Spring 5-15-2023

The influence of the gut microbiota on asthma in school-aged children and adults

Naomi Grace Wilson

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Computational & Systems Biology

Dissertation Examination Committee:
Andrew L. Kau, Chair
S. Joshua Swamidass, Co-Chair
Megan Baldridge
Gautam Dantas
Makedonka Mitreva

The Influence of the Gut Microbiota on Asthma in School-Aged Children and Adults

by
Naomi Grace Wilson

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2023
St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgments

"Pause you who read this, and think for a moment of the long chain of iron or gold, of thorns or flowers, that would never have bound you, but for the formation of the first link on one memorable day." – Charles Dickens

I would like to first thank the educators that have been links in the long chain of events that brought me to this moment. My high school teachers Andrew Johnson and Heather Boyum were instrumental in sparking my fascination with science which had until then been starved for the knowledge they skillfully communicated. My many mentors at SUNY Geneseo were pivotal models of rigorous curiosity wrapped in humility and care for their students. Thank you, Professors Harold Hoops, Wendy Pogozelski, Jeffrey Peterson, Gregg Hartvisgen, Barnabas Gikonyo, Eric Helms, Dave & Cristina Geiger, Kevin Millitello, George Reuter, and Jim Tiller. Thank you to my co-mentor, Dr Josh Swamidass for rapidly taking me up on my offer (read: plea) to work in his lab as an undergraduate so that I could have summer research experience. Thank you for the connections you helped me to make in both the St Louis science and faith communities, and especially in the community found at the overlap between the two. Thank you to my thesis chair, Dr Gautam Dantas, for being one of the first CSB-affiliated faces I met and who gave timely advice to a anxiety-ridden PhD student throughout her graduate school journey. Thank you Drs Megan Baldridge and Makedonka Mitreva for the insightful questions and encouragements offered freely that helped me grow as a scientist during my time as a student. Lastly, I would like to thank Andy Kau for his warmth, energy, and good humor that sustained our tight-knit team through the challenges that come with doing high quality research. His unending willingness to try new things and be available for his people are contagious qualities I hope to emulate.

Each of my Kau Lab contemporaries have a special place in my grateful heart. Ryan, Natalia, Anne, Ariel, Jesus, Allie, Michael, and even our new youngest sibling, Christopher: y'all are stupid good to me and we have spent way more time together than any of us would like to admit. Thanks for tolerating both 3pm Nerma and hangry Nerma.

My time in graduate school has been made unimaginably sweet by the classmates that have suffered, studied, encouraged, laughed, and read books with me. Victoria Ismail, Huiming Xia, Jhullian Alston, Jasmine Cubuk, Kellan Weston, Matthew Cruz, Mickey Ward, Paige Hall, and Rose dos Reis Silva, thank you for being a constant from day one until now and, I am certain, for a long time into the future. We maintained our sanity by clinging to each other and now we celebrate the people we have become. Also, I am the real imposter. To the Good Sumerians: Ariel Hernandez-Leyva, Hunter Patterson, Ryan Friedman, Kai Loell, Gowri Kalugotla, Victoria Ismail, Kellan Weston, Shashank Anand, Rebecca Brady, Michael Nehls, and Melvin King, thank you for making mine and Ariel's dream of reading anything else besides another academic paper true- for 57 books and counting! Our weekly book discussions of science fiction gems and flops (and the enigmas that are both at once) have been a supreme delight beyond imagination. I'm still shocked that all of you brilliant people kept showing up.

There are several people that took special care of me in ways I did not know I even needed in the past 6 years. Amber, Joben, & Seth Lewis, you have modeled true gentleness towards others and oneself that I would not have believed was achievable but that you're living proof. Jeni Fehr, Karly Nelson, Gracie & Kevin York, BethAnna Jones, Emily Johnson, Patrick Mattia, Harrison & Amy Hartsough, Kevin & Rachel Hainline, and Olivia & Scott Johnson: thank you for seeing me, literally chasing waterfalls with me, and being so nerdy that even I felt out of my depth at times. Love you, goofs. Thank you to many others in The Journey Hanley

Road community, including Sarah Gentry, Gregg & Lindsey Fox, Kristin Guilliams, Jason Magers, Mark Goins, and the worship & production team for discussions that enriched my mind and heart, music-making that filled my soul, and Comet coffee dates that rescued my tenuous sanity. To my Karis House counselor, thank you for being excellent. This is your success, too.

To Lydia, thanks for being the bestest spider sister. I'm forever glad we have each other. Caleb, thanks for letting me crash your home and drink your bourbon so often since you married my best friend. You're pretty okay. Kevin & Karen, your snuggles are everything. Millie, my COVID dog, you saved me. To Mom and Dad: thanks for the unquestioning, unending support including, but not limited to hugs, homemade food, mutual dog sitting, your Spectrum login, hot tub access, and your toaster. Thank you for the determined efforts to understand the opaque process that is doctoral training, and for the lifelong sacrifices made to prioritize my education and wellbeing from a very young age.

Soli Deo Gloria.

Naomi Wilson

Washington University in St. Louis

May 2023

Dedicated to
Thelma McQuiston
and
Eleanor Joy Wilson

ABSTRACT OF THE DISSERTATION

The influence of the gut microbiota on asthma in school-aged children and adults

for Arts & Sciences Graduate Students

by

Naomi Grace Wilson

Doctor of Philosophy in Biology and Biomedical Sciences

Computational & Systems Biology

Washington University in St. Louis, 2023

Professor Andrew L. Kau, Chair

Professor S. Joshua Swamidass, Co-Chair

Asthma is a common respiratory disease with a highly heterogeneous pathophysiology. The human gut microbiota, comprising of all the microorganisms that inhabit the gastrointestinal tract, is linked to the development of asthma and can alter airway inflammation in animal models. The idea that the gut microbiota can have bidirectional cross-talk with the lung, such as gut dysbiosis affecting lung disease, is termed the gut-lung axis. While the gut microbiota of early life has been an area of particular interest for asthma pathogenesis research, the effect of its taxonomic and functional composition after asthma diagnosis is less clear. This dissertation employs both amplicon sequencing followed by gnotobiotic mouse models (Chapter 2) and whole metagenomic shotgun sequencing (Chapter 3) to characterize human fecal microbiomes from school-aged children and adults with asthma.

Investigation of the gut-lung axis in asthma later in life requires clinical studies with well-defined asthma inclusion criteria and strategic gnotobiotic experiments guided by

taxonomic profiling of properly handled human samples. In Chapter 2, amplicon sequencing was used to identify population shifts between allergic moderate-to-severe asthma and healthy cohorts. Taxonomic shifts associated with asthma were observed even when accounting for other microbiome-modifying covariates such as age and race. Additionally, statistical modeling and gnotobiotic mouse models were used to identify taxa that could affect lung inflammation *in vivo*. A Naïve Bayes Classifier fit to a mixture model that accounts for the sparsity inherent to compositional data was built to optimize selection of samples from the human cohorts that would best represent a asthma-associated microbial community differences. The selected stool samples were then used to inoculate, or "humanize", germ-free mouse gastrointestinal tracts before administration of allergen sensitization and challenge. Immunophenotyping, IgA-Seq, gut permeability assays, and whole genome sequencing of human fecal bacterial isolates uncovered an enterotoxigenic *Bacteroides fragilis* that affected lung inflammation in the context of intact human fecal communities as well as on its own in a monocolonization experiment. A PCR screen for the *B. fragilis* toxin (*bft*) across all human participants revealed that *bft* was more prevalent in the stool of people with asthma compared to that of healthy individuals. These findings suggest that the gut microbiota affects lung inflammation even after the diagnosis of asthma.

While discovery of disease-modifying taxa is invaluable, taxonomic profiling by amplicon sequencing skips the genetic material that encodes a wealth of functional information about gut microbes. In Chapter 3, whole metagenomic shotgun sequencing is utilized on the human fecal samples from Chapter 2 to describe the genetic content of the entire gut microbiota, also called the "metagenome". Read-based annotation revealed a shift in genetic content attributable to asthma even when accounting for covariates such as age and race. Metabolic pathway annotation suggested that fatty acid metabolism pathways, particularly those that result

in long-chain fatty acid synthesis, are differentially abundant in the asthma cohort. Antibiotic resistance is a growing concern among physicians who provide asthma care and patients with asthma tend to require more antibiotic prescriptions than usual, particularly macrolide antibiotics. Antibiotic usage was tallied for all participants in this study and a higher proportion of the asthma cohort was found to have taken antibiotics in the past year compared to the healthy cohort. Subsequent profiling of antibiotic resistance genes (ARGs) in the gut metagenomes revealed an increased richness of ARGs in the asthma cohort while the total abundance of ARGs was not increased. Additionally, macrolide resistance markers were differentially abundant in the asthma cohort. Interestingly, the *B. fragilis* toxin, found to be more prevalent in the same asthma cohort in Chapter 2, was more likely to co-occur in the samples with *ermF* in the asthma cohort compared to the healthy cohort. Co-occurrence analysis of all ARGs and all virulence factors revealed a unique set of VF-ARG pairs in the asthma cohort compared to the healthy, together suggesting that the asthma gut microbiota offers opportunities for virulence factors and ARGs to co-occur that do not co-occur in healthy gut microbiota. The *ermF-bft* pair is particularly concerning given that *bft* has the potential to affect airway inflammation and macrolide resistance is already becoming a clinical problem for patients with asthma.

In summary, this work characterizes metagenomic shifts in the gut microbiota associated with asthma, identifies a gut pathobiont that can alter lung inflammation, and reveals accumulation of antibiotic resistance genes in populations suffering from asthma. These findings provide needed insights into the gut-lung axis of asthma beyond diagnosis, and will guide development of gut-directed therapy for a frustratingly common disease.

Chapter 1: Introduction to the Gut-Lung Axis in Asthma

1.1 The ABCs of Wheeze: Asthma and bacterial communities

Naomi G. Wilson[1,2,*], Ariel Hernandez-Leyva[1,2,*] and Andrew L. Kau[1,2,§]

[1] Department of Medicine, Division of Allergy and Immunology, Washington University School of Medicine, St. Louis, MO, 63110

[2] Center for Women's Infectious Disease Research, Washington University School of Medicine, St. Louis, MO, 63110

[§] These authors contributed equally

[§] Correspondence: akau@wustl.edu

1.1.1 Is asthma a problem of microbial ecology?

Asthma is a common respiratory illness affecting approximately 8% of Americans and is characterized by symptoms of wheeze, cough, and shortness of breath. These symptoms are caused by an inappropriate sensitization to an environmental antigen, or allergen, that leads to airway inflammation upon reexposure. These responses are typically mediated by allergic T cells (T-helper 2 [Th2]) cells, which trigger eosinophilic inflammation characteristic of allergic asthma. Multiple factors contribute to the clinical risk for asthma. A genetic predisposition to allergy and asthma, called atopy, is a well-established feature in its development. Additionally, environmental factors including exposure to allergens, birth delivery mode, diet, and childhood surroundings may also modify asthma risk.

More recently, the role played by the body's endogenous microbial communities (microbiota) has emerged as a new component to this puzzle. Healthy environmental microbial exposures not only help the neonatal immune system, including humoral and innate components, to mature into an adult-like state after birth but also likely influence immune function throughout life [1]. Maladaptive microbiota–host interactions have likewise been linked to the development of disease. A number of studies across a wide variety of illnesses have changed the way that we regard "pathologic" microbe–host interactions to include the concept of "dysbiosis," or dysfunctional microbial communities that contribute to pathology. These dysbiotic communities fail to adequately perform the normal functions of a healthy microbiota in educating the immune system, processing dietary components, producing bioactive compounds, and more, which can profoundly impact host health and disease susceptibility. Dysbiotic gut and airway microbial communities have been implicated in asthma pathogenesis, indicating that a comprehensive

understanding of the disease will require a detailed understanding of resident bacterial–host interactions.

1.1.2 How could microbial community dynamics promote allergy?

Our microbiota can shape immune responses in numerous ways, including interacting directly with host immune components, altering metabolic functions, or preventing or promoting pathogen invasion. Understanding microbial ecology in the context of the host environment is therefore an increasingly important objective as we seek to understand how microbes may facilitate allergic inflammation. Even in a stable, nondiseased state, the mucosal environment in which most commensal microbes reside presents a highly dynamic and complex habitat with distinct niches and available resources. Further complicating this interaction, microbes in the mammalian host have the capacity to alter their environment through their interaction with the host immune system. These microbe-stimulated changes to the environment may be an opportunity for bacteria to alter their habitat in their favor, either producing a unique niche or creating a barrier to competing microbes. *Bacteroides fragilis*, for example, has recently been reported to take advantage of the host intestinal immunoglobulin A (IgA) response to generate a unique niche that enables resistance to displacement by other microbes [2]. Although beneficial to a specific microbe, such immunomodulation may have consequences for the host.

Like asthma, atopic dermatitis (AD), also called eczema, is a chronic inflammatory skin disease triggered by Th2 allergic inflammation and is characterized by itchy, dry, and red skin rashes. Interestingly, AD is associated with a defined skin microbial signature, dominated by *Staphylococcus aureus*, that likely plays a role in AD manifestations. *S. aureus* colonization in AD has been associated with disease severity [3,4] and may directly promote AD by activating

mast cells through the production of δ-toxin [5] and contributing to Th2 activation [6] (Figure 1.1F). Inflamed skin, in turn, down-regulates the expression of antimicrobial peptides [7], which permits further proliferation and persistence of *S. aureus*, potentially amplifying the course of disease. Although it is not yet clear what the specific role of such interactions is in allergic responses, these examples underscore how the factors that govern the microbial ecology of our microbiota may contribute to asthma and emphasize the potential importance of microbial communities residing in close association with allergic inflammation in the airway.

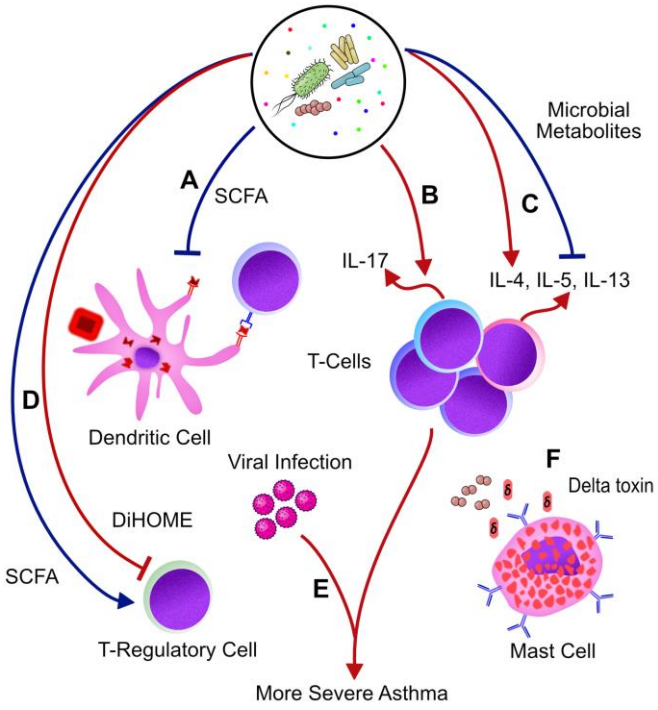

Figure 1.1. Bacterial communities and metabolites impact multiple stages in the well-established pathway of allergic inflammation, both offering potentially protective effects (blue arrows) and exacerbating allergic inflammation (red arrows).

(A) SCFAs produced by gut microbiota enhance bone marrow production of dendritic cells and macrophages with increased phagocytic capacity but reduced ability to stimulate Th2 responses in the lung [8]. (B) Bacteria found in the airway have been associated with increased markers of Th17 inflammation [9–11]. (C) Microbially derived metabolites may directly stimulate or inhibit [12] Th2 development [13]. (D) Microbially derived SCFAs found in the healthy gut promote Treg differentiation [14], whereas metabolites like DiHOME produced by asthmatic gut microbiota inhibit Treg proliferation [13]. (E) Viral respiratory tract infections are associated with increased risk for asthma and exacerbations [15]. (F) *Staphylococcus aureus* colonizes the skin of AD patients depleted of antimicrobial peptides by

allergic inflammation and carries virulence factors that influence the host immune system to sustain the inflamed state. *S. aureus* delta toxin, for example, acts on pathways not associated with immunoglobulin E to stimulate mast-cell degranulation near the site of colonization [3]. Investigating the precise mechanisms by which bacteria can worsen or mitigate asthma and allergic inflammation will offer novel probiotic candidates and biomarkers with therapeutic potential. AD, atopic dermatitis; DiHOME, dihydroxy-9Z-octadecenoic acid; IL, interleukin; SCFA, short-chain fatty acid; Th2, T-helper 2; Th17, T-helper 17; Treg, T regulatory cell. This figure was created by an integral contributing author of this work, Ariel Hernandez-Leyva.

1.1.3 Is microbial community composition in the upper airway a risk factor for asthma?

Early-childhood viral respiratory tract infections have long been recognized to increase the risk of asthma and contribute to asthma exacerbations [15] (Figure 1.1E). Bacterial infections have likewise been implicated in promoting asthma exacerbations, but the roles of airway-colonizing bacteria in shaping allergic airway responses are only recently being explored. Similar to the gastrointestinal tract [16], the anatomy of the airway plays an important role in shaping microbial communities along the respiratory tract. The nasal vestibule is perpetually exposed to the environment and encounters a constant barrage of debris, microbes, and microbial products, which presumably contribute to the relatively high density of microbes in the upper airway. Beyond the nasal vault, air, remaining particles, and microbes aspirated from the oropharynx can pass into the trachea, which is often defined as the beginning of the lower respiratory tract. Particles and microbes that are deposited in the lower airway are entrapped in mucus and returned back up to the esophagus by mucociliary clearance. These natural clearance and relocation mechanisms contribute to shared members but marked differences in abundance between the upper- and lower-airway bacterial communities [17].

In the upper respiratory tract, specific bacterial taxa have been associated with the development of asthma. Multiple studies have reported that neonates colonized with *Streptococcus pneumoniae, Haemophilus influenzae,* or *Moraxella catarrhalis* are at an increased risk for later diagnosis of asthma compared with uncolonized children [18,19]. The

presence of these asthma-associated microbes in infancy does not reflect mere chance exposure, because their appearance is associated with acute viral respiratory infections as well as multiple other environmental factors including antibiotic usage and daycare attendance [20]. These findings have been extended to early childhood through 16S rRNA sequencing of nasopharyngeal aspirates from a cohort of children carefully monitored for respiratory infections during the first 5 years of life. Confirming earlier reports, these studies found that the same three taxa were predictors of future diagnosis of asthma [21]. Additionally, the high temporal resolution of sampling enabled the observation that illness-associated taxa bloom in abundance preceding viral infection. Perhaps most intriguingly, this study found a time period during early childhood in which colonization with *Streptococcus*, *Moraxella*, and *Haemophilus* in the upper airway predicted later wheeze (a precursor to asthma) in children [21]. A similar "critical window" of microbial exposure important for later asthma development has also been described in the gut microbiota and could be scientifically and clinically valuable because it marks a time frame for potential intervention.

1.1.4 Is there a lung microbiota, and does it play a role in asthma?

The lung itself was regarded as a sterile environment in healthy individuals, maintained by intrinsic clearance mechanisms, until culture-independent techniques revealed the presence of a lung microbiota, though at a markedly decreased abundance compared with the upper airway [22]. Although studying microbial communities in the lungs is technically challenging because of low bacterial biomass and the difficulty in acquiring samples [23], differences in the lung microbiota are associated with asthma [24] and other pulmonary diseases [25]. Airway microbes may also exert an important effect on asthma even after it is established by modulating response to

therapy [26] or altering the character of inflammatory response [9,10]. Moreover, specific alterations to the lung microbiota have been associated with distinct asthma phenotypes. For instance, Proteobacteria from lower-airway samples have been associated with neutrophilic asthma, an endotype of asthma associated with T-helper 17 (Th17) responses [9,10] (Figure 1.1B). This ability of airway microbes to alter clinical features of asthma parallels our evolving understanding of asthma as a highly heterogeneous condition, with many unique pathophysiologies (referred to as endotypes) that all ultimately lead to asthma's characteristic features: airway hyperresponsiveness and obstruction. Whereas the endotype associated with "classical" allergic asthma is mediated by Th2 inflammation, multiple endotypes have been described [27].

Although there are multiple clinical studies validating the association between upper- and lower-airway microbes and asthma, further efforts will be needed to unravel how these bacteria modify allergic responses. The ability of upper-airway microbes to shape allergic responses has been investigated in animal models. Inoculation of neonatal mice with nontypeable *Haemophilus* results in transient infection and later susceptibility to worsened allergic airway inflammation [28]. Exposure to *M. catarrhalis* has likewise been found to intensify allergic inflammation in mice through the induction of interleukin 17 (IL-17), resulting in more recruitment of neutrophils and eosinophils to the lungs [11]. Using these animal models, we have additionally gleaned that early colonization of the airway by bacteria helps shield mice from asthma through programmed death ligand 1 (PD-L1)-dependent induction of T regulatory (Treg) cells [29]. Microbial colonization in the first 2 weeks of life induces PD-L1 expression on dendritic cells, which signal Tregs to persist in the immune system. The specific host–bacterial interactions that promote allergic airway inflammation remain enigmatic, however, and require additional investigation.

1.1.5 How do gut microbes shape allergic inflammation in the lung?

The vast community of microbes inhabiting the gut has been documented to influence a surprisingly diverse set of diseases, even illnesses that typically involve pathologies at anatomically distant sites, like multiple sclerosis [30,31] and arthritis [32]. The influence of gut microbes on pulmonary diseases has been referred to as the "gut–lung axis," and several studies have implicated intestinal microbial communities in the pathogenesis of asthma, particularly in early childhood. Although multiple pathways probably underlie the communication between gut microbial communities and the lung in asthma [33,34], microbiota-derived metabolites are emerging as a particularly compelling example.

For instance, the composition of neonatal gut microbial communities can acquire configurations with decreased representation of *Akkermansia*, *Bifidobacterium*, and *Faecalibacterium*, which predict later development of allergy and asthma [13]. Coculture of sterile fecal extracts generated from these dysfunctional microbial communities with human lymphocytes promotes the development of allergy-inducing Th2 cells while inhibiting the differentiation of Treg cells, directly implicating microbially derived products in conferring susceptibility to allergic disease (Figure 1.1C,D).

