
MYOM2	-1.6	●	Structural protein
TITIN	-1.6	●	Structural protein
ANGT	-2.0	●	Protease inhibitor
PRS29	-2.0	●	Serine protease
UN45B	-2.3	●	Hsp90/chaperone binding
CAH3	-4.3	●	Dehydratase

Log2 FoldChange

down ← | → up   Baseline

-4 -2 0 2 4

● = not significantly DE

C. IL-4 p=0.020

D. IL-13 p=0.050

E. IL-33 p=0.036

Normalized Counts

PBS_RSV LactoX_RSV

Figure 4.9. (A) Down- and (B) upregulated genes by *Lactobacillus* during RSV infection. All significant DEs are indicated as blue (downregulated) and red (upregulated) cells as determined by their Log2 FoldChange to uninfected controls. Dots indicate the gene was not DE compared to controls. (C-E) Normalized expression levels (TMM) for Th2 cytokines; statistics performed using Student's T-test in Prism 8.

To confirm differentially expressed genes determined by RNA-seq, RT-qPCR was performed on the same lung RNA that was sequenced. Eight genes were randomly selected for RT-qPCR confirmation of relative expression (**Figure 4.10A**). RT-qPCR analysis confirmed an increase in *CADH3* and *CLD15* expression in PBS+RSV animals but not in *LactoX*+RSV animals compared to uninfected controls, which confirms changes seen by RNA-Seq (**Figure 4.9B**). RT-qPCR also confirmed higher expression of *CXCL9* in *Lacto*+RSV animals but not PBS+RSV animals compared to controls, as shown by RNA-seq data (**Figure 4.9A-B**). *VIPR2* expression was higher in PBS+RSV and *Lacto*+RSV animals (RT-qPCR), but RNA-seq only showed upregulation of *VIPR2* in PBS+RSV animals. Although insignificant, RT-qPCR showed an increase in expression of *MX2* in both PBS+RSV and *LactoX*+RSV groups compared to uninfected controls; RNA-Seq showed a significant increase. *ISG15* did show increased expression in PBS+RSV and *Lacto*+RSV (p=0.06) animals, and RNA-seq data suggested a significant upregulation in any group. Lung expression of Suppressor Of Cytokine Signaling 3 (*SOCS3*) and Heparin-binding EGF-like growth factor (*HB-EGF*) were not shown to be significantly upregulated in any group as determined by RT-qPCR and RNA-seq.

To confirm changes in Log2FoldChange (L2FC) between PBS+RSV and *Lacto*+RSV, ΔL2FC values were calculated by the formula ΔL2FC=(L2FC PBS+RSV)–(L2FC *LactoX*+RSV) for both RT-qPCR (2^ddCT) and RNA-seq data (DESeq2 L2FC values) (**Figure 4.10B**). Correlation analysis between RT-qPCR and RNA-seq L2FC revealed a significant correlation between the two methods (Spearman's rho p=0.0025, R^2=0.9096).

A.

B.

	RNA-Seq		RT-qPCR	
	PBS+RSV	*LactoX*+RSV	PBS+RSV	*LactoX*+RSV
CADH3	1.9***	ns	2.4***	ns
VIPR2	1.2**	ns	2.5***	1.7*
CXCL9	ns	4.6**	ns	1.6*
CLD15	1.8*	ns	1.9*	ns
ISG15	1.2****	1.9*	2.0*	1.5*
MX2	1.9****	2.1***	1.4^{ns}	1.2^{ns}
SOCS3	ns	ns	ns	ns
HBEGF	ns	ns	ns	ns

C. Change in Expression Confirmation: PBS+RSV vs. *LactoX*+RSV
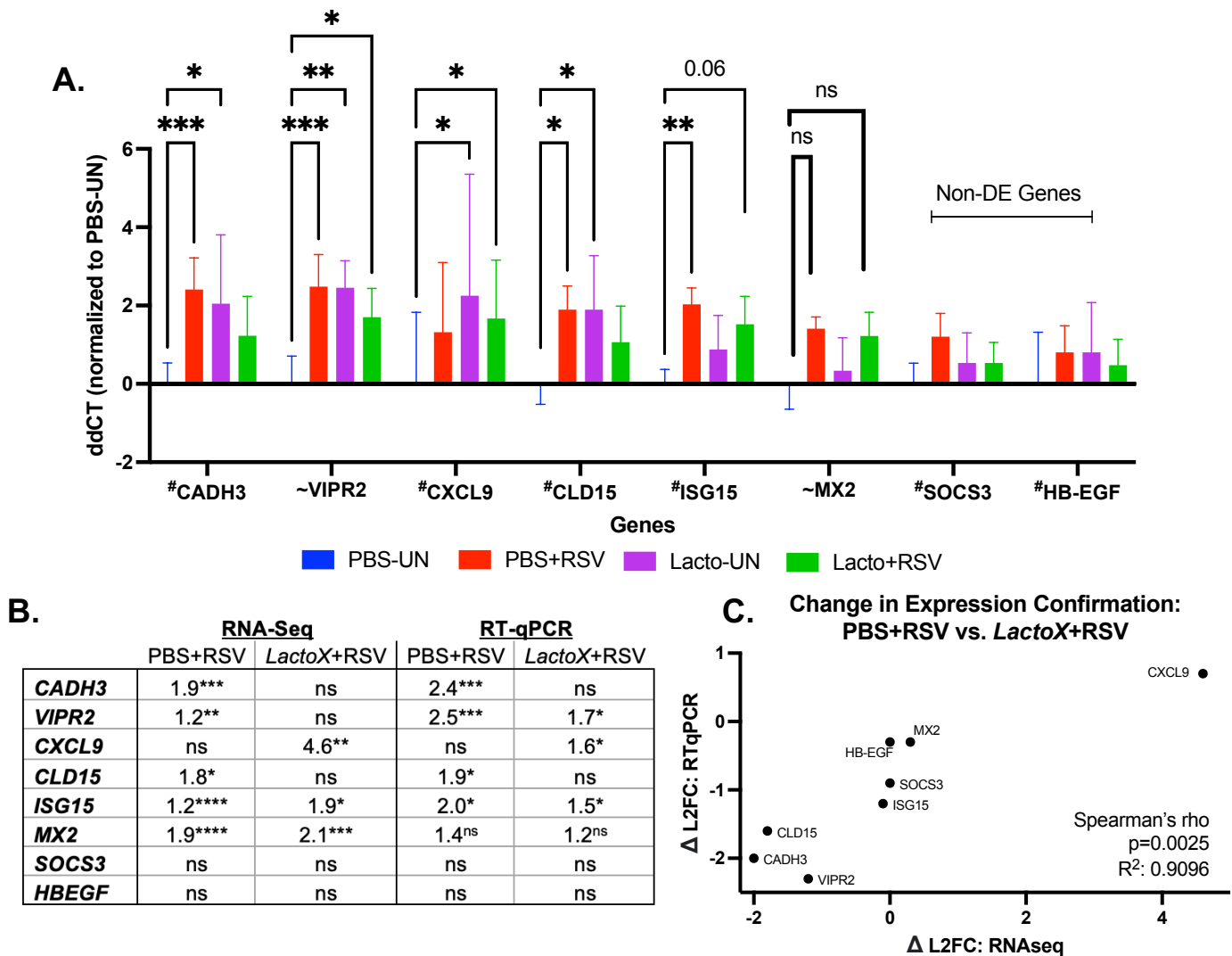
Figure 4.10. Confirmation of RNA-seq data by RT-qPCR. (**A**) 8 genes were selected for confirmation by RT-qPCR, measured by ddCT method (normalizing each sample to both B-actin and GAPDH, followed by normalization to Uninfected controls). 6 of the 8 genes directly correlate with RNA-seq data (indicated by [#]). 2 of the 8 genes were directionally but not statistically correlated (indicated by ~). (**B**) Fold changes between groups to Uninfected Controls for RNA-Seq and RT-qPCR. Significance stars: RNA-Seq=q values (FDR), RT-qPCR=Tukey multiple comparisons p values. (**C**) Correlation analysis between RNA-seq and RT-qPCR log2 fold change (L2FC) results from the same RNA sample. Spearman's correlation analysis shown. *CADH3*: Cadherin-3; *VIPR2:* Vasoactive intestinal polypeptide receptor 2; *CXCL9:* CXC motif chemokine ligand 9; *CLD15:* Claudin-15; *SOCS3:* Suppressor of Cytokine Signaling 3; *HB-EGF:* Heparin-binding EGF-like growth factor, *MX2:* Interferon-induced GTP-binding protein Mx2. *=P≤0.05, **=P≤0.01, ***=P≤0.001, ****=P≤0.0001.

The data presented in this chapter suggest a significant role of gut *Lactobacillus* in protecting from severe RSV infection in cotton rats for the first time. Not only did oral *Lactobacillus* protect animals from more severe lung histopathology, but transcriptomic analysis also indicated significant modulation of gene expression in the lung. These modulated pathways suggest several mechanisms involved in the gut-lung axis. Additionally, this data presents the cotton rat as a novel model for microbiome-viral-host interactions.

Gut *Lactobacillus* abundance was only increased after *LactoX* gavage in *S. fulviventer* but not *S. hispidus*. Gut *Lactobacillus* was also of higher abundance at baseline (Day 0) in *S. hispidus than* in *S. fulviventer*. Strickland et al. previously showed that the healthy gut flora of *S. hispidus* is dominated by several strains of *Lactobacillus*– specifically the two strains used in this study–compared to *Lactobacillus*-deficient *S. fulviventer* (181). Interestingly, this study shows that only *Lactobacillus*-deficient *S. fulviventer* was significantly populated with *Lactobacillus* upon oral gavage compared to *S. hispidus*. While oral *Lactobacillus* did not alter the gut microbiome community structure in either species, the presence and increased abundance of *Lactobacillus* in the healthy gut did not allow for further modulation of *Lactobacillus* by oral gavage. This suggests

that such probiotic regimens may not affect animals with an increased abundance of resident Lactobacillus.

This study also showed a significant drop in microbiome richness upon RSV infection, confirming similar outcomes in both humans (124) and mice (319-321). Oral *Lactobacillus* protected infected animals from such changes in the microbiome (**Figure 4.5-4.6**), which could be indicative of reduced systemic inflammation and gut-lung axis mechanisms. While no significant change in richness was seen in uninfected animals receiving oral *Lactobacillus*, it is still unclear if this change is due strictly to oral gavage of bacteria or protection from a viral phenotype.

Many studies have reported many host-transcriptional changes upon *Lactobacillus* supplementation. Similarly, oral *LactoX* was shown to modulate many genes in both the healthy intestine and lung. Of these genes, *MUC16, TIG1*, and *ARMD4* were upregulated in both the lung and intestine. *MUC16*, or Mucin-16, is an extracellular-associated mucin that fortifies the glycocalyx to protect the cell surface in normal epithelia. Increased expression can improve mucosal barrier function and protection against infection, and *Lactobacillus* has been previously shown to upregulate other associated extracellular mucins (322). *TIG1*, or Tazarotene-induced gene 1, plays a role in retinoic acid-mediated cellular differentiation and tumor growth suppression, which can help promote the epithelial barrier (323). *MZB1* (Marginal zone B- and B1-cell-specific protein) was also upregulated by *LactoX* in the healthy intestine. *MZB1* is a molecular chaperone facilitating the formation of polymeric IgA (324), an immunoglobulin previously shown to be induced by *L. rhamnosus GG* (325) and promote RSV immunity (326). Additionally, *MZB1* has been shown to increase IgA secretion in the gut and induce anti-inflammatory effects in

the mucosa (327), indicating that this may be a mechanism by which *Lactobacillus* upregulates IgA. TGF-β was downregulated in the healthy lung but not modulated in the intestines (**Figure 4.4 B,E**). Gut *Lactobacillus* has been shown to upregulate TGF-β to promote regulatory T cells in the intestines (328, 329). However, gut *Lactobacillus* seems to exert TGF-β-suppressive effects in other tissues (liver and heart) (330) and breast milk (331) due to the fibrotic tissue response to TGF-β in these tissues (332). The suppressive effect of gut *Lactobacillus* on lung TGF-β suggests a protective gut-lung axis mechanism of protecting against collagen accumulation and tissue fibrosis in the lung that can contribute to long-term RSV outcomes (333). One of the most interesting upregulated gene in the intestine is *HB-EGF* due to previously published literature on the *Lactobacillus*-secreted protective protein *p40* in increasing *HB-EGF* production in the intestines. These genes could be potential mediators of gut-lung axis activity that warrants further experimentation.

Supplementation of *LactoX* into healthy cotton rats confirmed several genes previously shown to be upregulated in the murine gut by oral *Lactobacillus.* These include Immunoglobulin G1 (334) (*IGG1*, upregulated in intestines), Phospholipase A2 (335) (*PA2GA*, upregulated in intestines), TRAF-interacting protein with FHA domain-containing protein A that reduced colitis inflammation (336) (*TIFA*, upregulated in intestines), C-C motif chemokine 5 (337) (*CCL5*, upregulated in intestines), T-cell-specific guanine nucleotide triphosphate-binding protein 2 (338) (*TGTP2*, upregulated in intestines), and Cadherin-1 in mice (339) (*CADH1*, upregulated in intestines). However, further experimentation is needed to further confirm *Lactobacillus*'s direct role in upregulation of these genes.

Oral *LactoX* administration prior to RSV infection significantly reduced histopathology scores and gene expression profiles compared to infected animals receiving PBS vehicle. These effects were not mediated by inhibition of viral replication, as lung and nasal viral titers were not reduced at 5 days post-infection in *Lactobacillus*-gavaged animals (**Figure 4.7**). Contrarily, Eguchi et al. reported a reduction in lung RSV titers in mice treated with oral *Lactobacillus gasseri* SBT2055 that correlated with protection from weight loss and inflammatory cytokines (133). However, it is unclear whether the *Lactobacillus* groups were more protected from RSV due to the action of *Lactobacillus* or due to lower viral titers during infection. Our study supports the significant role of *Lactobacillus* in the protective outcomes as there was no difference in viral titers between groups. Additionally, since cotton rats are more susceptible to RSV infection than mice, these results may be more translatable for outcomes hypothesized in humans.

Several genes were differentially expressed in the lung upon RSV infection in both *LactoX* and PBS groups. Comparison of differentially expressed genes (DEs) in lungs between PBS+RSV and *LactoX*+RSV animals were investigated, as this would suggest a *Lactobacillus*-mediated role in changing the expression of these genes. *CXCL9* and *CXL10* are upregulated in *Lacto*+RSV animals but not PBS+RSV animals, indicating a potential role of these chemokines in reducing RSV severity. These IFN-y-inducible chemokines function to recruit Th1 and NK cells to the site of infection, suggesting a less severe and less-Th2-biased response to RSV (340). Upregulation of *CXCL9* and *CXL10* have been previously shown in RSV Th1-associated expression profiles in humans (341). Additionally, data from the INSPIRE birth cohort at VUMC found that RSV-infected infants with decreased expression of *CXCL9* and *CXL10* were at a higher risk for developing

wheeze at 1-2 years (342). Previous literature has also shown *CXCL9* to be upregulated by *L.plantarum* in rats, further confirming this *Lactobacillus*-induced phenotype (343). This increase in expression was also confirmed by RT-qPCR (**Figure 4.10**). IL-4, IL-13, and IL-33 were also downregulated in the lung in *LactoX*-gavaged, RSV-infected animals. These type-2 inflammatory cytokines secreted by mast cells/eosinophils/Th2 (IL-4, IL-13) and damaged epithelia (IL-33) suggest decreased activation and recruitment of both ILC2 and Th2 cells.

There were no RSV-related receptors differentially expressed in this study. However, several host factors involved in restricting RSV pathogenesis were found to be upregulated. Interferon-induced protein 44 (*IFI44*) was upregulated in the lung by *Lactobacillus* during infection, which has been previously shown to restrict RSV infection in mice and associated with decreased RSV severity in children (344). Guanylate-binding protein 5 (*GBP5*) was also upregulated by *Lactobacillus* during infection, which has been associated with reduced RSV susceptibility in children (345). *GBP5* can restrict the replication of RSV by decreasing intracellular SH protein necessary for capsid construction, although some RSV G proteins can stimulate the uniquitin-mediated degradation of GBP5 (345, 346).

Another gene upregulated in the lung by *Lactobacillus* during infection is insulin-like growth factor (IGF) binding protein 1 (*IGFBP1*). *IGFBP1* has not been associated with RSV infection, but the IGF1 receptor (*IGFR1*) has been recently described in the receptor-mediated entry of RSV into host cells (41). RSV-F can bind the back of *IGFR1* without disrupting the *IGF1* binding site, which allows viral entry into the cell once *IGF1* ligand binds to its receptor (this claim is based on soon-to-be-published data from the Marchant

laboratory, conveyed through correspondence with the supervising author). *IGFBP1* can bind *IGF1* with equal or greater affinity than the receptor to block IGF signaling (347), which would effectively block potential binding and subsequent *IGFR1*-mediated entry of RSV into the cell. *IGFBP1* was also upregulated in the lung upon *LactoX* supplementation in healthy animals, further confirming its activity in *Lactobacillus*-mediated protection phenotypes (**Figure 4.4**). However, transcriptional evidence of *IGFBP1* upregulation is not enough to suggest it as a mechanism of *Lactobacillus*-mediated protection. Additional experimentation of the stimulation of *IGFBP1* by *Lactobacillus* and the effect on RSV infection is necessary to further elucidate this mechanism.

Genes downregulated by *Lactobacillus* during infection include *AGRB1* (adhesion G protein-coupled receptor B1), a phosphatidylserine receptor that enhances the engulfment of apoptotic cells (348). *CLD15* (claudin-15), which was previously confirmed to be upregulated during RSV infection in cotton rats (**Chapter 2**), is a major constituent of the tight junction complexes that regulate the permeability of epithelia and was downregulated by *Lactobacillus*. *I27L2* (Interferon Alpha Inducible Protein 27 Like 2) is a proapoptotic gene upregulated in PBS+RSV but not *LactoX*+RSV animals. *I27L2* was previously found to be the top differentially expressed gene in preterm infants with severe RSV infection (286). This gene was also confirmed in the previous RSV transcriptomic analysis in **Chapter 2**.

There are several strengths to this study. Probiotic supplementation protocols used in previous studies have used a non-host source for the bacterial strain. Because bacteria rapidly adapt to their host environment, supplementing a food- or human-derived bacterial strain to a non-host environment may affect colonization, survival, and probiotic potential.

Based on previous literature, this is the first *Lactobacillus*-induced protection phenotype from severe respiratory viral infections using a host-adapted strain of bacteria.

There are also limitations to this study that must be recognized. Because the long-term effects (> 12 days) of oral *Lactobacillus* in these animals were not assessed, it is unknown how long *Lactobacillus* remains present in the gut upon oral supplementation. Additionally, although there were no significant differences in alpha and beta diversity upon *Lactobacillus* gavage, the abundance of 4 taxa in *S. fulviventer* and 3 taxa in *S. hispidus* indicate small-scale changes to the microbiome environment that will need to be further tested through longitudinal testing or competition assays between *Lactobacillus* and these taxa. Furthermore, alterations in gene expression as identified by transcriptomic analysis do not indicate that these genes are implicated in Lactobacillus-induced protection from severe RSV. Further experimentation in knockout models will further aid in narrowing these mechanisms of action. Lastly, the mechanisms of *Lactobacillus* in RSV protection are not fully addressed in this chapter, although one proposed mechanism is discussed in **Chapter 5**.

LactoX preparation and administration

Lactobacillus paragasseri BRTN and *Limosilactobacillus reuteri* SMN were previously isolated from *S. hispidus* stool as described in **Appendix A**. Frozen glycerol stocks were streaked on De Man, Rogosa, and Sharpe (MRS) agar plates, and a single clonal isolate was individually incubated in 5mL MRS broth at 37ºC with no shaking. After overnight growth, *L. gasseri* and *L. reuteri* reached OD600 of 1.179 and 1.244 respectively (measured in triplicate, subtracted blank MRS broth). Each culture was diluted to an OD600 of 0.27 in 20mL and grown at 37ºC with no shaking for 6hrs. Equal volumes of both cultures were mixed 50/50 with an equal volume of 40% glycerol (final conc 20%). These stocks (referred to as *LactoX*) were aliquoted into sterile screw cap vials stored at -80ºC.

For administration into cotton rats, *LactoX* was thawed on ice, gently mixed by pipetting, and centrifuged at 14,000 RPM for 30 seconds at 4ºC. The supernatant was decanted, and the bacterial pellet was resuspended/centrifuged 2x in 1mL sterile 1x PBS to remove residual glycerol and MRS media. The washed pellet was resuspended in 1mL sterile 1x PBS, and plating assay on MRS agar determined bacterial concentration to be 5×10^8 CFU/1mL tube. 1mL *LactoX* or PBS was orally gavaged into each animal for 7 or 12 days as specified in the Results.

Animal husbandry and sample collection

Four- to six-week-old cotton rats (~100 g) were obtained from the inbred colony maintained at Sigmovir Biosystems, Inc. Cotton rats in the colony were seronegative by ELISA to adventitious respiratory viruses (i.e., Pneumonia Virus of Mice, Rat parvovirus, Rat coronavirus, Sendai virus). Each species was randomly split into 2 groups: those receiving *LactoX* and PBS vehicle controls. To avoid fighting, all animals were individually housed in large polycarbonate cages and fed a diet of standard rodent chow and water ad libitum.

For assessment of *LactoX* effect on the cotton rat microbiome, *S. hispidus* (n=8 male, 1 female) and *S. fulviventer* (n=8 male, 1 female) were examined. For assessment of *LactoX* effect on RSV outcomes, the following groups contained the following number of male/female *S. fulviventer*: +*LactoX*+RSV (7 male, 3 female), +*LactoX*–RSV (4 male, 1 female), +PBS+RSV (7 male, 3 female), +PBS–RSV (5 male).