Furthermore, investigators studying the Canadian Healthy Infant Longitudinal Development cohort identified four taxa in neonates—*Faecalibacterium*, *Lachnospira*, *Veillonella*, and *Rothia* (FLVR)—that corresponded to later protection from asthma. When these four bacteria were supplemented to an asthmatic gut microbial community in gnotobiotic mice, animals receiving FLVR taxa experienced protection from allergic airway inflammation compared with controls that received no supplementation. Colonization with FLVR also

corresponded to increases in the amounts of fecal acetate [35], a short-chain fatty acid (SCFA) known to protect from asthma [8]. Differences in SCFAs have been observed in 3-month-olds who displayed atopy and wheeze at 1 year [35] – symptoms that are predictive of later asthma. Moreover, SCFAs are known to influence Treg differentiation and function (e.g., [14]), which may contribute to asthma. Additionally, SCFAs generated by the gut microbiota have also been demonstrated to alter dendritic-cell recruitment and function within the lungs and reduce Th2 inflammation [8,12] (Figure 1.1A). Intriguingly, FLVR were protective only when present during the first 100 days of life, which suggests a critical window of gut colonization during which exposure to key organisms can prevent later disease. This phenomenon is similar to the critical window observed for airway colonization [21], suggesting that exposure to particular microbial communities during a crucial moment in early life results in a lasting restructuring of the host's immune system.

1.1.6 Concluding remarks

Our growing appreciation for the role of commensal microbes in human health has placed the microbiota among the important factors contributing to asthma pathogenesis. Although the mechanisms by which endogenous bacterial communities affect asthma are still being elucidated, the enticing potential for therapeutic interventions will ensure continued interest and exploration on the influence of microbial communities on allergic diseases. Perhaps most compellingly, unlike host genetics and environmental exposures, which are known risk factors for allergic diseases, the microbiota presents a more easily modifiable feature in asthma development. One potential approach is to design probiotics that alter the composition or function of respiratory or gastrointestinal microbial communities in at-risk children. The promise of this approach relies on

identifying individuals who would benefit from a probiotic intervention and exploiting the "critical window" in early life, during which microbial exposures can shape later asthma [13,21,35]. However, before we can benefit from microbiota-directed therapeutics in asthma, it is necessary to further define the mechanism(s) by which gut and airway microbes protect from the disease and to assist the identification of therapeutic candidates and biomarkers. Together, these new insights signify the beginning of a novel paradigm to understand the etiology of asthma as an emergent condition resulting from the codevelopment of microbial communities and host immunity. Innovations from this research may lead to transformative advances in our understanding and treatment of asthma.

1.1.7 Acknowledgements

This work was co-authored by **Naomi G Wilson*, Ariel Hernandez-Leyva*, and Andrew L. Kau** (*authors contributed equally)[36]. This work was supported by the National Institutes of Health (K08 AI113184 to A.L.K., T32 GM007200 to A.H-L) and the AAAAI (Foundation Faculty Development Award to A.L.K.).

1.2 The gut-lung axis after asthma diagnosis

In addition to the gut microbiota being an area of keen interest for understanding the etiology of allergy and asthma in order to prevent its onset, it is a largely unexplored source of potential therapeutic targets for those already suffering from their asthma diagnosis. Asthma is not only a frustratingly common disease but can be complex to manage due to its interpersonal heterogeneity. Under the label of "asthma" you find patients with a spectrum of disease triggers, molecular pathophysiology, and clinical features[37]. "Endotypes" of asthma refer to these subsets of patients, often categorized by the main immune component that infiltrates the airways. For

example, many children with early-onset allergic asthma have a T-helper 2 cell driven disease that tends to be receptive to corticosteroid treatment whereas another subset of patients tend to be adults that exhibit an eosinophil-dominated phenotype that is resistant to corticosteroids[27]. No definitive cause for susceptibility to one type of asthma over another has yet been determined. However, since the gut microbiota is already known to co-develop with the immune system[38] and is highly variable between individuals, it has been hypothesized that perhaps some of the heterogeneity of asthma can be explained by the interpersonal variability in the gut microbiota.

A handful of studies have focused on patients after asthma diagnosis and found taxonomic and functional shifts in the gut microbiota that were associated with disease[39]. Notable gut taxa that are reduced in abundance in school-aged and/or adult asthma include *Akkermansia muciniphila*[40], *Faecalibacterium prausnitzii*[41] as well as *Roseburia* species[42]. Even fewer studies have utilized whole metagenomic sequencing to identify microbial functions associated with asthma. One such study sequenced stool from a population of women with asthma[41] and found an increase in lipid and amino acid metabolism pathways and a decrease in the production of short-chain fatty acids (SCFAs) compared to healthy controls[41]. Similarly, microbial SCFA biosynthesis and tryptophan metabolism were associated with improved asthma symptoms in a clinical study of probiotic supplementation to adults[43]. Interestingly, histamine production in the gut and has been linked to asthma as well. One study found an increase in histidine decarboxylase, an enzyme that converts histidine to histamine, as well as the bacteria that often encodes it, *Morganella morganii,* in the gut microbiota of adults with asthma[44]. In a follow-up study, the same group found that histamine produced by bacteria in the gut can affect allergic airway inflammation in a mouse model[45]. Each of these studies reveal that the gut

microbiota has the potential to shape lung inflammation in children and adults with asthma. However, more studies using metagenomics sequencing and mouse models are necessary to characterize the gut microbiota and mechanisms of the gut-lung axis in asthma

One hypothesis raised by this dissertation is that transfer across a dysfunctional gut barrier may be one route connecting microbial products to lung immune components. The gut barrier consists of a monolayer of epithelial cells coated in mucus that is essential for keeping pathogens and bacterial toxins from leaving the gut lumen, while allowing ions and nutrients to access the circulation. Further, the immune cells of the lamina propria are exposed to only the microbes and products that the gut barrier allows, which means that this one layer of cells is an immensely important part of immune tolerance. A more permissible, i.e. permeable, gut barrier has been implicated in intestinal diseases, food allergy, and metabolic diseases[46–50]. The barrier becomes more permeable when the gut is damaged by inflammation and when the tight junction and adherens proteins that form the seal between epithelial cells are under-expressed[51]. The gut microbiota is an integral part of preventing inappropriate permeability and maintaining homeostasis of the gut barrier[52–54]. For example, gut dysbiosis is associated with altered permeability, especially in disease states[55,56]. Since gut barrier function is important for communication with the immune system and maintaining homeostasis, it is plausible that gut barrier dysfunction could be involved in the pathogenesis of asthma. In fact, studies have demonstrated increased gut permeability in both adults[57] and children[58] with asthma compared to healthy cohorts. No study has yet tested if the gut microbiota alone is sufficient to cause gut permeability in asthma.

The negative association of SCFAs with asthma[8,12,41,59] may itself be indicative of the role of the gut barrier to prevent asthma. One of the ways the gut microbiota affects barrier homeostasis is via SCFAs which guide maturation of the epithelium by increasing tight junction protein expression[60,61], strengthening the epithelial seal. Seminal work by Arrieta and colleagues found that 3-month old infants who were at higher risk of asthma symptoms at 1 year had less fecal acetate content than their healthy counterparts[35]. The lack of SCFA-producing bacteria seen in studies of high-risk infants[13,35] supports this result. These data suggest that SCFAs may play a role in asthma susceptibility early in life. No study has measured the fecal content of SCFAs in adults or older children with asthma, so it is still unknown whether SCFA levels are differentially abundant later in life. Functional annotation of the gut metagenome and metabolomics of human microbiota samples are necessary to reveal the capacity of asthmatic microbiota to produce SCFAs and change the trajectory of the disease.

Gnotobiotic models of allergic airway inflammation are another method of uncovering mechanisms of the gut microbiota that could be life-changing therapeutic targets in humans. Several seminal gnotobiotic experiments using neonatal stool microbiota have been previously discussed and have found metabolites such as SCFAs and DiHOME-12,13 as potential mediators of asthma development[13,35,62], but there are very few studies focused on what effect the adult or pediatric gut microbiota has on asthma after diagnosis, if any. One such study of school-aged children found no difference in markers of allergy between germ-free mice colonized with fecal microbiota from five subjects with asthma compared to that from seven healthy subjects[63]. However, another study demonstrated that supplementing *Akkermansia muciniphila,* a species found decreased in a population with obesity-associatetd asthma, to non-germ-free mice protects

against allergic airway inflammation[40]. Considering these studies do not rule out the effect of the gut microbiota on asthma in older populations, mouse studies to address this question are still needed.

Additionally, in order to further pinpoint actionable differences between the gut microbiota of individuals in older populations with and without asthma, state-of-the-art characterization of the human gut metagenome is required. Functional annotation of the gut microbiota in asthma research has previously been done using marker gene identification of taxa followed by inferred metabolic profiling from those taxonomic classifications with tools such as PICRUSt[35,64]. These approaches suffer from false positives and false negatives due to arbitrary marker gene similarity cutoffs and incomplete reference databases. Shotgun metagenomic sequencing improves on this method by directly measuring total gene content and is widely used to profile metabolic functions. Due to the rapid decrease in cost, shotgun metagenomics sequencing has become more feasible on large scales, and has begun to be used to study asthma in metagenome-wide association studies (MWAS)[41].

A feature of the gut microbiota that only functional annotation can measure is the library of antibiotic resistance markers that are carried by gut microbes. The gut microbiota is now a recognized reservoir for antibiotic resistance genes, the sum of which is called the antibiotic "resistome" that is shaped by antibiotic usage[65,66]. Characterization of the gut resistome could disentangle why pre- and post-natal antibiotic usage has been associated with an increased risk of asthma later in life[67,68]. One recent study found an increase in richness of antibiotic resistance genes carried by infants whose gut microbiota showed signatures associated with asthma risk[69]. No study has yet, to my knowledge, characterized the gut resistome of older pediatric and adults

with asthma to determine whether this pattern persists. Antibiotic resistance is already a growing concern among physicians that offer asthma care and is thought to be rising due to the high amount of antibiotic drug prescriptions that patients with asthma require compared to a given healthy patient. This increased need is likely linked to the observation that many patients with asthma are prone to airway infections that are often treated with antibiotics[70]. In some studies, long-term low dose macrolide antibiotic usage prevented asthma exacerbations[71–73], but not in others[74], and establishing such a treatment as standard practice has remained controversial, not least because of the risk of antibiotic resistance[75]. Interestingly, antibiotic usage can also affect the gut barrier, largely thought to be due to the antibiotic-induced gut dysbiosis. Antibiotics have been associated with increased gut permeability, decreasing tight junction gene expression, and affecting zonula-occludin-1 morphology[76]. Profiling the gut resistome of patients with asthma would further inform physicians concerned with antibiotic resistance and may contextualize the functions affecting the gut-lung axis.

Chapter 2: The gut microbiota from people with asthma influences lung inflammation in gnotobiotic mice

Naomi G. Wilson[1,#], Ariel Hernandez-Leyva[1,#], Anne L. Rosen[1], Natalia Jaeger[1], Ryan T. McDonough[1], Jesus Santiago-Borges[1], Michael A. Lint[1], Thomas R. Rosen[1], Christopher P. Tomera[1], Leonard B. Bacharier[2], S. Joshua Swamidass[3], Andrew L. Kau[1,*]

[1] Division of Allergy and Immunology, Department of Medicine and Center for Women's Infectious Disease Research, Washington University School of Medicine, St. Louis, MO, 63110, USA

[2] Division of Allergy, Immunology and Pulmonary Medicine, Department of Pediatrics, Monroe Carell Jr Children's Hospital at Vanderbilt University Medical Center, Nashville, TN 37232, USA

[3] Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

[#] These authors contributed equally.

* Lead Contact; correspondence: akau@wustl.edu

2.1 Abstract

The composition of the gut microbiota in early childhood is linked to asthma risk, but the gut microbiota may continue to play a role in older patients with established asthma. In this study, we profile the gut microbiota of 38 school-aged children (19 with asthma, median age 8) and 57 adults (17 with asthma, median age 28) by 16S rRNA sequencing and found evidence that individuals with asthma harbored compositional differences from healthy controls in both adults and children. We develop a model that aids in the design of mechanistic experiments in gnotobiotic mice and show that enterotoxigenic *Bacteroides fragilis* (ETBF) is more prevalent in the gut microbiota of patients with asthma over 6 years of age compared to healthy controls. In mice colonized with ETBF experiencing allergic airway inflammation (AAI), ETBF, modulated by community context, can increase oxidative stress in the lungs. Our results provide evidence that ETBF affects the phenotype of airway inflammation in a subset of patients with asthma outside of early childhood which suggests that therapies targeting the gut microbiota may be helpful tools for asthma control.

2.2 Introduction

Asthma is a common respiratory disease characterized by airway inflammation triggered by an allergic response to environmental antigens. While asthma is predominately associated with T helper (Th)-2 inflammation associated with high levels of cytokines IL-4, IL-5, and IL-13, detailed characterization from clinical studies has revealed substantial heterogeneity in the immunopathology of this disease, referred to as an asthma endotype [77]. The majority of children tend to have a Th2-dominant endotype and often experience virus-induced exacerbations. In contrast, a significant fraction of adults have a non-Th2 driven endotype and experience a wider spectrum of immunopathologies[27]. For example, Th17-associated inflammation is more common

in adults and is associated with corticosteroid resistance and increased frequency of exacerbations[78]. Furthermore, greater intracellular oxidative stress increases inflammation and hyperreactivity in the airways of patients with asthma[79] and has been associated with IL-17A mediated inflammation in mouse models[80]. Although disease heterogeneity plays an important role in the prognosis and treatment of asthma, the specific factors that drive the endotype of asthma are not well understood.

A potential source of disease variability in the lung lies in the diverse immunologic and metabolic activities of the gut microbiota. Gut microbes have been implicated in the pathology of a range of lung diseases including chronic obstructive pulmonary disease [81], fungal [82] and bacterial pneumonia [83,84], and, notably, asthma [35]. One mechanism by which this phenomenon, termed the "gut-lung" axis [85], affects asthma is through the synthesis of bioactive metabolites. For example, microbe-derived molecules like short-chain fatty acids [12,35] and 12,13-diHOME [13,62] are thought to permanently alter the immune system during infancy. Other metabolites produced by the gut microbiota have also been shown to influence oxidative stress in distal tissues including the brain and kidneys later in life [86,87].

Metagenomic surveys have revealed a "critical window" [88] where microbiome and immune co-development [1] during the first year of life has an exaggerated effect on the development of asthma later in life. [21,35,89,90]. Additionally, gnotobiotic animal studies have shown that asthma-associated microbes from early childhood can modulate susceptibility to experimental models of AAI [35]. However, it remains unclear whether the mechanisms influencing the development of asthma in infants continue to affect the disease in older children and adults, or whether functions of the gut microbiota modulate the characteristics of asthma

after it has been established. Among older children and adults that are beyond the critical window, individuals with asthma harbor differences in microbial diversity and composition [40,44,91,92] compared to healthy controls, emphasizing the need to further investigate the influence the microbiota might have on asthma throughout life.

Few studies have been devoted to identifying and defining the effects of the gut microbiota on lung inflammation in older individuals with asthma. One of these studies has demonstrated that the gut commensal, *Akkermansia muciniphila,* which is reduced in individuals with obesity-associated asthma, can mitigate AAI in mice [40]. In contrast, a study of AAI severity in gnotobiotic mice colonized with gut microbiota from school-aged children did not find a difference in markers of allergy between mice colonized with microbiota from healthy children and children with asthma [63]. Given these variable results, there remains a need to assess the effect of the gut microbiota on established asthma and to find testable mechanisms for human clinical studies.

In this study, our goal was to analyze the fecal microbiota from a clinical study of healthy and asthmatic subjects from which we could select representative human samples to thoroughly characterize in a gnotobiotic mouse model of AAI. Follow up experiments would then test the prevalence of the discovered immune phenotypes in additional fecal microbiota donors. We hypothesized that there would be effector gut microbes, more prevalent in people with asthma, that could promote airway inflammation in mice. Our results show that while typical markers of allergy were not affected by the gut microbiota, asthma-associated microbiomes were more likely to harbor enterotoxigenic *Bacteroides fragilis* (ETBF) which was associated with gut barrier dysfunction, as well as oxidative stress in the lungs of gnotobiotic mice.

2.3 Results

2.3.1 The composition of the gut microbiota differs between individuals with and without asthma

To investigate if asthma is associated with distinct gut microbial signatures outside of early childhood, we recruited 17 adults and 19 school-aged children with physician-diagnosed, moderate-to-severe atopic asthma along with 40 adult and 19 school-aged healthy controls into the previously described Microbiome and Asthma Research Study [93] (MARS; see demographic summary in Table S1A, B and Figure 5.1A; see also Methods). The adult cohort was 18-40 years old with a median age of 28 years (s.d. 6.2 years) and the pediatric cohort was 6-10 years old with a median age of 8 years (s.d. 1.6 years). We performed V4-16S rRNA amplicon sequencing of participant stool samples obtained at the patient's baseline and identified amplicon sequence variants (ASVs) using DADA2 [94] (Figure 5.1B). There was a slight increase in alpha diversity in subjects as measured by Simpson's metric within individuals with asthma compared to healthy controls ($p = 0.035$), even after accounting for differences between age group ($p = 0.030$) and read depth ($p = 0.063$, Figure 2.1A). The significance of this trend is unclear however, as the same was not true for other metrics of alpha diversity which were more susceptible to read depth (Figure 5.1C). To determine how demographic features affected community composition, we performed a PERMANOVA on Bray-Curtis dissimilarity distances (Figure 2.1B and Figure 5.1D) between fecal microbiomes. First, we analyzed demographic factors reported to influence microbiome composition as independent terms using a sequential PERMANOVA and found that asthma status ($p=0.00009$; $R^2=0.025$), age ($p=0.00001$; $R^2=0.032$), and race ($p=0.001$; $R^2=0.02$) significantly contributed to the variation in subject gut microbiota composition, but adiposity ($p=0.3$; $R2=0.03$), sex ($p=0.4$; $R2=0.009$), smoking history ($p=0.4$; $R2=0.009$), and antibiotic

exposure within the past year (p=0.7; R2=0.008) did not (Table S2). We then ran another sequential PERMANOVA that included interaction terms between all the factors that had a p-value of less than 0.05 in the first PERMANOVA and found that asthma status (p=0.02; R^2=0.015), age (p=0.0017; R^2=0.019), and race (p=0.0041; R^2=0.018) remain contributors to the variation in beta diversity, even when accounting for sequencing batch effect and interaction terms (Figure 2.1C). Although age is an important factor in determining asthma phenotype, we did not find the interaction of age and asthma to be significant. Guided by the PERMANOVA

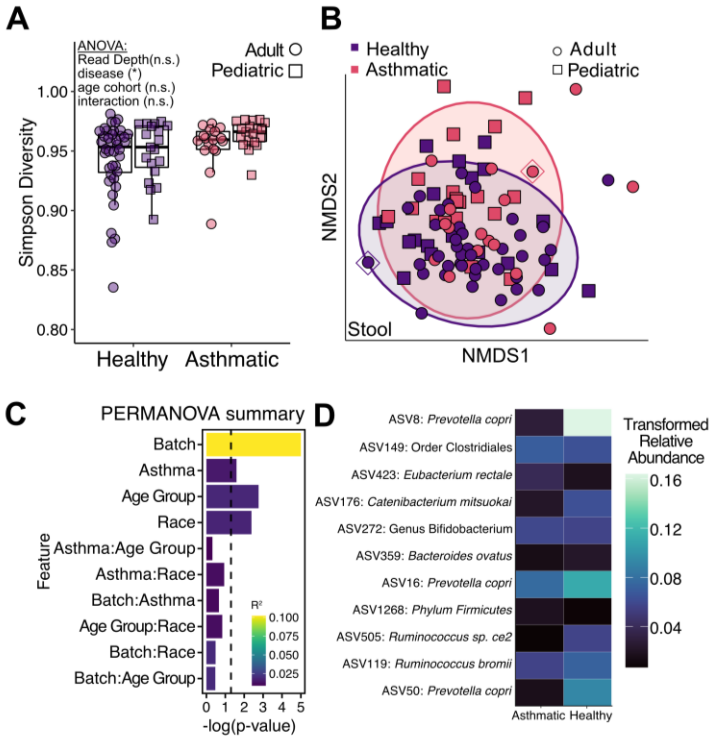

Figure 2.1. V4-16S rRNA profiling of stool from MARS cohort identifies gut microbiome differences in patients with asthma.

A) Simpson alpha diversity based on ASVs in stool samples from the MARS Cohort. Read depth was included as a variable to control for differences in library size. B) Non-metric multidimensional scaling (NMDS) on Bray-Curtis dissimilarity of MARS gut microbiomes. Ellipses represent 95% confidence intervals; diamonds indicate donor dyad (MARS0022/MARS0043). C) Bar plot summarizing results of PERMANOVA analysis on the beta diversity in (B). We also tested the homogeneity of the variance using PERMDISP2 and found no detectable difference in dispersion associated with age group or asthma status. Color represents R2 value and the length of the bar represents -log(p-value). Dashed line indicates a p-value threshold of 0.05. D) Heatmap of average relative abundances of differentially abundant taxa identified between patients with asthma and healthy controls using DESEQ2. Average relative abundances have been rescaled using the arcsine-square-root transformation. See also Figure 5.1.

results, we performed differential abundance analysis and identified alterations in taxa corresponding to disease status (11 ASVs), age (13 ASVs), and race (18 ASVs; see Table S3) after accounting for sequencing batch effect. Taxa differentially abundant between healthy and asthmatic individuals included several that have been previously reported to discriminate between healthy and atopic individuals including *Prevotella copri* (ASV8, ASV50) [95,96] and *Ruminococcus bromii* [97] (ASV119; Figure 2.1D).

2.3.2 Germ-free mice humanized with fecal microbiota from an adult with asthma showed an increase in lung oxidative stress and Th17 responses

We next sought to test whether the asthmatic gut microbiota could affect pulmonary inflammation in a mouse model of AAI. To explore this question, we selected a pair ("dyad") of human fecal microbiota samples from a healthy and asthmatic subject to "humanize" gnotobiotic mice by oral gavage. We initially selected an unrelated healthy-asthma dyad where the individuals were demographically similar (Subjects 0022 and 0043; matched for age, sex, BMI, race, and smoking history, see Table S1C). We constructed a Naïve Bayes Classifier (NBC) to generate several metrics that would help us evaluate the suitability of the selected dyad for characterization in a gnotobiotic animal model (see Methods for additional details). First, to quantify how similar each sample was to its respective cohort, we calculated a Sample Score that ranged from 0 (typical of healthy) to 1 (typical of asthma) and found that both candidate donor microbiomes were typical of their respective disease cohorts (Figure 2.2A). Second, we visualized samples by Feature Score (Figure 2.2B, Figure 5.2A-C, see Methods) to confirm our selected samples cluster with their respective cohorts. Third, to evaluate the testable microbial relationships in the selected dyad relative to all other possible selections, we counted the number of ASVs, for all possible dyads agnostic to host demographics, whose relative abundance was

consistent with the NBC's learned differences between the asthma and healthy cohorts. We defined this metric as a Pairwise Feature Score (PFS) which is greater than zero for every taxon in a dyad whose relative abundances are concordant with our model's predictions (see Methods). We compared the number of "model concordant taxa" (PFS>0) between all possible dyads (Figure 2.2C, Figure 5.2D) and found that our selected dyad contained a greater than average proportion of model concordant taxa. This indicates that the number of testable microbial comparisons mirroring the relationship between the asthma and healthy cohorts within our selected dyad is typical among all dyads. Together, these results support the idea that the demographically well-matched MARS0022-0043 dyad is characteristic of the microbial differences between cohorts and showcases a new tool for characterizing microbiome dyads.

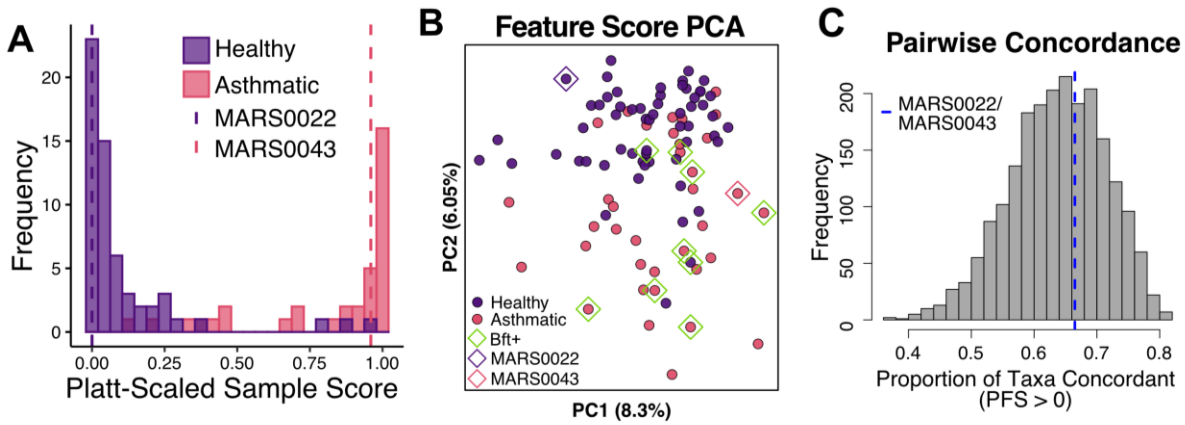

Figure 2.2. Selected donors capture discriminatory microbial relationships.
A) Histogram of Platt-scaled MARS NBC Sample Scores. A score of 0 is consistent with a "healthy appearing" sample and a score of 1 is consistent with an "asthma appearing" sample. Vertical dashed lines denote the sample scores for MARS0022 (purple) and MARS0043 (pink). **B)** PCA of NBC Feature Scores, calculated as the log likelihood that the relative abundance for a given taxon would occur in the healthy or asthma cohorts. MARS0022 (purple diamond) and MARS0043 (pink diamond) are the donor samples used in subsequent experiments in gnotobiotic mice. Green diamonds denote *bft* positive samples. **C)** Histogram of the proportion of pairwise concordant taxa across all possible healthy-asthma donor dyads. Vertical dashed line denotes the MARS0022 and MARS0043 dyad. See also Figure 5.2, Figure 5.8, and STAR Methods.

Mice humanized with the "healthy" (originating from MARS0022) or "asthmatic" (MARS0043) fecal microbiota underwent ovalbumin sensitization and challenge (OSC) or were sacrificed after 4 weeks without OSC, resulting in four groups of mice we will denote as follows: HO, mice colonized with a healthy microbiota undergoing OSC; AO, mice colonized with the asthmatic microbiota undergoing OSC; HN, mice colonized with a healthy microbiota that remained naïve; AN, mice colonized with an asthmatic microbiota that remained naïve (Figure 2.3A).

We performed V4-16S rRNA profiling of humanized gnotobiotic mice to better understand how gut microbial community ecology influences immune responses in the lung. This analysis affirmed that the gut microbial communities present in the gnotobiotic mice strongly resembled the human donors from which they had originated (Figure 5.3A-B). While human fecal transplantation into gnotobiotic mice can result in some reconfigurations of the original community [98], our analysis showed that many pairwise concordant taxa identified by our NBC colonized AO and HO gnotobiotic mice. These taxa included *B. uniformis*, *B. fragilis*, and Erysipelotrichaceae (*Longicatena caecimuris*), of which the latter two have been previously implicated in asthma pathogenesis[8,99,100] (Figure 5.3C).

To demonstrate that we successfully induced AAI in these mice, we performed bulk RNA-Seq on whole lungs and measured Th2 cytokines, *Il4*, *Il5* and *Il13* as well as serum anti-ovalbumin IgE. Although the degree of sensitization to ovalbumin and expression of Th2-related cytokines (*Il4*, *Il5*, or *Il13*) were markedly different between naïve germ-free mice and both OSC treated groups, consistent with successful induction of AAI, we found no difference between AO and HO mice (Figure 5.4A,B). As expected, humanized mice undergoing OSC had

transcriptome profiles that were distinct from naïve humanized mice, reflecting upregulation of genes and pathways related to Type 2 and eosinophilic inflammation and demonstrating that we successfully induced AAI in HO and AO mice (Figure 5.4C-E, Table S4, and Table S5).

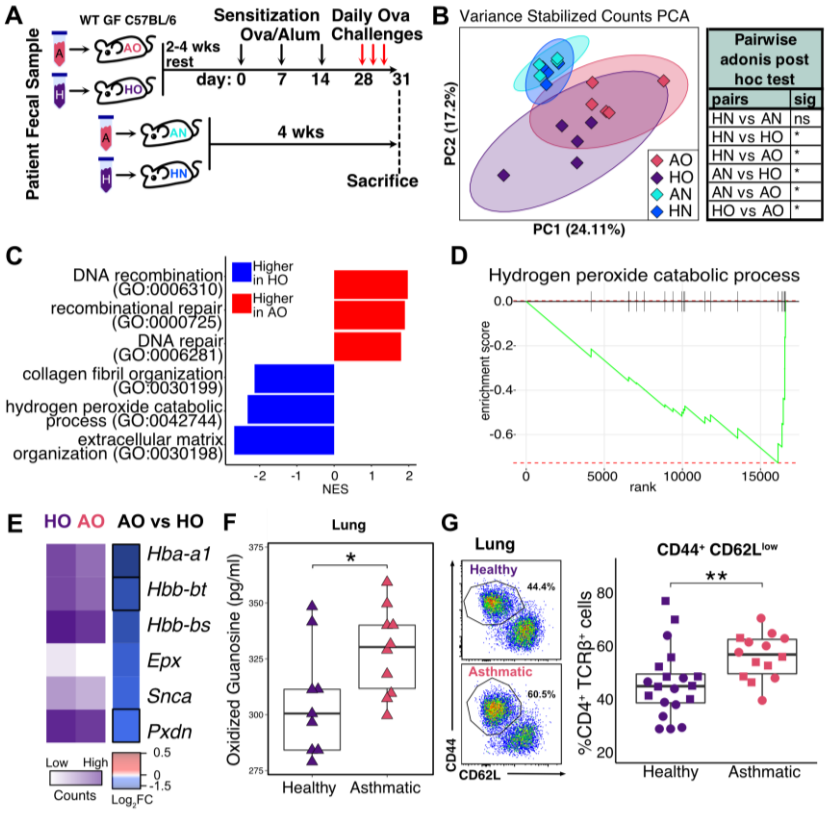

Figure 2.3 An asthmatic microbiota alters allergic airway inflammation phenotype in humanized gnotobiotic mice. The panels combine four different experiments shown denoted by shape. **A)** Overview of experimental paradigm. O: recipients of ovalbumin sensitization and challenge (OSC); N: no ovalbumin sensitization or challenge; A: recipient of fecal sample from human donor with asthma; H: recipient of healthy donor fecal sample. **B)** Principal Components Analysis of variance-stabilized gene counts from lung RNA-Seq (n=4,5). (PERMANOVA, 999 permutations: OSC ($p = 0.001$), donor ($p = 0.022$), and donor:OSC interaction ($p=0.037$)**;** Results of a post-hoc test are also shown in the table. **C)** Bar plot showing normalized enrichment score (NES) of select GO pathways upregulated (red) or down-regulated (blue) in AO mice compared to HO mice (n=5) (p-adjusted<0.05). **D)** GSEA enrichment plot of the hydrogen peroxide catabolic process pathway upregulated in the lungs of AO mice compared to HO mice. **E)** Heatmap of individual hydrogen peroxide catabolic process genes in AO and HO mice. In purple scale: log-transformed size factor normalized counts. In blue-red scale: Log-2 fold change of AO compared to HO. Genes outlined in black represent a $p < 0.05$. **F)** Oxidized guanosine in AO and HO mouse lungs (n=9-10 mice per group; Wilcoxon, one-tailed). **G)** Flow cytometry of the lungs of AO and HO mice comparing effector T-cells ($CD44+CD62L^{low}$, $CD4+TCRb^{+}$) from the lungs of AO and HO mice (n=6-10 mice/group). All experiments include 2-5 males and 2-5 females per group. Shapes denote four separate experiments and are consistent with subsequent figure. See also Figure 5.3, Figure 5.4, and Figure 5.5.

We next evaluated the effect of the gut microbiota on overall lung transcriptome profiles.

In naïve mice, the gut microbiota appeared to have little or no effect on the overall transcriptome profiles since the difference between HN and AN mice was negligible (Figure 2.3B). In contrast, mice undergoing OSC colonized with different microbiota had distinct lung transcriptome profiles (see HO vs. AO mice in Figure 2.3B), suggesting the gut microbiota may be important in the context of AAI. Notably, expected allergy-related genes (*Il4, Il5, Il13*) and pathways (Type 2 immune response, hyperreactivity, eosinophil, and neutrophil pathways) that were upregulated by the induction of AAI were not differentially expressed between HO and AO mice (Figure 5.4A-C). These data suggest that Th2-related responses do not make up the transcriptomic differences between HO and AO mice.

Differentially abundant pathways between HO and AO mice included several involved in DNA repair and recombination (Figure 2.3C, Figure 5.4F-G, Table S4 and S5). We also saw enrichment of a pathway associated with the breakdown of hydrogen peroxide (GO:0042744) in HO compared to AO mice (Figure 2.3C, D). Many of the genes in this pathway have been implicated in protective responses to oxidative stress. These include hemoglobin synthesis genes (*Hba-a1, Hbb-bt,* and *Hbb-bs*), which are known to be upregulated and protective during oxidative stress in extra-erythropoietic tissues [101,102], and peroxidasin (*Pxdn*), an enzyme which is likewise known to play a protective role during oxidative conditions in tissues (Figure 2.3E)[103]. When considered with the known roles of oxidative stress in upregulating DNA damage and repair machinery in AAI [104,105], we hypothesized that increased oxidative stress led to DNA damage in AO mice. To test this idea, we measured oxidized guanosine, a marker for oxidative stress[106], in the lung tissue, and found it to be increased in AO compared to HO groups (Figure 2.3F). Together, these data reflect increased oxidative stress in the lungs of mice that received

the asthmatic microbiota in the context of airway inflammation.

We further investigated the impact of the microbiota on immune cell subsets by performing immunophenotyping on tissues from HO and AO mice. Flow cytometry of lung tissue demonstrated an increase in effector T cells in AO mice compared to HO mice ($TCRb^{+}CD4^{+}CD62LloCD44^{+}$ cells; Figure 2.3G). Consistent with RNA-Seq results, we did not observe a difference in neutrophils or eosinophils recovered from the lung tissue (Figure 5.5A,B). However, we observed an increase in CD4+ T cells expressing IL-17A after restimulation and intracellular staining (Figure 5.5C). Coupled with the increased transcription of *Il17a* in the lung tissues (Figure 5.4B), these findings are consistent with enhanced Th17 cell recruitment to the lungs of AO mice compared to HO mice. This increased Th17 response could also be detected in the mesenteric lymph nodes of AO mice but we did not observe systemic increases in Th17 cells in the spleen or increased serum IL-17A protein (Figure 5.5C-E), suggesting the presence of Th17 cells in the gut and the lung. Taken together, these results support the idea that the gut microbiota from our selected dyad did not change the expression of allergy-associated pathways but did demonstrate alterations in Th17 responses and increased markers of oxidative stress in the lungs of OSC treated mice.

2.3.3 IgA-Seq identifies enterotoxigenic *Bacteroides fragilis* as a potential effector taxon during AAI

Based on other studies finding that IgA-coated bacteria modulate the host mucosal immune response [107], we used a technique called IgA-seq to identify microbes from the fecal microbiota that are coated with IgA. Our IgA-seq analysis showed that *Bacteroides* species were more likely to be coated by IgA in AO mice compared to HO mice (Figure 2.4A, B). In AO

mice, these IgA-coated bacteria included *B. uniformis* and most prominently, *B. fragilis*.

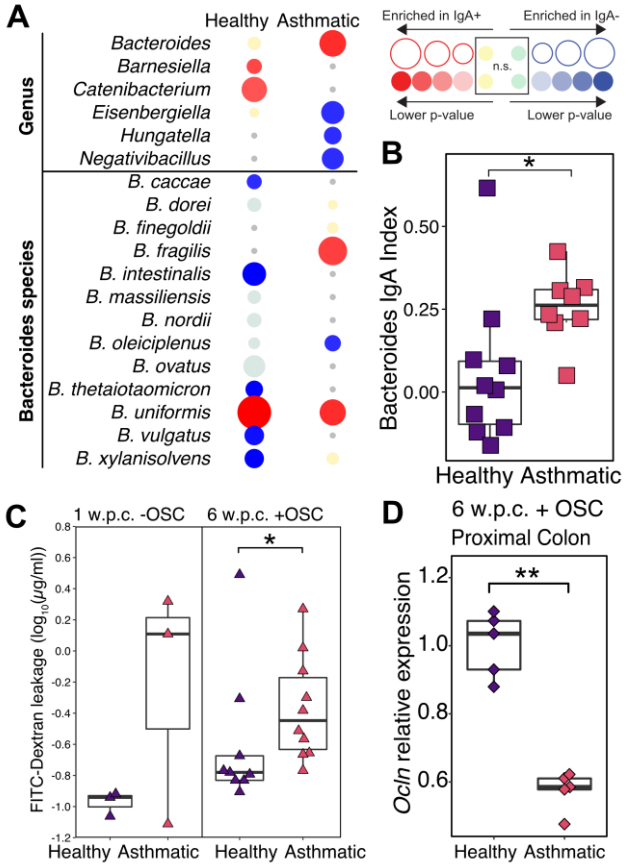

Figure 2.4. Enterotoxigenic Bacteroides fragilis from human donor with asthma is linked to increased gut barrier permeability in humanized gnotobiotic mice.
A) Bubble plot of IgA-Seq results from AO and HO mice at the genus level (n=8-10 mice/group), and specific species of *Bacteroides*. Bubble color indicates significant enrichment (red) or depletion (blue) in the IgA coated fraction. Bubble size indicates the magnitude of enrichment as determined by the IgA index. **B)** IgA Index of Bacteroides genus between AO mice and HO mice colonized with asthmatic or healthy microbiota (Wilcoxon p=0.012; n=8,10 mice/group). **C)** Intestinal permeability of mice colonized with asthmatic and healthy microbiota following a 1-week (Wilcoxon p=0.7; n=3 females/group) and 6-week colonization (Wilcoxon p=0.03; n=8-10 mice/group) on a log10 scale. **D)** qPCR of occludin gene (*Ocln*) in the proximal colon of humanized mice (Wilcoxon p=0.008; n=5 mice/group). All experiments include 4-5 males and 4-5 females per group unless otherwise stated. Shapes denote separate experiments and are consistent with previous figure. Two-sided Wilcoxon test for all boxplots. See also Figure 5.3 and Figure 5.4.

Supported by the observation that *B. fragilis* was in the top 25% of the most discriminatory taxa in the NBC, we isolated and sequenced the *B. fragilis* strain found in the asthmatic microbiota, referred to here as BFM04319. We found that its genome encoded for the *B. fragilis* toxin gene, *bft*, also called fragilysin.

Fragilysin is a *B. fragilis* diarrheal toxin [108] whose proteolytic activity is directed against the adherens junction protein E-cadherin and disrupts the intestinal epithelial barrier [109,110]. We reasoned that gut barrier function would be impaired in AO mice, so we assessed FITC-Dextran leakage in the humanized gnotobiotic mice and found a trend towards increased intestinal permeability in mice colonized with the asthmatic microbiota as early as one week after gavage that persists after OSC (Figure 2.4C). Similar to previous studies [111], increased FITC-Dextran leakage was also accompanied by downregulation of the gene encoding for occludin (*Ocln*), an important protein involved in the regulation of intestinal tight junctions, in the colons of AO mice (Figure 2.4D).

Since gut barrier dysfunction has been previously implicated in inducing systemic oxidative stress [111,112], we hypothesized that ETBF in our asthmatic stool sample was responsible for the increased oxidative stress in the lungs of AO mice by disrupting the gut barrier. We directly tested whether ETBF alone is sufficient to modulate oxidative stress after OSC by examining three groups of gnotobiotic mice: 1) mice that remained germ-free (GF); 2) mice monocolonized with BFM04319 (ETBF), and 3) mice monocolonized with a non-toxigenic *B. fragilis* strain VPI2553 (NTBF) (Figure 2.5A) [113]. We performed OSC on all three groups and confirmed that all groups had increases in anti-ovalbumin IgE and Th2 cytokine expression, with the exception of *Il4,* compared to naïve germ-free controls providing evidence that we successfully induced AAI (Figure 5.6A-B). We note that IL-4 from whole lung tissues in C57 BL6 mice is a less sensitive measure of AAI than other markers such as IL-5 or IL-13 [114,115]. There were no differences in these markers between OSC groups. As expected, mice colonized with ETBF had greater intestinal permeability than either GF controls or NTBF colonized mice

(Figure 2.5B). We observed higher levels of oxidized guanosine in the lungs of ETBF colonized mice compared to GF controls (p = 0.018), a trend for NTBF colonized mice to have greater lung oxidized guanosine than GF mice (p = 0.15), as well as a trend for ETBF colonized mice to have higher levels of lung oxidized guanosine than NTBF colonized mice (p = 0.1, Figure 2.5C).

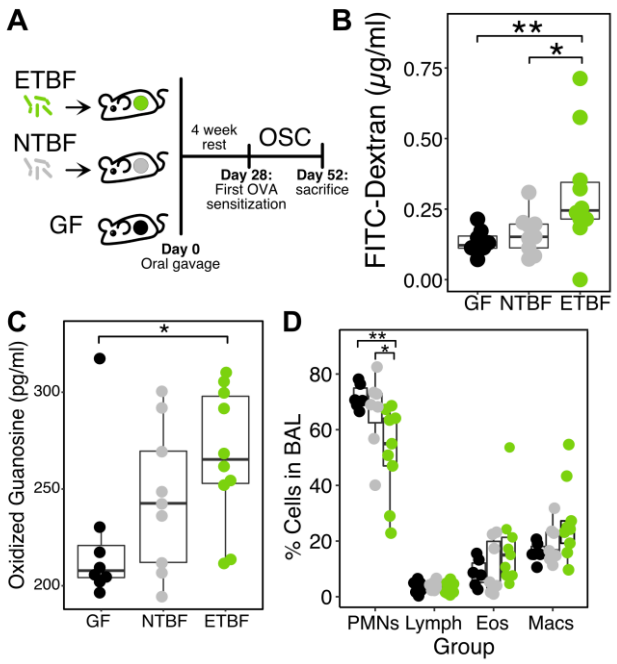

Figure 2.5. Monocolonization with ETBF increases gut barrier permeability and lung oxidative stress in gnotobiotic mice.
A) Overview of the experimental paradigm testing the ability of *bft* carrying enterotoxigenic *B. fragilis* (ETBF) and *bft* lacking non-toxigenic *B. fragilis* (NTBF) to affect gut barrier function in mice. **B)** Intestinal permeability of mice colonized with either ETBF or NTBF and germ-free controls (GF) (n=8-10 mice/group; Kruskal-Wallis, one-tailed Dunn post hoc with Benjamini-Hochberg correction). **C)** Oxidized guanosine in lungs of ETBF, NTBF, and GF mice (n=8-10 mice/group; Kruskal-Wallis, one-tailed Dunn post hoc with Benjamini-Hochberg correction). **D)** Cytospin cell counts from bronchoalveolar lavage collected from mice colonized with ETBF, NTBF or germ-free controls. PMN=polymorphonuclear cells (i.e. neutrophils), Lymph=lymphocytes, Eos=eosinophils, Mac=macrophages (n=8-10 mice/group; Kruskal-Wallis, one-tailed Dunn post hoc with Benjamini-Hochberg correction). This experiment includes 3-4 males and 5-7 females per group. See also Figure 5.6.

Further, mice colonized with ETBF had lower neutrophil counts in their bronchoalveolar lavage compared to NTBF-colonized mice or GF controls, although eosinophils, lymphocytes, alveolar macrophages and *Il17a* transcription were not significantly different (Figure 2.5D, Figure 5.6B). These results provide evidence that colonization with ETBF is sufficient to cause gut barrier

dysfunction, modulate airway inflammatory profile, and increase oxidative stress following OSC.

2.3.4 Expression of *bft* in mice humanized with ETBF+ donor microbiota is highly dependent on microbial community context

Following our identification of ETBF as an important driver of oxidative stress in the lungs of gnotobiotic mice undergoing OSC, we then sought to investigate the importance of ETBF in the context of different microbial communities. While exploring multiple donors does not guarantee translatability to humans [98], it does reveal the strength and robustness of the ETBF-oxidative stress phenotype. Therefore, we selected multiple dyads made up of one ETBF+ donor with asthma and one ETBF- healthy donor for humanization and measured gut barrier permeability, lung cytokine gene expression, and pulmonary oxidative stress after OSC. To select donor samples, we screened MARS stool samples for *bft* using PCR and qPCR [116,117] and identified five asthmatic fecal samples harboring ETBF and five healthy fecal samples lacking ETBF, each pair matched by both age group and community composition metrics derived from our NBC (see Methods; Figure 5.7A-B and Table S1C). Using the selected communities, we humanized germ-free mice and performed OSC. We evaluated humanization in recipient gnotobiotic mice by comparing the Bray-Curtis dissimilarity scores from 16S rRNA sequencing of stool from human donors and recipient mice at the time of sacrifice. This led us to exclude three groups of humanized mice (2 ETBF-, 1 ETBF+) whose engrafted fecal microbiome did not best reflect their donors, leaving seven humanized mouse groups (Figure 2.6A, Figure 5.7C; see Methods). As previously observed, all experimental mice had increases in Th2 cytokine expression, except *Il4*, and anti-ovalbumin IgE compared to naïve germ-free controls and that there were no differences in these markers between OSC groups (Figure 5.7D-E).