For infection, RSV A/Long (ATCC Cat. # VR-26) was propagated in HEp-2 cells (ATCC), and stock titer was determined using plaque assay. Cotton rats were intranasally inoculated under isoflurane anesthesia with 10^5 plaque-forming units in 100 uL of either RSV suspension or PBS vehicle.

For stool collection, cage beddings were changed in the late afternoon for each animal one day before *LactoX* administration. Stool samples were collected between 10 am and 1 pm with sterile forceps before daily *LactoX* administration. On average, 10-15 feces pellets were collected from each animal daily. Immediately after collection, samples were frozen at −80°C. All animal procedures followed NIH and USDA guidelines

144

approved by the Sigmovir Biosystems, Inc. IACUC. Study design, analysis, and reporting of methods and results are presented in accordance with the ARRIVE guidelines.

Animals were sacrificed by carbon dioxide inhalation on day 5 of infection. Whole lung, large intestine, heart, spleen, and kidney were dissected from all animals and frozen in RNA-later (Invitrogen AM7021) for processing and sequencing at VUMC.

Lactobacillus Quantification

Plating assay and qPCR to quantify *Lactobacillus* in the stool were directly adapted from methods described in **Chapter 3: Methods**. Figures and statistical analysis were performed using either T-test or ANOVA/Tukey's multiple comparisons test as appropriate in Prism 8.

16S Sequencing, RNA-Seq, and Data Analysis

Methods for nucleic acid extraction, library preparation (16S rRNA sequencing, RNA-seq), and data analysis were directly adapted from methods described in **Chapter 3: Methods** and in (264). For DESeq2 testing of differentially abundant taxa (16S data), sex was accounted for as a covariate.

Lung histopathology and viral titers

Lungs (right lobe) were dissected and prepared for histopathology as described in **Chapter 3: Methods.** Viral titers were determined using Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) from RNA extracted from lung (lingular lobe; same RNA as sequenced) and nasal homogenates. Following RNA

extraction, cDNA was generated using 1ug of total RNA and the SuperScript™ III Reverse Transcriptase (Invitrogen™) kit according to the manufacturer's instructions. cDNA was diluted 1:5, and 3uL was added to qPCR reactions using iQ™ SYBR® Green Supermix (10uL total). Reactions for each target were performed in duplicate for each sample using primers (IDT, 250nm final concentration) targeting the NS1 protein of RSV and β-actin, which have been previously published and validated for use in both species of cotton rats (306). No-template-controls and a positive control (RNA extracted from RSV A/Long viral stock used for infection) were run on each plate. Duplicate CT values for NS1 were averaged and normalized to β-actin. Results were calculated using the $2^{-\Delta\Delta CT}$ method (307).

Viral titers were also determined by viral plaque assay. Lungs (left lobe) were homogenized, and serial dilutions (10^{-0}-10^{-5}) of supernatants were prepared. HEp-2 cells (ATCC) were seeded to a concentration of 5×10^{5} per well in 12-well plates in DMEM+10% FBS. Cell media was aspirated, and monolayers were infected in 50uL of each dilution in triplicate for 2 hours at 37ºC and shaking every 15 minutes. Cells were overlayed with warm DMEM+10%FBS+0.75%Methylcellulose for 4-5 days until significant syncytia appeared under the microscope. Methylcellulose media was aspirated, and cells were washed 2x with DPBS. Washed cell layers were fixed and stained in 0.067%crystal violet+1.25%glutaraldehyde at room temperature for 20min. Plates were washed 4x with dH_2O and left to air dry overnight. Plaques were calculated as PFU/gram of tissue. Figures and statistical analysis were performed using ANOVA/Tukey's multiple comparisons test in Prism 8.

Transcriptome analysis

RNA-seq data was parsed into individual samples by barcode. Adaptor sequences, low quality (minimum Phred 33), and short (<75bp) reads were removed using Trimmomatic (version 0.36, "ILLUMINACLIP: NEB_multiplexoligos.fa:2:30:10 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75") (308). Only paired-end reads that passed the quality threshold described in (309) were retained. Transcripts were indexed using Salmon v1.8.0 (314). Trinotate v3.2.2 annotation pipeline was adapted for transcript quantification at the gene level (raw and TPM, transcripts per million) *abundance_estimates_to_matrix.pl* script with Salmon input parameters. Each transcript was annotated using the *S. fulviventer* transcriptome reference presented in **Chapter 2** using the Trinotate script *Trinotate_get_feature_name_encoding_attributes.pl* (to generate annotation reference index) and *rename_matrix_feature_identifiers.pl* (to add annotations to the matrix file). DESeq2 package (274) was used within the pipeline by comparing the experimental group containing biological replicates with the corresponding control group. Genes with a P<0.05, adjusted P < 0.05 ("q"/false discovery rate/Benjamini-Hochberg), and a $\log_2$ fold change > |1| were treated as differentially expressed. We used the goseq package in Bioconductor to detect differentially abundant GO terms (278). GOs with a P<0.05 and adjusted P < 0.05 ("q"/false discovery rate/Benjamini-Hochberg) were treated as differentially expressed. PANTHER Protein Classes and Gene Ontology terms were determined by mapping significant differentially expressed genes' UniProtKB-ID to the PANTHER v17.0 Gene List Analysis mapper (available at http://pantherdb.org).

Validation of differentially expressed genes

RNA extracted from the lung tissue was used for qRT-PCR assays. qRT-PCR was performed in duplicate using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosciences) and PowerSYBR Green Master Mix (Applied Biosciences). qRT-PCR primers were designed for 4 randomly selected up-regulated genes and 4 randomly selected non-differentially expressed genes. Primer sequences were designed with Primer3Plus (315), and sequences are provided in **Table 4.1**. No template controls were run on each plate. Technical replicate CT values were averaged for each gene and normalized to an average CT of β-actin and GAPDH housekeeping gene. Results were calculated as fold change induction over uninfected lungs using the $2^{-\Delta\Delta CT}$ method (307).

Gene	Transcriptome ID	Forward Primer	Reverse Primer	Tm (ºC)	Product Size (bp)
SOCS3	Sfulv_DN3570_c0_g3	CAGATGTTGGCGGTCGTGAA	CTCAGTCGTAGCCCCTTGC	60	136
CCL5	Sfulv_DN3711_c1_g1	GTGCTCCAACTTTGCAGTCG	TCCCCAAGCTGGTTAGGACT	60	184
MX2	Sfulv_DN430_c0_g1	GCTGAGCAGCTGCATAAGGG	GTCTGGACAGCCGCTCCTTC	60	105
VIPR2	Sfulv_DN147146_c1_g1	AGCATACTCAAGCCAGTGCC	GGCCATTGGGTATTGGGACA	60	128
HBEGF	Sfulv_DN29394_c0_g1	GTTCCAGCAAACACCAAGGC	TACATCTGGCTTGGGCATCG	60	81
CADH3	Sfulv_DN8120_c0_g1	CACCCATGTATCGTCCTCGG	GAGGCAGCATCAGAACCACT	60	151
CXCL9	Sfulv_DN160651_c0_g1	CCTAGGCTTGGTGACATGGG	ATCACGCTCTTGAGCACAGT	60	140
CLD15	Sfulv_DN28362_c1_g1	ATGGGAACGTCATCACCACC	AGAGGGCCAGCATAGAAGGA	60	111
MYH6	Sfulv_DN142091_c0_g1	GCCTAAATAAGCCCGGTCCA	AACTTGGGTCTTGGGCTGAC	60	183
ISG15	Sfulv_DN6453_c0_g3	CCCCAATGGCTGATTGTCCT	GAAGTGCGACAACCCTCTGA	60	187
GAPDH	Sfulv_DN171591_c0_g1	GGTGCCCAGTACATTGTGGA	GCATCAACAGAAGGGCCTGA	60	113
B-actin	Prev. published (282)	GGCCAACCGTGAAAAGATGACTC	GTCCGCCTAGAAGCATTTGCG	60	150

Table 4.1. RT-qPCR primers for differentially expressed gene validation.

E. <u>ACKNOWLEDGMENTS AND DATA AVAILABILITY</u>

CHAPTER 5

Lactobacillus-secreted protein p40 protects against
Respiratory Syncytial Virus infection in vitro

Lactobacillus has been shown to protect against several severe disease states in both the gut and the lung. Several protective mechanisms of *Lactobacillus* have been studied in the gut, but there is a lack of literature pertaining to how it induces lung protection. Because of similarities in cell biology/immunology between epithelia across the body, protection mechanisms previously studied in the gut may apply to the lungs or the gut-lung axis.

The soluble, *Lactobacillus*-secreted protein *p40* is an influential mediator of protection against inflammatory damage to gut epithelial cells during colitis (349). *p40* was originally derived from *Lactobacillus rhamnosus* GG and is used by bacteria as a cell wall hydrolase. The 40kDa protein is secreted in high abundance from some but not all *Lactobacillus* species (350). As soluble p40 traverses the epithelial membrane, it transactivates EGFR (epidermal growth factor receptor) by catalyzing the activity of ADAM17 (a disintegrin and metalloproteinase 17) to selectively cleave HB-EGF (heparin-bound epidermal growth factor). HB-EGF activation of EGFR stimulates the downstream Akt pathway through a PI3K-dependent manner (192, 349, 351). This signaling cascade results in upregulated mucin production, increased barrier function, and inhibition of cytokine-induced apoptosis (194, 349). Other mechanisms of action include increased IgA class switching and production in B cells (325). The host colon epithelia have also

been shown to communicate with *Lactobacillus* via extracellular vesicles containing HSP90 to upregulate *p40* production (352). Through these mechanisms, *p40* supplementation has been shown to protect against severe colitis *in vitro* and *in vivo* (349, 353).

Colitis is induced by several inflammatory mechanisms involving toll-like receptors (TLR), type 1 interferons, macrophages, CD4+ T cells, and TNF-α (354, 355). Like colitis, viral respiratory infections have a similar immune cascade that leads to cytokine-induced apoptosis of the host epithelia. If *Lactobacillus* can inhibit these mechanisms in the gut, the host could possibly utilize the same mechanism in the lung. As with colitis, RSV disease mechanisms are highly dependent on TLR signaling, type 1 interferons, macrophages, CD4+ T cells, and TNF-a production to progress disease (356, 357). Many studies have already attributed commensal *Lactobacillus* in protecting against severe RSV disease, as discussed in **Chapter 1**. Furthermore, EGFR has been shown to be a necessary host-entry cofactor in RSV viral pathogenesis (358). In this chapter, I tested whether *Lactobacillus* or p40 could reduce RSV severity in lung epithelia and if p40 interacts with viral replication.

i. Purified p40, Lactobacillus paragasseri BRTN, and Limosilactobacillus reuteri SMN increases cell viability in a dose-dependent manner

Two strains of *Lactobacilli* previously discussed in **Chapter 4** were assessed for *p40* secretion. Normalized, bacteria-free *Lactobacillus paragasseri* BRTN and *Limosilactobacillus reuteri* SMN supernatants were prepared via sterile filtering (**Figure 5.1A**). Supernatants and purified p40 isolated from *L. rhamnosus GG* were analyzed by western blotting using an anti-p40 antibody. Both species produced p40, with *L. reuteri* producing less *p40* than *L. paragasseri* (**Figure 5.1B**).

To assess the effect of p40 on cell viability during RSV infection, viable cells were quantified following RSV infection by measuring available ATP using the luminescent-based CellTiterGlo (Promega) assay. Two immortalized lines of lung epithelial cells (HEp-2, A549) were pre-treated with increasing concentrations of *p40* (with heat-inactivated negative control) and *L. paragasseri/reuteri* supernatants before RSV infection. *p40* significantly increased the amount of viable cells at 48hrs post-infection compared to controls in a dose-dependent manner; however, 50ng of boiled *p40* did not increase viability (**Figure 5.1 B-C**). 1/100 and 1/250 dilutions of *Lactobacillus gasseri* supernatant significantly increased cell viability in both HEp-2 and A549 cells, and 1/500 dilutions significantly increased viability in only A549. Only the 1/100 dilution of *L. reuteri* supernatant significantly increased cell viability, but only in A549 cells. (**Figure 5.1 B-C**). No significant effect was seen at 12hrs or 24hrs timepoints (data not shown).
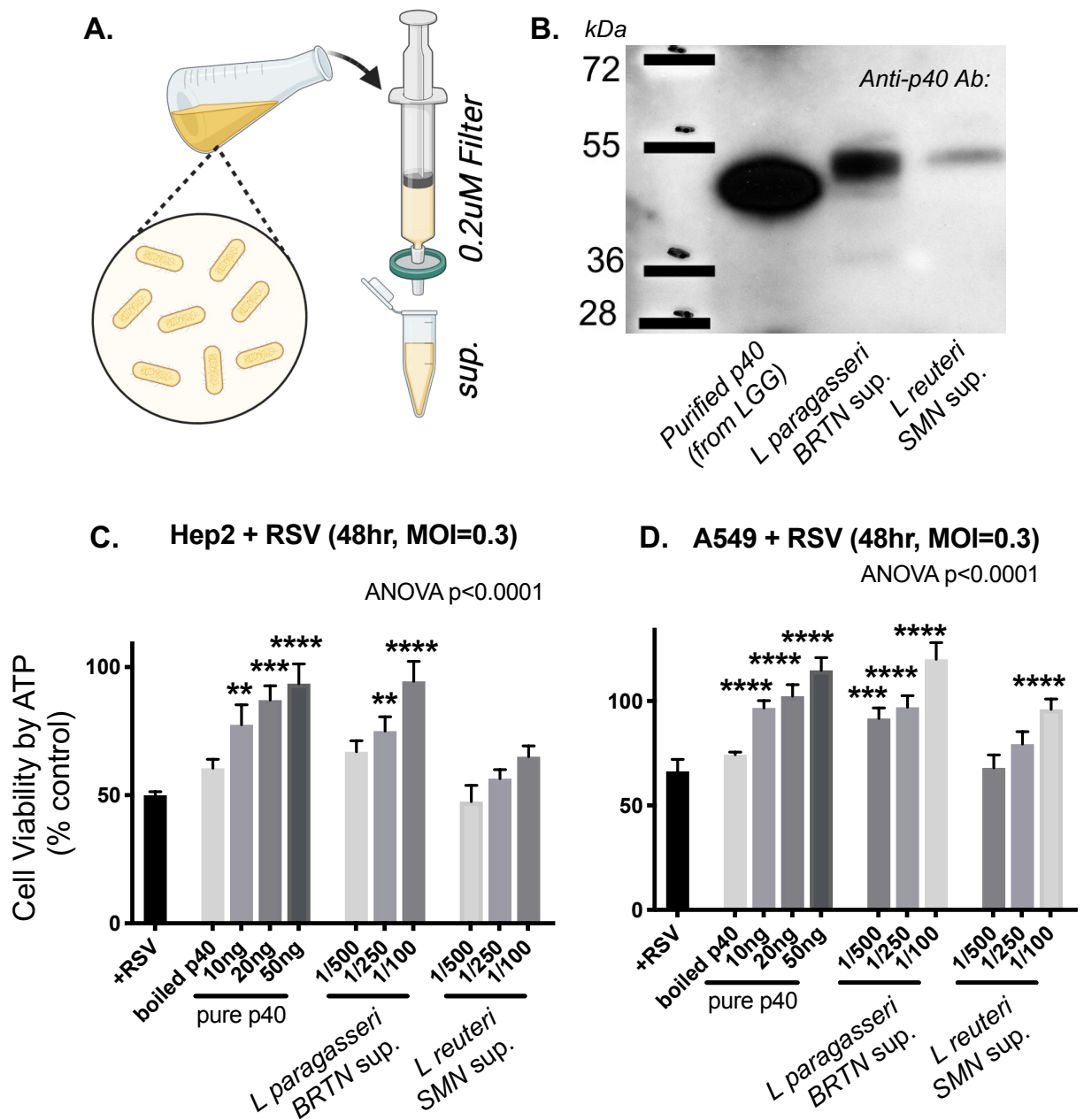
Figure 5.1. (A) Supernatant (sup.) preparation by filter sterilization using 0.2uM filter. (B) Western blot using an anti-*p40* antibody to quantify the presence of p40 in the supernatant of *Lactobacillus paragasseri BRTN* and *Limosilactobacillus reuteri SMN*. (C-D) Cellular viability assays (which quantifies the number of live cells based on available ATP) in both (C) HEp2 and (D) A549 cells pretreated with increasing concentrations of purified *p40* (boiled control, 10-50ng) and *L. paragasseri/L.reuteri* bacterial-free supernatants. Percent viability indicates percentage ATP compared to uninfected controls, as measured by luminescent assay (CellTiterGlo) at 48hr RSV A2 infection. Statistics using Repeated Measures ANOVA with Sidak Multiple Comparison Test.

ii. Both p40 and RSV activate EGFR of lung epithelia in a dose-dependent manner

Because *p40* has been shown to stimulate EGFR in a PI3K-dependent Akt manner, activation of Y-1068 EGFR (residue upstream of Akt) was examined in lung epithelial cells upon *p40* treatment. Serum-starved Hep-2 cells were treated with decreasing concentrations of purified *p40* for 2 hours and blotted for percent EGFR activation (**Figure 5.2A**). Normalization of P-EGFR band intensity to total EGFR and B-actin revealed significant activation of EGFR by *p40* in a dose-dependent manner (highest in 50ng, Tukey Post Hoc p=0.0083) (**Figure 5.2B**).

As RSV has also been shown to activate EGFR through a similar mechanism, Hep-2 cells were infected with RSV/A2 (MOI=3) for 12 and 24 hours to measure EGFR activation. RSV was able to significantly activate EGFR (Y-1068) at 24hrs but not 12hrs post-infection compared to mock-infected controls (**Figure 5.2C**).

iii. p40 reduces RSV viral titers in a dose-dependent manner

To see if viral titers were affected by blockage of EGFR activation, Hep-2 cells were pretreated with the EGFR inhibitor AG1478 for 2hrs. EGFR-blockage reduced viral titers at 48hrs post-infection compared to vehicle controls but not 12 and 24hrs (**Figure 5.2D**). Hep-2 cells were then pretreated with *p40* for 2hrs then infected with RSV/A2 for 12, 24, and 48hrs. At 12hrs post-infection, when viral titers are low (~10^3), 50ng but not 5ng *p40* was able to reduce viral titers slightly but insignificantly (**Figure 5.2E**). At 48hrs

post-infection, when viral titers are high ($>10^6$), 50ng but not 5ng p40 was able to significantly reduce viral titers (**Figure 5.2F**). No effect was seen at 24hrs.