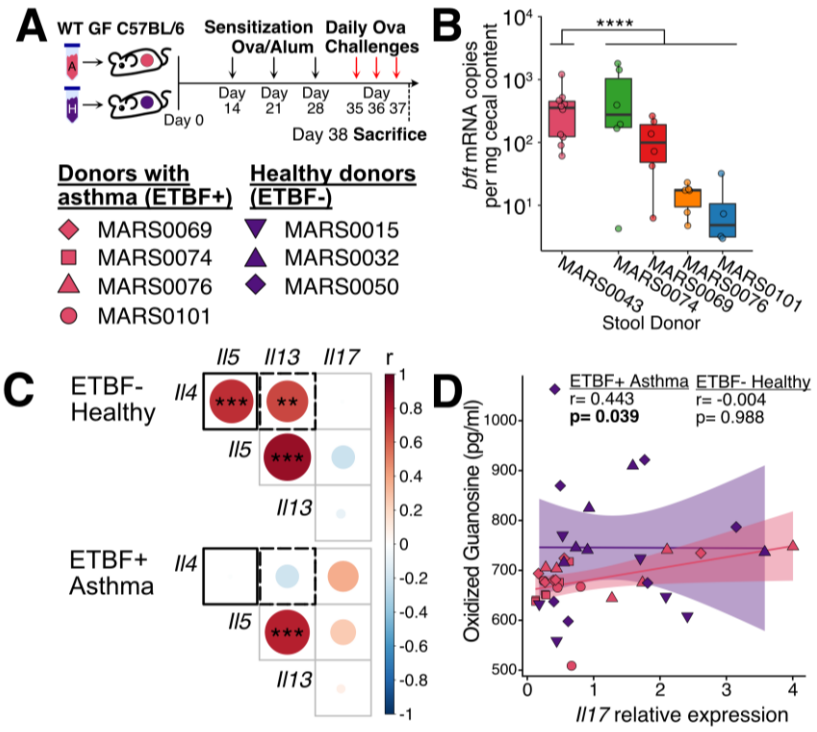

Figure 2.6. Humanization with multiple microbiota suggests ETBF produced bft can alter lung inflammation in a community context dependent manner.
A) Overview of experimental paradigm. Five healthy and five asthmatic donor microbiota were used to humanize germ-free mice (4 - 7 mice per donor). Two weeks later, pulmonary inflammation was induced by OSC. After evaluating humanization, only seven mouse groups were used in subsequent analyses. **B)** qPCR measurements of *bft* in the cecal content of humanized mice (t-test against initial MARS0043 microbiome). **C)** Pearson correlations between RT-qPCR measures of lung cytokines from mice humanized by ETBF- healthy (top) and ETBF+ asthmatic (bottom) microbiota. Area and color of the circle represent the absolute value of the Pearson correlation coefficient with corresponding color key. Asterisks represent the p-value from the Pearson correlation. Solid (Steiger's test; p=0.0044) and dotted (Steiger's test; p = 0.0026) black squares indicate a statistically significant difference between the two outlined correlations. **D)** Pearson correlation between oxidized guanosine in the lungs of humanized mice and RT-qPCR measures of lung *Il17*. All experimental groups included 1-4 female and 2-4 male mice. See also Figure 5.7.

Overall, we found no differences in gut barrier leakage, pulmonary oxidative stress, or gene expression of lung cytokines between mice colonized with an asthmatic ETBF+ microbiota and those receiving an ETBF- healthy microbiota (Figure 5.7E-G). To understand the differences between these experiments and our original donor experiment, we measured copies of *bft* RNA and DNA in the cecal contents of ETBF+ humanized mice and found that the median expression levels of the mice with newly selected human samples were 1.2- to 67.5-fold lower than that of

the original donor sample, MARS0043 (Figure 2.6B, t-test p=0.01). These results are consistent with previous research showing that ETBF vary their enterotoxin expression based on the microbial community context [118] and may explain the lack of an ETBF-oxidative stress phenotype seen in these mice.

Given that the expression of *bft* was far more variable than observed in the original donor dyad, we tested if there was any correlation between *bft* expression and markers of Th17 or AAI. As expected, the expression of Th2 cytokine encoding genes *Il5* and *Il13* were highly correlated with each other in both experimental groups. However, we found that *Il5* and *Il13* were correlated with *Il4* in mice colonized with ETBF- microbiota from healthy individuals, but not in mice colonized with ETBF+ microbiota from individuals with asthma (Figure 2.6C). The *Il4*/*Il5* and *Il4*/*Il13* correlations observed in the healthy group were significantly stronger than that in the asthma group (Steiger's test p=0.0044 and 0.0026 respectively), suggesting different patterns of Th2 cytokine expression in the lungs. Previous studies have identified a correlation between *Il17* expression and oxidative stress in the lungs[80]. We found that ETBF+ humanized mice displayed the expected correlation between lung *Il17a* expression and oxidative stress (Figure 2.6D, Pearson = 0.443, p = 0.039), but the ETBF- healthy humanized mice did not (Pearson = -0.004, p = .987). Together, these results suggest that ETBF+ microbiota from donors with asthma may link oxidative stress to *Il17a*.

2.3.5 Gut colonization with ETBF is more prevalent among individuals with asthma compared to healthy controls

Finally, we investigated the importance of *bft* in humans by examining its prevalence in the stool of MARS participants. Using PCR and qPCR [116,117], we screened all human fecal specimens in which *B. fragilis* was detected by 16S rRNA sequencing for *bft* and found a total of

eight individuals with asthma and two healthy subjects with detectable *bft*. Though prevalent in only 22% of the asthma cohort, *bft* is significantly enriched in subjects with asthma compared to healthy individuals (Figure 2.7A and Figure 2.2B). There was no statistically significant difference in Asthma Control Test score [119] between patients with or without ETBF (Table S1D, p=0.3). We then asked whether higher rates of ETBF could be a consequence of an increased prevalence of *B. fragilis* in the asthma cohort but found no difference in the frequency of *B. fragilis* colonization (Figure 2.7B). Even amongst only individuals colonized with *B. fragilis*, patients with asthma were still more likely to be colonized with ETBF than NTBF compared to healthy individuals (Figure 2.7C). We next examined fecal calprotectin levels from remaining subjects to determine whether asthma was associated with gut barrier permeability. Overall, we found that calprotectin was either low or undetectable in all samples but tended to be more detectable in samples originating from a donor with asthma (Fisher's test p-value = 0.053, Figure 2.7D). While the low levels of calprotectin are likely due to long-term sample storage, these results tentatively suggest a degree of barrier dysfunction in patients with asthma.

To check if *bft* presence alone associated with a shift in the microbiota as a whole, we tested *bft* presence as a feature in a sequential PERMANOVA on Bray-Curtis beta diversity (Figure 2.7E) after accounting for the important demographics presented in Figure 2.1C (sequencing batch, race, age, and asthma status). We found no statistically significant effect of *bft* presence alone on global shift in the beta diversity (p=0.24; Figure 2.7F). However, we suspected that other taxa in the gut may be associated with ETBF, so we performed a co-occurrence analysis with *bft* presence and our 16S rRNA sequencing results (Figure 2.7G). In

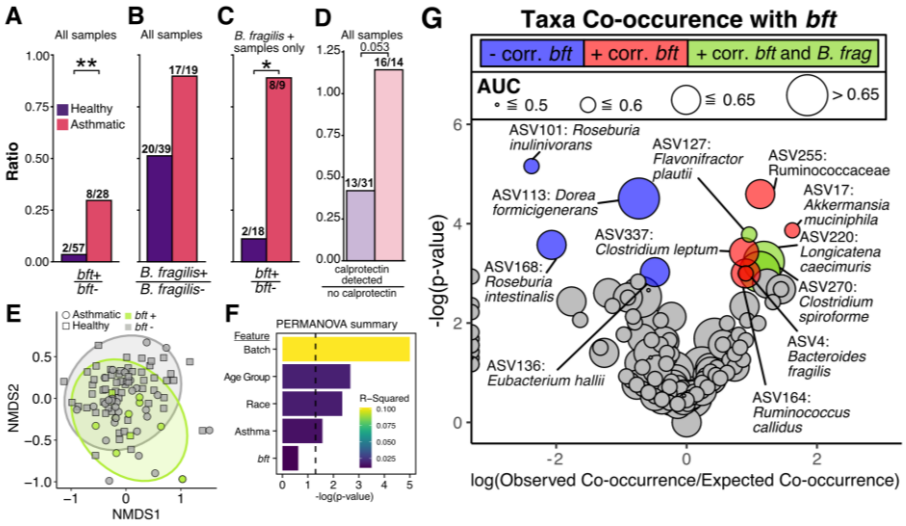

Figure 2.7. bft is enriched in individuals with asthma.

A-D) Ratios of **(A)** *bft* positive over *bft* negative samples (n=59 Healthy, 36 Asthmatic). *bft* positive samples were characterized by PCR and qPCR screening among the 37 total subjects with *B. fragilis*. The 58 remaining subjects without *B. fragilis* by 16S rRNA sequencing were not screened and were considered *bft* negative. **B)** *B. fragilis* positive samples over *B. fragilis* negative samples, characterized by V4 16S rRNA community profiling (n=59 Healthy, 36 Asthmatic). **C)** *bft* positive samples with *B. fragilis* over *bft* negative samples with *B. fragilis* (n=20 Healthy, 17 Asthmatic). **D)** Stool samples with detectable calprotectin via ELISA over those with no detectable calprotectin. **E)** Non-metric multidimensional scaling (NMDS) on Bray-Curtis dissimilarity of MARS donor microbiomes at ASV-level. Ellipses represent 95% confidence intervals. Shape denotes asthma (circle) or healthy (square). Color denotes a positive (green) or negative (grey) hit for *bft* in fecal gDNA. **F)** Bar plot summarizing results of a sequential PERMANOVA analysis on the beta diversity in (E). Color represents R^2 value and the length of the bar represents -log(p-value). Dashed line indicates a p-value threshold of 0.05. **G)** Volcano plot describing co-occurrence between *bft* and taxa in human stool samples, calculated as the log ratio of the number of samples in which *bft* was observed in the same sample as the taxon compared to the number of times *bft* and the taxon would be expected to occur in the same sample by chance alone[73]. Color represents a significant positive (red) or negative (blue) co-occurrence with *bft* after multiple hypothesis correction. Taxa positively correlated with *bft* and *B. fragilis* (ASV4) are shown in green. Gray represents a non-significant relationship. Size represents the area under the ROC curve (AUC) of the taxa calculated from the NBC. (For all panels except C and D: n=59 Healthy, 36 Asthmatic). See also Figure 5.1.

addition to *B. fragilis*, we identified 7 taxa positively correlated and 4 negatively correlated with ETBF colonization. Among the positively correlated taxa were 4 taxa that co-occurred with *bft* but not *B. fragilis* (ASV4), suggesting that ETBF may influence gut microbiota composition in a different manner from NTBF species. Notably, *bft* co-occurs with *Dorea formicigenerans* (ASV113, AUC = 0.663, Rank 11), *Longicatena caecimuris* (ASV220, AUC = 0.662, Rank 12), and *Clostridium spiroforme* (ASV270, AUC = 0.657, Rank 13), which are highly discriminatory between asthma and healthy individuals according to our NBC (Figure 5.8A). Additionally, our

differential abundance analysis of ETBF+ microbiota against ETBF- microbiota, identified enrichment of *C. spiroforme* (ASV270, AUC = 0.657, Rank 13) and *Ruminococcacae* (ASV255, AUC = 0.628, Rank 29) in the ETBF+ donors and *R. intestinalis* (ASV168, AUC = 0.602, Rank 65) in the ETBF- donors, supporting the findings from the co-occurrence analysis (Table S3). These results imply that the presence of *bft* reinforces asthma-associated changes in the gut microbiota.

2.4 Discussion

Despite our growing understanding of the origins of asthma, the heterogeneous nature of the disease remains a barrier to treatment. In addition to factors such as age [120], sex [121], smoking status [122], and microbial exposures [123], the gut microbiota is increasingly appreciated as a determinant of asthma risk [8,35]. Here we leverage a cross-sectional human clinical study of 95 patients with and without asthma and use humanized gnotobiotic mice to show that, in the right community context, ETBF can increase Th17 inflammation and oxidative stress in the lungs of mice with AAI, potentially by disrupting the gut barrier. This finding is of potential clinical relevance since we found that ETBF is more prevalent in the gut microbiomes of people with asthma.

Humanization of gnotobiotic mice is a powerful tool to study the human microbiota and its effects on host phenotypes, but has many well-established caveats. First, the human microbiota does not perfectly maintain community and functional structure between donors and mouse recipients[98,124]. Second, virulence factors adapted to human hosts may not affect mice to the same degree. Third, a gut microbiota may influence the host phenotype by multiple mechanisms simultaneously, such as in asthma where multiple microbial metabolites are known

to alter the course of the disease, each by their own respective mechanisms [12,13,35,62]. Together, these issues can complicate the interpretation of results from gnotobiotic experiments and affect the translatability of those results back to humans.

Designing experiments to address these limitations requires the selection of donor samples that both control for confounding factors that modify the microbiome apart from the disease of interest as well as capture the relevant microbial relationships between healthy and disease-affected populations. Control of confounding factors is often accomplished by matching the clinical demographics of donors [98,125] (Figure 2.1B-C) but there is no standardized practice for identifying samples that capture important microbial relationships. Given the immense interpersonal variability of the microbiome [126–128], no single pair of samples captures all the discriminatory microbial features from a population but we can still select pairs enriched in these relationships to study. We believe that the approach presented here for sample selection will be useful in identifying microbial drivers of disease in future clinical studies.

Our in-depth profiling of a single dyad revealed an ETBF+ microbiome from a donor with asthma increased the Th17 response and oxidative stress within the lungs of humanized gnotobiotic mice in the context of AAI. While previous studies have implicated airway microbes in inducing Th17 responses in AAI [11,93], our work provides evidence that this phenotype can be mediated by a gut microbiota-expressed factor. In follow-up monocolonization experiments, we confirmed that the ETBF isolated from the donor with asthma causes increased gut barrier permeability and pulmonary oxidative stress in the context of pre-existing inflammation. Intriguingly, we observed that neutrophils were lower in ETBF colonized mice, implying that the increased oxidative stress is not a result of reactive oxygen species generation from neutrophils.

On the other hand microbial products, including LPS and fragilysin, have been shown to directly induce oxidative damage in lung and gut epithelial cells, respectively [129,130], suggesting an alternative source of oxidative stress. Notably, we did not measure an increased Th17 response in the lungs of ETBF monocolonized mice compared to NTBF or GF mice. This could mean that other community members in addition to ETBF are important for inducing systemic [131] and lung Th17 responses.

We also explored the effect of community context on ETBF by colonizing groups of mice with ETBF+ communities from multiple donors with asthma. We did not detect increased gut permeability, pulmonary oxidative stress, or lung Th17 responses in the ETBF+ humanized mice. However, we did observe that overall expression of *bft* was much lower in this experiment compared to the MARS0043. Despite this, ETBF+ asthmatic microbiota caused an altered pattern of Th2 cytokine expression compared to ETBF- healthy microbiota. Taken with our ETBF monocolonization data, these results imply that *bft* may influence inflammation in the lungs, but that its expression and impact on asthma may be modulated by community context. Based on these results, we propose that *bft* and ETBF may contribute to the clinical heterogeneity of asthma in humans and raise the question of whether additional gut microbes may contribute to asthma endotype.

While gnotobiotic experiments offer control over many features of the gut microbiota, they ultimately remain a proxy of asthma in humans. Confirming the findings from our humanized mouse models will require a longitudinal human clinical study including ETBF colonized subjects with asthma and controls that undergo rigorous phenotypic characterization that incorporates measuring gut barrier permeability, circulating microbial products, and

immunophenotyping of the blood and bronchial lavage. The results of this study would establish whether ETBF could adversely influence asthma in people and define the environmental and microbial circumstances that enable ETBF to cause these phenotypes. Additionally, the outcomes of such a clinical study could also link two previously disparate features of asthma: first, that increased gut permeability [57,58] and increased gastrointestinal symptoms [132] have been observed in patients with asthma, and secondly, that increased oxidative stress is associated with more severe asthma with a higher rate of exacerbations and corticosteroid resistance[133,134].

We acknowledge that ETBF is unlikely to be a universal mechanism contributing to asthma and would probably only apply to a subset of people living with asthma (22% from our study). We also observed in our mouse models that the effects of ETBF were highly variable and caused barrier permeability and increased pulmonary oxidative stress in only one of the six ETBF+ microbiota samples we tested from individuals with asthma. These results came from four different colonization experiments (Figure 2.3 and Figure 2.4) using the first human donor dyad before the experiment that employed ten human donors (Figure 2.6), strongly supporting a distinct effect of the initial donor dyad on AAI. The first set of experiments replicated an effect on lung inflammation associated with ETBF colonization in AAI, while the second set of experiments demonstrates that this phenotype is not generalizable to all of the ETBF+ donors. It could be the case that, like ETBF-induced weight loss, the gut and lung phenotypes caused by the one ETBF+ microbiota sample opportunely emerge in particular community and environmental contexts (Figure 2.6 and Figure 5.7) [118]. Here, we suspect that other members of the gut microbiota are important for *bft* expression and substantially modulate the penetrance of the ETBF phenotype on AAI. For example, *Roseburia* species, linked to protective effects in

AAI [90], tend not to co-occur with ETBF in our study while *Ruminococcus* species including *R. callidus* that co-occur with ETBF have been found to be enriched in the stool of patients with asthma [135].

On the other hand, implicating a barrier-disrupting organism in asthma pathology could lead to unique interventions to improve asthma control. For instance, targeting ETBF colonization by vaccination, antibiotics, or phage therapy could offer a novel means of manipulating the immune manifestations of asthma. Alternatively, candidate therapeutics designed to improve gut barrier function through their effects on tight junctions for other diseases including arthritis [136] and celiac disease [137] could be repurposed to modify the gut-lung axis in asthma. Unlike early childhood interventions that aim to alter the development of asthma, these therapies may provide benefit to patients with established disease.

2.4.1 Limitations of the study

Since our cross-sectional study focused on patients with established disease, we cannot demonstrate a causal relationship between ETBF or the gut microbiota and asthma. Rather, our findings demonstrate that ETBF modifies the immunological characteristics of the lung in gnotobiotic mice experiencing AAI. We identified these findings by contrasting microbiota sourced from healthy donors against stool sourced from patients with asthma, but our findings are limited in that we cannot separate the effects of the rest of the microbiota. While this could be accomplished by testing microbiota from subjects with asthma that contain ETBF against microbiota from asthma subjects without ETBF, based on our data from Figure 2.6, we estimate that the number of samples we need to separate the effect of ETBF from other microbes exceeds the number available from our clinical study. In our experiments, we intentionally chose the

ovalbumin-alum model since it is a well-characterized antigen that is easily translated into gnotobiotic models. Ovalbumin does not evoke the strongest AAI response in C57 BL6 mice[114,115], as we observed for *Il4* in our study, and future gnotobiotic experiments will need to include a broader range of biologically relevant antigens such as house dust mite or cockroach antigen to better contextualize our findings in acute mouse AAI to human asthma, a chronic disease. Additionally, humanized gnotobiotic experiments are limited by the level of engraftment of donor microbes and the differences between human and mouse host biology. While we could reduce variation due to engraftment in some experiments by performing multiple replicates (e.g. Figure 2.3), it was not technically feasible for us to perform this on all experiments given the challenges inherent to gnotobiotic experiments. Future studies extrapolating our results to other strains and microbiota would further contextualize our findings, but ultimately the mechanistic insights revealed by these models must be confirmed in clinical studies.

2.5 Materials and Methods

Data and code availability

RNASeq and V4 16S rRNA and genome sequencing data have been deposited at the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) and are publicly available as of the date of publication under project accession number PRJEB45298. All original code to run the mixture distribution-based NBC is available via zenodo (doi: 10.5281/zenodo.7522060). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Human Subjects

The Microbiome and Asthma Research Study (MARS) was designed to investigate the role of intra- and interpersonal variation in the gut microbiome on human asthma pathogenesis. This cohort has been described in a previous manuscript [93]. Briefly, all individuals were either healthy or had moderate-to-severe asthma and had not received oral corticosteroids or antibiotics in the 30 days prior to enrollment. Individuals were included in the asthma cohort if they had at least one positive skin prick test or serum aeroallergen-specific IgE. Two age cohorts were recruited representing a pediatric (aged 6-10 years) and an adult (aged 18-40 years) population. Of 104 patients initially enrolled in the study, 103 were asked to provide samples and fill out the questionnaire. Of the 103 asked, 6 were excluded due to insufficient medical documentation and 2 did not provide stool samples (**Error! Reference source not found.**). As summarized in Table S 1A, a total of 95 patients provided stool and relevant demographic data, including antibiotic and asthma medication use over the past year, and were ultimately included in this study. Fecal collection tubes with spoons and toilet hats were provided to patients after enrollment. Patients either provided a fecal specimen at the recruitment visit or collected and stored a fecal sample at home at -20° C for no more than 24 hours before returning the sample to the study site where it was stored at -80° C until processing.

MARS was approved by the Washington University Institutional Review Board (IRB ID# 201412035). Written informed consent documents were obtained from all MARS subjects or their legal guardians. All animal studies conformed to ethics of animal experimentation and were approved by the Institutional Animal Care and Use Committee (IACUC Protocol ID #: 20180286 and 21-0394).

Experimental Animals and Ethics

Germ-free C57BL/6 mice were bred and maintained in sterile flexible vinyl isolators. Sterility was assured by monthly monitoring of mouse stools by 16S rRNA gene PCR amplification as well as aerobic and anaerobic culture. Germ-free mice were maintained on a strict 12 hour light cycle and a diet of autoclaved mouse chow (LabDiet: Standard Diet 5021 - Autoclavable Mouse Breeder). For each experiment no more than five mice were housed in a cage. Mice were randomly assigned to experimental groups, with the exception of age and gender matching mice in each experimental group. Investigators were not blinded to experimental groups. Male and female mice within a group were caged separately but housed and handled within the same flexible vinyl isolator. Within all experiments, mice of the same sex and receiving the same human microbiota (or isolate) were caged together. Sex was tested, but not a significant contributor to the phenotypes described unless otherwise shown.

Isolation and Growth Bacteroides fragilis strains

B. fragilis strain BFM04319 was isolated from MARS0043 (subject with asthma) stool using anaerobic culture methods. Briefly, a 10 mg/mL stock of homogenized stool in 10% glycerol/PBS was thawed in a Coy anaerobic chamber, plated on BHI/mucin agar containing 0.1% mucin, 0.1% Resazurin, and 0.05% L-Cysteine (HCl) and 1.2 mg/L histidine/hematin and incubated at 37˚C for 6 days. Single colonies were isolated in BHI/mucin liquid media and stocked in 10% glycerol. Culture purity and identity was confirmed by V4 16S sequencing. Non-toxigenic *B. fragilis* VPI2553 was provided as a generous gift from Dr. Jeffrey I. Gordon. All strains were grown anaerobically in BHI/mucin media as above at 37˚C for 24 hours before diluting 1:2 with 20% glycerol, freezing, and gavaging into GF mice.

Humanized Gnotobiotic Mouse Model

We began with frozen pulverized fecal samples brought into an anaerobic chamber and transferred into reduced 1X PBS supplemented with 0.1% Resazurin and 0.05% L-Cysteine (HCl) to a concentration of 10 mg/mL. Fecal samples were then thoroughly homogenized using a sterilized probe homogenizer. Resuspended fecal samples were diluted 1:2 in 20% glycerol in PBS/Cysteine and stored in sealed HPLC vials at -80° C until use. Sample viability was confirmed by outgrowth of 100 uL of prepped gavage material on BHI/mucin agar containing 0.1% mucin, 0.1% Resazurin, and 0.05% L-Cysteine (HCl) and 1.2 mg/L histidine/hematin. A minimum of $1x10^4$ CFU/mg was used for colonization. Germ-free mice were humanized by oral gavage of 200 mL of a thawed homogenized stool sample. The microbiota was allowed to stabilize over two to four weeks prior to further intervention. All experiments were performed with at least one asthmatic and one healthy microbiota and separate donor groups were maintained in separate isolators to prevent cross-contamination. For the additional humanization experiments seen in Figure 2.6, we selected asthmatic microbiota containing ETBF, of which 5 were viable in culture outgrowth. For each donor with asthma, we selected a healthy sample within the same age group – either adult or pediatric – that did not contain ETBF but had the highest pairwise feature score (Figure 5.7C). We assessed the quality of human microbiota transplantation by calculating the Bray-Curtis dissimilarity of 16S rRNA sequencing data between human donors and murine recipients. We then ranked the average dissimilarity between recipients and donor and only included recipient groups who were the top ranked in similarity to their human donor.

Allergic Airway Inflammation Model

Allergic airway inflammation was induced in mice using chicken egg ovalbumin as previously described [93,138]. The ovalbumin model of sensitization and challenge is convenient for gnotobiotic experiments since it is a pure, sterile protein and has previously been used in other gnotobiotic systems[35]. Germ free mice were sensitized on days 0, 7, and 14 by intraperitoneal injections of 200 mL of OVA/alum: ovalbumin (50 mg, Sigma grade V) combined with Imject Alum (Thermo Scientific) as per the manufacturer's recommendations. Mice were challenged on days 20 - 22 by intranasal introduction of 1 mg ovalbumin suspended in 50 mL sterile PBS while under anesthesia. Ovalbumin lot number SLBK645SV was used for experiments shown in Figure 2.3, Figure 2.4, and Figure 2.5. Ovalbumin lot number SLBQ9036V was used for experiments shown in Figure 2.6. We controlled for this by only directly comparing experimental groups of mice treated with the same lot of ovalbumin. Control mice were neither sensitized nor challenged unless otherwise noted.

Processing of Stool and 16S rRNA Profiling Library Preparation

Human stool was pulverized in a biosafety cabinet with liquid nitrogen using a pestle and mortar and aliquoted into 50 - 100 mg samples and stored at -80° C prior to use. For both human and mouse fecal specimens, crude DNA was extracted using phenol:chloroform:isoamyl alcohol and homogenized with a bead beater using sterilized zirconium and steel beads as previously described [107]. The aqueous layer was then purified with a 96-well QIAGEN PCR Clean up kit and quantitated by measuring the absorbance at 260/280 nm. DNA concentrations were normalized to 5 ng/mL using a broad range Quant-iT™ dsDNA Assay and 10 ng of DNA was used to PCR amplify the V4 16S rRNA region using barcoded primers as previously described [139]. PCR-amplified DNA was pooled to equal concentration and the library purified using

AMPure XP SPRI beads. DNA was sequenced using a MiSeq with 2x250 bp chemistry. All samples had a minimum read-depth of 5000.

Analysis of 16S rRNA data

Fastq files were demultiplexed and binned into amplicon sequence variants (ASVs; Table S6) using DADA2 as previously described[93]. Taxonomic determination of ASV sequences to the lowest possible level was performed with RDP Classifier[140] using a database built to permit species level identification[107] with a minimum bootstrap support of 80%.

ASVs were normalized using total sum scaling. Diversity analysis of 16S data was carried out with vegan (v2.5-7) and phyloseq (v1.28.0) in R. Richness was estimated as the average count of observed taxa after rarefying to 5000 reads using the vegan rarefy function. Sequential PERMANOVA was carried out for 100,000 iterations to achieve a minimum p-value of 10^{-5} using adonis2 in vegan. PERMDISP2 was performed with the betadisper function in vegan. Differential abundance analysis was carried out with DESEQ2 (version 1.24.0) as previously described[141,142]. Random permutation of samples revealed a DESEQ2 positivity rate of 6.7% for asthma and 6.8% for age, consistent with previous reports of DESEQ2 performing with a slightly higher than the expected 5% positivity rate, but less than the positivity rate of real data [143]. Batch effect in 16S data resulting from sequencing run differences was assessed using PERMANOVA and found to contribute significantly to the variance in the data, but did not otherwise affect our results (see Table S2). Co-occurrence was calculated using the cooccur package[144].

IgA-Seq

We performed IgA-seq on mouse fecal samples as previously described [107]. In brief, fecal samples were prepared in reduced PBS, stained with polyclonal goat anti-human IgA (Abcam #ab96998) or goat anti-mouse IgA (Abcam #ab97104) fluorescently labeled with DyLight649. Following antibody staining, samples were washed with PBS and resuspended in 100 mM HEPES and 150 mM NaCl containing a 1:4000 dilution of Syto-BC (Invitrogen). Samples were run and acquired on a BD Aria II maintained in a laminar flow biosafety cabinet. Input, IgA- and IgA+ positive fractions were acquired and sequenced as previously described [145]. To analyze these data we used a previously reported IgA index [107].

Whole Genome Sequencing of BFM04319

Genomic DNA was extracted from *B. fragilis* strain BFM04319 by phenol/chloroform extraction. We then used an adaptation of the Nextera Library Prep kit (Illumina, cat. FC-121-1030/1031) [146] and sequenced on a MiSeq to achieve ~80X coverage of the 5Mbp ETBF genome. Reads were trimmed by quality and adapter content with bbtools (sourceforge.net/projects/bbmap/). Scaffolds were created with SPAdes [147] and annotated with prokka [148]. Our assembly had an N50 of 432688 and an L50 of 5. BFM04319 had an average nucleotide identity of 98.84% to the genome of the *B. fragilis* type strain VPI2553 (NCBI reference sequence: CR626927.1) [149].

Immune Cell Isolation from Tissues

Cells were extracted from tissues as described[93]. Briefly, lungs were minced and incubated in digestion buffer (0.2 U/ml Liberase DL (Roche Applied Sciences) and 0.2 mg/ml DNase (Sigma) in Hank's Buffered salt Solution (without Ca^{2+}/Mg^{2+}) for 25 min at 37°C before

being passed through a 70μm cell strainer[150]. Spleen and lymph nodes were dissociated manually and passed through a 70 μM cell strainer. Red blood cells were removed from lung and spleen samples by treatment with ACK lysis buffer.

Flow Cytometry of Isolated Immune Cells

Data were acquired on a FACSCanto II (BD Biosciences) equipped for the detection of eight fluorescent parameters. The following antibodies were used: PE anti-mouse SiglecF (Clone E50-2240; BD Pharmigen™), FITC anti-mouse CD4 (Clone GK1.5, Biolegend), FITC anti-mouse CD11c (Clone N418, Biolegend), PE anti-mouse CD44 (Clone IM7, BD Pharmigen™), PE anti-mouse IL-17A (Clone TC11-18H0.1, Biolegend), PerCP-Cy™5.5 anti-mouse TCR β chain (Clone H57-597, BD Pharmigen™), PE-Cy™7 anti-mouse CD11b (Clone M1/70, BD Pharmigen™), PE/Cyanine7 anti-mouse CD62L (Clone MEL-14, Biolegend), APC anti-mouse Ly6G (Clone 1A8, Biolegend), APC anti-mouse CD45 (Clone 30-F11, Biolegend), PerCP anti-mouse CD45 (Clone 30-F11, Biolegend), APC/Cyanine7 anti-mouse I-A/I-E (Clone M5/114.15.2, Biolegend), eFluor450 anti-mouse FoxP3 (Clone FJK-16s, eBioscience™), eFluor450 anti-mouse IL-13 (Clone 13A, eBioscience™), Brilliant Violet 421™ anti-mouse F4/80 (Clone BM8, Biolegend), APC/Cyanine7 anti-mouse TCR β chain (Clone H57-597, Biolegend), PE/Cyanine7 anti-mouse IFNγ (Clone XMG1.2, Biolegend), APC anti-mouse TNFα (Clone MP6-XT22, BD Pharmigen™). Intracellular staining of cytokines was conducted as previously described[93]. Briefly, cells were stimulated for 4 h at 37°C with PMA (10ng/mL), ionomycin (200ng/mL), monensin (1:1000), and brefeldin A (1:1000). LIVE/DEAD Fixable Aqua Dead Cell Stain Kit was used to assess cell viability in all panels. Data analysis was

performed using FlowJo version 10 or higher software (Treestar, Ashland, OR). Gating strategies are summarized in Figure 5.5F-H.

Transcriptional Profiling of Mouse Lungs

We isolated total RNA from mouse lungs and performed transcriptomic analysis as previously described[93]. Whole lungs were removed from mice and homogenized in 2 mL of TRIzol reagent. Crude RNA was extracted from 0.3 mLs of homogenized tissue using the QIAgen RNAeasy kit following manufacturer's protocol. Reads were mapped to the mouse genome using bowtie2 (v2.3.4.1)[151], quantified at the gene level using htseq (v 0.9.1)[152], and differentially expressed genes were identified using DESeq2 (v1.24.0)[142]. Functional pathways altered during colonization and/or OSC were identified using gene set enrichment analysis FGSEA (fgsea R package; v1.10.1)[153] with KEGG and GO databases. PERMANOVA was performed using adonis (vegan R package; v2.5-7) and the post hoc test was performed with pairwise.adonis (pairwiseAdonis, v0.3).

Lung RT-qPCR

RNA from flash-frozen, pulverized lung tissue or lung tissue stored in RNALater (Invitrogen cat. AM7021) was extracted by probe homogenization in TRIzol Reagent followed by chloroform phase separation. DNase I (Qiagen) was then used to degrade DNA according to the manufacturer and this reaction was carried onto Qiagen's RNeasy Mini kit for purification of RNA. We then quantified the RNA with a Quanti-iT Ribogreen RNA Assay kit (Invitrogen cat. R11490) and synthesized cDNA with a high-capacity RNA-to-cDNA kit (AB cat. 4387406). RT-qPCR was performed on a Biorad CFX96 Real-Time System using Power SYBR Green PCR

Master Mix (AB 4367659). Primer pairs used in this paper are shown in Table S8 including GAPDH as the reference gene. Samples were run in triplicate and were excluded if the range of raw Ct values for target or reference exceeded 2.

Protein Quantification

Mouse Serum IgE specific to ovalbumin was quantified by sandwich ELISA [93]. Briefly, plates were coated with 10 mg/mL purified ovalbumin overnight at 4˚ C and then blocked with PBS 1x with 1% BSA. Sera were diluted 1:10 and plated alongside purified OVA-specific IgE (Clone 2C6, AbD Serotec) as a standard curve. The plate was incubated for 2 hours at room temperature. Bound IgE was detected using goat anti-mouse IgE-HRP (Clone RME-1, Biolegend). Serum IL-17A was measured as part of a LEGENDPlex Multiplex protein assay (Biolegend) following the manufacturer's protocol.

DNA/RNA Oxidative Damage ELISA

DNA was extracted from the lungs of mice by homogenization and ethanol precipitation or using the QIAGEN Dneasy Blood & Tissue Kit. To estimate pulmonary oxidative stress we measured oxidized guanosine (8- hydroxyguanosine, 8-hydroxy-2'-deoxyguanosine, and 8-hydroxyguanine) from extracted lung DNA by ELISA (Cayman Chemical cat. 589320) following the manufacturer's instructions and analysis template.

Intestinal Permeability Assay

In order to assess intestinal permeability, FITC-Dextran gut-to-serum absorption was measured at the time of sacrifice as previously described [56,154]. Briefly, a baseline blood sample

(150 - 250 uL) was taken from each mouse via facial vein puncture. After one hour of fasting, the mice were orally gavaged with 200 mL of 40 mg/ml FITC-Dextran (4 kDa; Sigma Aldrich) in sterile PBS. Food and water were then withheld for an additional 45 minutes. After three hours mice were sacrificed and blood was collected by cardiac puncture. Serum was separated from blood samples using serum separator tubes according to manufacturer's instructions (BD microtainer). Fluorescence of pre- and post-gavage serum were measured (at a 5-fold dilution in PBS) using an excitation wavelength of 485nm and an emission wavelength of 528nm. A standard curve from 0 to 40 mg/ml FITC-Dextran read on the same plate was used to convert the RFU values to concentration of FITC-Dextran. Finally, pre-gavage serum FITC-Dextran concentrations were subtracted from post-gavage serum FITC-Dextran concentrations to quantify leakage from the gastrointestinal tract to the circulation.

Cytospin

Bronchoalveolar lavage was collected by flushing the mouse lungs using 1 mL of 0.1% bovine serum albumin (BSA) in sterile PBS. $5x10^5$ cells of each sample were loaded onto a slide using a Shandon Cytospin 2. Slides were then methanol fixed and stained with eosin and methylene blue following kit directions (ThermoScientific Shandon KwikDiff Stains). Cells were counted using a bright-field microscope at 400X magnification. 300 non-red blood cells were counted per sample, with careful scanning to ensure no repetition between high power fields.

Screening for bft in Subject Stool Samples

Purified DNA from stool samples containing *Bacteroides fragilis* based on V4-16S rRNA community profiling were normalized to 5 ng/mL. PCR was used to amplify the constant c-terminal region of the *bft* gene (Table S8)[116]. The results of the reaction were assessed using a 2% agarose gel stained with GelRed (Biotium). Previous literature has identified this method suffers from low sensitivity, so we also conducted qPCR on the same samples[117]. qPCR primers were designed based on the *bft* sequence identified in BFM04319 and verified by confirming amplification against purified BFM04319 genome, and a complete fecal community harboring BFM04319 (AO stool gDNA), but not in a community lacking BFM04319 (HO stool gDNA). A single band of the expected size was detected by electrophoresis of BFM04319 genome amplification and AO stool gDNA amplification, but not HO stool gDNA. A dilution series of *B. fragilis* genome was used to demonstrate primers were sufficiently sensitive to detect *bft* from less than $2*10^{-5}$ ng of *B. fragilis* genome. The presence of a band or a Cq value of less than 35 was considered a positive result and indicated the presence of *bft* in patient stool.

Quantitation of bft expression in mouse cecal contents

Mouse cecal content RNA and DNA was extracted using the QIAgen AllPrep PowerFecal DNA/RNA kit as per kit instructions (cat. 80244). Nucleic acids were quantitated and cDNA synthesis was performed on 450 ng of cecal RNA using the Lambda Biotech EasyScriptPlus cDNA Synthesis kit (cat. G236) and a primer specific to *bft* (see Table S8). qPCR was performed as previously described. Copies of *B. fragilis* genome were estimated by comparing cycle number against a dilution series of purified BFM04319 genome and then normalized to the amount of nucleic acid per milligram of cecal content.

Quantitation of calprotectin in human stool samples

Human fecal calprotectin was measured using the Calprotectin ELISA Assay Kit (Eagle Biosciences cat. CAL35-K01) following manufacturer's directions. Between 50 and 100 mg of pulverized human stool was used for each assayed sample.

Statistics

Statistics and analysis were all performed in R Version 3.6.3. Data are presented as mean with SEM. Statistical significance was conducted using an unpaired Wilcoxon test or Kruskal Wallis test with a post hoc Dunn test where appropriate. Adjustment of p-values for multiple hypotheses was performed using Benjamini-Hochberg correction. Boxplots display IQR and whiskers display 1.5*IQR. The following symbols were used to designate significance: *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Mixture Distribution Naïve Bayes' Classifier for 16S Profiling of Asthma vs. Healthy Patients

Input Data:

Raw ASV counts data were normalized to total counts (relative abundance). To be included in the NBC, a taxon had to appear in at least 7 samples. This number was determined by calculating how many samples are required to identify enrichment of a taxa in either the healthy or asthmatic cohort with 95% confidence, based on presence/absence alone (Binomial Test). In total, 392 ASVs between 95 human stool subjects were included in this analysis.

Algorithm:

The algorithm is summarized in Figure 5.8B. By Bayes' theorem, the probability that a microbiome composed of many taxa belongs to an individual with asthma is $P(Asthma|X) = \frac{P(X|Asthma)*P(Asthma)}{P(X)}$, where $P(X|Asthma)$ is the likelihood of a microbiome sample (X) occurring given all the microbiome data from the asthmatic cohort, $P(Asthma)$ is the prior probability of any MARS patient having asthma, and $P(X)$ is the probability of the microbiome data (X) occurring given the entire MARS microbiome dataset. To accommodate the sparsity and zero-inflation inherent to microbiome data, we built our NBC to fit relative abundance data to a mixture distribution. We model P(X|Asthma) and $P(X|Healthy)$ (see Figure 2.2A: pink and purple, respectively) as mixtures of (1) a beta distribution of relative abundance when the taxon is present and (2) a binary distribution when the taxon is not detected,

$$P(X|Class) = \begin{cases} P(Not\ Detected|Class)\ ,\ x = 0 \\ \big(1 - (Not\ Detected|Class)\big)Beta(X, \alpha, \beta|Class),\ x \geq 0 \end{cases}.$$

By modeling the frequency of non-detection separately from relative abundance we increase the sensitivity of our model to learn differences in sparse taxa. The prior probability of a patient having asthma $P(A)$ was determined based on the proportion of patients with asthma used in the training data. The beta distribution for each taxon was fit by Maximum a Posteriori estimation using Newton's Method, given a Gaussian prior. Hyperparameters of the prior distributions were optimized by a grid search. Using mixture distributions generated for each taxon in the model, we constructed a Naïve Bayes' Classifier in R to predict patient asthma status based on microbiome composition. The NBC produces metrics that are useful for further analysis, including a Feature Score which can be described as the log likelihood ratio of a taxon

occurring at a given relative abundance in stool from an asthma donor compared to from a healthy donor. The Feature Score per taxon (i) per sample (j) is $F_{ij} = \log\left(\frac{P(X_i=x_{ij}|Asthma)}{P(X_i=x_{ij}|Healthy)}\right)$.

For each sample, the Feature Scores for all taxa (n) in a sample are summed to calculate the Sample Score, which can be described as the log likelihood of the sample being from the asthma population rather than the healthy. The Sample Score per sample is $S_j = \sum_i^n F_{ij}$. Samples were classified as asthmatic if the Sample Score was positive, and healthy if negative. We identified taxa concordant with the model in any given healthy-asthma dyad by calculating the Pairwise Feature Score between the two samples for each ASV. The Pairwise Feature Score for a taxon in a dyad is then calculated as the difference between the Feature Score for the taxon in the sample with asthma and the Feature Score of the same taxon in the healthy control is $PFS_{ij_Aj_H} = F_{ij_{Asthmatic}} - F_{ij_{Healthy}}$. ASVs in a dyad with a positive pairwise feature score were considered concordant with the model. See examples in Figure 5.2A.

Random Forest:

The random forest model (RF) was created in R using the randomForest package[155] (v4.6-14). All forests included 1000 trees (ntrees=1000) with 30x30 tree sampling with replacement (sampsize=c(30,30)) and were built on the same 392 ASVs used in the NBC.

Model Evaluation:

All AUC and ROC curve values were calculated in R using pROC (v1.16.2). ROC curve and AUC values in Figure 2.2C were calculated based on the sample scores for the NBC and tree classification votes for the RF. Platt-scaling was performed using the glm (family =

binomial(logit)) and predict functions in R. We performed leave-one-out cross-validation (LOOCV) and repeated this process 100 times to estimate an average LOOCV classification rate. The NBC achieved a leave-one-out cross validation (LOOCV) accuracy of 75.8% (Figure 5.8C,D), which is similar to studies of comparable size[40,156], and performed as well as a Random Forest classifier (LOOCV accuracy: 75.1%), another tool commonly used to classify disease-associated microbiome data. Taxa identified as highly discriminatory by the NBC were highly correlated with those found to be important by Random Forest (Figure 5.8A, Table S7, rho = 0.4439, p-value < 0.0001, Spearman Correlation).

2.6 Acknowledgements

This work was co-authored by **Naomi G. Wilson***, **Ariel Hernandez-Leyva***, **Anne L. Rosen, Natalia Jaeger, Ryan T. McDonough, Jesus Santiago-Borges, Michael A. Lint, Thomas R. Rosen, Christopher P. Tomera, Leonard B. Bacharier, S. Joshua Swamidass, and Andrew L. Kau** (*Authors contributed equally)[157]. N.G.W., A.H.-L., and A.L.K. conceptualized the work. L.B.B. and A.L.K. planned the clinical study. N.G.W., A.H.-L., A.L.R., N.J., R.T.M., J.S.-B., M.A.L., C.P.T. and A.L.K. contributed to the design and conduct of experiments. N.G.W., A.H.-L., A.L.R., N.J., R.T.M., and A.L.K. analyzed the data. N.G.W., A.H.-L., T.R.R., and S.J.S. developed gnominator. N.G.W., A.H.-L., and A.L.K. drafted the manuscript. All authors interpreted the data and contributed to revising the manuscript. We would like to thank our clinical study coordinators Tarisa Mantia, Caitlin O'Shaughnessy, and Shannon Rook; the physicians of Washington University Pediatric and Adolescent Ambulatory Research Consortium, especially Dr. Jane Garbutt; the Volunteer for Health registry; and the MARS participants and their families. We would also like to acknowledge Dr. Devesha

Kulkarni, for her technical expertise as well as the Center for Genome Sciences Sequencing Core and Genome Technology Access Center at the McDonnell Genome Institute for sequencing services. The authors also thank Drs. Brain Laidlaw and Leyao Wang for critical reading of the manuscript and Drs Philip Ahern, Neelendu Dey, and Ansel Hsiao for thoughtful feedback on our work. This work was supported by the National Institutes of Health (K08 AI113184 to A.L.K., T32 GM007200 & F30 DK127584 to A.H-L., T32 GM007067 to J.S-B.) and the AAAAI (Foundation Faculty Development Award to A.L.K.).