iv. Intranasal p40 administration in mice did not affect animal health

To further test this hypothesis, *in vivo* studies in animals were designed. First, a toxicity experiment measured by weight loss (a major indicator of RSV disease severity in mice) was conducted on 20 female BALB/c mice (8wks old) intranasally inoculated with 20ug purified p40 or PBS-vehicle for 7 days. No significant change in weight loss occurred throughout the administration (**Figure 5.3**).
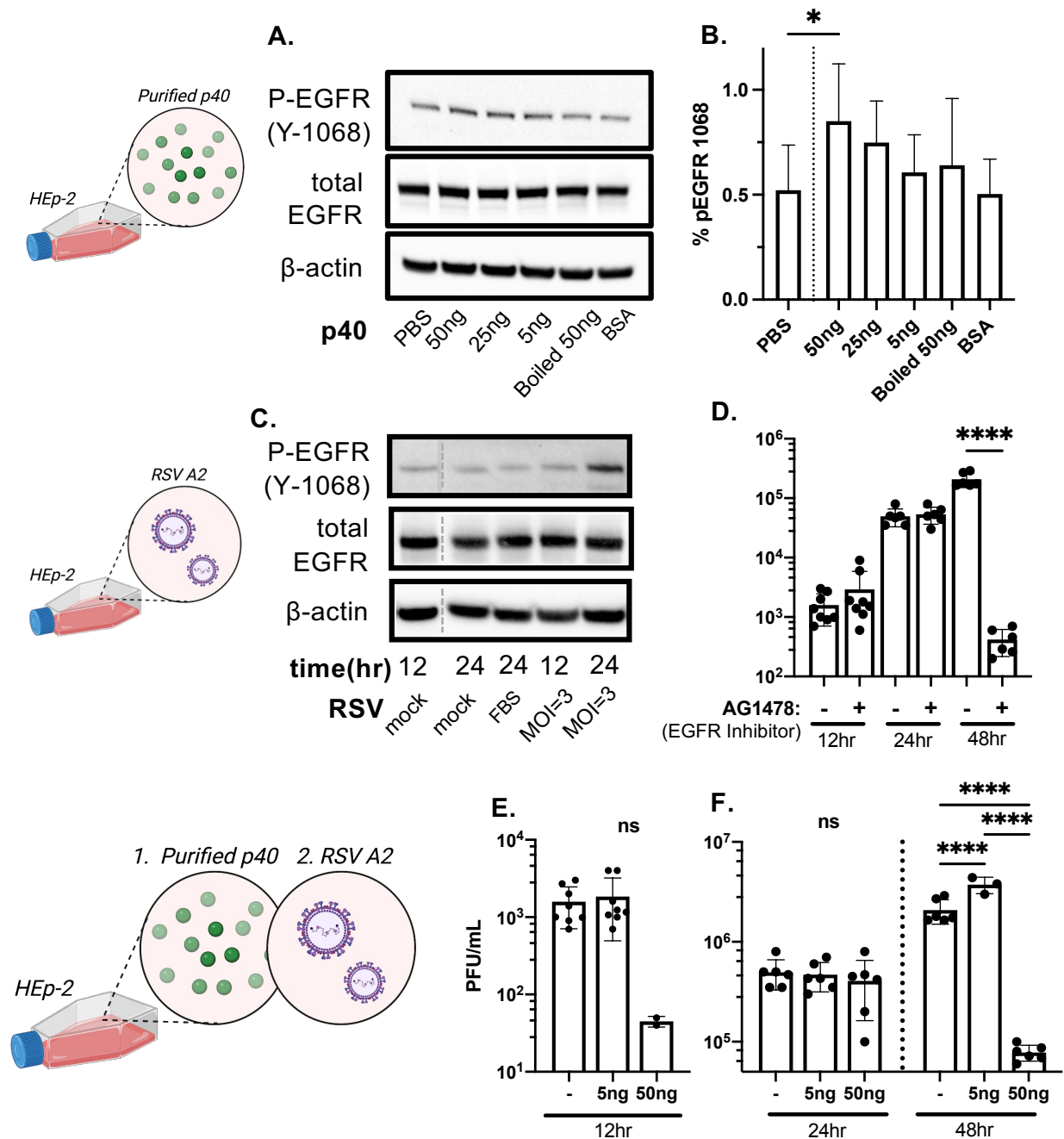
Figure 5.2. *p40* signaling via EGFR affects RSV viral replication. (A) EGFR phosphorylation by p40 at Y-1068 for downstream Akt activation. (B) Quantification of P-EGFR/totalEGFR band intensity measured using ImageJ. (C) Time-dependent activation of P-EGFR by RSV A2 from 12-24 hours (MOI=multiplicity of infection). All samples were run on the same gel, but 12hr/mock sample was cropped for special arrangement. Results mirror those shown by Currier et al., 2016 (39). (D) Viral load at 12-48 hours following pretreatment EGFR inhibitor AG1478 or (E-F) 5ng-50ng p40. Results reflect experiment n=3. Statistics using Repeated Measures ANOVA with Sidak Multiple Comparison Test.
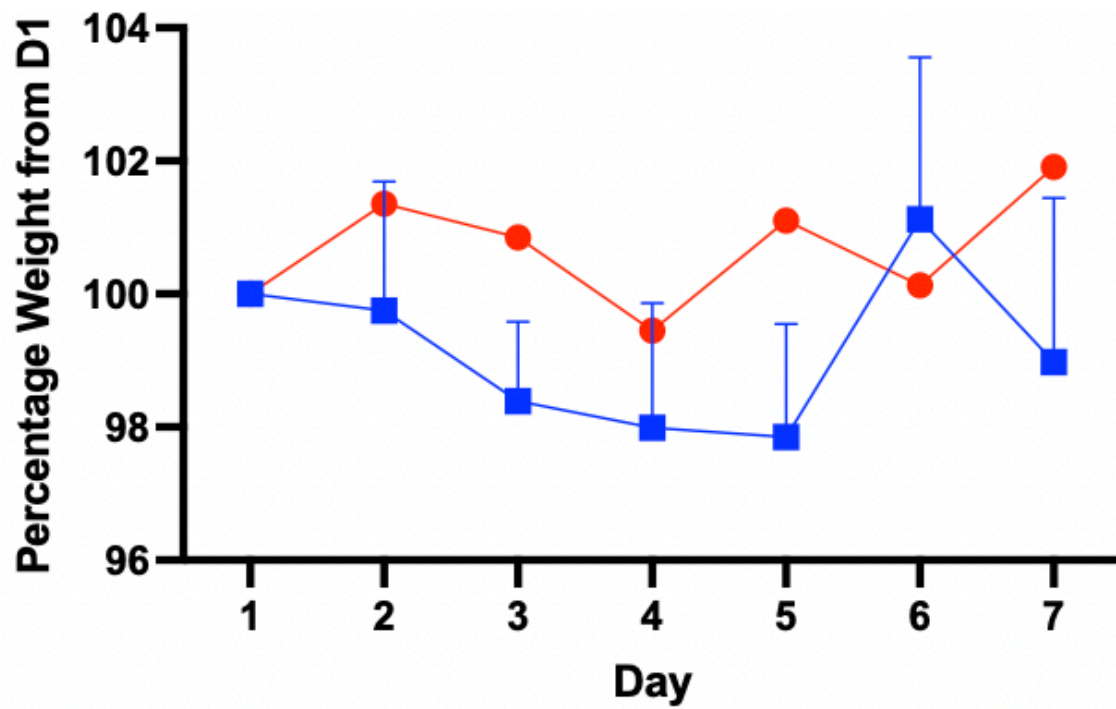
Figure 5.3. Weight monitoring following 7 days of consecutive intranasal administration of 20ug *p40*. N=10 mice per group.

This chapter proposes soluble *p40* as a potent mechanism of *Lactobacillus*-mediated reduction of inflammatory signaling and pathology during Respiratory Syncytial Virus (RSV) infection. Luminescent ATP assays determined that *p40* significantly enhanced the viability of RSV-infected cells in a dose-dependent manner. Supernatants of *Lactobacillus paragasseri BRTN* and *Limosilactobacillus reuteri SMN* were also able to enhance viability during infection, which correlated to differential amounts of *p40* produced during growth (**Figure 5.2**). However, these results are limited, as the assay can only infer cells with proliferative/metabolic activity based on available ATP. To further interpret p40's role in other cell death pathways, assays measuring cytotoxicity (metabolic activity with MTT) and apoptosis (early DNA breaks with TUNNEL, alterations to cellular morphology with Annexin V) may be utilized in further studies. Additionally, the effectivity of p40 as a protective immunobiotic could be further confirmed using additional negative controls, such as a non-p40 producing *Lactobacilli*, a purified protein control (e.g., BSA), or *Lactobacillus*-supernatant depleted of p40 by immunoprecipitation. These conditions have been applied in the "Future Directions" section of **Chapter 6**.

As previously described in the colon (351), *p40* interacts similarly with lung epithelia through transactivation of EGFR Y-1068 (**Figure 5.2A-B**). This further activates the downstream Akt serine-threonine kinases, which are critical in regulating cell

proliferation, survival, and metabolism (359). These mechanisms are hypothesized to enhance cell survival and reducing viral titers during RSV infection. RSV has also been shown to interact with EGFR, which is required as a cofactor for entry and infection (39). *p40*, which transactivates EGFR, was able to reduce viral titers at 48 hours post-infection, acting as an immunobiotic blockade of viral entry and pathogenesis. This was confirmed by blocking EGFR signaling using Tyrphostin AG1478, which also reduced viral titers 48 hours post-infection.

The biological mechanisms of ligand:EGFR binding suggests that pre-activation of EGFR in the lung may reduce the amount of unbound EGFR necessary for RSV attachment and entry. As EGFR binds ligand, receptor dimerization results in subsequent phosphorylation. Phosphorylated EGFR is then internalized and undergoes proteasomal degradation (360-364). The remaining unbound EGFR can also be internalized in response to the surrounding activation (365). Therefore, higher doses of *p40* can effectively block viral entry and replication by removing the viral cofactors necessary for entry.

These effects correlate with the literature on RSV and EGFR interactions. Thomas et al. first identified RSV to interact with PI3K, a downstream signal of EGFR activation. Inhibition of PI3K in cell culture resulted in earlier cytotoxicity at 12hrs vs 24-48hrs, but there was no mention of EGFR or impact on viral replication (366). Monick et al. later found that RSV activates EGFR to prolong viral survival through the upregulation of ERK and inflammatory IL-8. Researchers tested this phenotype in both A549 and HTBE primary cells using RSV A2. Blocking EGFR and ERK reduced IL-8 and cell death but did not affect viral replication. However, titers were only examined using western blot for viral

protein expression, which is not as quantitative as other methods (plaque assay, RT-qPCR). Although this study suggests ERK signaling, they used the Y-1068 P-EGFR antibody specific to Akt signaling and not ERK. Although the researchers misinterpreted their results, their data support the overlap of RSV and *p40* signaling interactions (367).

Currier et al. later linked EGFR as a cofactor of RSV infection and increased mucin production. Researchers used co-immunoprecipitation to reveal the RSV F protein's binding to EGFR, which facilitates entry into the cell (39). Furthermore, they found that inhibition of EGFR phosphorylation diminished RSV infectivity, which mimics the results of this study (**Figure 5.2**). Kalinowski et al. followed up to the Currier 2016 study and found RSV-induced EGFR activation suppressed *IRF1* antiviral response and increased infection via F protein binding. EGFR inhibition augmented IRF1, interferon-lambda, and viral titers (358).

To test these hypotheses *in vivo,* future studies have been designed to examine the use of intranasal *p40* on RSV outcomes. Daily administration of purified *p40* in the nose did not result in any significant weight loss across 7-days. Future directions of this study will be discussed in **Chapter 6**.

The effect of *p40 in vitro* also reflects the effect of bacterial-free supernatants from *Lactobacillus paragasseri* strain *BRTN* and *Limosilactobacillus reuteri* strain *SMN*. As *L. reuteri SMN* produced less *p40 than L. paragasseri BRTN*, the effect on viability was reduced, which supports a dose-dependent role. In **Chapter 4**, these two *p40*-producing strains effectively reduced RSV disease severity in cotton rats, suggesting *p40* as a potential mechanism for *Lactobacillus*-mediated protection against severe RSV infection and outcomes.

Viral stocks, infection, and plaque assay

RSV A2 master stocks were acquired from Dr. Stokes Peebles at VUMC, and working stocks were prepared using HEp-2 cells (ATCC). HEp-2 cells were propagated in sterile Nunc EasYFlask 175-cm^2 (Thermo) and DMEM+10%FBS. Upon 70% cell confluency, cells were washed 3x with PBS without calcium and magnesium (Corning). 5mL of serum-free DMEM with MOI=0.01 RSV A2 (~50uL) was added to the flask and incubated at 37ºC+5%CO$_2$ for 2hrs with shaking every 15min. An uninfected flask was also used as "mock infection" control. 25mL of DMEM+10%FBS was added to flask, and cells were incubated for 4-5 days until significant syncytia appeared under the microscope. Cells were scraped into 50mL conical tubes and sonicated (1 pulse per mL at 50% power) on ice. Sonicated suspensions were centrifuged at 2000 RPM for 10min at 4º to pellet cell debris, and supernatants were pooled and aliquoted in 1mL screw-cap tubes for -80ºC storage.

For experimental infections with *Lactobacillus* supernatant or purified p40, cells in 96-well (Cell Viability Assays) or 6-well (Western Blot) plates were infected with RSV A2 viral stock or mock control at the indicated MOI at 37ºC for 2hrs with shaking every 15min. Cells were then overlayed with warm media and incubated for 12-48hrs.

Viral titers for both stocks and experimental samples were determined by plaque assay as described in **Chapter 4: Methods**.

Cell Culture

A549 and HEp-2 cells (ATCC) were seeded at 1×10^6 cells per well in 6-well plates in DMEM+10%FBS. Before confluency, cells were serum-starved overnight. Cells were treated (2hrs for western blot, overnight for viability assays) with media supplemented with purified p40 (5-50ng, along with 10min@95ºC boiled control) or 0.2uM-filtered *L. gasseri/reuteri* supernatants (cultures grown according to **Chapter 3: Methods**). AG1478 was added to cells using a 1:1000 dilution of 150uM stock. Purified p40 stocks were acquired from Dr. Fang Yan, which were purified from *Lactobacillus rhamnosus* GG (ATCC) culture supernatant and stored at -80°C as previously reported (191). Cells were then infected and analyzed as described earlier.

Western Blot

Treated or infected cells in 6-well plates were washed 2x with ice-cold 1x PBS and lysed in 50uL RIPA buffer + 10% phosphatase inhibitor for 1hr on ice. Cells were scraped, vortexed for 1min, and centrifuged at 13k RPM for 10min at 4ºC. Protein concentration was measured by Pierce™ BCA Protein Assay Kit following manufacturer's protocol. Lysates were mixed with 4xLDS Loading Buffer and 10X Reducing agent (Invitrogen) and heated at 95ºC for 5min to denature protein. Samples were loaded into a NuPAGE™ 4-12% Bis-Tris gel (Invitrogen) in an XCell SureLock MiniCell running tank with 1x MES SDS Running Buffer at 200v for 45min. Gels were transferred to PVDF membrane using iBLOT2 Dry Blotting System using the P0 protocol (20V 1min, 23V 44min, 25V 2min). Membranes were blocked for 1hr in 5% milk+1xTBS+0.1%Tween. Blot was probed with overnight 1º antibody overnight at 4ºC , washed 3x with TBS-Tween, and 2º antibody for

1hr at room temperature. Blots were stripped in between each targeted antibody with 20mL Restore Western Blot Stripping Buffer (Thermo 21059) for 10min followed by re-blocking. Probed blots were developed using SuperSignal West Pico Plus Chemiluminescent Substrate (Thermo 34580) for 5min. Signal was captured using Chemidoc MP in the MCBR Core at VUMC.

Antibodies: Activated EGFR was measured using 1° P-EGFR Y-1068 D7A5 XP Rabbit mAB (1:500, Cell Signaling 3777) and 2° Goat AntiRabbit IgG Hrp-linked Ab (1:2000, Cell Signaling 7074). Total EGFR was measured using 1° EGFR D38B1 XP Rabbit mAb (1:1000 Cell Signaling 4267) and 2° Goat AntiRabbit IgG Hrp-linked Ab (1:2000, Cell Signaling 7074). B-actin was measured with 1° Mouse B-actin mAb (1:1000, Sigma A1979) and 2° HRP Goat AntiMouse IgG Ab (1:10,000, MP Bio 55550).

Cell Viability Assays

Upon pre-treatment with *Lactobacillus* supernatant or purified p40 and infection with RSV (MOI=0.3), viability of A549 and HEp-2 cells (performed in triplicate) were measured using Cell Titer Glo 2.0 (Promega) according to the manufacturer's protocol. In summary, cells were washed 2x PBS and overlaid with 50uL Cell Titer Glo reagent to measure available ATP. Plate was rocked for 10min, and luminescence was measured with a GloMax® Luminometer in the MCBR Core at VUMC. Luminescence values for technical replicates were averaged and standardized to an ATP standard curve.

Animals: 20 Balb/C mice were intranasally administered 100uL 20uG purified p40 or PBS under anesthesia. Weight and behavior were monitored daily for 10 days.

E. <u>ACKNOWLEDGMENTS AND AVAILABILITY OF DATA</u>

CHAPTER 6

Discussion

In **Chapter 2**, I comprehensively characterized multiple sites of the cotton rat microbiome using 16S rRNA and whole genome sequencing. This data established key differences in microbiome community structure and function between body sites in *S. hispidus* and *S. fulviventer,* which were housed in the same facility for multiple generations and fed with the same diet. This analysis also uncovered sex as a significant variable in gut microbiome composition. This foundational study establishes a microbiome sequencing and analysis pipeline for future hypothesis testing in understanding the role of the microbiome in viral pathogenesis, especially for RSV. Furthermore, it revealed a powerful *Lactobacillus*-deficient animal model to be utilized in future studies with RSV. Additionally, this study adds to the small but expanding literature in understanding the role of host genetics on microbiome structure and composition.

In **Chapter 3**, I developed two transcriptome references for *S. hispidus* and *S. fulviventer*, as well as a multi-use transcriptomic pipeline that will be used at VUMC and Sigmovir Biosystems. This study improved on the previously published cotton rat lung transcriptome by using multiple body sites and including an additional species of animal (*S. fulviventer*) (264). The results of this chapter highlight the utility of cotton rat transcriptome in understanding infection-induced transcriptomic changes in the absence of available genome references. Both *S. hispidus* and *S. fulviventer* transcriptome

167

references have been made publicly available at BioProject PRJNA816878 and at https://doi.org/10.1038/s41598-022-19810-4 to be used in future research efforts. Additionally, identifying gene sequences allows for better understanding of cotton rat genetics and the generation of molecular tools, such as qRT-PCR primers and probes, recombinant proteins, antibodies, and other assays. Differential gene expression analysis also revealed host species-specific differences in response to RSV. As the development of RSV therapeutics calls for a well-developed, robust pre-clinical model for RSV, these transcriptome references and gene associations with RSV can accelerate biomedical interventions against this pathogen of significant public health importance.

In **Chapter 4**, cotton rat-isolated *Lactobacillus* was orally administered to *Lactobacillus*-deficient animals, and probiotic supplementation with *LactoX* significantly reduced the severity of RSV disease. These results incorporated pipelines developed in **Chapters 2 and 3** to determine the effect of probiotic *Lactobacillus* on gut microbiome and lung transcriptomic environment. Many of these genes were related to decreased inflammation and promotion of cell function and immunity. *Lactobacillus* also significantly increased *CXCL9* and *CXCL10* during infection, which has been shown to reduce Th2 bias and severe long-term outcomes. A reduction of Th2-inducing proinflammatory (IL-4, IL-13, IL-33) cytokines reflected these results. Additionally, upregulation of both *IFI44, GBP5,* and *IGFBP1* in the lung suggests potential mechanisms of RSV restriction by oral *Lactobacillus* supplementation.

In **Chapter 5**, *p40* was proposed as a potent postbiotic mechanism of *Lactobacillus*-mediated protection against RSV. These effects were mediated by transactivation of cell proliferation signaling pathways and inhibition of viral entry and

replication. As both strains of *Lactobacilli* within the *LactoX* cocktail produce *p40*, this is a potential mechanism of action in reducing RSV disease severity in cotton rats described in **Chapter 4**. These results must be interpreted separately because *p40* was administered directly to the lung epithelia while *LactoX* was administered orally to the gut. However, gut *LactoX* was able to significantly modulate the lung transcriptome in healthy animals, which suggests a role of the gut-lung axis.

Several studies in humans and mice have found that gut *Lactobacillus* can protect the lung against respiratory diseases (rhinovirus, influenza, RSV). Others have described this phenomenon as the gut-lung axis, where distal immune responses are catalyzed by trafficking of metabolites or immune-mediators (cytokines, immune cells) from the gut to the lung. The results in **Chapter 4** support this, as supplementation of gut *Lactobacillus* in cotton rats successfully reduced RSV disease severity in the lung. The two bacterial isolates utilized in this study (*LactoX*) secrete *p40* (**Figure 5.1A**), and oral supplementation to healthy animals increased expression of intestinal HB-EGF in cotton rats (**Figure 4.3B**). Yan et al. found that gut *p40* supplementation could effectively stimulate the release of intestinal HB-EGF into the serum, which may then transactivate EGFR via ADAM17 at distal sites (351). Based on these results, I investigated the potential of HB-EGF as a protective gut-lung axis mediator.

HB-EGF has been previously proposed as a lung protectant in multiple diseases by stimulating mitosis and restoration of alveolar type II lung epithelia after injury (368). Su et al. supplemented mice suffering from COPD lung inflammation with recombinant HB-EGF to reduce lung injury, inflammatory cells, cytokine infiltration, decreased caspase 3 expression, and increased EGFR expression (369). Intestinal HB-EGF has also been

shown to induce protection in the lung. Lutmer et al. used intestinal administration of HB-EGF to protect against lung injury associated with severe burns. Mice pretreated with enternal HB-EGF before being scalded saw decreased lung myeloperoxidase damage, pulmonary apoptosis, and bronchial reactivity/airway resistance (370), suggesting distal HB-EGF-dependent signaling in the lung. HB-EGF has also been shown to reduce acute lung injury and increase survival caused by intestinal ischemia/reperfusion (I/R) (371). Clinically, gut I/R can cause remote injury to the lung via tissue hypoxia and circulating leukocytes that induce inflammation. James et al. used intra-luminal injection of HB-EGF to reduce the detrimental lung damage caused by gut I/R in a mouse model (371). If gut *Lactobacillus* or *p40* can increase HB-EGF expression and subsequent secretion into circulation, this may be one mechanism by which gut *Lactobacillus* is facilitating lung protection.

Based on this literature and data presented in this thesis, I propose these conclusions and hypotheses:

- Gut *Lactobacillus* species (*L. paragasseri, L. reuteri*) exert a protective effect on respiratory disease outcomes in cotton rats.

- Gut *Lactobacillus* reduces RSV epithelial damage and inflammation by enhancing antiviral defense mechanisms, metabolite interconversion, and transmembrane signaling. These include promotion of *CXCL9* and *CXCL10* to enhance Th1 polarization, reduction of Th2-promoting inflammatory cytokines, and upregulation of *IGFBP1*, *GPB5*, and *IFI44* to restrict viral entry and pathogenesis.

- *p40* effectively enhances cellular proliferation and reduces RSV viral titers *in vitro* by an EGFR-dependent mechanism previously described in the colon (351)

170

- I hypothesize that airway *Lactobacillus* reduces RSV titers by *p40*-mediated transactivation of EGFR and subsequent receptor internalization, which removes the infection cofactor needed for RSV and cellular interaction/entry.

- I hypothesize that gut *Lactobacillus* secretes *p40* to increase gut HB-EGF production that is transported through the blood into the lung, where it can reduce RSV disease and titers through decreased inflammatory cytokines, epithelial apoptosis, and airway resistance.

- I hypothesize that *Lactobacillus* species or *Lactobacillus*-secreted *p40* may be an effective live biotherapeutic product in averting RSV disease severity, episodes of wheezing, and asthma outcomes following RSV infection in children. This effect may also expand to other respiratory viruses and infectious diseases.

B. <u>CLINICAL IMPLICATIONS OF THE WORK</u>

Several clinical trials are exploring the possibility of infant probiotics and vaginal seeding for babies born via Cesarean section to populate the neonatal microbiome with protective bacteria. Many companies have included pre-, pro-, and post-biotics into infant formula and supplements. This is especially important for infants breastfed for less than 6 months. Other postbiotics, like *p40*, are already used in infant formulas, which reduces the challenges with probiotic dosing and side effects (372). The experimental pipelines presented in this thesis can be applied to other postbiotic mechanisms for pre-clinical

testing. The translational value of cotton rats in RSV research suggests it is an excellent preclinical model to test for *Lactobacillus*-mediated protection against respiratory viral infection.