Chapter 3: The gut metagenome harbors metabolic and antibiotic resistance signatures of moderate-to-severe asthma

Naomi G. Wilson[1], Ariel Hernandez-Leyva[1], Drew J. Schwartz[2], Leonard B. Bacharier[3], Andrew L. Kau[1,*]

[1] Division of Allergy and Immunology, Department of Medicine and Center for Women's Infectious Disease Research, Washington University School of Medicine, St. Louis, MO, 63110, USA

[2] Division of Infectious Diseases, Department of Pediatrics and Center for Women's Infectious Disease Research, Washington University School of Medicine, St. Louis, MO, 63110, USA

[3] Division of Allergy, Immunology and Pulmonary Medicine, Department of Pediatrics, Monroe Carell Jr Children's Hospital at Vanderbilt University Medical Center, Nashville, TN 37232, USA

* Correspondence: akau@wustl.edu

3.1 Abstract

Asthma is a common allergic airway disease that develops in association with the human microbiome early in life. Both the composition and function of the infant gut microbiota have been linked to asthma risk, but functional alterations in the gut microbiota of older patients with established asthma remain an important knowledge gap. Here, we performed whole metagenomic shotgun sequencing of 95 stool samples from 59 healthy and 36 subjects with moderate-to-severe asthma to characterize the metagenomes of gut microbiota in children and adults 6 years and older. Mapping of functional orthologs revealed that asthma contributes to 2.9% of the variation in metagenomic content even when accounting for other important clinical demographics. Differential abundance analysis showed an enrichment of long-chain fatty acid (LCFA) metabolism pathways which have been previously implicated in airway smooth muscle and immune responses in asthma. We also observed increased richness of antibiotic resistance genes (ARGs) in people with asthma. One differentially abundant ARG was a macrolide resistance marker, *ermF*, which significantly co-occurred with the *Bacteroides fragilis* toxin, suggesting a possible relationship between enterotoxigenic *B. fragilis*, antibiotic resistance, and asthma. Lastly, we found multiple virulence factor (VF) and ARG pairs that co-occurred in both cohorts suggesting that virulence and antibiotic resistance traits are co-selected and maintained in the fecal microbiota of people with asthma. Overall, our results show functional alterations via LCFA biosynthetic genes and increases in antibiotic resistance genes in the gut microbiota of subjects with moderate-to-severe asthma and could have implications for asthma management and treatment.

Importance

Asthma is an airway disease that affects the everyday lives of millions of people and

accounts for approximately 1.5 million emergency room visits yearly in the US[158]. Both antibiotic usage and gut microbiota dysbiosis have been linked to the development of asthma, however, little is known about the specific gut microbial functions associated with asthma, particularly in older populations. In this study, we characterize the gut microbiota of school-aged children and adults with moderate-to-severe asthma to uncover asthma-associated microbial functions that may contribute to disease features. We find that people with asthma have an increase in gut microbial genes associated with long-chain fatty acid metabolism as well as an accumulation of antibiotic resistance genes, both of which may have practical consequences for monitoring and treatment of asthma.

3.2 Introduction

Asthma is a common respiratory disease characterized by symptoms of airway obstruction including wheeze, cough, and shortness of breath. In most cases, asthma onsets in early childhood with the development of sensitization to environmental allergens. Ongoing environmental exposures lead to airway inflammation and ultimately result in asthma symptoms manifesting within the first few years of life. Recent findings support the notion that asthma develops in association with the human gut microbiome composition early in life[35,39]. This finding is supported by 16S rRNA sequencing surveys demonstrating that alterations in the gut microbiota precede asthma development within the first few months of life[13,35].

Early childhood gut microbial communities have been proposed to contribute to asthma by several mechanisms. Epoxide hydrolases encoded by enterococci and other gut bacteria produce the lipokine 12,13-diHOME that predisposes towards atopic sensitization and asthma[13,62]. Similarly, short-chain fatty acids (SCFAs), produced by the metabolism of dietary

fibers by diverse members of the gut microbiota, are thought to protect from asthma through their effect on the host G-protein coupled receptor GPR41, shaping immune cell differentiation in the lungs, and ameliorating allergic airway inflammation[8,12,35,159,160].

In addition to microbially-encoded metabolic features, carriage of antibiotic resistance genes (ARGs) within the gut microbiota, termed the resistome, has been associated with asthma risk. In infants, microbial signatures associated with the development of asthma are also associated with increased richness of ARGs in the gut microbiome[69]. These differences in ARG carriage were found to be driven primarily by *E. coli*, which is a common colonizer in the first days of life[69]. These findings are important in understanding the origins of asthma since antibiotic exposure correlates both to the number of ARGs within the gut microbiome[161] and the later development of asthma and other allergic diseases[162–164]. This association between antibiotic exposure and asthma is supported by animal models that found antibiotic treatment worsens allergic airway inflammation (AAI)[165–167].

While there is an abundance of data supporting the idea that asthma susceptibility is associated with features of the gut microbiota in early childhood, the potential effect of gut microbial functions on asthma later in life remains an important knowledge gap. Since asthma often begins in infancy when the gut microbiota composition is highly unstable, disease-causing microbial functions may not persist into older children and adults. Nevertheless, the gut microbiota in older individuals could underlie the variable manifestations of asthma[27] and may hold valuable prognostic and therapeutic significance.

Asthma-associated differences in later childhood and adult gut microbial communities have already been noted in several reports. Studies in preschool-aged children have noted distinct taxonomic composition of gut microbial communities in subjects with asthma compared to healthy controls[39]. These differences are reported to include reductions in *Akkermansia muciniphila*[40], *Faecalibacterium prausnitzii*[41] as well as *Roseburia* species[42]. Functional characterization of microbial communities by whole metagenomic sequencing from an older population of women with asthma[41] has shown that pathways related to lipid and amino acid metabolism, as well as carbohydrate utilization were enriched compared to healthy controls. In contrast, microbial pathways involved in the production of SCFAs, like butyrate, were enriched in the healthy cohort of the same study[41]. These findings are supported by a complementary study designed to test the effect of probiotic supplementation on asthma that found an association of improved asthma symptoms with SCFA biosynthesis as well as tryptophan metabolism pathways in the adult gut microbiota[43].

In this study, we describe an analysis of whole metagenomic sequencing data from a cohort of 36 subjects with physician-diagnosed, moderate-severe asthma along with a matched cohort of 59 healthy controls. We test the hypothesis that the gut metagenome harbors signatures of asthma after the disease has been established. Our results identify global differences in metagenomic functions between the asthma and healthy cohorts and reveal an enrichment in the asthma cohort for long-chain fatty acid biosynthesis pathways. We also find increased richness of ARGs associated with asthma and co-occurrence of ARGs with known bacterial virulence factors, suggesting a potential relationship between antibiotic exposure and pathogen colonization in people with asthma.

3.3 Results

3.3.1 Whole metagenomic shotgun sequencing of fecal samples from adults and children with asthma and healthy controls

We performed whole metagenomic sequencing on fecal samples from subjects with asthma and healthy controls taking part in the Microbiome & Asthma Research Study (MARS), which we have previously described[93,157]. MARS participants were recruited from the St. Louis, Missouri area and included pediatric (6-10 years) and adult (18-40 years) age groups. All asthma cohort patients had a physician diagnosis of moderate-to-severe asthma, and history of allergic sensitization as evidenced by positive skin testing or serum specific-IgE to one or more common aeroallergens. In total, we analyzed 95 patient stool samples including 17 adults and 19 school-aged participants with asthma, and 40 adults and 19 school-aged participants without asthma.

NovaSeq S4 sequencing of our libraries yielded 1.69 billion paired-end reads translating to a total of approximately 500 Gigabases (Gb). After filtering for read quality, dropping host contaminants, and trimming adaptor content, we achieved 1.23 billion paired-end reads and an average 3.4 Gb per stool sample with a range of 0.4-9.9 Gb/sample (Figure 5.9A). Neither host contamination nor sequencing depth differed between asthma and healthy cohorts (t-test p=0.2 and 0.7, Table S9). All samples achieved an estimated average metagenomic coverage of at 89% (range of 61-98%) with the annotation-free redundancy-based metagenome coverage estimator, Nonpareil[168] (Figure 5.9B). Further, estimated metagenome coverage was not different between the asthma and healthy cohorts, although we noted coverage was slightly reduced in the pediatric cohort (Figure SB, Table S9). We also found that the most abundant functional pathways (Figure 5.9C) across all MARS participants are involved in essential processes of gut microbes such as starch degradation and glycolysis, demonstrating that our sequencing captured core functions of

the gut metagenome as expected. Taken together, we concluded that our sequencing is of sufficient depth and quality to be used for further analyses.

3.3.2 Gut taxonomic composition differs between people with and without asthma

We first leveraged the clade marker annotation tool, MetaPhlAn[169], to analyze the taxonomic composition of the study participants. We found dominate genera typical in gut microbiota communities including *Bacteroides* (phylum Bacteroidota) and *Faecalibacterium* (phylum Bacillota) (Figure 5.9D). Simpson alpha diversity was slightly higher in the asthma cohort even when taking read depth and age group into account (Figure 5.9E). Bray-Curtis dissimilarity (Figure 5.9F) was shifted between the asthma and healthy cohorts ($p<0.0004$, $R^2=0.029$) even when accounting for other covariates including age ($p<0.001$, $R^2=0.032$), race ($p=0.0006$, $R^2=0.026$), recent antibiotic usage ($p=0.9$, $R^2=0.006$), read depth ($p=0.2$, $R^2=0.013$), obesity ($p=0.7$, $R^2=0.008$), sex ($p=0.4$, $R^2=0.011$), and tobacco exposure ($p=0.2$, $R^2=0.012$) by sequential PERMANOVA (Figure 5.9G). There was also no significant interaction between asthma status and age group ($p=0.8$, $R^2=0.007$), or between asthma status and recent antibiotic usage ($p=0.6$, $R^2=0.009$) (Figure 5.9G). To determine differentially abundant taxa, we tested the fixed effect of asthma along with the random effects of age group and race in a general linear model[170] and found *Eubacterium rectale* and *Prevotella copri* were enriched in the healthy cohort (Figure 5.9H, Table S10). All of these findings are consistent with 16S rRNA sequencing performed in a previous study[157] which lent us further confidence that our sequencing data was suitable for functional profiling.

3.3.3 Fatty acid metabolism pathways are enriched in the gut metagenomes of people with asthma

Given that our samples had adequate coverage to capture expected taxonomic shifts, we

started interrogating the differences in metagenomic functions of the gut microbiota attributable to asthma status. The alpha diversity of genes (UniRef90 clusters) was neither different between the asthma and healthy cohorts nor between the pediatric and adult cohorts, suggesting that our gene profiling reached a similar total number of genes in both cohorts (Figure 3.1A). Using PERMANOVA, we noted that, even while accounting for significant covariates of age ($p<0.001$, $R^2=0.029$), race ($p<0.001$, $R^2=0.024$), and read depth ($p=0.03$, $R^2=0.015$), asthma status also significantly impacted gut microbiome functional composition ($p=0.008$, $R^2=0.017$; Figure 3.1B, C). We note that age group's interaction term with asthma did not significantly contribute to the variance in beta diversity, suggesting that the influence of asthma and age on beta diversity is non-overlapping. These findings support the idea that the gut metagenomic content of people with asthma is different than that of healthy individuals, even when accounting for other clinical sources of interpersonal gut microbiome variation.

We next considered which metagenomic functions and metabolic pathways may be involved in the differences between asthma and healthy cohorts. We first examined a list of specific metagenomic functions previously implicated in asthma, including genes related to histamine production, 12-13 diHOME biosynthesis, and tryptophan metabolism, but we were unable to identify a difference between cohorts (Figure 5.10A). To identify pathways that differed between asthma and healthy subjects, we performed a Wilcoxon Rank Sum test with a false discovery rate $q<0.2$ on the relative abundance of all pathways annotated by the MetaCyc database that were above 10% prevalence within the population. Using these criteria, we found seven pathways that were enriched in asthma and one that was enriched in the healthy cohort out of 312 total pathways (Figure 3.1D). To determine if these findings were robust to other analysis

methods, we performed additional differential abundance approaches on the 312 MetaCyc pathways, including a Wilcoxon test on centered log-transformed counts and ALDEX2, both of which demonstrated that these pathways differed between healthy and asthma cohorts (See Table S11). All differentially abundant pathways enriched in patients with asthma were involved in fatty acid synthesis, and included the production of oleate, palmitoleate, (5Z)-dodecenoate, 8-amino-7-oxononanoate, biotin, and octanoyl acyl-carrier protein, as well as saturated fatty acid elongation. In the healthy cohort, only a single L-lysine biosynthesis pathway was enriched.

Using taxonomically tiered functional mapping, we determined which taxa were driving the observed differences in asthma-associated pathways. For the L-lysine biosynthesis III pathway which was more abundant in healthy subject, we found that it primarily originated from *Blautia obeum*, Figure 5.10B). In the case of the asthma-enriched pathways, we found that *Bacteroides vulgatus* and *Alistipes finegoldii* account for the largest fraction of complete fatty acid biosynthesis pathways (Figure 3.1E, Figures S3C). However, the differential abundance of these asthma-associated pathways was probably not due solely to an enrichment of *B. vulgatus* or *A. finegoldii* in asthma stool since neither species was differentially abundant (maaslin2 q-value=0.58 and 0.25, respectively; See Table S10). Further, the majority of mapped pathways were not attributable to any single species and these unmapped pathway counts made up more of the overall pathway richness than *B. vulgatus* (Wilcoxon q values < 0.05 for all seven pathways; see "Community" stratification in Figure 5.10C). Taken together, these findings indicate that the differences may be either driven by community-level effort (i.e. distinct steps of the pathway are encoded across more than one species), or that current databases are insufficiently granular to identify the key taxa responsible for these differences.

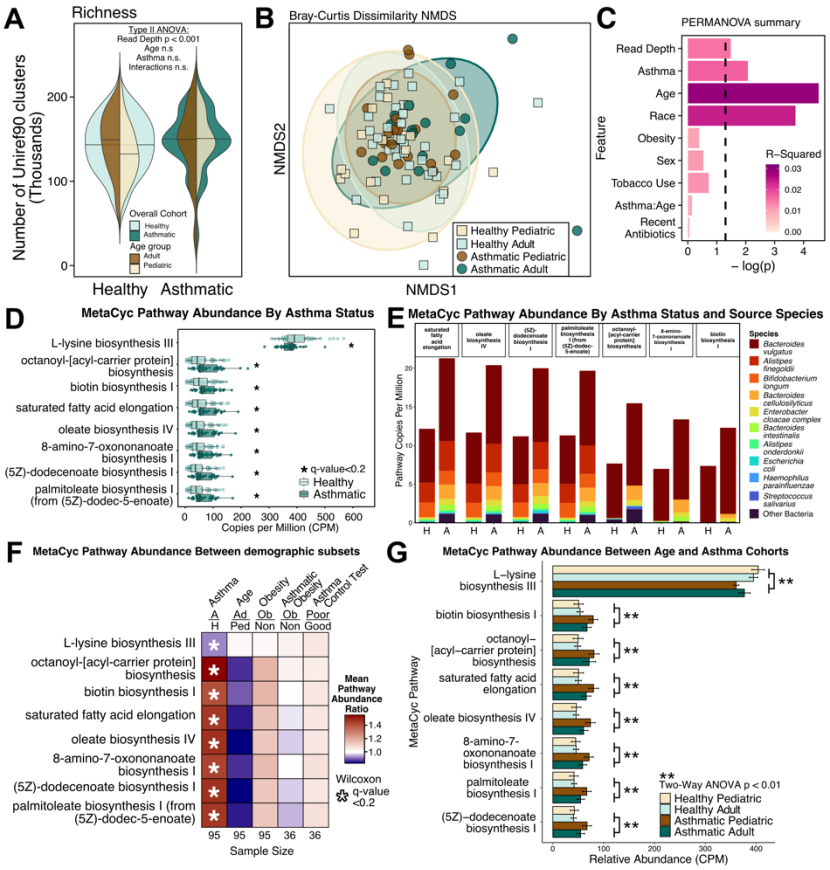

Figure 3.1. Gut metagenomes from individuals with asthma show increased genes encoding fatty acid metabolism.

A) Stacked violin plots of Uniref90 cluster richness (unique Uniref90 cluster with CPM>0) grouped by either healthy and asthma cohort (blue green colors in background) or age (brown colors in foreground. **B)** Non-metric multidimensional scaling plot of Bray-Curtis Dissimilarity distance between Uniref90 (copies per million) profiles. Axis 1 and 2 of five total are shown of an NMDS with stress value 0.09. **C)** Sequential PERMANOVA of Bray-Curtis dissimilarities between Uniref90 profiles. Input order of terms to the test is identical to the order of the barplot from top to bottom. **D)** Relative abundance of MetaCyc pathways that were differentially abundant given a Wilcoxon q value below 0.2 (p-value after FDR correction). **E)** Stacked bar plot of differentially abundant fatty acid metabolism pathways mapped to respective taxa by MetaPhlAn3.0/HUMAnN3.0, averaged within asthma or healthy cohorts. **F)** Heatmap of MetaCyc pathway abundance ratios between groups in important clinical demographics: Asthma vs. Healthy, Adult vs. Pediatric, Obese vs Non-Obese, and Well-Controlled Asthma vs. Poorly-Controlled Asthma. Asterisk denotes a significant differential abundance (*q<0.2) according to Wilcoxon tests controlled for multiple comparison testing within each demographic category. **G)** Differentially abundant MetaCyc pathways plotted as four cohorts: asthma by age with respective Two-Way ANOVAs. Only statistically significant p values shown.

We reviewed the enzymatic steps of each of the eight pathways represented in Figure 3.1D and found that, of the 78 total reactions in these pathways, only 11 reactions were shared between 2 pathways (Figure 5.11). The 8-amino-7-oxononanoate biosynthesis I pathway consists

of the first 11 reactions of the larger biotin biosynthesis pathway and the latter only has four additional reaction steps past synthesizing 8-amino-7-oxonanoate to produce biotin. Additionally, the (5Z)-dodecenoate pathway can feed directly into the palmitoleate biosynthesis pathway, and that the octanoyl acyl carrier protein pathway shares an upstream substrate (acetoacetyl-acyl carrier protein) with the saturated fatty acid elongation pathway (Figure 5.11). Together, our findings indicate that long chain fatty acid biosynthesis is differentially abundant in the asthma gut metagenome via related but largely non-redundant pathways.

Given the association between obesity with fatty acid metabolism[171] as well as asthma[172–174], we next wanted to determine whether obesity (which we define here as a BMI greater than 30 in adults or a BMI-for-age percentile of greater than 95% in children) confounds the association of microbial fatty acid metabolism with asthma. We compared the abundance of the differentially abundant fatty acid pathways between all non-obese and obese patients and found no significant difference (Figure 3.1F). Within the asthma cohort, there was similarly no statistically significant difference between the patients with and without obesity, suggesting that obesity is not a confounder for the difference we observed in fatty acid metabolism. To determine whether fatty acid metabolism is related to the intensity of asthma symptoms and their effect on everyday life activities, we utilized a validated survey of asthma control (The Asthma Control Test; ACT)[119]. None of the fatty acid pathways were differentially abundant between patients with well-controlled and poorly-controlled asthma (Figure 3.1F). We tested if age group affects the differentially abundant metabolic pathways and found that these pathways were not differentially abundant between age groups alone (Figure 3.1F). We also tested the impact of asthma and age as independent variables to differentially abundant metabolic pathways using a

Two-way ANOVA. We found that, even while taking age into account, these pathways are differentially abundant between asthma and healthy cohorts, but are not different by age or an interaction between asthma and age (Figure 3.1G, 2-Way ANOVA). Given that the effect of asthma status on differentially abundant metagenomic functions was distinct from that of age, we primarily focused our subsequent analyses on the asthma and healthy cohorts overall, combining age groups.

3.3.4 Richness of antibiotic resistance genes is increased in the gut metagenomes of people with asthma

Since people with asthma tend to be prescribed antibiotics frequently[175] and oral antibiotic exposure is a risk factor for the acquisition of ARGs in the gut[161], we wanted to determine if the members of our asthma cohort were more likely to have received antibiotics. To test this, we counted how many subjects had taken a course of antibiotics within one year of their participation in the study. As part of the study design, participants could not take antibiotics in the month prior to fecal donation. We found that a greater proportion of the asthma cohort received antibiotics in the past year compared to that of healthy participants (42% of asthma cohort versus 15% of the healthy cohort, Fisher's test, $p=0.011$, Figure 3.2A). This finding represents evidence of increased antibiotic exposure amongst subjects with asthma in our study.