C. <u>FUTURE DIRECTIONS</u>

The translational value and utility of the cotton rat model for RSV infection highlights the need for additional reagents and assays for future studies. Assay development, such as RT-qPCR, has been limited by the lack of gene sequences needed to create targeted primers for genes of interest. The *de novo* transcriptome references generated in **Chapter 3** provide a plethora of information for targeted transcriptomic analysis via digital/RT-qPCR and gene editing via CRISPR. These references will be utilized for primer design and recombinant protein production, specifically for inflammatory mediators such as cytokines and interferon-stimulated genes. One valuable tool that would greatly increase the specificity of these results is antibody production for use in cytometric assays, immunohistochemical staining, and western blot analysis. For example, these resources are necessary to investigate the contribution of *Lactobacillus* to immune cell types such as Th1/Th2 bias and inflammatory monocytes. The *de novo* transcriptome references have provided many gene and protein sequences that can be used to test cross-reactivity of human and rodent antibodies currently available (to which an 80% sequence homology of protein-antibody is indicative of specificity and

performance). Additionally, targeted immunoglobulins for cotton rats can be generated by synthesizing or purifying antigen and immunizing animals for polyclonal antibody production. Generation of such reagents will be continually investigated by the Das Lab and Sigmovir Biosystems for future studies.

This work also encompasses published and preliminary data for a project to be continued at VUMC upon NIH R01 funding by Suman Das, Ph.D. Further experiments have been designed based on this work to serve as a roadmap for future studies. The team of interdisciplinary scientists employed in this grant proposal includes experienced investigators spanning multiple career stages in virology, RSV immunologic/virologic outcome assessment using animal models, bacteriology and *Lactobacillus*/*p40* biology, genomics, and bioinformatics.

Project 1: Determine the role of the *Lactobacillus*-secreted protein *p40* in RSV pathogenesis and immunopathology using primary epithelial cells. Several studies have linked the gut-lung axis to how the microbiome mediates the reduction of respiratory disease severity (110, 373). We hypothesize that *Lactobacillus* via p40 secretion transactivates EGFR and the consequent activation of PI3K/AKT significantly reduces RSV replication, severity, and allergic and immunological outcomes. While we have robust preliminary data to support this aim, these experiments were performed using immortalized cell lines (Hep2 and A549 cells). In collaboration with Co-Investigator Julie Bastarache, M.D., we will use primary human lung epithelial cell cultures established by her lab as a more relevant human model (374-377). Alveolar type II epithelial cells will be isolated by digesting the tissue using dispase & collagenase, followed by filtering and

173

achieving single-cell suspensions. Alveolar epithelial type II cells will be isolated and purified using Miltenyi Biotech magnetic microbeads and LS columns for a positive cell selection (374-377). *p40* has been previously shown to be non-toxic to epithelial cell culture and mice (**Chapter 5**). To test if *p40* transactivates EGFR and Akt pathway in primary human lung epithelial cells, we will test activation of EGFR, PI3K, and Akt phosphorylation by Western blotting. To test if *p40* treatment directly reduces viral replication *in vitro*, lung epithelial cells after overnight treatment with *p40* or *p40*-expressing *Lactobacillus* culture supernatant will be infected with RSV with an MOI of 0.1. Additional negative controls will also be utilized in parallel to cell experiments: p40-negative *Lactobacillus* supernatant (*L. murinus*, previously isolated from *S. fulviventer* and shown not to produce p40 (350)), a non-*Lactobacillus* supernatant (*Enterococcus* isolated from *S. hispidus*), BSA as a negative protein control, and heat-inactivated p40 and supernatants. Initially, we will test with three virus strains: RSV A2, RSV A/Long, and RSV 01/2-20 (clinical isolate from Tennessee) (378). To identify if the effect of *p40* is strain/genotype-specific, we will also test some of the currently circulating isolates of the ON1 or BA2 clades isolated and sequenced by us from clinical infant cohorts (379). In addition to titration of the virus by plaque assay for viral replication kinetics, we will test for p40-induced protection using both a luminescent cell viability assay (Cell Titer Glo), apoptosis assays (Annexin V), and confocal microscopy (using DAPI for DNA, phalloidin for actin, and a fluorescently-tagged *p40* antibody for amounts within the cells). All the *in vitro* experiments will be done in triplicate and repeated at least three times.

Expected outcomes: Similar to our preliminary data, we anticipate a reduction in apoptosis in the primary lung epithelial cells supplemented with *p40* (*L. paragasseri-* and

pure *p40*-treated) compared to control-treated (BSA and heat-treated protein) cells in a concertation-dependent manner. We also anticipate early activation of EGFR and reduction of viral titer at 48hrs.

Project 2: Test if the purified *p40* supplementation in the gut or lung of *S. fulviventer* protects from RSV-induced immunopathology, modulates host cytokine gene expression, and increases viral clearance. Fang Yan, M.D., Ph.D. (Co-Investigator) has successfully developed a pectin/zein hydrogel bead system to specifically deliver *p40* to the mouse colon and activate EGF receptors in colon epithelial cells. Using this method of administration of the *p40*-containing beads, her team has shown that it reduces intestinal epithelial apoptosis and inflammation disruption of barrier function in the colon epithelium in an EGF receptor-dependent manner, thereby preventing intestinal inflammation in mouse models of colitis (325). We will test the efficacy of the p40-pectin/zein hydrogel bead delivering system in *S. fulviventer* and *S. hispidus*. We will orally treat animals with recombinant purified *p40* coupled to pectin/zein hydrogel beads using three different doses (10-50ug). We anticipate that these concentrations are in the range of proper delivery to the intestines, preventing potential degradation during digestion and avoiding adverse outcomes such as weight loss. Control groups include *Lactobacillus rhamnosus GG* bacterial-free supernatant as a positive control and PBS vehicle only as a negative control. Following *p40* administration for 7 days, cotton rats will be challenged intranasally with RSV. Activation levels of EGFR, PI3K, and AKT will also be measured in lung epithelia using western blotting to test if *p40* initiates distal EGFR activation in the lung. Bronchoalveolar lavage fluid (BAL) will be collected each day

following infection under anesthesia. Two animals will be sacrificed after 4- and 6-days post-infection. Viral titers will be measured by plaque assay in nasal turbinates and harvested lung homoginates. Histopathology and transcriptomic analysis (RNA-seq) will be performed on lung tissues. Daily levels of inflammatory cytokines will be measured by RT-qPCR in BAL fluid samples. Additionally, we will repeat this experiment using intranasal administration of purified *p40* without pectin/zein hydrogel bead delivery.

RNA extracted from lung tissue will be used for gene expression analysis using the cotton rat transcriptome references described in **Chapter 3**. Additionally, qRT-PCR will be performed to measure host cytokine responses (INF-γ, TNF-α, TNF-β, IL-1β, IL-6, IL-10, growth-regulated protein (GRO), and RANTES). β-Actin and GAPDH will be used as housekeeping genes to control for differences in cDNA for each treatment, as described in earlier chapters. Statistical analysis between groups will be performed using ANOVA followed by the Tukey post hoc test. Lung pathology will be scored as described in **Chapters 3 and 4**.

Expected outcomes: We anticipate that both *p40* supplementation through the gut and directly in lung to cotton rats will positively modify host response to RSV infection, resulting in lower viral titer, improved histopathology scores, and lower inflammatory host gene expression pathways. Specifically, we expect RSV disease severity (alveolitis, interstitial pneumonia, and inflammatory cytokines) will be reduced in cotton rats following *p40* supplementation. If either histology or cytokine responses are unchanged when *p40* is supplied orally, cotton rats will be intranasally treated with recombinant purified *p40* protein in PBS. We do not anticipate any technical problems measuring the specified endpoints, as Jorge Blanco, Ph.D. (Co-Investigator) has substantial published experience

using the cotton rat model of RSV infection (233, 380-385). Delivery of purified protein into cotton rat colon is an issue due to potential degradation by digestive enzymes in the stomach. However, prior work by Dr. Yan shows optimal delivery of the *p40* to the colon in mice when coated with pectin/zein hydrogel beads (325). Success of this project will present a viable postbiotic mechanism for RSV protection that avoids challenges with bioavailability and bacterial survival, as seen in *S. hispidus* (**Figure 4.1E-H**).

Project 3: Determine if the *Lactobacillus* protein *p40* protects against RSV-induced illness, immunopathology, viral clearance, and allergen sensitization in mice. Unfortunately, despite being the best model for RSV, there are limited immunological reagents to probe cellular immunity in cotton rats, and no tractable genetic modification tools are available. Thus, to further test this hypothesis, we will assess acute changes in the murine immunopathologic response upon Lactobacillus-*p40* priming. Precisely, we will quantify Th1 and Th2 bias and bronchiolitis using an established mice model, as an imbalance of these outcomes is linked to wheezing and subsequent asthma in children. If these experiments are successful, it will translate to human intervention, as our bronchiolitis mouse model is closest to mimicking the phenotype detailed below.

Mouse model for assessment of the immunopathological outcome of acute RSV infection: Our team has extensive experience with the mouse model of RSV infection and has the capability of determining physiologic and immunologic endpoints of RSV infection in mice. Using this infection model, Stokes Peebles, M.D. (Co-Investigator) has found increased lung IL-13 expression, induction of airway mucus, and airway reactivity to methacholine upon infection by RSV 01/2-20 (clinical isolate from Tennessee) (386, 387). In our retrospective birth cohort studies using the INSPIRE cohort, we have identified

strong evidence that bronchiolitis severity in infancy is a significant risk factor for the development of childhood asthma (96, 388). However, one of the significant drawbacks of the mouse model of RSV infection is that airway obstruction resulting from epithelial cell desquamation in human bronchiolitis is not present in the mouse. Preliminary data show that mouse infection with RSV 01/2-20 results in the deposition of desquamated epithelial cells and airway mucus expression into the airway, miming the pathology seen in infants with severe bronchiolitis. Mucus production has been associated with the activation of CD4$^+$ Th2 cells (389). In this mouse model of RSV, Th2 cytokines cannot be detected in the lungs before day 5 post-infection (390). Further, group 2 innate lymphoid cells (ILC2) are an early, significant source of IL-13 in several other pulmonary diseases (391-401). Work from Dr. Peebles's team and others have revealed that RSV induced ILC2 activation during ARI (402, 403).

To test our hypothesis, there will be 16 groups of BALB/C mice. After acclimatization in our facility for two weeks, mice will be intranasally and orally primed for 7 days with *L. reuteri* or *L. gasseri* culture supernatant, purified *p40*, and culture supernatant depleted of *p40* by antibody as a negative control. For delivery to the gut, purified *p40* coupled with the protein to pectin/zein hydrogel beads will be orally administered. After priming, viruses will be infected intranasally. We will infect mice with 1 x 10^6 plaque-forming units (PFU) of each RSV stain in all experiments. 24hrs after RSV infection, a group of mice (5 mice each) will be sacrificed for lung innate immune response to RSV infection. On Day 4, a group of five mice will be sacrificed to collect lungs for both peak viral titer and BAL cells. On day 6, a group of five mice will be sacrificed for viral titer clearance, BAL for cells, and peak IFN-γ. On day 8, lungs from 5 mice from each group

will be harvested, digested and single-cell suspensions will be made for surface and intracellular cytokine staining to identify ILC2s, RSV-specific CD8+ T Cells, histopathology, viral clearance, and airway responsiveness measurements by methacholine challenge. Additionally, based on our preliminary data above and the role of alveolar macrophages in p40-mediated activation, we will also look for M1 and M2 macrophages and their infiltration in the lung. On day 28, antibody titers to RSV infection will be quantified.

Expected outcomes: Based on our preliminary data in infant cohort and *in vitro* studies, we anticipate a reduction in lung viral titer and lung histopathologic scores in animals that are treated with *p40* in a dose-dependent manner. We further expect that *p40* supplementation will result in reduced recruitment of NK cells, increased infiltration of M2 vs. M1 macrophages, decreased Th2 bias in the lungs, and a decrease in tetramer-positive CD8$^+$ cells that are RSV-specific CTLs in lungs, spleen, and lymph nodes. We also anticipate that there will be a significant reduction in the number of IL-13$^+$ ILC2 in the lungs of mice primed with *p40* and infected with both RSV A2 or RSV 01/2-20 that induces a Th2 immune response with pathology that has features of asthma. A reduction in the number of IL-13$^+$ ILC2 in the lungs of mice supplemented with *Lactobacillus* p40 would indicate that *Lactobacillus* is protective against severe RSV disease outcomes. Suppose we confirm that p40 is protective for RSV disease outcomes. In that case, we will further confirm the role of gut EGFR activation downstream effect on lung using Egfr$^{fl/fl}$-Vil-Cre mice that have EGFR knock out specifically in villus and crypt epithelial cells of the small and large intestines and Egfr$^{fl/fl}$ as the control (325). This experiment will confirm if activation of EGFR in the gut positively impacts lung outcomes (the gut-lung axis).

Given our experience with plaque assays, ELISA, intracellular cytokine staining, tetramer analysis, histopathology, airway responsiveness, and cytokine protein determination, we do not anticipate any technical difficulties with the proposed experiments. Suppose we find lung RSV titers are still detectable at day 8. In that case, we will extend the experimental timeline for all endpoints to days 10 and 12 to fully assess the effect of the RSV 01/2-20 on antiviral immunity. One potential problem is that day 8 is not the optimal time point for measuring these endpoints. Based on our published results, RSV 01/2-20 induction of IL-13 peaks in the lung on day 8 and is accompanied by an increase in airway restriction and mucus production.

APPENDIX A

Complete genome sequence of Lactobacillus paragasseri BRTN and Limosilactobacillus reuteri SMN, isolated from Sigmodon hispidus stool

The gut microbiome contributes to many aspects of host health and disease. Members of the genus *Lactobacillus* have been shown to protect against severe allergy and respiratory infections (55, 126, 133). The cotton rat (genus *Sigmodon*) is a susceptible model for many human respiratory viruses (300). Our comprehensive microbiome characterization of two cotton rat species (*S. hispidus*, *S. fulviventer*) found that *Lactobacillus paragasseri* and *Limosilactobacillus* (formerly *Lactobacillus*) *reuteri* were significantly more abundant in the *S. hispidus* gut (181). Upon isolation from *S. hispidus* stool, we have sequenced, assembled, and annotated the draft genome of these two unique isolates: *Lactobacillus paragasseri BRTN* and *Limosilactobacillus reuteri SMN.*

For isolation, frozen stool pellets were collected from a male *S. hispidus* and resuspended in PBS. Serial dilutions were plated on *Lactobacilli* MRS agar (BD Biosciences) and incubated at 37ºC for 48h. Colony PCR screen was performed to identify *Lactobacilli* from stool, using *Lactobacillus*-specific primers (F: GAGGCAGCAGTAGGGAATCTTC, R: GGCCAGTTACTACCTCTATCCTTCTTC) (257) and MyTaq HS Red (Bioline®) with the following cycling conditions: 95°C/2 min; 30x (95°C/20sec, 50°C/15sec, 72°C/60sec); 72°C/10min. PCR products were run on 1% agarose gel to verify *Lactobacillus*-positive colonies. PCR was then repeated using universal 16s rRNA primers 27F-NoDeg (AGAGTTTGATCCTGGCTCAG)/1492R-NoDeg (GGTTACCTTGTTACGACTT) (243) (cycling conditions listed above) and purified PCR products were sequenced via Sanger (IDT). Most colonies were taxonomically classified as *Lactobacillus paragasseri* and *Limosilactobacillus reuteri* (NCBI BLAST (276)).

Clonal isolates of *L. paragasseri* and *L. reuteri* were incubated in MRS broth for 16hrs at 37ºC with no shaking, and genomic DNA was extracted from bacterial pellets (washed 2x with PBS) using DNeasy PowerSoil Kit (Qiagen) with manufacturer's protocol. Genomic libraries were prepared using the NEBNext® Ultra™ IS DNA Library Prep Kit (Illumina). Library quality was assessed on an Agilent 2100 Bioanalyzer System using the Agilent High Sensitivity DNA Kit (5067-4626). Libraries were sequenced using NovaSeq 6000 2x150 (Illumina) to get 40 million reads. FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to examine data quality. Trimmomatic v0.39 (308) was used to remove adaptors with parameters TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75. The package *ASA3P* v 1.3.0 (404) was used for bacterial genome assembly, scaffolding, and annotation following default

parameters. CheckM v1.4.0 (405) and BUSCO v5.0.0 (406) were used to assess the completeness of assembly and annotation respectively with default parameters.

Using *de novo* assembly, one isolate was then identified as *Lactobacillus paragasseri*, with 98.79% average nucleotide identity and 82.64% conserved DNA against *L. paragasseri* strain JCM 5343 [GenBank accession number AP018549]. The final assembly yielded a 2,096,736bp genome consisting of 22 contigs with an N50 value of 590,019 and GC content of 35%. From the final assembly, 2,087 genes were identified with 1,991 protein-coding genes, of which 85.9% were annotated with biological function. The genome contained 295 hypothetical proteins, 33 ncRNAs, 3 rRNAs, 59 tRNAs, and 1 tmRNA. BUSCO indicated the annotation as 91.9% complete against the bacteria_odb10 reference database (Complete:91.9%, Fragmented:1.6%,Missing:6.5%,n BUSCOs:124). CheckM indicated the assembly was 99.22% complete based on 449 markers within 58 *Lactobacillus* genomes. Verification of scaffolds was performed by complete alignment to *Lactobacillus paragasseri* (strain JCM 5343 [GenBank accession number AP018549]). The associated sequencing data and assembled/annotated genome are deposited under BioProject PRJNA816882.

Using *de novo* assembly, the other isolate was identified as *Limosilactobacillus reuteri*, with 97.06% average nucleotide identity and 74.45% conserved DNA against *L. reuteri* subspecies *reuteri* [RefSeq accession number AP018549]). The final assembly yielded a 2,180,873bp genome consisting of 17 scaffolds of 33 contigs with an N50 value of 115,069 and GC content of 38%. 2,169 genes were identified with 2,051 protein-coding genes, of which 87.8% were annotated with biological function. In addition, the genome contained 265 hypothetical proteins, 46 ncRNAs, 4 rRNAs, 67 tRNAs, and 1 tmRNA.

BUSCO indicated the annotation as 99.2% complete against the bacteria_odb10 reference database (Complete:98.4%, Fragmented:0.8%,Missing:0.0%,n BUSCOs:124). CheckM indicated the assembly was 99.46% complete based on 336 markers within 58 *Lactobacillus* genomes. The associated sequencing data and assembled/annotated genomes are deposited under BioProject PRJNA816885.

These are the first genome sequences of any bacteria isolated from cotton rats and the only *Lactobacillus reuteri* and *Limosilactobacillus reuteri* genomes from non-human sources except a dairy cow isolate (BioSample SAMN14239244) based on BioSample data acquired from https://www.ncbi.nlm.nih.gov/datasets/genomes as of August 11, 2022. The availability of this genome will help inform unique features of probiotic bacteria that have adapted to unique hosts.

1. Feldman AS, He Y, Moore ML, Hershenson MB, and Hartert TV. Toward Primary Prevention of Asthma. Reviewing the Evidence for Early-Life Respiratory Viral Infections as Modifiable Risk Factors to Prevent Childhood Asthma. *American Journal of Respiratory and Critical Care Medicine.* 2015;191(1):34-44.
2. Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, et al. The Burden of Respiratory Syncytial Virus Infection in Young Children. *New England Journal of Medicine.* 2009;360(6):588-98.
3. Li Y, Wang X, Blau DM, Caballero MT, Feikin DR, Gill CJ, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. *Lancet.* 2022;399(10340):2047-64.
4. Srikantiah P, Vora P, and Klugman KP. Assessing the Full Burden of Respiratory Syncytial Virus in Young Infants in Low- and Middle-Income Countries: The Importance of Community Mortality Studies. *Clin Infect Dis.* 2021;73(Suppl_3):S177-S9.
5. Borchers AT, Chang C, Gershwin ME, and Gershwin LJ. Respiratory syncytial virus--a comprehensive review. *Clin Rev Allergy Immunol.* 2013;45(3):331-79.
6. Hall CB, Weinberg GA, Blumkin AK, Edwards KM, Staat MA, Schultz AF, et al. Respiratory syncytial virus-associated hospitalizations among children less than 24 months of age. *Pediatrics.* 2013;132(2):e341-8.
7. Glezen WP, Taber LH, Frank AL, and Kasel JA. Risk of primary infection and reinfection with respiratory syncytial virus. *Am J Dis Child.* 1986;140(6):543-6.
8. Wu P, and Hartert TV. Evidence for a causal relationship between respiratory syncytial virus infection and asthma. *Expert Rev Anti Infect Ther.* 2011;9(9):731-45.
9. Sigurs N, Aljassim F, Kjellman B, Robinson PD, Sigurbergsson F, Bjarnason R, et al. Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life. *Thorax.* 2010;65(12):1045-52.
10. Rha B, Curns AT, Lively JY, Campbell AP, Englund JA, Boom JA, et al. Respiratory Syncytial Virus-Associated Hospitalizations Among Young Children: 2015-2016. *Pediatrics.* 2020;146(1).
11. Goldstein E, Finelli L, O'Halloran A, Liu P, Karaca Z, Steiner CA, et al. Hospitalizations Associated with Respiratory Syncytial Virus and Influenza in Children, Including Children Diagnosed with Asthma. *Epidemiology.* 2019;30(6):918-26.
12. Régnier S, and Huels J. Association between Respiratory Syncytial Virus Hospitalizations in Infants and Respiratory Sequelae. *The Pediatric Infectious Disease Journal.* 2013;32(8):1-.

13. Carroll KN, Wu P, Gebretsadik T, Griffin MR, Dupont WD, Mitchel EF, et al. The severity-dependent relationship of infant bronchiolitis on the risk and morbidity of early childhood asthma. *J Allergy Clin Immunol.* 2009;123(5):1055-61, 61 e1.

14. Schauer U, Hoffjan S, Bittscheidt J, Kochling A, Hemmis S, Bongartz S, et al. RSV bronchiolitis and risk of wheeze and allergic sensitisation in the first year of life. *Eur Respir J.* 2002;20(5):1277-83.

15. Peebles RS, Jr., Hashimoto K, and Graham BS. The complex relationship between respiratory syncytial virus and allergy in lung disease. *Viral Immunol.* 2003;16(1):25-34.

16. Hall CB, Douglas RG, Jr., Schnabel KC, and Geiman JM. Infectivity of respiratory syncytial virus by various routes of inoculation. *Infect Immun.* 1981;33(3):779-83.

17. Johnson JE, Gonzales RA, Olson SJ, Wright PF, and Graham BS. The histopathology of fatal untreated human respiratory syncytial virus infection. *Modern Pathology.* 2007;20(1):108-19.

18. Hall CB, Walsh EE, Long CE, and Schnabel KC. Immunity to and frequency of reinfection with respiratory syncytial virus. *J Infect Dis.* 1991;163(4):693-8.

19. American Academy of Pediatrics Committee on Infectious D, and American Academy of Pediatrics Bronchiolitis Guidelines C. Updated Guidance for Palivizumab Prophylaxis Among Infants and Young Children at Increased Risk of Hospitalization for Respiratory Syncytial Virus Infection. *PEDIATRICS.* 2014;134(2):e620-e38.

20. Andabaka T, Nickerson JW, Rojas-Reyes MX, Rueda JD, Bacic Vrca V, and Barsic B. Monoclonal antibody for reducing the risk of respiratory syncytial virus infection in children. *Cochrane Database of Systematic Reviews.* 2013(4):CD006602-CD.

21. Fulginiti VA, Eller JJ, Sieber OF, Joyner JW, Minamitani M, and Meiklejohn G. Respiratory virus immunization. I. A field trial of two inactivated respiratory virus vaccines; an aqueous trivalent parainfluenza virus vaccine and an alum-precipitated respiratory syncytial virus vaccine. *Am J Epidemiol.* 1969;89(4):435-48.

22. Chin J, Magoffin RL, Shearer LA, Schieble JH, and Lennette EH. Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. *Am J Epidemiol.* 1969;89(4):449-63.

23. Kapikian AZ, Mitchell RH, Chanock RM, Shvedoff RA, and Stewart CE. An epidemiologic study of altered clinical reactivity to respiratory syncytial (RS) virus infection in children previously vaccinated with an inactivated RS virus vaccine. *Am J Epidemiol.* 1969;89(4):405-21.

24. Endt K, Wollmann Y, Haug J, Bernig C, Feigl M, Heiseke A, et al. A Recombinant MVA-Based RSV Vaccine Induces T-Cell and Antibody Responses That Cooperate in the Protection Against RSV Infection. *Front Immunol.* 2022;13:841471.

25. Detalle L, Stohr T, Palomo C, Piedra PA, Gilbert BE, Mas V, et al. Generation and Characterization of ALX-0171, a Potent Novel Therapeutic Nanobody for the Treatment of Respiratory Syncytial Virus Infection. *Antimicrob Agents Chemother.* 2016;60(1):6-13.

26. Tang W, Li M, Liu Y, Liang N, Yang Z, Zhao Y, et al. Small molecule inhibits respiratory syncytial virus entry and infection by blocking the interaction of the viral fusion protein with the cell membrane. *FASEB J.* 2019;33(3):4287-99.

27. Shan J, Britton PN, King CL, and Booy R. The immunogenicity and safety of respiratory syncytial virus vaccines in development: A systematic review. *Influenza Other Respir Viruses.* 2021;15(4):539-51.

28. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet.* 2017;390(10098):946-58.

29. Blount RE, Jr., Morris JA, and Savage RE. Recovery of cytopathogenic agent from chimpanzees with coryza. *Proc Soc Exp Biol Med.* 1956;92(3):544-9.

30. Chanock R, and Finberg L. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiologic aspects of infection in infants and young children. *Am J Hyg.* 1957;66(3):291-300.

31. Ke Z, Dillard RS, Chirkova T, Leon F, Stobart CC, Hampton CM, et al. The Morphology and Assembly of Respiratory Syncytial Virus Revealed by Cryo-Electron Tomography. *Viruses.* 2018;10(8).

32. Alshaghdali K, Saeed M, Kamal MA, and Saeed A. Interaction of Ectodomain of Respiratory Syncytial Virus G Protein with TLR2/ TLR6 Heterodimer: An In vitro and In silico Approach to Decipher the Role of RSV G Protein in Pro-inflammatory Response against the Virus. *Curr Pharm Des.* 2021;27(44):4464-76.

33. Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, and Anderson LJ. CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nat Immunol.* 2001;2(8):732-8.

34. Feldman SA, Hendry RM, and Beeler JA. Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J Virol.* 1999;73(8):6610-7.

35. Green G, Johnson SM, Costello H, Brakel K, Harder O, Oomens AG, et al. CX3CR1 Is a Receptor for Human Respiratory Syncytial Virus in Cotton Rats. *J Virol.* 2021;95(16):e0001021.

36. Chirkova T, Boyoglu-Barnum S, Gaston KA, Malik FM, Trau SP, Oomens AG, et al. Respiratory syncytial virus G protein CX3C motif impairs human airway epithelial and immune cell responses. *J Virol.* 2013;87(24):13466-79.

37. Zhivaki D, Lemoine S, Lim A, Morva A, Vidalain PO, Schandene L, et al. Respiratory Syncytial Virus Infects Regulatory B Cells in Human Neonates via Chemokine Receptor CX3CR1 and Promotes Lung Disease Severity. *Immunity.* 2017;46(2):301-14.

38. Lo MS, Brazas RM, and Holtzman MJ. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness. *J Virol.* 2005;79(14):9315-9.

39. Currier MG, Lee S, Stobart CC, Hotard AL, Villenave R, Meng J, et al. EGFR Interacts with the Fusion Protein of Respiratory Syncytial Virus Strain 2-20 and Mediates Infection and Mucin Expression. *PLOS Pathogens.* 2016;12(5):e1005622-e.

40. Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, and Hegele RG. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. *Nat Med.* 2011;17(9):1132-5.

41. Griffiths CD, Bilawchuk LM, McDonough JE, Jamieson KC, Elawar F, Cen Y, et al. IGF1R is an entry receptor for respiratory syncytial virus. *Nature.* 2020;583(7817):615-9.

42. Behera AK, Matsuse H, Kumar M, Kong X, Lockey RF, and Mohapatra SS. Blocking intercellular adhesion molecule-1 on human epithelial cells decreases respiratory syncytial virus infection. *Biochem Biophys Res Commun.* 2001;280(1):188-95.

43. Chaiwatpongsakorn S, Epand RF, Collins PL, Epand RM, and Peeples ME. Soluble respiratory syncytial virus fusion protein in the fully cleaved, pretriggered state is triggered by exposure to low-molarity buffer. *J Virol.* 2011;85(8):3968-77.

44. Mehedi M, McCarty T, Martin SE, Le Nouen C, Buehler E, Chen YC, et al. Actin-Related Protein 2 (ARP2) and Virus-Induced Filopodia Facilitate Human Respiratory Syncytial Virus Spread. *PLoS Pathog.* 2016;12(12):e1006062.

45. Kidd P. Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. *Altern Med Rev.* 2003;8(3):223-46.

46. Graham BS, Rutigliano JA, and Johnson TR. Respiratory syncytial virus immunobiology and pathogenesis. *Virology.* 2002;297(1):1-7.

47. Welliver RC, Wong DT, Sun M, Middleton E, Jr., Vaughan RS, and Ogra PL. The development of respiratory syncytial virus-specific IgE and the release of histamine in nasopharyngeal secretions after infection. *N Engl J Med.* 1981;305(15):841-6.

48. Lanari M, Vandini S, Capretti MG, Lazzarotto T, and Faldella G. Respiratory syncytial virus infections in infants affected by primary immunodeficiency. *J Immunol Res.* 2014;2014:850831.

49. Siezen CL, Bont L, Hodemaekers HM, Ermers MJ, Doornbos G, Van't Slot R, et al. Genetic susceptibility to respiratory syncytial virus bronchiolitis in preterm children is associated with airway remodeling genes and innate immune genes. *Pediatr Infect Dis J.* 2009;28(4):333-5.

50. Welliver RC. Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. *J Pediatr.* 2003;143(5 Suppl):S112-7.

51. Abraha HY, Lanctot KL, and Paes B. Risk of respiratory syncytial virus infection in preterm infants: reviewing the need for prevention. *Expert Rev Respir Med.* 2015;9(6):779-99.

52. Kollmann TR, Crabtree J, Rein-Weston A, Blimkie D, Thommai F, Wang XY, et al. Neonatal innate TLR-mediated responses are distinct from those of adults. *J Immunol.* 2009;183(11):7150-60.

53. Kristjansson S, Bjarnarson SP, Wennergren G, Palsdottir AH, Arnadottir T, Haraldsson A, et al. Respiratory syncytial virus and other respiratory viruses during the first 3 months of life promote a local TH2-like response. *Journal of Allergy and Clinical Immunology.* 2005.

54. Fonseca W, Lucey K, Jang S, Fujimura KE, Rasky A, Ting HA, et al. Lactobacillus johnsonii supplementation attenuates respiratory viral infection via

metabolic reprogramming and immune cell modulation. *Mucosal Immunology.* 2017;10(6):1569-80.

55. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, et al. Nasopharyngeal Lactobacillus is associated with a reduced risk of childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy. *Journal of Allergy and Clinical Immunology.* 2018;0(0).

56. Aryana KJ, and Olson DW. A 100-Year Review: Yogurt and other cultured dairy products. *J Dairy Sci.* 2017;100(12):9987-10013.

57. Metchnikoff E, and Mitchell PC. *The prolongation of life.* New York & London,: G. P. Putnam's sons; 1910.

58. Weirich A, and Hoffmann GF. Ernst Moro (1874-1951)--a great pediatric career started at the rise of university-based pediatric research but was curtailed in the shadows of Nazi laws. *Eur J Pediatr.* 2005;164(10):599-606.

59. Fisberg M, and Machado R. History of yogurt and current patterns of consumption. *Nutr Rev.* 2015;73 Suppl 1:4-7.

60. Zheng J, Wittouck S, Salvetti E, Franz C, Harris HMB, Mattarelli P, et al. 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.* 2020;70(4):2782-858.

61. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol.* 2014;11(8):506-14.

62. Walter J. Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl Environ Microbiol.* 2008;74(16):4985-96.

63. Eiseman B, Silen W, Bascom GS, and Kauvar AJ. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery.* 1958;44(5):854-9.

64. Abrams GD, Bauer H, and Sprinz H. Influence of the normal flora on mucosal morphology and cellular renewal in the ileum. A comparison of germ-free and conventional mice. *Lab Invest.* 1963;12:355-64.

65. Janeway CA, Jr. The immune system evolved to discriminate infectious nonself from noninfectious self. *Immunol Today.* 1992;13(1):11-6.

66. Matzinger P. The danger model: a renewed sense of self. *Science.* 2002;296(5566):301-5.

67. Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, and Medzhitov R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell.* 2004;118(2):229-41.

68. Hooper LV, Falk PG, and Gordon JI. Analyzing the molecular foundations of commensalism in the mouse intestine. *Curr Opin Microbiol.* 2000;3(1):79-85.

69. Unger SA, and Bogaert D. The respiratory microbiome and respiratory infections. *J Infect.* 2017;74 Suppl 1:S84-S8.

70. Atherton JC, and Blaser MJ. Coadaptation of Helicobacter pylori and humans: ancient history, modern implications. *The Journal of Clinical Investigation.* 2009;119(9):2475-87.

71.   Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet.* 2016;48(11):1413-7.

72.   Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic Determinants of the Gut Microbiome in UK Twins. *Cell Host Microbe.* 2016;19(5):731-43.

73.   Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, and Bordenstein SR. Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLoS Biol.* 2016;14(11):e2000225.

74.   Gaulke CA, Arnold HK, Humphreys IR, Kembel SW, O'Dwyer JP, and Sharpton TJ. Ecophylogenetics Clarifies the Evolutionary Association between Mammals and Their Gut Microbiota. *mBio.* 2018;9(5).

75.   Markowitz RHG, LaBella AL, Shi M, Rokas A, Capra JA, Ferguson JF, et al. Microbiome-associated human genetic variants impact phenome-wide disease risk. *Proc Natl Acad Sci U S A.* 2022;119(26):e2200551119.

76.   Korach-Rechtman H, Freilich S, Gerassy-Vainberg S, Buhnik-Rosenblau K, Danin-Poleg Y, Bar H, et al. Murine Genetic Background Has a Stronger Impact on the Composition of the Gut Microbiota than Maternal Inoculation or Exposure to Unlike Exogenous Microbiota. *Applied and environmental microbiology.* 2019;85(18).

77.   Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature.* 2018;555(7695):210-5.

78.   Wopereis H, Oozeer R, Knipping K, Belzer C, and Knol J. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. *Pediatr Allergy Immunol.* 2014;25(5):428-38.

79.   Korpela K, and de Vos WM. Antibiotic use in childhood alters the gut microbiota and predisposes to overweight. *Microb Cell.* 2016;3(7):296-8.

80.   Bernstein CN, Burchill C, Targownik LE, Singh H, and Roos LL. Events Within the First Year of Life, but Not the Neonatal Period, Affect Risk for Later Development of Inflammatory Bowel Diseases. *Gastroenterology.* 2019;156(8):2190-7 e10.

81.   Cavalcante GG, Guimaraes AG, Queiroz-Glauss CP, Goncalves Pereira MH, Dias ASL, Horta LS, et al. Treatment with Distinct Antibiotic Classes Causes Different Pulmonary Outcomes on Allergic Airway Inflammation Associated with Modulation of Symbiotic Microbiota. *J Immunol Res.* 2022;2022:1466011.

82.   Ferretti P, Pasolli E, Tett A, Asnicar F, Gorfer V, Fedi S, et al. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. *Cell Host Microbe.* 2018;24(1):133-45 e5.

83.   Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A.* 2010;107(26):11971-5.

84.   Gronlund MM, Gueimonde M, Laitinen K, Kociubinski G, Gronroos T, Salminen S, et al. Maternal breast-milk and intestinal bifidobacteria guide the compositional

development of the Bifidobacterium microbiota in infants at risk of allergic disease. *Clin Exp Allergy.* 2007;37(12):1764-72.

85.   Yan W, Luo B, Zhang X, Ni Y, and Tian F. Association and Occurrence of Bifidobacterial Phylotypes Between Breast Milk and Fecal Microbiomes in Mother-Infant Dyads During the First 2 Years of Life. *Front Microbiol.* 2021;12:669442.

86.   Laursen MF, Sakanaka M, von Burg N, Morbe U, Andersen D, Moll JM, et al. Bifidobacterium species associated with breastfeeding produce aromatic lactic acids in the infant gut. *Nat Microbiol.* 2021;6(11):1367-82.

87.   Fujimura KE, Demoor T, Rauch M, Faruqi AA, Jang S, Johnson CC, et al. House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. *Proceedings of the National Academy of Sciences.* 2014;111(2):805-10.

88.   Ownby DR, Johnson CC, and Peterson EL. Exposure to dogs and cats in the first year of life and risk of allergic sensitization at 6 to 7 years of age. *JAMA.* 2002;288(8):963-72.

89.   Karmaus W, and Botezan C. Does a higher number of siblings protect against the development of allergy and asthma? A review. *J Epidemiol Community Health.* 2002;56(3):209-17.

90.   Stokholm J, Thorsen J, Blaser MJ, Rasmussen MA, Hjelmso M, Shah S, et al. Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Sci Transl Med.* 2020;12(569).

91.   Jakubczyk D, and Gorska S. Impact of Probiotic Bacteria on Respiratory Allergy Disorders. *Front Microbiol.* 2021;12:688137.

92.   Bloomfield SF, Rook GA, Scott EA, Shanahan F, Stanwell-Smith R, and Turner P. Time to abandon the hygiene hypothesis: new perspectives on allergic disease, the human microbiome, infectious disease prevention and the role of targeted hygiene. *Perspect Public Health.* 2016;136(4):213-24.

93.   Strachan DP. Hay fever, hygiene, and household size. *BMJ.* 1989;299:1259-60.

94.   Dominguez-Bello MG, De Jesus-Laboy KM, Shen N, Cox LM, Amir A, Gonzalez A, et al. Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. *Nat Med.* 2016;22(3):250-3.

95.   Chu DM, Ma J, Prince AL, Antony KM, Seferovic MD, and Aagaard KM. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med.* 2017;23(3):314-26.

96.   Wu P, Feldman AS, Rosas-Salazar C, James K, Escobar G, Gebretsadik T, et al. Relative Importance and Additive Effects of Maternal and Infant Risk Factors on Childhood Asthma. *PLoS One.* 2016;11(3):e0151705.

97.   Vital M, Harkema JR, Rizzo M, Tiedje J, and Brandenberger C. Alterations of the Murine Gut Microbiome with Age and Allergic Airway Disease. *J Immunol Res.* 2015;2015:892568.

98.   Abrahamsson TR, Jakobsson HE, Andersson AF, Bjorksten B, Engstrand L, and Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy.* 2014;44(6):842-50.

99. Zheng D, Liwinski T, and Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Res.* 2020;30(6):492-506.

100. Yan F, and Polk DB. Probiotics and Probiotic-Derived Functional Factors-Mechanistic Insights Into Applications for Intestinal Homeostasis. *Front Immunol.* 2020;11:1428.

101. Kamdar K, Johnson AMF, Chac D, Myers K, Kulur V, Truevillian K, et al. Innate Recognition of the Microbiota by TLR1 Promotes Epithelial Homeostasis and Prevents Chronic Inflammation. *J Immunol.* 2018;201(1):230-42.

102. Wells JM, Rossi O, Meijerink M, and van Baarlen P. Epithelial crosstalk at the microbiota-mucosal interface. *Proceedings of the National Academy of Sciences.* 2010;108(Supplement_1):4607-14.

103. Mazmanian SK, Liu CH, Tzianabos AO, and Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell.* 2005;122(1):107-18.

104. Alvarez CA, Jones MB, Hambor J, and Cobb BA. Characterization of Polysaccharide A Response Reveals Interferon Responsive Gene Signature and Immunomodulatory Marker Expression. *Front Immunol.* 2020;11:556813.

105. Antunes KH, Stein RT, Franceschina C, da Silva EF, de Freitas DN, Silveira J, et al. Short-chain fatty acid acetate triggers antiviral response mediated by RIG-I in cells from infants with respiratory syncytial virus bronchiolitis. *EBioMedicine.* 2022;77:103891.

106. Antunes KH, Fachi JL, de Paula R, da Silva EF, Pral LP, Dos Santos AA, et al. Microbiota-derived acetate protects against respiratory syncytial virus infection through a GPR43-type 1 interferon response. *Nat Commun.* 2019;10(1):3273.

107. Abreu NA, Nagalingam NA, Song Y, Roediger FC, Pletcher SD, Goldberg AN, et al. Sinus microbiome diversity depletion and Corynebacterium tuberculostearicum enrichment mediates rhinosinusitis. *Sci Transl Med.* 2012;4(151):151ra24.

108. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al. Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proceedings of the National Academy of Sciences of the United States of America.* 2011;108(13):5354-9.

109. Barcik W, Boutin RCT, Sokolowska M, and Finlay BB. The Role of Lung and Gut Microbiota in the Pathology of Asthma. *Immunity.* 2020;52(2):241-55.

110. Dang AT, and Marsland BJ. Microbes, metabolites, and the gut-lung axis. *Mucosal Immunol.* 2019.

111. Marsland BJ, Trompette A, and Gollwitzer ES. The Gut-Lung Axis in Respiratory Disease. *Ann Am Thorac Soc.* 2015;12 Suppl 2:S150-6.

112. Wedgwood S, Warford C, Agvatisiri SR, Thai PN, Chiamvimonvat N, Kalanetra KM, et al. The developing gut-lung axis: postnatal growth restriction, intestinal dysbiosis, and pulmonary hypertension in a rodent model. *Pediatr Res.* 2020;87(3):472-9.

113. Yazar A, Atis S, Konca K, Pata C, Akbay E, Calikoglu M, et al. Respiratory symptoms and pulmonary functional changes in patients with irritable bowel syndrome. *Am J Gastroenterol.* 2001;96(5):1511-6.

114. Mirsepasi-Lauridsen HC, Vrankx K, Engberg J, Friis-Moller A, Brynskov J, Nordgaard-Lassen I, et al. Disease-Specific Enteric Microbiome Dysbiosis in Inflammatory Bowel Disease. *Front Med (Lausanne).* 2018;5:304.

115. Liu Q, Tian X, Maruyama D, Arjomandi M, and Prakash A. Lung immune tone via gut-lung axis: gut-derived LPS and short-chain fatty acids' immunometabolic regulation of lung IL-1beta, FFAR2, and FFAR3 expression. *Am J Physiol Lung Cell Mol Physiol.* 2021;321(1):L65-L78.

116. Pu Q, Lin P, Gao P, Wang Z, Guo K, Qin S, et al. Gut Microbiota Regulate Gut-Lung Axis Inflammatory Responses by Mediating ILC2 Compartmental Migration. *J Immunol.* 2021;207(1):257-67.

117. Wang J, Li F, Wei H, Lian ZX, Sun R, and Tian Z. Respiratory influenza virus infection induces intestinal immune injury via microbiota-mediated Th17 cell-dependent inflammation. *J Exp Med.* 2014;211(12):2397-410.

118. Bliss ES, and Whiteside E. The gut-brain axis, the human gut microbiota and their integration in the development of obesity. *Frontiers in Physiology.* 2018;9(JUL).