We next sought to characterize the gut antibiotic resistome in the asthma and healthy cohorts. To test if the increased antibiotic exposure in the asthma cohort was reflected in the gut resistome, we utilized the ShortBRED pipeline[176] to detect reads mapped to the Comprehensive Antibiotic Resistance Database (CARD)[177]. We first asked whether there were more ARGs in our asthma cohort by summarizing our dataset into richness (Total number of unique ARGs

detected

A

MARS Cohort	Antibiotics in the past year		Fisher's Exact test
	YES	NO	p
Asthmatic	15	21	0.011
Healthy	9	48	

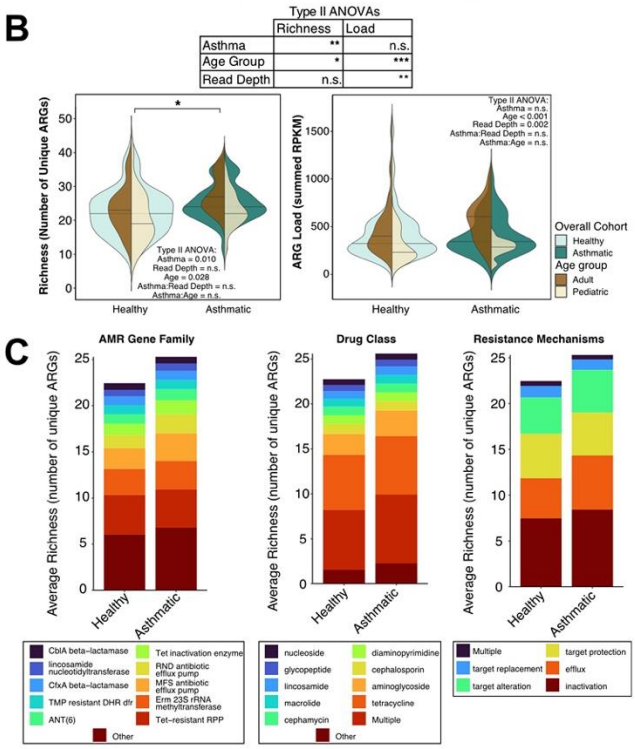

Figure 3.2. Gut metagenomes from individuals with asthma harbor an increased richness of antibiotic resistance genes.
A) Table describing short-term antibiotic usage in the MARS cohorts. **B)** Overlapping violin plots of ARG richness and load by grouped by either healthy and asthma cohort (blue green colors in background) or age (brown colors in foreground. **C)** Stacked bar plots of average ARG richness painted by antimicrobial family (AMR), drug class to which the ARG confers resistance, and ARG resistance mechanism.

per sample) and load (Total sum of ARG RPKM per sample). We found that ARG richness was higher in people with asthma even when accounting for differences due to age (p=0.03) and sequencing depth (p=0.09 while ARG load was not different between asthma and healthy cohorts (p=0.4) when accounting for age (p<0.001) and read depth (0.002) (Figure 3.2B). We note that *E coli* was not differentially abundant between asthma and healthy cohorts (p=0.52, Table S10), so the richness increase we observe in the asthma cohort is not due solely to an increase in *E. coli*

relative abundance. These results suggest that there are a higher number of unique ARGs, or a higher diversity, in asthma compared to healthy controls.

From our 95 stool samples, we detected 71 unique ARGs, comprising 32 antimicrobial resistance families, 29 drug classes, and 7 mechanisms of resistance, with 26 ARGs (37% of the total) conferring multi-drug resistance (Figure 3.2C). Similar to previous studies of gut resistomes, we found that tetracycline resistance markers were the most commonly detected ARGs and inactivation is the most common mechanism of resistance followed by efflux pumps[69] (Figure 3.2C). Using the abundance data of each detected ARG, we determined that asthma ($p=0.005$, $R^2=0.028$) and age ($p<0.001$, $R^2=0.053$) were the strongest factors contributing to the variance in ARG beta diversity even when accounting for important technical and demographic covariates (Figure 3.3A and Figure 3.3B). We next wanted to ascertain to what degree the

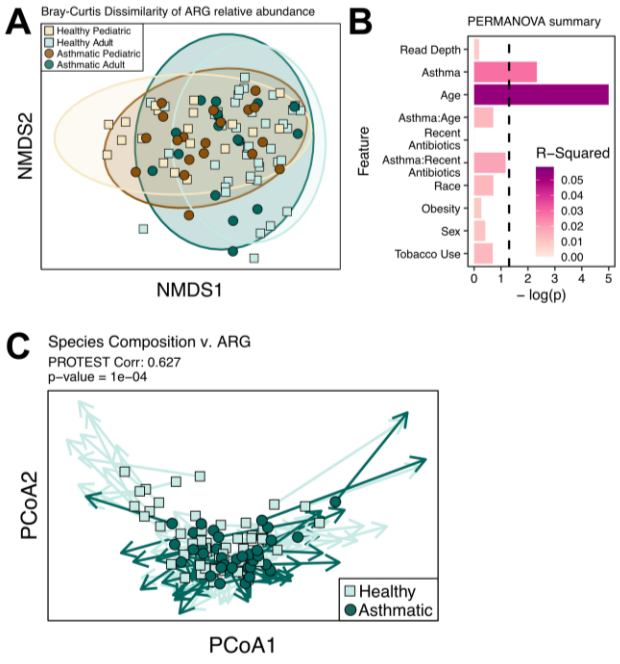

Figure 3.3. The gut antibiotic resistome is altered in asthma patients.
A) Non-metric Multidimensional Scaling (NMDS) plot of antibiotic resistome with units in Bray-Curtis dissimilarity of total-sum scaled RPKM, labeled by asthma and age cohorts. Showing two axes out of five with stress value=0.1. **B)** Effect of demographic categories on antibiotic resistome data in A (sequential PERMANOVA). **C)** Procrustes and PROTEST analysis between MetaPhlAn species-level Bray-Curtis dissimilarity distances and

CARD ShortBRED Bray-Curtis dissimilarity distances. Arrows connect the two data points belonging to identical samples.

resistome profile was determined by microbial composition. We used a Procrustes analysis[178] to compare compositional data generated from MetaPhlAn[169] to the antibiotic resistome profile derived from ShortBRED and found that the microbiome composition correlated to the resistome profile (Figure 3.3C, PROTEST corr = 0.627, p-value < 0.0001), indicating that ARG profiles are directly related to bacterial species composition.

3.3.5 Macrolide resistance markers are differentially abundant in asthma

To determine gut-associated ARGs that are differentially abundant between patients with and without asthma, we applied negative binomial tests to the abundance of all ARGs detected in at least 7 samples. This prevalence cutoff was chose because it is the minimum number of samples needed to detect a difference using a negative binomial distribution. We found that genes encoding resistance to macrolides (*ermF, ermB* and *ermA*), vancomycin (*vanRO*), tetracycline (*tet(45)*), as well as multi-drug efflux pumps (*smeB, mdtO*, and *oqxA*) were enriched in the asthma cohort (Figure 3.4A, Table S12). Prominent amongst these was the 23S rRNA methyltransferase *ermF*, which is typically encoded by *Bacteroides* species and confers resistance to macrolides.

Next, we explored the genomic context of *ermF* by assembling metagenomic sequencing reads into contigs with metaSPAdes[147] and annotating open reading frames with Prokka[148] and BLAST. We detected full-length *ermF* with 98% or higher identity in 53 out of 95 samples. Out of 53 contigs, the vast majority originated from members of the Bacteroidota, 75.4% originated from the *Bacteroides* genus and 60.3% of them were likely from *B. fragilis* based on the top BLAST homology. Of the contigs that encoded *ermF,* 68% occurred on scaffolds with at least

one other open reading frame within ten kilobases (Figure 3.4B). We found that many *ermF genes* are co-located with genes associated with mobile genetic elements such as transposases, mobilization genes, and toxin/antitoxin systems, as well as with other ARGs like *btgA* which encodes clindamycin resistance (Figure 3.4B,C). This indicates that *ermF* occurs in multiple different genomic contexts within our cohort and suggests that its presence is not strictly due to propagation of a single *B. fragilis* strain.

3.3.6 People with asthma have a distinct set of co-existing pairs of antibiotic resistance genes and virulence factors in the gut metagenome

In our prior work on this same cohort of patients, we found that, compared to healthy subjects, a greater portion of asthma subjects were colonized with *B. fragilis* strains harboring the virulence factor *B. fragilis* toxin (*bft*), which we showed has the potential to shape inflammation in the lung[157]. Given that our resistome analysis pointed to an enrichment of a *B. fragilis* ARG, we wanted to test whether the *ermF* gene is associated with *bft* in the asthma cohort. We found that metagenomes harboring both *ermF* and *bft* were more prevalent in individuals with asthma compared to those without (Figure 3.4D). In our MARS samples, we did not find any instances where *bft* and *ermF* occurred on the same scaffold, so it remains unclear whether these two genes are encoded within the same *B. fragilis* strain or within two separate strains. Nevertheless, the enrichment of *ermF* and *bft* in adults and older children with asthma could suggest that the intestinal habitat of individuals with asthma presents opportunities or niches for macrolide ARGs and virulence factors such as *bft*.

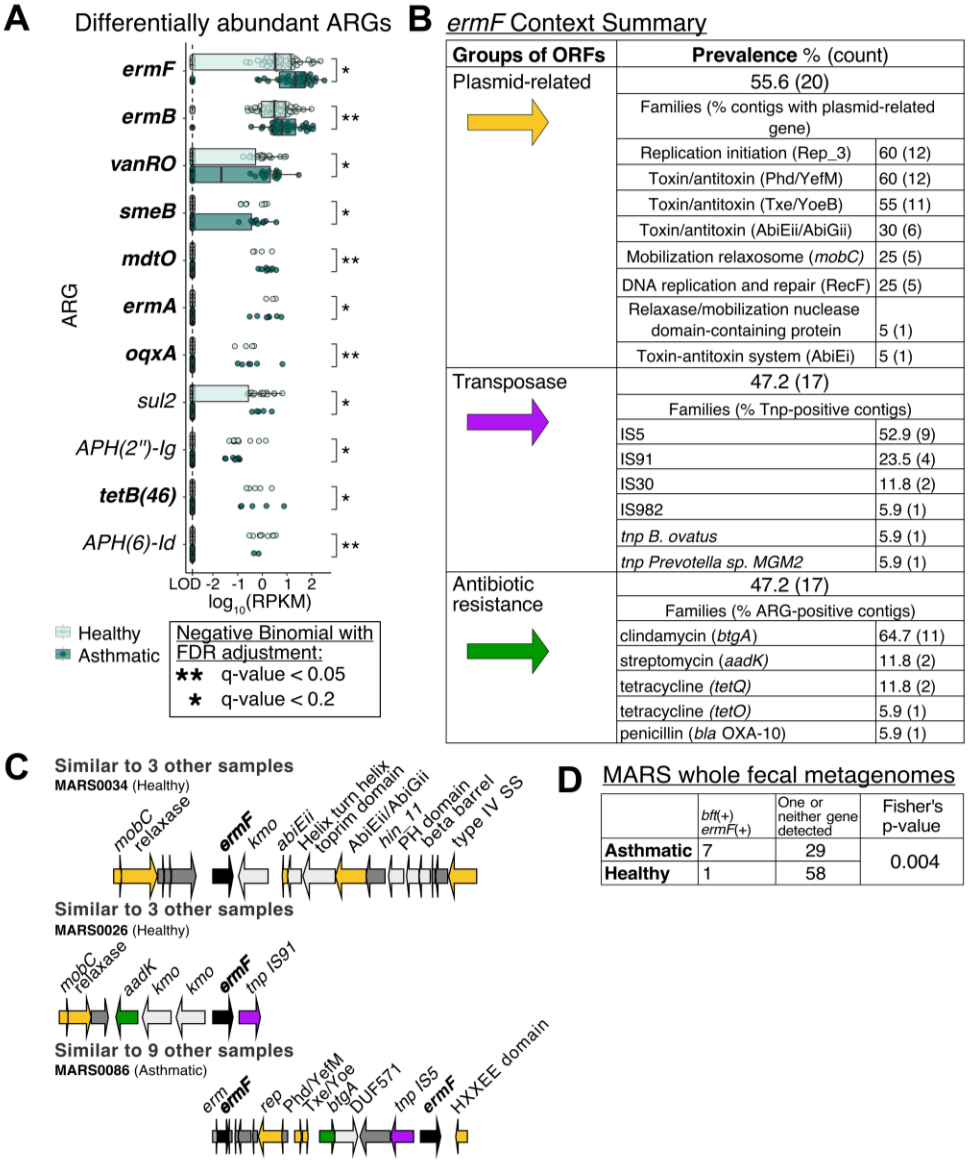

Figure 3.4. Resistance gene ermF is differentially abundant in diverse genomic contexts of gut resistomes belonging to individuals with asthma.

A) Boxplots of antibiotic resistance gene (ARG) abundance by cohort on log-scale. Showing only ARGs present in at least 7 out of 95 samples and have q-values less than 0.2. A pseudocount of 0.0015 RPKM (designated as the limit of detection "LOD") was used for the negative binomial tests. Bolded genes are enriched in the asthma cohort while non-bolded are enriched in the healthy cohort. B) Summary of ermF contexts on contigs from metagenomic assemblies that had at least one detectable open reading frame flanking the ermF within 10 kilobases. C) Three representative ermF context maps generated in GeneSpy. D) Count table of fecal metagenomes with co-detection of bft+ and ermF+ vs. detection of one or neither of ermF and bft, split by donor asthma status. Fisher's Exact two-sided p-value shown.

To explore the possibility that virulence traits and ARGs are linked in the gut microbiota, we characterized virulence factor (VF) content of all samples using the Virulence Factor

Database[179] and compared these data to the antibiotic resistome profiles. We did not find the same overall shift in the virulence factor beta diversity between asthma and healthy that we observed with the resistomes (Figure 5.12A-C), but we did find differentially abundant VFs belonging to capsule and peritrichous flagella VF families (Table S13, q values<0.2). Further, we found that microbiota composition is highly correlated with virulence factor profile (Figure 5.12D, Protest correlation coefficient=0.61, p<0.0001). Given that microbiota composition strongly affects both VF and ARG content, we used a partial correlation between VF and ARG richness to test our hypothesis while removing the effect of total metagenomic content. We found a positive partial correlation between VF and ARG richness in both the asthma and healthy cohorts (Figure 3.5A). Similarly, virulence factor and resistome beta diversity profiles were also positively correlated (Figure 3.5B, Protest correlation coefficient=0.574, p=1e-4). Together, our results suggest that these two microbial features, virulence and antibiotic resistance, are closely linked within the gut metagenome.

We next performed a co-occurrence analysis to uncover other linked virulence and antibiotic resistance traits that could be important in gut ecology. We found numerous co-occurring VF-ARG pairs in MARS gut metagenomes (Figure 3.5C, p<0.05). Several of these positively co-occurring pairs were shared between the two cohorts (yellow), suggesting that these relationships are not dependent on asthma status. In contrast, many pairs specifically co-occur in one cohort and may indicate microbial interactions important in asthma but not healthy gut metagenomes (Figure 3.5C). In summary, we found that VF and ARG presence is linked in the gut metagenome and that people with asthma have a distinct set of co-occurring functions compared to healthy people.

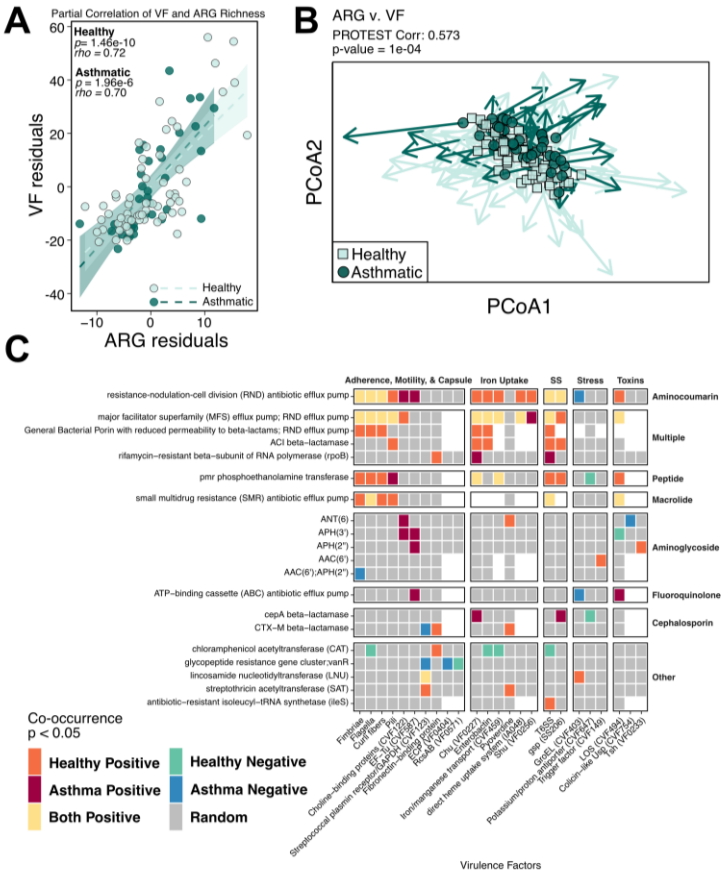

Figure 3.5. Asthma patients have unique sets of virulence factor and antibiotic resistance gene associations. **A)** Partial correlations split by asthma status between virulence factor richness and ARG richness after accounting for species richness. **B)** Procrustes and PROTEST analysis between Bray-Curtis dissimilarity distances of virulence factors and CARD resistomes. Arrows connect the two data points belonging to identical samples. **C)** Heatmap of statistically significant (cooccur R package $p<0.05$) co-occurrence relationships between all VFs and ARGs. Colors indicate direction of co-occurrence and in which cohort(s) the respective effect was detected. Grey squares mark pairs with no statistically significant co-occurrence. White squares were pairs filtered out due to a lack of observed co-occurrence.

While our co-occurrence analysis between VFs and ARGs demonstrated multiple examples of virulence and antibiotic resistance traits found in the same gut metagenome, this analysis does not indicate if these genes are present in a single organism. To obtain a more granular view of VF-ARG co-occurrence, we limited our analysis to look for VF-ARG pairs that could be encoded by the same species. This analysis showed that the asthma cohort had a greater number of ARGs (p=0.007 and 0.01) and VFs (p=0.005 and 0.09) annotated as coming from *Klebsiella pneumoniae* and *Escherichia coli*, respectively (Figure 5.13A). Individual co-

occurrences attributable to each of these species are summarized in Figure 5.13B and show that *cepA*, encoding a beta-lactamase, and *chuU*, a VF involved in iron acquisition, are both putatively encoded by *E. coli* and co-occur in patients with asthma, suggesting that the metagenome-wide co-occurrence of CepA and Chu families observed in Figure 3.5C may be due to enrichment within one or more *E. coli* strains harboring these VF/ARG pairs. Together, our co-occurrence analyses show that there appear to be multiple co-occurring VFs and ARGs, similar to *B. fragilis*-encoded *bft* and *ermF*, in the gut metagenome and within putative individual species that could be important for asthma. The cohort-specific co-occurring VF-ARG pairs found here could serve as candidates for future studies of asthma gut microbiome ecology.

3.4 Discussion

In this study, we present an exploratory analysis of fecal whole metagenomic sequencing contrasting subjects with moderate-to-severe asthma to a group of healthy controls to identify disease-associated microbial genes with the strongest likelihood of affecting disease. Our sequencing and subsequent analyses revealed that the functional content of individuals with asthma differed significantly from that of healthy controls. We found an enrichment of functions associated with saturated and mono-unsaturated fatty acids, including oleate, palmitoleate, 5(Z)-dodecenoate, biotin, 8-amino-oxononanoate, saturated fatty acid elongation, and octanoyl acyl carrier protein pathways. Currently, the functional significance of gut bacterial synthesis of these long-chain fatty acids (LCFA) to asthma has not been well defined. Excess LCFAs, usually studied in the context of dietary fat intake, have been associated with metabolic diseases including diabetes, obesity, and atherosclerosis risk[173] but is also linked to asthma risk in adults[172–174,180]. Increasing recognition that obesity predisposes to asthma has motivated

investigation of the impact of fatty acids on airway biology and has shown that LCFA signaling through free fatty acid receptor 1 (FFAR1, also called GPR40) induces airway smooth muscle cell contraction and proliferation, both of which are important components of asthma pathophysiology[173,181]. Notably, a study that sequenced airway microbes in children with cystic fibrosis implicated a similar list of LCFA production pathways during exacerbations, suggesting that microbially produced LCFAs may influence airway physiology[182]. To our knowledge, the potential for gut microbes to contribute to the amount of free fatty acids available to the lung has not yet been defined, however, LCFAs are readily absorbed into the circulation[183] and could plausibly reach the airways. Further, previous studies have shown the effect of SCFA (e.g. acetate, butyrate, propionate) produced by gut microbes to directly alter lung inflammation via GPR41 (FFAR3)[8,160]. While our study did not find a direct enrichment of SCFA production pathways in the healthy cohort as has been previously reported[41], we did observe that lysine biosynthesis was enriched. Since lysine may serve as a precursor to the SCFA butyrate[184], SCFAs may still be more abundant in our healthy cohort but may be subject to transcriptional regulation that would not be detected by metagenomic DNA sequencing. Together, our metabolic pathway analyses of the gut metagenome demonstrate a positive association between LCFAs produced by gut microbes and asthma, in contrast to the negatively associated SCFAs.

In addition to metabolic alterations, analysis of the gut resistome demonstrated that subjects with asthma had a distinct ARG composition. In a recently published prospective gut metagenomic study of infants, asthma-associated taxonomic signatures were associated with a higher number of ARGs[69]. These differences in the resistome were largely driven by a single species of bacteria, *E. coli*, and reveals that acquisition of ARGs in subjects with asthma may

begin in early childhood and could affect asthma development. In our study of older subjects with established asthma, we similarly found a higher richness of ARGs that is associated with asthma in both school-aged children and adults, supporting the idea that increased ARG carriage may persist in people with asthma throughout life. Based on our resistome annotation, however, ARGs in our cohort were likely from a diverse assemblage of bacteria in contrast to what was observed in infants. This is likely due to differences in gut dynamics between age groups. The infant microbiome is heavily shaped by limited available niches in the developing gut, which favor transient, facultative anaerobes like *E. coli* [69], whereas the gut resistome in older subjects reflects selective pressures experienced over a lifetime. One important consequence of increased richness of ARGs in people with asthma is that it may promote persistence of some bacterial strains[185,186] and contribute to the taxonomic differences in the gut microbiota between asthma and healthy people[39,157].

While asthma was among the important factors accounting for a significant amount of the variance in ARG beta diversity, we found that recent antibiotic exposure (within the past year) was not. Notably, no participant in our cohort received a course of antibiotics in the month prior to fecal sampling since this could have confounded our analyses on asthma-associated microbial community changes. Previous studies have shown that the gut microbiota recovers in approximately a month after perturbation from antibiotics in healthy adults[66]. We interpret these findings to mean that recent exposure (within 1 - 12 months) to antibiotics does not drastically change the resistome, whereas repeated exposures over time may be more important for driving the population-wide shifts we observed in our cohort[186].

Of the ARGs found to be enriched within asthma resistomes, the ARG *ermF*, encoding

resistance to macrolide antibiotics, was especially prominent amongst the cohort with asthma. While we did not collect data on the antibiotic drug classes, number of courses and their duration, or the reason for prescription of antibiotics, our subjects received, it is likely that our asthma population has been exposed to macrolides. Macrolide antibiotics, including clarithromycin and azithromycin, are commonly prescribed for upper and lower airway infections which disproportionately affect people with asthma[70]. This class of antibiotics, particularly azithromycin, have been a focus of special concern for driving antibiotic resistance due to their frequent usage and pharmacological properties[187–189]. Nevertheless, azithromycin has been noted to have beneficial effects in asthma, and some[72], but not all[74], studies suggest that azithromycin may prevent exacerbations in patients with asthma. Given the interest in azithromycin as a treatment modality in asthma, there will be an urgent need for additional studies to determine the robustness of the association between asthma and macrolide ARG accumulation in the gut to inform parameters for antibiotic selection and prescription in people with asthma.