119. Zaiss MM, Joyce Wu HJ, Mauro D, Schett G, and Ciccia F. The gut-joint axis in rheumatoid arthritis. *Nat Rev Rheumatol.* 2021;17(4):224-37.

120. Jang MJ, Kim YJ, Hong S, Na J, Hwang JH, Shin SM, et al. Positive association of breastfeeding on respiratory syncytial virus infection in hospitalized infants: a multicenter retrospective study. *Clin Exp Pediatr.* 2020;63(4):135-40.

121. Nishimura T, Suzue J, and Kaji H. Breastfeeding reduces the severity of respiratory syncytial virus infection among young infants: a multi-center prospective study. *Pediatr Int.* 2009;51(6):812-6.

122. Kristensen K, Fisker N, Haerskjold A, Ravn H, Simoes EA, and Stensballe L. Caesarean section and hospitalization for respiratory syncytial virus infection: a population-based study. *Pediatr Infect Dis J.* 2015;34(2):145-8.

123. Miller JE, Wu C, Pedersen LH, de Klerk N, Olsen J, and Burgner DP. Maternal antibiotic exposure during pregnancy and hospitalization with infection in offspring: a population-based cohort study. *Int J Epidemiol.* 2018;47(2):561-71.

124. Harding JN, Siefker D, Vu L, You D, DeVincenzo J, Pierre JF, et al. Altered gut microbiota in infants is associated with respiratory syncytial virus disease severity. *BMC Microbiol.* 2020;20(1):140.

125. Hasegawa K, Linnemann RW, Mansbach JM, Ajami NJ, Espinola JA, Petrosino JF, et al. The Fecal Microbiota Profile and Bronchiolitis in Infants. *Pediatrics.* 2016;138(1).

126. Fonseca W, Lucey K, Jang S, Fujimura KE, Rasky A, Ting HA, et al. Lactobacillus johnsonii supplementation attenuates respiratory viral infection via metabolic reprogramming and immune cell modulation. *Mucosal Immunol.* 2017;10(6):1569-80.

127. Fonseca W, Malinczak CA, Fujimura K, Li D, McCauley K, Li J, et al. Maternal gut microbiome regulates immunity to RSV infection in offspring. *J Exp Med.* 2021;218(11).

128. Forsythe P, Inman MD, and Bienenstock J. Oral treatment with live Lactobacillus reuteri inhibits the allergic airway response in mice. *American Journal of Respiratory and Critical Care Medicine.* 2007;175(6):561-9.

129. Ji JJ, Sun QM, Nie DY, Wang Q, Zhang H, Qin FF, et al. Probiotics protect against RSV infection by modulating the microbiota-alveolar-macrophage axis. *Acta Pharmacol Sin.* 2021;42(10):1630-41.

130. Chiba E, Tomosada Y, Vizoso-Pinto MG, Salva S, Takahashi T, Tsukida K, et al. Immunobiotic Lactobacillus rhamnosus improves resistance of infant mice against respiratory syncytial virus infection. *International Immunopharmacology.* 2013;17(2):373-82.

131. Garcia-Castillo V, Tomokiyo M, Raya Tonetti F, Islam MA, Takahashi H, Kitazawa H, et al. Alveolar Macrophages Are Key Players in the Modulation of the Respiratory Antiviral Immunity Induced by Orally Administered Lacticaseibacillus rhamnosus CRL1505. *Front Immunol.* 2020;11:568636.

132. Villena JC, Salva MS, Nuñez MS, Corzo J, Tolaba R, Faedda J, et al. Probiotics for everyone! The novel immunobiotic Lactobacillus rhamnosus CRL1505 and the beginning of Social Probiotic Programs in Argentina. 2012.

133. Eguchi K, Fujitani N, Nakagawa H, and Tadaaki M. Prevention of respiratory syncytial virus infection with probiotic lactic acid bacterium Lactobacillus gasseri SBT2055. *Scientific Reports.* 2019(December 2018):1-2.

134. de Steenhuijsen Piters WA, Heinonen S, Hasrat R, Bunsow E, Smith B, Suarez-Arrabal MC, et al. Nasopharyngeal Microbiota, Host Transcriptome, and Disease Severity in Children with Respiratory Syncytial Virus Infection. *Am J Respir Crit Care Med.* 2016;194(9):1104-15.

135. Ederveen THA, Ferwerda G, Ahout IM, Vissers M, de Groot R, Boekhorst J, et al. Haemophilus is overrepresented in the nasopharynx of infants hospitalized with RSV infection and associated with increased viral load and enhanced mucosal CXCL8 responses. *Microbiome.* 2018;6(1):10.

136. Mansbach JM, Hasegawa K, Piedra PA, Avadhanula V, Petrosino JF, Sullivan AF, et al. Haemophilus-Dominant Nasopharyngeal Microbiota Is Associated With Delayed Clearance of Respiratory Syncytial Virus in Infants Hospitalized for Bronchiolitis. *J Infect Dis.* 2019;219(11):1804-8.

137. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Chappell JD, Larkin EK, Nelson KE, et al. Nasopharyngeal Microbiome in Respiratory Syncytial Virus Resembles Profile Associated with Increased Childhood Asthma Risk. *American Journal of Respiratory and Critical Care Medicine.* 2016;193(10):1180-3.

138. Hasegawa K, Mansbach JM, Ajami NJ, Espinola JA, Henke DM, Petrosino JF, et al. Association of nasopharyngeal microbiota profiles with bronchiolitis severity in infants hospitalised for bronchiolitis. *Eur Respir J.* 2016;48(5):1329-39.

139. De Boeck I, van den Broek MFL, Allonsius CN, Spacova I, Wittouck S, Martens K, et al. Lactobacilli Have a Niche in the Human Nose. *Cell Rep.* 2020;31(8):107674.

140. Clua P, Kanmani P, Zelaya H, Tada A, Humayun Kober AKM, Salva S, et al. Peptidoglycan from immunobiotic lactobacillus rhamnosus improves resistance of infant Mice to respiratory syncytial viral infection and secondary pneumococcal pneumonia. *Frontiers in Immunology.* 2017;8(AUG).

141. Tomosada Y, Chiba E, Zelaya H, Takahashi T, Tsukida K, Kitazawa H, et al. Nasally administered Lactobacillus rhamnosus strains differentially modulate

respiratory antiviral immune responses and induce protection against respiratory syncytial virus infection. *BMC Immunology.* 2013;14(1):40-.

142. Youn H-N, Lee D-H, Lee Y-N, Park J-K, Yuk S-S, Yang S-Y, et al. Intranasal administration of live Lactobacillus species facilitates protection against influenza virus infection in mice. *Antiviral Research.* 2012;93(1):138-43.

143. Wang B, Hylwka T, Smieja M, Surrette M, Bowdish DME, and Loeb M. Probiotics to Prevent Respiratory Infections in Nursing Homes: A Pilot Randomized Controlled Trial. *J Am Geriatr Soc.* 2018;66(7):1346-52.

144. Luoto R, Ruuskanen O, Waris M, Kalliomaki M, Salminen S, and Isolauri E. Prebiotic and probiotic supplementation prevents rhinovirus infections in preterm infants: a randomized, placebo-controlled trial. *J Allergy Clin Immunol.* 2014;133(2):405-13.

145. Lehtoranta L, Kalima K, He L, Lappalainen M, Roivainen M, Narkio M, et al. Specific probiotics and virological findings in symptomatic conscripts attending military service in Finland. *J Clin Virol.* 2014;60(3):276-81.

146. Sugimura T, Takahashi H, Jounai K, Ohshio K, Kanayama M, Tazumi K, et al. Effects of oral intake of plasmacytoid dendritic cells-stimulative lactic acid bacterial strain on pathogenesis of influenza-like illness and immunological response to influenza virus. *Br J Nutr.* 2015;114(5):727-33.

147. Davidson LE, Fiorino AM, Snydman DR, and Hibberd PL. Lactobacillus GG as an immune adjuvant for live-attenuated influenza vaccine in healthy adults: a randomized double-blind placebo-controlled trial. *Eur J Clin Nutr.* 2011;65(4):501-7.

148. Turner RB, Woodfolk JA, Borish L, Steinke JW, Patrie JT, Muehling LM, et al. Effect of probiotic on innate inflammatory response and viral shedding in experimental rhinovirus infection - a randomised controlled trial. *Benef Microbes.* 2017;8(2):207-15.

149. Tapiovaara L, Kumpu M, Makivuokko H, Waris M, Korpela R, Pitkaranta A, et al. Human rhinovirus in experimental infection after peroral Lactobacillus rhamnosus GG consumption, a pilot study. *Int Forum Allergy Rhinol.* 2016;6(8):848-53.

150. Kumpu M, Kekkonen RA, Korpela R, Tynkkynen S, Järvenpää S, Kautiainen H, et al. Effect of live and inactivated <i>Lactobacillus rhamnosus</i> GG on experimentally induced rhinovirus colds: randomised, double blind, placebo-controlled pilot trial. *Beneficial Microbes.* 2015;6(5):631-9.

151. Makino S, Ikegami S, Kume A, Horiuchi H, Sasaki H, and Orii N. Reducing the risk of infection in the elderly by dietary intake of yoghurt fermented with Lactobacillus delbrueckii ssp. bulgaricus OLL1073R-1. *Br J Nutr.* 2010;104(7):998-1006.

152. Jespersen L, Tarnow I, Eskesen D, Morberg CM, Michelsen B, Bugel S, et al. Effect of Lactobacillus paracasei subsp. paracasei, L. casei 431 on immune response to influenza vaccination and upper respiratory tract infections in healthy adult volunteers: a randomized, double-blind, placebo-controlled, parallel-group study. *Am J Clin Nutr.* 2015;101(6):1188-96.

153. Kumpu M, Lehtoranta L, Roivainen M, Ronkko E, Ziegler T, Soderlund-Venermo M, et al. The use of the probiotic Lactobacillus rhamnosus GG and viral findings

in the nasopharynx of children attending day care. *J Med Virol.* 2013;85(9):1632-8.

154.   Shi HY, Zhu X, Li WL, Mak JWY, Wong SH, Zhu ST, et al. Modulation of gut microbiota protects against viral respiratory tract infections: a systematic review of animal and clinical studies. *Eur J Nutr.* 2021;60(8):4151-74.

155.   Lehtoranta L, Latvala S, and Lehtinen MJ. Role of Probiotics in Stimulating the Immune System in Viral Respiratory Tract Infections: A Narrative Review. *Nutrients.* 2020;12(10).

156.   Garber K. First microbiome-based drug clears phase III, in clinical trial turnaround. *Nat Rev Drug Discov.* 2020;19(10):655-6.

157.   Boukhvalova MS, and Blanco JCG. Springer, Berlin, Heidelberg; 2013:347-58.

158.   Prince GA, Jenson AB, Horswood RL, Camargo E, and Chanock RM. The pathogenesis of respiratory syncytial virus infection in cotton rats. *The American journal of pathology.* 1978;93(3):771-91.

159.   Gitiban N, Jurcisek JA, Harris RH, Mertz SE, Durbin RK, Bakaletz LO, et al. Chinchilla and Murine Models of Upper Respiratory Tract Infections with Respiratory Syncytial Virus. *Journal of Virology.* 2005;79(10):6035-42.

160.   Sacco RE, Durbin RK, and Durbin JE. Animal models of respiratory syncytial virus infection and disease. *Curr Opin Virol.* 2015;13:117-22.

161.   Grieves JL, Yin Z, Durbin RK, and Durbin JE. Acute and Chronic Airway Disease After Human Respiratory Syncytial Virus Infection in Cotton Rats (Sigmodon hispidus). *Comparative medicine.* 2015;65(4):315-26.

162.   Pletneva LM, Haller O, Porter DD, Prince GA, and Blanco JC. Interferon-inducible Mx gene expression in cotton rats: cloning, characterization, and expression during influenza viral infection. *J Interferon Cytokine Res.* 2006;26(12):914-21.

163.   Jin HK, Takada A, Kon Y, Haller O, and Watanabe T. Identification of the murine Mx2 gene: interferon-induced expression of the Mx2 protein from the feral mouse gene confers resistance to vesicular stomatitis virus. *J Virol.* 1999;73(6):4925-30.

164.   Ottolini MG, Porter DD, Hemming VG, and Prince GA. Enhanced pulmonary pathology in cotton rats upon challenge after immunization with inactivated parainfluenza virus 3 vaccines. *Viral Immunol.* 2000;13(2):231-6.

165.   Ottolini MG, Porter DD, Blanco JC, and Prince GA. A cotton rat model of human parainfluenza 3 laryngotracheitis: virus growth, pathology, and therapy. *J Infect Dis.* 2002;186(12):1713-7.

166.   Blanco JC, Pletneva LM, Wan H, Araya Y, Angel M, Oue RO, et al. Receptor characterization and susceptibility of cotton rats to avian and 2009 pandemic influenza virus strains. *J Virol.* 2013;87(4):2036-45.

167.   Ottolini MG, Blanco JCG, Eichelberger MC, Porter DD, Pletneva L, Richardson JY, et al. The cotton rat provides a useful small-animal model for the study of influenza virus pathogenesis. *Journal of General Virology.* 2005;86(10):2823-30.

168.   Pfeuffer J, Puschel K, Meulen V, Schneider-Schaulies J, and Niewiesk S. Extent of measles virus spread and immune suppression differentiates between wild-type and vaccine strains in the cotton rat model (Sigmodon hispidus). *J Virol.* 2003;77(1):150-8.

169. Hamelin ME, Yim K, Kuhn KH, Cragin RP, Boukhvalova M, Blanco JC, et al. Pathogenesis of human metapneumovirus lung infection in BALB/c mice and cotton rats. *J Virol.* 2005;79(14):8894-903.

170. Patel MC, Wang W, Pletneva LM, Rajagopala SV, Tan Y, Hartert TV, et al. Enterovirus D-68 Infection, Prophylaxis, and Vaccination in a Novel Permissive Animal Model, the Cotton Rat (Sigmodon hispidus). *PLoS One.* 2016;11(11):e0166336.

171. Rodriguez WJ, Gruber WC, Welliver RC, Groothuis JR, Simoes EA, Meissner HC, et al. Respiratory syncytial virus (RSV) immune globulin intravenous therapy for RSV lower respiratory tract infection in infants and young children at high risk for severe RSV infections: Respiratory Syncytial Virus Immune Globulin Study Group. *Pediatrics.* 1997;99(3):454-61.

172. Prince GA, Curtis SJ, Yim KC, and Porter DD. Vaccine-enhanced respiratory syncytial virus disease in cotton rats following immunization with Lot 100 or a newly prepared reference vaccine. *Journal of General Virology.* 2001;82(12):2881-8.

173. Prince GA, Jenson AB, Hemming VG, Murphy BR, Walsh EE, Horswood RL, et al. Enhancement of respiratory syncytial virus pulmonary pathology in cotton rats by prior intramuscular inoculation of formalin-inactiva ted virus. *Journal of virology.* 1986;57(3):721-8.

174. Ottolini MG, Curtis SR, Mathews A, Ottolini SR, and Prince GA. Palivizumab is highly effective in suppressing respiratory syncytial virus in an immunosuppressed animal model. *Bone Marrow Transplant.* 2002;29(2):117-20.

175. Martinez ME, Harder OE, Rosas LE, Joseph L, Davis IC, and Niewiesk S. Pulmonary function analysis in cotton rats after respiratory syncytial virus infection. *PLoS One.* 2020;15(8):e0237404.

176. Armstrong C. The Experimental Transmission of Poliomyelitis to the Eastern Cotton Rat, Sigmodon hispidus hispidus. *Public Health Reports (1896-1970).* 1939;54(38):1719-.

177. Piazza FM, Johnson SA, Ottolini MG, Schmidt HJ, Darnell ME, Hemming VG, et al. Immunotherapy of respiratory syncytial virus infection in cotton rats (Sigmodon fulviventer) using IgG in a small-particle aerosol. *J Infect Dis.* 1992;166(6):1422-4.

178. Porter DD, Prince GA, Hemming VG, and Porter HG. Pathogenesis of human parainfluenza virus 3 infection in two species of cotton rats: Sigmodon hispidus develops bronchiolitis, while Sigmodon fulviventer develops interstitial pneumonia. *Journal of virology.* 1991;65(1):103-11.

179. Henson DD, and Bradley RD. Molecular systematics of the genus Sigmodon: results from mitochondrial and nuclear gene sequences. *Can J Zool.* 2009;87(3):211-20.

180. Petersen MK. Interactions Between the Cotton Rats, Sigmodon fulviventer and S. hispidus *The American Midland Naturalist.* 1973 90, No. 2, Oct., 1973(2):319-33.

181. Strickland BA, Patel MC, Shilts MH, Boone HH, Kamali A, Zhang W, et al. Microbial community structure and composition is associated with host species and sex in Sigmodon cotton rats. *Anim Microbiome.* 2021;3(1):29.

182. Zmora N, Zilberman-Schapira G, Suez J, Mor U, Dori-Bachash M, Bashiardes S, et al. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell.* 2018;174(6):1388-405 e21.

183. Han S, Lu Y, Xie J, Fei Y, Zheng G, Wang Z, et al. Probiotic Gastrointestinal Transit and Colonization After Oral Administration: A Long Journey. *Front Cell Infect Microbiol.* 2021;11:609722.

184. Vahabnezhad E, Mochon AB, Wozniak LJ, and Ziring DA. Lactobacillus bacteremia associated with probiotic use in a pediatric patient with ulcerative colitis. *J Clin Gastroenterol.* 2013;47(5):437-9.

185. Wegh CAM, Geerlings SY, Knol J, Roeselers G, and Belzer C. Postbiotics and Their Potential Applications in Early Life Nutrition and Beyond. *Int J Mol Sci.* 2019;20(19).

186. Steed AL, Christophi GP, Kaiko GE, Sun L, Goodwin VM, Jain U, et al. The microbial metabolite desaminotyrosine protects from influenza through type I interferon. *Science.* 2017;357(6350):498-502.

187. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, et al. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity.* 2013;39(2):372-85.

188. Singh N, Gurav A, Sivaprakasam S, Brady E, Padia R, Shi H, et al. Activation of Gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity.* 2014;40(1):128-39.

189. Singh R, Chandrashekharappa S, Bodduluri SR, Baby BV, Hegde B, Kotla NG, et al. Enhancement of the gut barrier integrity by a microbial metabolite through the Nrf2 pathway. *Nat Commun.* 2019;10(1):89.

190. Morita N, Umemoto E, Fujita S, Hayashi A, Kikuta J, Kimura I, et al. GPR31-dependent dendrite protrusion of intestinal CX3CR1(+) cells by bacterial metabolites. *Nature.* 2019;566(7742):110-4.

191. Yan F, Cao H, Cover TL, Whitehead R, Washington MK, and Polk DB. Soluble Proteins Produced by Probiotic Bacteria Regulate Intestinal Epithelial Cell Survival and Growth. *Gastroenterology.* 2007;132(2):562-75.

192. Yan F, and Polk DB. Probiotic bacterium prevents cytokine-induced apoptosis in intestinal epithelial cells. *The Journal of biological chemistry.* 2002;277(52):50959-65.

193. Bäuerl C, Fang P-m, and Polk DB. Functional Analysis of the p40 and p75 Proteins from Lactobacillus casei BL23. 2010:231-41.

194. Wang L, Cao H, Liu L, Wang B, Walker WA, Acra SA, et al. Activation of epidermal growth factor receptor mediates mucin production stimulated by p40, a Lactobacillus rhamnosus GG-derived protein. *Journal of Biological Chemistry.* 2014;289(29):20234-44.

195. Rutigliano JA, and Graham BS. Prolonged Production of TNF-α Exacerbates Illness during Respiratory Syncytial Virus Infection. *The Journal of Immunology.* 2014;173(5):3408-17.

196. Nguyen TH, Maltby S, Simpson JL, Baines KJ, Gibson PG, and Foster PS. TNF-α and Macrophages Are Critical for Respiratory Syncytial Virus − Induced Exacerbations in a Mouse Model of Allergic Airways Disease. 2019.

197. Kotelkin A, Prikhod'ko EA, Cohen JI, Collins PL, and Bukreyev A. Respiratory syncytial virus infection sensitizes cells to apoptosis mediated by tumor necrosis factor-related apoptosis-inducing ligand. *Journal of Virology.* 2003;77(17):9156-72.

198. Atherton JC, and Blaser MJ. Coadaptation of Helicobacter pylori and humans: ancient history, modern implications. *J Clin Invest.* 2009;119(9):2475-87.

199. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 2014;505(7484):559-63.

200. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. *Science.* 2008;320(5883):1647-51.

201. Fierer N, Hamady M, Lauber CL, and Knight R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A.* 2008;105(46):17994-9.

202. Jernberg C, Lofmark S, Edlund C, and Jansson JK. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J.* 2007;1(1):56-66.

203. Raman AS, Gehrig JL, Venkatesh S, Chang HW, Hibberd MC, Subramanian S, et al. A sparse covarying unit that describes healthy and impaired human gut microbiota development. *Science.* 2019;365(6449).

204. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, et al. A catalog of the mouse gut metagenome. *Nat Biotechnol.* 2015;33(10):1103-8.

205. Limborg MT, and Heeb P. Special Issue: Coevolution of Hosts and Their Microbiome. *Genes (Basel).* 2018;9(11).

206. Foster KR, Schluter J, Coyte KZ, and Rakoff-Nahoum S. The evolution of the host microbiome as an ecosystem on a leash. *Nature.* 2017;548(7665):43-51.

207. O'Brien PA, Webster NS, Miller DJ, and Bourne DG. Host-Microbe Coevolution: Applying Evidence from Model Systems to Complex Marine Invertebrate Holobionts. *mBio.* 2019;10(1).

208. Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV, Muller MN, et al. Cospeciation of gut microbiota with hominids. *Science.* 2016;353(6297):380-2.

209. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Chappell JD, Larkin EK, Nelson KE, et al. Nasopharyngeal Microbiome in Respiratory Syncytial Virus Resembles Profile Associated with Increased Childhood Asthma Risk. *Am J Respir Crit Care Med.* 2016;193(10):1180-3.

210. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al. Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc Natl Acad Sci U S A.* 2011;108(13):5354-9.

211. Prince GA, Hemming VG, Horswood RL, Baron PA, and Chanock RM. Effectiveness of topically administered neutralizing antibodies in experimental immunotherapy of respiratory syncytial virus infection in cotton rats. *J Virol.* 1987;61(6):1851-4.

212. Ottolini MG, Blanco JCG, Eichelberger MC, Porter DD, Pletneva L, Richardson JY, et al. The cotton rat provides a useful small-animal model for the study of influenza virus pathogenesis. *J Gen Virol.* 2005;86(Pt 10):2823-30.

213. Blanco JC, Core S, Pletneva LM, March TH, Boukhvalova MS, and Kajon AE. Prophylactic Antibody Treatment and Intramuscular Immunization Reduce Infectious Human Rhinovirus 16 Load in the Lower Respiratory Tract of Challenged Cotton Rats. *Trials Vaccinol.* 2014;3:52-60.

214. Burian M, Rautenberg M, Kohler T, Fritz M, Krismer B, Unger C, et al. Temporal expression of adhesion factors and activity of global regulators during establishment of Staphylococcus aureus nasal colonization. *J Infect Dis.* 2010;201(9):1414-21.

215. Carrara AS, Coffey LL, Aguilar PV, Moncayo AC, Da Rosa AP, Nunes MR, et al. Venezuelan equine encephalitis virus infection of cotton rats. *Emerg Infect Dis.* 2007;13(8):1158-65.

216. Rollin PE, Ksiazek TG, Elliott LH, Ravkov EV, Martin ML, Morzunov S, et al. Isolation of black creek canal virus, a new hantavirus from Sigmodon hispidus in Florida. *J Med Virol.* 1995;46(1):35-9.

217. Holsomback TS, McIntyre NE, Nisbett RA, Strauss RE, Chu YK, Abuzeineh AA, et al. Bayou virus detected in non-oryzomyine rodent hosts: an assessment of habitat composition, reservoir community structure, and marsh rice rat social dynamics. *J Vector Ecol.* 2009;34(1):9-21.

218. Winn WC, Jr., and Murphy FA. Tamiami virus infection in mice and cotton rats. *Bull World Health Organ.* 1975;52(4-6):501-6.

219. Dietrich G, Montenieri JA, Panella NA, Langevin S, Lasater SE, Klenk K, et al. Serologic evidence of west nile virus infection in free-ranging mammals, Slidell, Louisiana, 2002. *Vector Borne Zoonotic Dis.* 2005;5(3):288-92.

220. Ehlen L, Todtmann J, Specht S, Kallies R, Papies J, Muller MA, et al. Epithelial cell lines of the cotton rat (Sigmodon hispidus) are highly susceptible in vitro models to zoonotic Bunya-, Rhabdo-, and Flaviviruses. *Virol J.* 2016;13:74.

221. Prince GA, Curtis SJ, Yim KC, and Porter DD. Vaccine-enhanced respiratory syncytial virus disease in cotton rats following immunization with Lot 100 or a newly prepared reference vaccine. *J Gen Virol.* 2001;82(Pt 12):2881-8.

222. Prince GA, Jenson AB, Hemming VG, Murphy BR, Walsh EE, Horswood RL, et al. Enhancement of respiratory syncytial virus pulmonary pathology in cotton rats by prior intramuscular inoculation of formalin-inactiva ted virus. *J Virol.* 1986;57(3):721-8.

223. Chiba E, Tomosada Y, Vizoso-Pinto MG, Salva S, Takahashi T, Tsukida K, et al. Immunobiotic Lactobacillus rhamnosus improves resistance of infant mice against respiratory syncytial virus infection. *Int Immunopharmacol.* 2013;17(2):373-82.

224. Hyde ER, Petrosino JF, Piedra PA, Camargo CA, Jr., Espinola JA, and Mansbach JM. Nasopharyngeal Proteobacteria are associated with viral etiology and acute wheezing in children with severe bronchiolitis. *J Allergy Clin Immunol.* 2014;133(4):1220-2.

225. Tomosada Y, Chiba E, Zelaya H, Takahashi T, Tsukida K, Kitazawa H, et al. Nasally administered Lactobacillus rhamnosus strains differentially modulate respiratory antiviral immune responses and induce protection against respiratory syncytial virus infection. *BMC Immunol.* 2013;14:40.

226. Casero D, Gill K, Sridharan V, Koturbash I, Nelson G, Hauer-Jensen M, et al. Space-type radiation induces multimodal responses in the mouse gut microbiome and metabolome. *Microbiome.* 2017;5(1):105.

227. Piccolo BD, Graham JL, Stanhope KL, Nookaew I, Mercer KE, Chintapalli SV, et al. Diabetes-associated alterations in the cecal microbiome and metabolome are independent of diet or environment in the UC Davis Type 2 Diabetes Mellitus Rat model. *Am J Physiol Endocrinol Metab.* 2018;315(5):E961-E72.

228. Chaves-Moreno D, Plumeier I, Kahl S, Krismer B, Peschel A, Oxley AP, et al. The microbial community structure of the cotton rat nose. *Environ Microbiol Rep.* 2015;7(6):929-35.

229. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

230. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011;12(6):R60.

231. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods.* 2018;15(11):962-8.

232. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* 2018;46(D1):D633-D9.

233. Boukhvalova MS, Prince GA, and Blanco JC. The cotton rat model of respiratory viral infections. *Biologicals.* 2009;37(3):152-9.

234. Sadowski W, Semkow R, Wilczynski J, Krus S, and Kantoch M. [The cotton rat (Sigmodon hispidus) as an experimental model for studying viruses in human respiratory tract infections. I. Para-influenza virus type 1, 2 and 3, adenovirus type 5 and RS virus]. *Med Dosw Mikrobiol.* 1987;39(1):33-42.

235. Hufeldt MR, Nielsen DS, Vogensen FK, Midtvedt T, and Hansen AK. Variation in the gut microbiota of laboratory mice is related to both genetic and environmental factors. *Comp Med.* 2010;60(5):336-47.

236. Korach-Rechtman H, Freilich S, Gerassy-Vainberg S, Buhnik-Rosenblau K, Danin-Poleg Y, Bar H, et al. Murine Genetic Background Has a Stronger Impact on the Composition of the Gut Microbiota than Maternal Inoculation or Exposure to Unlike Exogenous Microbiota. *Appl Environ Microbiol.* 2019;85(18).

237. Org E, Parks BW, Joo JW, Emert B, Schwartzman W, Kang EY, et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res.* 2015;25(10):1558-69.

238. Tabrett A, and Horton MW. The influence of host genetics on the microbiome. *F1000Res.* 2020;9.

239. Kumpu M, Kekkonen RA, Korpela R, Tynkkynen S, Jarvenpaa S, Kautiainen H, et al. Effect of live and inactivated Lactobacillus rhamnosus GG on experimentally induced rhinovirus colds: randomised, double blind, placebo-controlled pilot trial. *Benef Microbes.* 2015;6(5):631-9.

240. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, et al. Nasopharyngeal Lactobacillus is associated with a reduced risk of

childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy. *J Allergy Clin Immunol.* 2018;142(5):1447-56 e9.

241. Kuss SK, Best GT, Etheredge CA, Pruijssers AJ, Frierson JM, Hooper LV, et al. Intestinal microbiota promote enteric virus replication and systemic pathogenesis. *Science.* 2011;334(6053):249-52.

242. Hiippala K, Jouhten H, Ronkainen A, Hartikainen A, Kainulainen V, Jalanka J, et al. The Potential of Gut Commensals in Reinforcing Intestinal Barrier Function and Alleviating Inflammation. *Nutrients.* 2018;10(8).

243. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, and Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol.* 2013;79(17):5112-20.

244. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol.* 2009;75(23):7537-41.

245. Edgar RC, Haas BJ, Clemente JC, Quince C, and Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics.* 2011;27(16):2194-200.

246. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 2009;37(Database issue):D141-5.

247. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 2007;35(21):7188-96.

248. ari Oksanen FGB, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner. vegan: Community Ecology Package. R package version 2.5-6. . https://CRAN.R-project.org/package=vegan. Accessed 09/24/2020, 2020.

249. Hill MO. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology; Ecological Society of America.* 1973.

250. Nicolai Meinshausen PB. Stability selection. *Journal of the Royal Statistical Society.* 2010;72(Part 4):417–73.

251. Boulesteix AL, and Slawski M. Stability and aggregation of ranked gene lists. *Brief Bioinform.* 2009;10(5):556-68.

252. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 09/24/2020.

253. Ewels P, Magnusson M, Lundin S, and Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016;32(19):3047-8.

254. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.

255. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, and Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods.* 2012;9(8):811-4.

256. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, and UniProt C. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics.* 2015;31(6):926-32.

257. Delroisse JM, Boulvin AL, Parmentier I, Dauphin RD, Vandenbol M, and Portetelle D. Quantification of Bifidobacterium spp. and Lactobacillus spp. in rat fecal samples by real-time PCR. *Microbiol Res.* 2008;163(6):663-70.

258. Nadkarni MA, Martin FE, Jacques NA, and Hunter N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology (Reading).* 2002;148(Pt 1):257-66.

259. Barman M, Unold D, Shifley K, Amir E, Hung K, Bos N, et al. Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract. *Infect Immun.* 2008;76(3):907-15.

260. The I-RSVSG. Palivizumab, a Humanized Respiratory Syncytial Virus Monoclonal Antibody, Reduces Hospitalization From Respiratory Syncytial Virus Infection in High-risk Infants. *Pediatrics.* 1998;102(3):531-7.

261. Xing Y, and Proesmans M. New therapies for acute RSV infections: where are we? *Eur J Pediatr.* 2019;178(2):131-8.

262. Prince GA, Jenson AB, Horswood RL, Camargo E, and Chanock RM. The pathogenesis of respiratory syncytial virus infection in cotton rats. *Am J Pathol.* 1978;93(3):771-91.

263. Graham BS, Perkins MD, Wright PF, and Karzon DT. Primary respiratory syncytial virus infection in mice. *J Med Virol.* 1988;26(2):153-62.

264. Rajagopala SV, Singh H, Patel MC, Wang W, Tan Y, Shilts MH, et al. Cotton rat lung transcriptome reveals host immune response to Respiratory Syncytial Virus infection. *Scientific Reports.* 2018;8(1):11318-.

265. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 2013;8(8):1494-512.

266. Shen W, Le S, Li Y, and Hu F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS One.* 2016;11(10):e0163962.

267. Fu L, Niu B, Zhu Z, Wu S, and Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28(23):3150-2.

268. Gilbert D. Gene-omes built from mRNA seq not genome DNA. *7th annual arthropod genomics symposium, Notre Dame.* 2013.

269. Moreno-Santillan DD, Machain-Williams C, Hernandez-Montes G, and Ortega J. De Novo Transcriptome Assembly and Functional Annotation in Five Species of Bats. *Sci Rep.* 2019;9(1):6222.

270. Thunders M, Cavanagh J, and Li Y. De novo transcriptome assembly, functional annotation and differential gene expression analysis of juvenile and adult E. fetida, a model oligochaete used in ecotoxicological studies. *Biol Res.* 2017;50(1):7.

271. Rajagopala SV, Singh H, Patel MC, Wang W, Tan Y, Shilts MH, et al. Cotton rat lung transcriptome reveals host immune response to Respiratory Syncytial Virus infection. *Scientific Reports.* 2018;8(1):1-12.

272. Langmead B, and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-9.

273. Manni M, Berkeley MR, Seppey M, Simao FA, and Zdobnov EM. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol Biol Evol.* 2021;38(10):4647-54.

274. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology.* 2014;15(12):550-.

275. UniProt C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480-D9.

276. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421.

277. Kanehisa M, and Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27-30.

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

279. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 2019;47(D1):D309-D14.

280. Langley RJ, Prince GA, and Ginsberg HS. HIV type-1 infection of the cotton rat (Sigmodon fulviventer and S. hispidus). *ProcNatlAcadSciUSA.* 1998;95(0027-8424):14355-60.

281. Young MD, Wakefield MJ, Smyth GK, and Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010;11(2):R14.

282. Blanco Jorge CG, Richardson Joann Y, Darnell Miriam ER, Rowzee A, Pletneva L, Porter David D, et al. Cytokine and Chemokine Gene Expression after Primary and Secondary Respiratory Syncytial Virus Infection in Cotton Rats. *The Journal of Infectious Diseases.* 2002;185(12):1780-5.

283. Cardoso-Silva CB, Costa EA, Mancini MC, Balsalobre TW, Canesin LE, Pinto LR, et al. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. *PLoS One.* 2014;9(2):e88462.

284. Jumat MR, Huong TN, Ravi LI, Stanford R, Tan BH, and Sugrue RJ. Viperin protein expression inhibits the late stage of respiratory syncytial virus morphogenesis. *Antiviral Res.* 2015;114:11-20.

285. Zaas AK, Chen M, Varkey J, Veldman T, Hero AO, 3rd, Lucas J, et al. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe.* 2009;6(3):207-17.

286. Gao J, Zhu X, Wu M, Jiang L, Wang F, and He S. IFI27 may predict and evaluate the severity of respiratory syncytial virus infection in preterm infants. *Hereditas.* 2021;158(1):3.

287. Lindell DM, Lane TE, and Lukacs NW. CXCL10/CXCR3-mediated responses promote immunity to respiratory syncytial virus infection by augmenting dendritic cell and CD8(+) T cell efficacy. *Eur J Immunol.* 2008;38(8):2168-79.

288. Vissers M, Schreurs I, Jans J, Heldens J, de Groot R, de Jonge MI, et al. Antibodies enhance CXCL10 production during RSV infection of infant and adult immune cells. *Cytokine.* 2015;76(2):458-64.

289. Li L, Ni YA, Song Z, Yi Z, and Wang F. Identification of pathogenic genes and transcription factors in respiratory syncytial virus. *BMC Pediatr.* 2021;21(1):27.

290. Behera AK, Kumar M, Lockey RF, and Mohapatra SS. 2'-5' Oligoadenylate synthetase plays a critical role in interferon-gamma inhibition of respiratory syncytial virus infection of human epithelial cells. *J Biol Chem.* 2002;277(28):25601-8.

291. Gonzalez-Sanz R, Mata M, Bermejo-Martin J, Alvarez A, Cortijo J, Melero JA, et al. ISG15 Is Upregulated in Respiratory Syncytial Virus Infection and Reduces Virus Growth through Protein ISGylation. *J Virol.* 2016;90(7):3428-38.

292. Fink K, Martin L, Mukawera E, Chartier S, De Deken X, Brochiero E, et al. IFNbeta/TNFalpha synergism induces a non-canonical STAT2/IRF9-dependent pathway triggering a novel DUOX2 NADPH oxidase-mediated airway antiviral response. *Cell Res.* 2013;23(5):673-90.

293. Guo X, Liu T, Shi H, Wang J, Ji P, Wang H, et al. Respiratory Syncytial Virus Infection Upregulates NLRC5 and Major Histocompatibility Complex Class I Expression through RIG-I Induction in Airway Epithelial Cells. *J Virol.* 2015;89(15):7636-45.

294. Shirato K, Ujike M, Kawase M, and Matsuyama S. Increased replication of respiratory syncytial virus in the presence of cytokeratin 8 and 18. *J Med Virol.* 2012;84(2):365-70.

295. Edwards KM, Snyder PN, and Wright PF. Complement activation by respiratory syncytial virus-infected cells. *Arch Virol.* 1986;88(1-2):49-56.

296. Hashimoto K, Graham BS, Geraci MW, FitzGerald GA, Egan K, Zhou W, et al. Signaling through the prostaglandin I2 receptor IP protects against respiratory syncytial virus-induced illness. *J Virol.* 2004;78(19):10303-9.

297. Zhuang X, Rambhatla SB, Lai AG, and McKeating JA. Interplay between circadian clock and viral infection. *J Mol Med (Berl).* 2017;95(12):1283-9.

298. Edgar RS, Stangherlin A, Nagy AD, Nicoll MP, Efstathiou S, O'Neill JS, et al. Cell autonomous regulation of herpes and influenza virus infection by the circadian clock. *Proc Natl Acad Sci U S A.* 2016;113(36):10085-90.

299. Majumdar T, Dhar J, Patel S, Kondratov R, and Barik S. Circadian transcription factor BMAL1 regulates innate immunity against select RNA viruses. *Innate Immun.* 2017;23(2):147-54.

300. Boukhvalova MS, Prince GA, and Blanco JCG. The cotton rat model of respiratory viral infections. *Biologicals : journal of the International Association of Biological Standardization.* 2009;37(3):152-9.

301. Pletneva LM, Haller O, Porter DD, Prince GA, and Blanco JCG. Induction of type I interferons and interferon-inducible Mx genes during respiratory syncytial virus infection and reinfection in cotton rats. *J Gen Virol.* 2008;89(Pt 1):261-70.

302. Uthaiah RC, Praefcke GJ, Howard JC, and Herrmann C. IIGP1, an interferon-gamma-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. *J Biol Chem.* 2003;278(31):29336-43.

303. Taylor GA, Collazo CM, Yap GS, Nguyen K, Gregorio TA, Taylor LS, et al. Pathogen-specific loss of host resistance in mice lacking the IFN-gamma-inducible gene IGTP. *Proc Natl Acad Sci U S A.* 2000;97(2):751-5.

304. De Groote MA, Sterling DG, Hraha T, Russell TM, Green LS, Wall K, et al. Discovery and Validation of a Six-Marker Serum Protein Signature for the Diagnosis of Active Pulmonary Tuberculosis. *J Clin Microbiol.* 2017;55(10):3057-71.

305. Blanco JC, Boukhvalova MS, Pletneva LM, Shirey KA, and Vogel SN. A recombinant anchorless respiratory syncytial virus (RSV) fusion (F) protein/monophosphoryl lipid A (MPL) vaccine protects against RSV-induced replication and lung pathology. *Vaccine.* 2014;32(13):1495-500.

306. Boukhvalova MS, Prince GA, and Blanco JC. Respiratory syncytial virus infects and abortively replicates in the lungs in spite of preexisting immunity. *J Virol.* 2007;81(17):9443-50.

307. Livak KJ, and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-8.

308. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.

309. Andrews S. 2010.

310. Li B, and Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12(1):323-.

311. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* 2021;49(D1):D412-D9.

312. Krogh A, Larsson B, von Heijne G, and Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001;305(3):567-80.

313. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* 2019;37(4):420-3.

314. Patro R, Duggal G, Love MI, Irizarry RA, and Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods.* 2017;14(4):417-9.

315. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3--new capabilities and interfaces. *Nucleic Acids Res.* 2012;40(15):e115.

316. Tran DM, Tran TT, Phung TTB, Bui HT, Nguyen PTT, Vu TT, et al. Nasal-spraying Bacillus spores as an effective symptomatic treatment for children with acute respiratory syncytial virus infection. *Sci Rep.* 2022;12(1):12402.

317. Herrera P, Schuster L, Wentrup C, Konig L, Kempinger T, Na H, et al. Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont. *Proc Natl Acad Sci U S A.* 2020;117(35):21658-66.

318. Green MG, Huey D, and Niewiesk S. The cotton rat (Sigmodon hispidus) as an animal model for respiratory tract infections with human pathogens. *Lab Anim (NY).* 2013;42(5):170-6.

319. Groves HT, Cuthbertson L, James P, Moffatt MF, Cox MJ, and Tregoning JS. Respiratory Disease following Viral Lung Infection Alters the Murine Gut Microbiota. *Front Immunol.* 2018;9:182.

320. Wang J, Lu H, Yu L, Cheng W, Yan W, and Jing X. Aggravation of airway inflammation in RSV-infected asthmatic mice following infection-induced alteration of gut microbiota. *Ann Palliat Med.* 2021;10(5):5084-97.

321. Groves HT, Higham SL, Moffatt MF, Cox MJ, and Tregoning JS. Respiratory Viral Infection Alters the Gut Microbiota by Inducing Inappetence. *mBio.* 2020;11(1).

322. Mack DR, Ahrne S, Hyde L, Wei S, and Hollingsworth MA. Extracellular MUC3 mucin secretion follows adherence of Lactobacillus strains to intestinal epithelial cells in vitro. *Gut.* 2003;52(6):827-33.

323. Shyu RY, Wang CH, Wu CC, Chen ML, Lee MC, Wang LK, et al. Tazarotene-Induced Gene 1 Enhanced Cervical Cell Autophagy through Transmembrane Protein 192. *Mol Cells.* 2016;39(12):877-87.

324. Suzuki K, Vogelzang A, and Fagarasan S. MZB1 folding and unfolding the role of IgA. *Proc Natl Acad Sci U S A.* 2019;116(27):13163-5.

325. Wang Y, Liu L, Moore DJ, Shen X, Peek RM, Acra SA, et al. An LGG-derived protein promotes IgA production through upregulation of APRIL expression in intestinal epithelial cells. *Mucosal Immunol.* 2017;10(2):373-84.

326. Hijano DR, Siefker DT, Shrestha B, Jaligama S, Vu LD, Tillman H, et al. Type I Interferon Potentiates IgA Immunity to Respiratory Syncytial Virus Infection During Infancy. *Sci Rep.* 2018;8(1):11034.

327. Xiong E, Li Y, Min Q, Cui C, Liu J, Hong R, et al. MZB1 promotes the secretion of J-chain-containing dimeric IgA and is critical for the suppression of gut inflammation. *Proc Natl Acad Sci U S A.* 2019;116(27):13480-9.

328. Sakai F, Hosoya T, Ono-Ohmachi A, Ukibe K, Ogawa A, Moriya T, et al. Lactobacillus gasseri SBT2055 induces TGF-β expression in dendritic cells and activates TLR2 signal to produce IgA in the small intestine. *PLoS ONE.* 2014;9(8).

329. Huang IF, Lin IC, Liu PF, Cheng MF, Liu YC, Hsieh YD, et al. Lactobacillus acidophilus attenuates Salmonella-induced intestinal inflammation via TGF-beta signaling. *BMC Microbiol.* 2015;15:203.

330. Ting WJ, Kuo WW, Hsieh DJ, Yeh YL, Day CH, Chen YH, et al. Heat Killed Lactobacillus reuteri GMNL-263 Reduces Fibrosis Effects on the Liver and Heart in High Fat Diet-Hamsters via TGF-beta Suppression. *Int J Mol Sci.* 2015;16(10):25881-96.

331. Bottcher MF, Abrahamsson TR, Fredriksson M, Jakobsson T, and Bjorksten B. Low breast milk TGF-beta2 is induced by Lactobacillus reuteri supplementation and associates with reduced risk of sensitization during infancy. *Pediatr Allergy Immunol.* 2008;19(6):497-504.

332. Kim KK, Sheppard D, and Chapman HA. TGF-beta1 Signaling and Tissue Fibrosis. *Cold Spring Harb Perspect Biol.* 2018;10(4).

333. Wang L, Cheng W, and Zhang Z. Respiratory syncytial virus infection accelerates lung fibrosis through the unfolded protein response in a bleomycin-induced pulmonary fibrosis animal model. *Mol Med Rep.* 2017;16(1):310-6.

334. Maassen CB, Boersma WJ, van Holten-Neelen C, Claassen E, and Laman JD. Growth phase of orally administered Lactobacillus strains differentially affects IgG1/IgG2a ratio for soluble antigens: implications for vaccine development. *Vaccine.* 2003;21(21-22):2751-7.

335. Uribe G, Villeger R, Bressollier P, Dillard RN, Worthley DL, Wang TC, et al. Lactobacillus rhamnosus GG increases cyclooxygenase-2 expression and prostaglandin E2 secretion in colonic myofibroblasts via a MyD88-dependent mechanism during homeostasis. *Cell Microbiol.* 2018;20(11):e12871.

336. Martin R, Chamignon C, Mhedbi-Hajri N, Chain F, Derrien M, Escribano-Vazquez U, et al. The potential probiotic Lactobacillus rhamnosus CNCM I-3690 strain protects the intestinal barrier by stimulating both mucus production and cytoprotective response. *Sci Rep.* 2019;9(1):5398.

337. Albarracin L, Kobayashi H, Iida H, Sato N, Nochi T, Aso H, et al. Transcriptomic Analysis of the Innate Antiviral Immune Response in Porcine Intestinal Epithelial Cells: Influence of Immunobiotic Lactobacilli. *Front Immunol.* 2017;8:57.

338. Weiss G, Rasmussen S, Zeuthen LH, Nielsen BN, Jarmer H, Jespersen L, et al. Lactobacillus acidophilus induces virus immune defence genes in murine dendritic cells by a Toll-like receptor-2-dependent mechanism. *Immunology.* 2010;131(2):268-81.

339. Laval L, Martin R, Natividad JN, Chain F, Miquel S, Desclee de Maredsous C, et al. Lactobacillus rhamnosus CNCM I-3690 and the commensal bacterium Faecalibacterium prausnitzii A2-165 exhibit similar protective effects to induced barrier hyper-permeability in mice. *Gut Microbes.* 2015;6(1):1-9.

340. Groom JR, Richmond J, Murooka TT, Sorensen EW, Sung JH, Bankert K, et al. CXCR3 chemokine receptor-ligand interactions in the lymph node optimize CD4+ T helper 1 cell differentiation. *Immunity.* 2012;37(6):1091-103.

341. Thakar J, Qian Y, Benoodt L, Roumanes D, Qiu X, Laniewski N, et al. Unbiased analysis of peripheral blood mononuclear cells reveals CD4 T cell response to RSV matrix protein. *Vaccine X.* 2020;5:100065.

342. Turi KN, Shankar J, Anderson LJ, Rajan D, Gaston K, Gebretsadik T, et al. Infant Viral Respiratory Infection Nasal Immune-Response Patterns and Their Association with Subsequent Childhood Recurrent Wheeze. *American Journal of Respiratory and Critical Care Medicine.* 2018;198(8):1064-73.

343. Elshaer AM, El-Kharashi OA, Hamam GG, Nabih ES, Magdy YM, and Abd El Samad AA. Involvement of TLR4/ CXCL9/ PREX-2 pathway in the development of hepatocellular carcinoma (HCC) and the promising role of early administration of lactobacillus plantarum in Wistar rats. *Tissue Cell.* 2019;60:38-47.

344. Busse DC, Habgood-Coote D, Clare S, Brandt C, Bassano I, Kaforou M, et al. Interferon-Induced Protein 44 and Interferon-Induced Protein 44-Like Restrict Replication of Respiratory Syncytial Virus. *J Virol.* 2020;94(18).

345. Li Z, Qu X, Liu X, Huan C, Wang H, Zhao Z, et al. GBP5 Is an Interferon-Induced Inhibitor of Respiratory Syncytial Virus. *J Virol.* 2020;94(21).

346. Zhang R, Li Z, Tang YD, Su C, and Zheng C. When human guanylate-binding proteins meet viral infections. *J Biomed Sci.* 2021;28(1):17.

347. Allard JB, and Duan C. IGF-Binding Proteins: Why Do They Exist and Why Are There So Many? *Front Endocrinol (Lausanne).* 2018;9:117.

348. Das S, Sarkar A, Ryan KA, Fox S, Berger AH, Juncadella IJ, et al. Brain angiogenesis inhibitor 1 is expressed by gastric phagocytes during infection with Helicobacter pylori and mediates the recognition and engulfment of human apoptotic gastric epithelial cells. *FASEB J.* 2014;28(5):2214-24.

349. Yan F, Cao H, Cover TL, Washington MK, Shi Y, Liu L, et al. Colon-specific delivery of a probiotic-derived soluble protein ameliorates intestinal inflammation in mice through an EGFR-dependent mechanism. *Journal of Clinical Investigation.* 2011;121(6):2242-53.

350. Bauerl C, Abitayeva G, Sosa-Carrillo S, Mencher-Beltran A, Navarro-Lleo N, Coll-Marques JM, et al. P40 and P75 Are Singular Functional Muramidases Present in the Lactobacillus casei /paracasei/rhamnosus Taxon. *Front Microbiol.* 2019;10:1420.

351. Yan F, Liu L, Dempsey PJ, Tsai YH, Raines EW, Wilson CL, et al. A lactobacillus rhamnosus GG-derived soluble protein, p40, stimulates ligand release from intestinal epithelial cells to transactivate epidermal growth factor receptor. *Journal of Biological Chemistry.* 2013;288(42):30742-51.

352. Yang L, Higginbotham JN, Liu L, Zhao G, Acra SA, Peek RM, et al. Production of a Functional Factor, p40, by Lactobacillus rhamnosus GG Is Promoted by Intestinal Epithelial Cell-Secreted Extracellular Vesicles. *Infection and immunity.* 2019;87(7):e00113-19.

353. Deng Y, McDonald OG, Means AL, Peek RM, Jr., Washington MK, Acra SA, et al. Exposure to p40 in Early Life Prevents Intestinal Inflammation in Adulthood Through Inducing a Long-Lasting Epigenetic Imprint on TGFbeta. *Cell Mol Gastroenterol Hepatol.* 2021;11(5):1327-45.

354. Eichele DD, and Kharbanda KK. Dextran sodium sulfate colitis murine model: An indispensable tool for advancing our understanding of inflammatory bowel diseases pathogenesis. *World J Gastroenterol.* 2017;23(33):6016-29.

355. Saleh M, and Trinchieri G. Innate immune mechanisms of colitis and colitis-associated colorectal cancer. *Nat Rev Immunol.* 2011;11(1):9-20.

356. Nguyen TH, Maltby S, Simpson JL, Eyers F, Baines KJ, Gibson PG, et al. TNF-alpha and Macrophages Are Critical for Respiratory Syncytial Virus-Induced Exacerbations in a Mouse Model of Allergic Airways Disease. *J Immunol.* 2016;196(9):3547-58.

357. Rutigliano JA, and Graham BS. Prolonged production of TNF-alpha exacerbates illness during respiratory syncytial virus infection. *J Immunol.* 2004;173(5):3408-17.

358. Kalinowski A, Galen BT, Ueki IF, Sun Y, Mulenos A, Osafo-Addo A, et al. Respiratory syncytial virus activates epidermal growth factor receptor to suppress interferon regulatory factor 1-dependent interferon-lambda and antiviral defense in airway epithelium. *Mucosal Immunology.* 2018;11(3):958-67.

359. Xu N, Lao Y, Zhang Y, and Gillespie DA. Akt: a double-edged sword in cell proliferation and genome stability. *J Oncol.* 2012;2012:951724.

360. Fitch KR, McGowan KA, van Raamsdonk CD, Fuchs H, Lee D, Puech A, et al. Genetics of dark skin in mice. *Genes Dev.* 2003;17(2):214-28.

361. Scheving LA, Zhang X, Garcia OA, Wang RF, Stevenson MC, Threadgill DW, et al. Epidermal growth factor receptor plays a role in the regulation of liver and plasma lipid levels in adult male mice. *Am J Physiol Gastrointest Liver Physiol.* 2014;306(5):G370-81.

362. Tomas A, Futter CE, and Eden ER. EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol.* 2014;24(1):26-34.

363. Futter CE, Pearse A, Hewlett LJ, and Hopkins CR. Multivesicular endosomes containing internalized EGF-EGF receptor complexes mature and then fuse directly with lysosomes. *J Cell Biol.* 1996;132(6):1011-23.

364. Fox CF, Wrann M, Linsley P, and Vale R. Hormone-induced modification of EGF receptor proteolysis in the induction of EGF action. *J Supramol Struct.* 1979;12(4):517-31.

365. Tanaka T, Zhou Y, Ozawa T, Okizono R, Banba A, Yamamura T, et al. Ligand-activated epidermal growth factor receptor (EGFR) signaling governs endocytic trafficking of unliganded receptor monomers by non-canonical phosphorylation. *J Biol Chem.* 2018;293(7):2288-301.

366. Thomas KW, Monick MM, Staber JM, Yarovinsky T, Carter AB, and Hunninghake GW. Respiratory syncytial virus inhibits apoptosis and induces NF-kappa B activity through a phosphatidylinositol 3-kinase-dependent pathway. *J Biol Chem.* 2002;277(1):492-501.

367. Monick MM, Cameron K, Staber J, Powers LS, Yarovinsky TO, Koland JG, et al. Activation of the Epidermal Growth Factor Receptor by Respiratory Syncytial Virus Results in Increased Inflammation and Delayed Apoptosis. *Journal of Biological Chemistry.* 2005;280(3):2147-58.

368. Leslie CC, McCormick-Shannon K, Shannon JM, Garrick B, Damm D, Abraham JA, et al. Heparin-binding EGF-like growth factor is a mitogen for rat alveolar type II cells. *Am J Respir Cell Mol Biol.* 1997;16(4):379-87.

369. Su Y, Luo H, and Yang J. Heparin-binding EGF-like growth factor attenuates lung inflammation and injury in a murine model of pulmonary emphysema. *Growth Factors.* 2018;36(5-6):246-62.

370. Lutmer J, Watkins D, Chen C-L, Velten M, and Besner G. Heparin-binding epidermal growth factor–like growth factor attenuates acute lung injury and multiorgan dysfunction after scald burn. *Journal of Surgical Research.* 2013;185(1):329-37.

371. James IAO, Chen C-L, Huang G, Zhang H-Y, Velten M, and Besner GE. HB-EGF Protects the Lungs after Intestinal Ischemia/Reperfusion Injury. *Journal of Surgical Research.* 2010;163(1):86-95.

372. Alliet P, Vandenplas Y, Roggero P, Jespers SNJ, Peeters S, Stalens JP, et al. Safety and efficacy of a probiotic-containing infant formula supplemented with 2'-fucosyllactose: a double-blind randomized controlled trial. *Nutr J.* 2022;21(1):11.

373. Budden KF, Gellatly SL, Wood DL, Cooper MA, Morrison M, Hugenholtz P, et al. Emerging pathogenic links between microbiota and the gut-lung axis. *Nat Rev Microbiol.* 2017;15(1):55-63.

374. Lee JW, Fang X, Dolganov G, Fremont RD, Bastarache JA, Ware LB, et al. Acute lung injury edema fluid decreases net fluid transport across human alveolar epithelial type II cells. *J Biol Chem.* 2007;282(33):24109-19.

375. Shaver CM, Grove BS, Putz ND, Clune JK, Lawson WE, Carnahan RH, et al. Regulation of alveolar procoagulant activity and permeability in direct acute lung injury by lung epithelial tissue factor. *Am J Respir Cell Mol Biol.* 2015;53(5):719-27.

376. Shaver CM, Wickersham N, McNeil JB, Nagata H, Sills G, Kuck JL, et al. Cell-Free Hemoglobin-mediated Increases in Vascular Permeability. A Novel Mechanism of Primary Graft Dysfunction and a New Therapeutic Target. *Ann Am Thorac Soc.* 2017;14(Supplement_3):S251-S2.

377. Shaver CM, Wickersham N, McNeil JB, Nagata H, Miller A, Landstreet SR, et al. Cell-free hemoglobin promotes primary graft dysfunction through oxidative lung endothelial injury. *JCI Insight.* 2018;3(2).

378. Moore ML, Chi MH, Luongo C, Lukacs NW, Polosukhin VV, Huckabee MM, et al. A chimeric A2 strain of respiratory syncytial virus (RSV) with the fusion protein of RSV strain line 19 exhibits enhanced viral load, mucus, and airway dysfunction. *J Virol.* 2009;83(9):4185-94.

379. Schobel SA, Stucker KM, Moore ML, Anderson LJ, Larkin EK, Shankar J, et al. Respiratory Syncytial Virus whole-genome sequencing identifies convergent evolution of sequence duplication in the C-terminus of the G gene. *Sci Rep.* 2016;6:26311.

380. Blanco JC, Pletneva L, Boukhvalova M, Richardson JY, Harris KA, and Prince GA. The cotton rat: an underutilized animal model for human infectious diseases can now be exploited using specific reagents to cytokines, chemokines, and interferons. *J Interferon Cytokine Res.* 2004;24(1):21-8.

381. Blanco JC, Pletneva LM, Wieczorek L, Khetawat D, Stantchev TS, Broder CC, et al. Expression of Human CD4 and chemokine receptors in cotton rat cells confers permissiveness for productive HIV infection. *Virol J.* 2009;6:57.

382. Blanco JC, Richardson JY, Darnell ME, Rowzee A, Pletneva L, Porter DD, et al. Cytokine and chemokine gene expression after primary and secondary respiratory syncytial virus infection in cotton rats. *J Infect Dis.* 2002;185(12):1780-5.

383. Boukhvalova MS, Prince GA, and Blanco JC. Inactivation of respiratory syncytial virus by zinc finger reactive compounds. *Virol J.* 2010;7:20.

384. Angel S, Maero E, Blanco JC, Pzsenny V, Zala C, Gonzalez R, et al. Early diagnosis of toxoplasmic encephalitis in AIDS patients by dot blot hybridization analysis. *J Clin Microbiol.* 1992;30(12):3286-7.

385. Boukhvalova MS, Yim KC, and Blanco J. Cotton rat model for testing vaccines and antivirals against respiratory syncytial virus. *Antivir Chem Chemother.* 2018;26:2040206618770518.

386. Lukacs NW, Moore ML, Rudd BD, Berlin AA, Collins RD, Olson SJ, et al. Differential immune responses and pulmonary pathophysiology are induced by two different strains of respiratory syncytial virus. *Am J Pathol.* 2006;169(3):977-86.

387. Stokes KL, Chi MH, Sakamoto K, Newcomb DC, Currier MG, Huckabee MM, et al. Differential pathogenesis of respiratory syncytial virus clinical isolates in BALB/c mice. *J Virol.* 2011;85(12):5782-93.

388. Feldman AS, He Y, Moore ML, Hershenson MB, and Hartert TV. Toward primary prevention of asthma. Reviewing the evidence for early-life respiratory viral infections as modifiable risk factors to prevent childhood asthma. *Am J Respir Crit Care Med.* 2015;191(1):34-44.

389. Cohn L, Homer RJ, Marinov A, Rankin J, and Bottomly K. Induction of airway mucus production By T helper 2 (Th2) cells: a critical role for interleukin 4 in cell recruitment but not mucus production. *J Exp Med.* 1997;186(10):1737-47.

390. Tripp RA, Moore D, and Anderson LJ. TH(1)- and TH(2)-TYPE cytokine expression by activated t lymphocytes from the lung and spleen during the inflammatory response to respiratory syncytial virus. *Cytokine.* 2000;12(6):801-7.

391. Barlow JL, Bellosi A, Hardman CS, Drynan LF, Wong SH, Cruickshank JP, et al. Innate IL-13-producing nuocytes arise during allergic lung inflammation and contribute to airways hyperreactivity. *J Allergy Clin Immunol.* 2012;129(1):191-8 e1-4.

392. Barlow JL, Peel S, Fox J, Panova V, Hardman CS, Camelo A, et al. IL-33 is more potent than IL-25 in provoking IL-13-producing nuocytes (type 2 innate lymphoid cells) and airway contraction. *J Allergy Clin Immunol.* 2013;132(4):933-41.

393. Bartemes KR, Iijima K, Kobayashi T, Kephart GM, McKenzie AN, and Kita H. IL-33-responsive lineage- CD25+ CD44(hi) lymphoid cells mediate innate type 2 immunity and allergic inflammation in the lungs. *J Immunol.* 2012;188(3):1503-13.

394. Chang YJ, Kim HY, Albacker LA, Lee HH, Baumgarth N, Akira S, et al. Influenza infection in suckling mice expands an NKT cell subset that protects against airway hyperreactivity. *J Clin Invest.* 2011;121(1):57-69.

395. Doherty TA, Khorram N, Chang JE, Kim HK, Rosenthal P, Croft M, et al. STAT6 regulates natural helper cell proliferation during lung inflammation initiated by Alternaria. *Am J Physiol Lung Cell Mol Physiol.* 2012;303(7):L577-88.

396. Halim TY, Krauss RH, Sun AC, and Takei F. Lung natural helper cells are a critical source of Th2 cell-type cytokines in protease allergen-induced airway inflammation. *Immunity.* 2012;36(3):451-63.

397. Hong JY, Bentley JK, Chung Y, Lei J, Steenrod JM, Chen Q, et al. Neonatal rhinovirus induces mucous metaplasia and airways hyperresponsiveness through IL-25 and type 2 innate lymphoid cells. *J Allergy Clin Immunol.* 2014;134(2):429-39.

398. Jackson DJ, Makrinioti H, Rana BM, Shamji BW, Trujillo-Torralbo MB, Footitt J, et al. IL-33-dependent type 2 inflammation during rhinovirus-induced asthma exacerbations in vivo. *Am J Respir Crit Care Med.* 2014;190(12):1373-82.

399. Beale J, Jayaraman A, Jackson DJ, Macintyre JDR, Edwards MR, Walton RP, et al. Rhinovirus-induced IL-25 in asthma exacerbation drives type 2 immunity and allergic pulmonary inflammation. *Sci Transl Med.* 2014;6(256):256ra134.

400. Monticelli LA, Sonnenberg GF, Abt MC, Alenghat T, Ziegler CG, Doering TA, et al. Innate lymphoid cells promote lung-tissue homeostasis after infection with influenza virus. *Nat Immunol.* 2011;12(11):1045-54.

401. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, et al. MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia. *Sci Transl Med.* 2013;5(203):203ra125.
402. Stier MT, Bloodworth MH, Toki S, Newcomb DC, Goleniewska K, Boyd KL, et al. Respiratory syncytial virus infection activates IL-13-producing group 2 innate lymphoid cells through thymic stromal lymphopoietin. *J Allergy Clin Immunol.* 2016;138(3):814-24 e11.
403. Saravia J, You D, Shrestha B, Jaligama S, Siefker D, Lee GI, et al. Respiratory Syncytial Virus Disease Is Mediated by Age-Variable IL-33. *PLoS Pathog.* 2015;11(10):e1005217.
404. Schwengers O, Hoek A, Fritzenwanker M, Falgenhauer L, Hain T, Chakraborty T, et al. ASA3P: An automatic and scalable pipeline for the assembly, annotation and higher-level analysis of closely related bacterial isolates. *PLoS Comput Biol.* 2020;16(3):e1007134.
405. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, and Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25(7):1043-55.
406. Seppey M, Manni M, and Zdobnov EM. BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol Biol.* 2019;1962:227-45.

"There is a curious idea among unscientific men that in scientific writing there is a common plateau of perfectionism. Nothing could be more untrue. The reports of biologists are the measure, not of the science, but of the men themselves. . . It is usually found that only the little stuffy men object to what is called 'popularization', by which they mean writing with a clarity understandable to one not familiar with the tricks and codes of the cult. We have not known a single great scientist who could not discourse freely and interestingly with a child. Can it be that the haters of clarity have nothing to say, have observed nothing, have no clear picture of even their own fields? A dull man seems to be a dull man no matter what his field, and of course it is the right of a dull scientist to protect himself with feathers and robes, emblems and degrees, as do other dull men who are potentates and grand imperial rulers of lodges of dull men."

John Steinbeck*, The Log from the Sea of Cortez,* 1951