Additional exploration of the gut metagenomes revealed potential co-selection in people with asthma for *B. fragilis* genes *ermF* and *bft* (*B. fragilis* toxin), the latter of which is more prevalent in fecal samples from the asthma compared to healthy cohort[157]. Untargeted analysis of gut resistomes revealed multiple examples of virulence factor and ARG co-occurrence as well as positive correlations between ARG and VF richness in people with and without asthma**.** Our findings are consistent with previous reports that found correlations between VFs and ARG richness and VF-ARG cooccurrence relationships in both gut metagenomes[190] and human-associated bacterial genomes[191]. Our findings also add to these studies by demonstrating that,

while the correlation between VF and ARG richness does not appear to be any stronger in the asthma cohort after taking gene richness into account, the two MARS cohorts do not have identical sets of statistically significant co-occurring VF-ARG pairs. These data suggest that people with asthma may be experiencing different selection pressures from that of healthy people, leading to accumulation of a distinct set of virulence and antibiotic determinants. Given that antibiotics induce gut inflammation through the disruption of the gut microbiota[192], and strains encoding virulence factors such as *bft* are known to thrive in an inflammatory environment[193], one plausible model for the apparent accumulation of distinct VF-ARG pairs is that antibiotic treatment not only selects for ARGs[161,186], but simultaneously selects for VFs. Together with evidence that virulence determinants, such as *bft,* are associated with airway inflammation[157], our model implies that heightened antibiotic treatment may contribute to the manifestations of asthma via co-selection for VFs and ARGs. Considering that prenatal and early life antibiotic exposure is linked to asthma risk[163,192], this model could be used to test whether the initial events driving VF and ARG co-occurrence start with the first vertical transmission events in very early life.

Our study has several limitations that constrain the scope of our claims. First, MARS is an exploratory, cross-sectional study with only a moderate number of subjects, which is less ideal for identifying disease-associated microbiome differences[98]. As a result, our study had limited statistical power to detect less prevalent or abundant functions. Second, our study focused on school-aged and older subjects with moderate-to-severe asthma, and thus our findings may not be applicable to other younger populations or those with less severe disease. These population differences may explain why we were unable to identify statistically significant

differences in microbial metabolic pathways identified from other studies including bile acid metabolism[35], epoxide hydrolases[62], histamine metabolism[44,45], or tryptophan metabolism[194,195] (Figure 5.10A). Third, the factors driving the shift in gut bacterial metabolism to LCFA biosynthesis and whether gut microbiome enrichment of this pathway is sufficient to change the hosts' LCFA profile is not known. Collecting blood to interrogate host metabolism as well as dietary information at the time of fecal sample collection would have helped to disentangle the effects of diet on host and gut microbiota metabolism. Fourth, we lacked relevant subject information, such as diet, environment, infrastructure, stress level, and social relationships, needed to precisely disentangle the effects of social, environmental, and health disparities on the gut microbiome[196]. We recognize that our finding of subject-reported race as a statistically significant covariate in our analyses of the gut metagenome likely does not reflect a direct effect of race on biology[197]. Rather, we interpret this finding as a proxy for the biological consequences of active systemic disparities associated with race[197]. We included race in our models to account, albeit inadequately, for the impact that multi-faceted ecosocial factors underlying race are known to have on asthma and the microbiome[198,199]. Fifth, a record of the frequency and class of antibiotics administered to our participants would have allowed us to confirm whether macrolide administration associates with the enrichment of *ermF* in our asthma cohort and whether a higher diversity of antibiotic usage correlates with ARG richness. It is likely that antibiotic exposures accumulated throughout life contribute to the resistome, and a complete catalog of exposures is critical to determine patterns of antibiotic prescription most likely to account for the ARG associations to asthma found in this study. Lastly, as with all metagenomic sequencing studies, we are limited by annotation bias in existing databases. This is a concern for our virulence factor

and antibiotic resistance profiling especially, where we rely on the database to predict source species for ARGs and VFs. We also recognize that the databases we used for these two analyses are biased towards well-studied human pathogens rather than commensals or opportunistic pathogens. However, we note that other investigators have reported similar co-occurrence of ARGs and VFs[190,191], and co-selection of these features is biologically plausible.

Despite these constraints on the scope of our study, we provide evidence that there is an increased production of LCFA and an increased richness of ARGs encoded by the gut microbiota in people with asthma. These findings could have applications in the care of patients with asthma. If LCFA pathways are shown to play a causal role in airway inflammation in future studies, microbiota-directed therapeutics in the form of dietary interventions or probiotics, could be developed to modify gut microbial metabolism to protect against asthma. Additionally, our resistome findings add to the growing concern over antibiotic resistance in patients with asthma by suggesting that antibiotic administration may also contribute to gut carriage of virulence factors that can alter airway inflammation. Ultimately, our study shows that the gut microbiota of school-aged and older subjects with moderate-to-severe asthma harbor important functional alterations that could serve as a foundation for future studies investigating how gut microbial functions affect pulmonary diseases.

3.5 Materials and Methods

MARS Study Population

The Microbiome and Asthma Research Study (MARS) consisted of 104 subjects from the St Louis, MO USA area that are either healthy or had physician-diagnosed moderate-to-severe asthma. This study included an adult cohort (ages 18-40 years) and pediatric cohort (ages 6-10 years). As described in previous manuscripts[93,157], 9 patients were disqualified or did not donate

stool samples. The remaining 95 patients donated stool samples either at home or at the recruitment visit and were evaluated with a clinical questionnaire to gather relevant metadata. Stool samples were kept at -20°C and delivered within 24 hours to the study site, Kau Lab at Washington University School of Medicine, where they were stored at -80°C for no more than three years until processing for DNA isolation. This study was approved by the Washington University Institutional Review Board (IRB# 201412035). Written informed consent documents were obtained from all MARS subjects or their legal guardians. All recruitment, follow up, and sample acquisition occurred between November 2015 and December 2017.

Fecal DNA Isolation

Frozen human stool samples were pulverized in liquid nitrogen using a pestle and mortar. We then homogenized the stool in a mixture of phenol, chloroform, and isoamyl alcohol with a bead beater using sterilized zirconium and steel beads as previously described[107] to extract crude DNA. We purified the fecal DNA with a 96-well QIAGEN PCR Clean up kit and quantitated by measuring the absorbance at 260/280 nm. Sample DNA concentrations were normalized to 0.5 ng/mL. Neither depletion of human DNA sequence nor enrichment of microbial or viral DNA was performed. No experimental quantification like a spike-in were used.

Whole Metagenomic Sequencing of Fecal Communities

To generate fecal metagenomic sequencing data, we adapter-ligated libraries by tagmentation using an adaptation of the Nextera Library Prep kit (Illumina, cat. No. FC-121-1030/1031)[146]. Individual libraries were then purified with AMPure XP SPRI beads, quantitated using Quant-iT (Invitrogen, cat. Q33130), and then combined in an equimolar ratio. We confirmed that each library was adequately represented in the combined library by preliminary sequencing on a MiSeq instrument at the Washington University in St. Louis Center for Genome

Sciences to assess the evenness of the library. Once the quality of the library was assured, we sequenced the combined library on a NovaSeq 6000 S4 with 2x150 bp chemistry to achieve an average of 3.4 Giga-base-pairs (Gb) per sample. NovaSeq services and data demultiplexing were performed by the Genome Technology Access Center at the McDonnell Genome Institute (St Louis, MO). All samples were tagmented simultaneously and sequenced on the same run to avoid batch effects.

Processing of sequencing data

Metagenomic raw demultiplexed reads were then processed to (1) remove spurious human sequences (human reference database was hg37dec_v0.1.1), (2) remove low quality sequences, and (3) trim remaining adapter content using Kneaddata v. 0.10.0 (huttenhower.sph.harvard.edu/kneaddata) bypassing the tandem repeat finder step ("- -bypass-trf"). FastQC (fastqc v0.11.7) and MultiQC (multiqc v1.2) with default settings were used to create quality reports and visualize processing steps. See Figure 5.9A and Table S9 for number of reads dropped per processing step. After trimming and filtering, no samples had adaptor content, overrepresented sequences, or an average sequence quality score below Phred 24. Estimated metagenome coverage was calculated with Nonpareil[168,200] (version 3.4.1) via the online querying tool at http://enve-omics.ce.gatech.edu/nonpareil/submit.

Read-based metagenome profiling

To obtain functional information about the metagenomic contents of fecal samples, we processed samples using HUMAnN[169] v3.0.0 on filtered reads with default parameters. The marker gene database used by HUMAnN to identify taxonomic identities was ChocoPhlAn v201901b and the protein database used by HUMAnN to identify functions was the UniRef90 full database v201901b. Alpha diversity analysis of Uniref90 genes and two-sample tests of

KEGG orthologs were performed on respective genes that were present (>0 copies per million) in at least 16 out of 95 samples, which was the lowest prevalence cutoff that would allow for Bonferroni corrected Wilcoxon p-values below 0.0001. HUMAnN was used to determine the abundance of metagenomic pathways by mapping UniRef90 genes to the MetaCyc database. We performed differential abundance analysis using the Wilcoxon 2-sample tests on pathways that had a minimum of 10% prevalence.

To identify antibiotic resistance genes present in the fecal metagenomes of MARS stools, we used ShortBRED-identify[176] (v0.9.4) with the Comprehensive Antibiotic Resistance Database[177] (downloaded 2021-07-05 16:10:04.04555) and Virulence Factor Database[179] (downloaded Fri Jul 16 10:06:01 2021). ShortBRED-Quantify was run on the filtered reads with default parameters. ARGs or VFs that had an abundance greater than zero in less than 7 out of 95 samples were excluded from downstream analyses. This prevalence cutoff was determined using the binomial distribution to maintain a 95% confidence that enrichment was not due to random chance (using stats::binom in R). In the analyses that compared virulence factor profiles to antibiotic resistance gene profiles, any gene with the same name was excluded from the list of antibiotic resistance and considered a virulence factor only, to prevent spurious results due to co-correlations. Only one gene matched this criterion: *ugd* (UDP-glucose 6-dehydrogenase).

Microbial composition was determined with MetaPhlAn 3.0 which is included in the HUMAnN pipeline described. MaasLin[170] (Maaslin2_1.5.1) was used in R to find taxa of any taxonomic level that correlated with asthma by setting asthma as a fixed effect and setting age group and race as random effects.

For PERMANOVA analyses, BMI class refers to two stratifications: Non-obese (underweight, healthy, or overweight) and obese determined for adults by BMI cutoffs and for pediatric patients by BMI-for-age percentile as defined by the Centers for Disease Control and Prevention (see cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html). Self-reported race was used as a co-variate in our PERMANOVAs since precise variables that would better describe the many facets of racism and health disparities were absent (see Limitations). Subject-reported race was represented in our models as a dichotomous variable of either "Caucasian" or "Non-Caucasian", with 92% of the latter population having reported as "Black or African-American" and the remaining 8% as "Other". We chose to combine self-reported "Black or African-American" and "Other" populations into a single category of "Non-Caucasian" because there were insufficient numbers of reported "Other" to power a robust analysis of this group but we still wanted to account for potential health disparities associated with non-Caucasian races[201].

Metagenome Assemblies

Filtered reads were assembled into contigs using spades[147] (v3.14.0) with the "meta" flag and k-mers lengths as follows: -k 21,33,55,77. The resulting scaffolds achieved an average N50 of 3525 +/- 178 bp, an average L50 of 7192 +/- 372 and an average total length of 136.8 +/- 4.5 Mbp as measured by QUAST (v 4.5) [202,203] (see Table S9). Prokka (v1.14.5) was used to find open reading frames and annotate them, and manual BLAST was used to annotate "hypothetical protein" open reading frames for the contexts of *ermF*.

Statistics and Reproducibility

R version 3.6.3 was used for all analyses downstream of HUMAnN and ShortBRED, and for data visualization. Wilcoxon tests with false discovery rate multiple testing correction or

Type II ANOVAs were used to determine statistically significant differences with the car::Anova package in R. PERMANOVAs were performed in R using the vegan::adonis package with default settings and 100,000 iterations. The following symbols were used to designate significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and the following for q values (FDR-adjusted p-values): * $q < 0.2$, ** $q < 0.05$.

Data Availability

Whole metagenomic sequencing data without host contamination, low quality reads, and adapter sequences are available at European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) under project accession number PRJEB56741. Demographic data can be found in this manuscript (Table S9) and previous articles about the MARS study[93,157]. A STORMS (Strengthening The Organizing and Reporting of Microbiome Studies) checklist[204] is available at doi: 10.5281/zenodo.7492635.

3.6 Acknowledgements

This work was co-authored by **Naomi G. Wilson, Ariel Hernandez-Leyva, Drew J. Schwartz, Leonard B. Bacharier, and Andrew L. Kau** and is *in review*[205]. N.G.W. and A.L.K. conceptualized the work. L.B.B. and A.L.K. planned the clinical study. N.G.W., A.H-L., D.J.S., and A.L.K. analyzed the data and drafted the manuscript. All authors interpreted the data and contributed to revising the manuscript. We would like to thank our clinical study coordinators Tarisa Mantia, Caitlin O'Shaughnessy, and Shannon Rook; the physicians of Washington University Pediatric and Adolescent Ambulatory Research Consortium, especially Dr. Jane Garbutt; the Volunteer for Health registry; and the MARS participants and their families. We would also like to acknowledge the Center for Genome Sciences Sequencing Core and Genome Technology Access Center at the McDonnell Genome Institute for sequencing services. The

authors also thank Dr. Leyao Wang and Master Anne Rosen for critical reading of the manuscript. This work was supported by the National Institutes of Health (K08 AI113184 to A.L.K., T32 GM007200 & F30 DK127584 to A.H-L.) and the AAAAI (Foundation Faculty Development Award to A.L.K.).

Chapter 4: General Conclusions and Future Directions

4.1 Features of the gut microbiota could be therapeutic targets for asthma

While several studies report that the gut microbiota is involved in the development of asthma and in shaping lung immune responses[206], little is currently known about the effect of the gut microbiota on established asthma. This thesis work characterized gut microbiota samples from adults and school-aged children with moderate-to-severe asthma and found an asthma-specific shift in the taxonomic and metagenomic profiles of the gut microbiota even after diagnosis. This shift cannot be attributed to age group, recent antibiotic usage, race, obesity, or tobacco usage. Parallel to studies of the early life gut microbiota, these results suggest the gut microbiota later in life may directly affect the trajectory of asthma. It also likely reflects exposure to asthma medication and the heightened mucosal inflammation within patients with asthma that healthy individuals do not experience. Importantly, this finding emphasizes that research into the human microbiota for asthma treatment is worthwhile and potentially life-changing. Future clinical studies of longitudinal nature with large sample sizes and narrow asthma severity exclusion criteria similar to this thesis work are necessary to confidently confirm that the shift in the gut microbiota found here is generalizable to the global population.

Lung oxidative stress and Th17 responses have been linked to asthma endotypes in human populations[77,78,207,208] and increased gut permeability has been observed in children and adults with asthma[57,58]. In this thesis work, lung oxidative stress, gut barrier permeability, and T-helper 17 related responses in gnotobiotic mice were caused by one but not all tested asthma gut microbiota samples. Further, lung oxidative stress and gut permeability were at least partially caused by the presence of toxin-producing strain of the commensal bacterium, *Bacteroides fragilis*. This thesis work provokes the question of whether clinically recognized asthma

endotypes can be associated with particular gut community features such as the presence of a toxin or virulence factor. Such a question could be answered with a longitudinal study designed to characterize endotypes of controllable and uncontrolled asthma alongside the gut microbiota. The ultimate goal of this study would be to gut microbiota states associated with patients that experience less resistance to existing treatment. With these insights, one could design a probiotic or prebiotic to cause an existing community to mimic the non-resistant population's gut microbiota. Then, ideally, the gut microbiota of patients with previously highly resistant, uncontrolled asthma could be reshaped towards controllable asthma and manage their disease with existing treatments.

Enterotoxigenic *B. fragilis* (ETBF) are well-known to wreak havoc on the gut barrier by cleaving E-cadherin[110], one of the proteins that keep the gut barrier from allowing harmful molecules and organisms into the circulation. While increased gut permeability has been linked to asthma and metabolic diseases[209], little is currently known about whether ETBF play a role in this association. Here, a human isolate of ETBF from an adult with asthma was sufficient to cause gut permeability and oxidative stress in the lungs of mice undergoing allergic airway inflammation. Further, this thesis work found that samples containing ETBF from additional donors with asthma neither caused the same level of *bft* expression nor resulted in the increase in *Il17* expression or oxidative stress in the lungs as was caused by the first donor sample. This suggests that the AAI-ETBF phenotype is 1) ETBF strain-dependent, 2) community-context dependent, or 3) both. Future investigations will perform monocolonization experiments that include multiple human stool isolates of ETBF to assess if the phenotype is persistently induced between strains. To establish a stronger link to human health, many more human stool samples

from people with and without moderate-to-severe asthma that do and do not contain ETBF should be tested in humanized gnotobiotic experiments.

Gnotobiotic mouse models are powerful tools to assess the *potential* for human microbiota to affect human disease. However, the transfer of fascinating insights from mouse models to humans is a colossal task for the laboratory. Although this thesis work is built on experiments performed meticulously in line with recommendations from the field[98], it has only begun to link its findings back to human populations. With the resources available, the presence of ETBF and a marker of gut permeability in stool was found to be more prevalent in the asthma compared to healthy cohort. This result was limited both by small sample sizes and human sample depletion. Optimal translation of the ETBF-AAI phenotype to humans will require a large, longitudinal clinical study with strict asthma severity exclusion criteria, that collects appropriate gut permeability and lung oxidative stress measurements, blood samples for immune cell panels, bronchoalveolar lavages or brushings, and stool. As discussed at length among human microbiome researchers[98], large, longitudinal studies are best for the statistical power needed to discover mechanisms that will be actionable for human gut-directed therapies.

4.2 Antibiotics may select for both resistance genes and immunomodulatory virulence factors

Another highlight of this thesis that has implications for asthma patients is that of the gut antibiotic resistome. Concerns about antibiotic resistance in the asthma population have become more widespread in the wake of studies assessing the efficacy of macrolides for asthma management[70–75] as well as studies revealing the increased risk of airway infections for individuals with asthma[70]. Whole metagenomics sequencing revealed an increased richness of

antibiotic resistance genes in the gut microbiota of the asthma cohort. Further, macrolide resistance markers were differentially abundant in the asthma cohort and one of these genes tended to co-occur with the gene encoding *B. fragilis* toxin (*bft*) in asthma gut microbiota. A follow-up analysis suggested that several antibiotic resistance genes co-occurred with virulence factors in the gut and some of these co-occurring pairs were unique to the asthma cohort. Together, these findings suggest that the resistome differs between those with and without asthma. Although this study lacked information about specific drug class prescribed to the participants, there was a statistically significant association of asthma status and recent antibiotic usage (at least one prescription in the past 1 to 12 months). Considering this, it is reasonable to hypothesize that the signature of macrolide resistance observed in the gut may be a result of the increased need for macrolides in the asthma population. The co-occurrence of a ARGs with virulence factors is alarming and warns of a possible co-selection of toxins that can affect airway inflammation and resistance to dependable medications (see Figure 4.1). Future studies

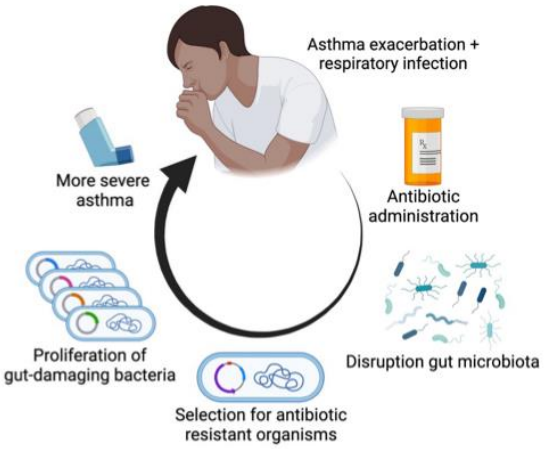

Figure 4.1. Hypothesized model of bft and macrolide resistance co-selection in patients with asthma
Created with BioRender.com

ought to sample stool from patients with moderate-to-severe asthma and controls over time while keeping detailed records of the participants' antibiotic usage: drug class, reason for prescription,

duration of treatment, multiple stool collections, and exacerbation frequency. Such studies could answer whether antibiotic usage in patients with asthma affects the gut resistome more than a given healthy person and whether antibiotic treatment is followed by virulence factor accumulation, vice versa, or if these are randomly co-selected. Further, one could follow how the resistome and the virulence factor profile change in response to perturbations caused by antibiotics and exacerbations. These studies would inform when is best to intervene with a potential gut-directed therapy or how to prescribe antibiotics for the asthma population in a way that will not select for virulence factors that might worsen symptoms.

4.3 Two understudied potential mechanisms of the gut-lung axis in asthma

This thesis work suggests a handful of promising avenues for mechanistic studies of the gut-lung axis in asthma. The finding that ETBF colonization in particular community contexts can alter lung inflammation and gut permeability, suggests that *B. fragilis* toxin is key to this avenue of the gut-lung axis (Chapter 2). One prevailing hypothesis for how this mechanism occurs is that the *B. fragilis* toxin reaches the gut lining, cleaves E-cadherin, disrupts gut barrier homeostasis, and thus allows the bloodstream and lymph system greater access to microbial products from the gut. The crosstalk of the immune components with microbial products then causes an amplified inflammatory response (i.e. oxidative stress and IL-17 production) upon sensitization and challenge with an allergen in the lungs. Animal models of sepsis have previously shown that circulating lipopolysaccharide – a microbial product – causes oxidative stress in the lungs[112]. Further, oxidative stress favors Th17 responses in asthma models[210] and one study found that blocking IL-17 decreases inflammation after induction of oxidative stress and lung injury with LPS[211]. Preliminary data generated during the course of this thesis work

could neither confirm nor deny that higher concentrations of microbial products were circulating in asthma recipient mice undergoing AAI compared to healthy recipient mice undergoing AAI, unfortunately. Focused experiments designed to test this hypothesis are needed to validate the proposed events, and measurement of circulating lipopolysaccharides or other microbial products may be illuminating for human clinical studies of asthma and the gut microbiota.

The metagenomics analysis outlined in Chapter 3 offers a detailed description of the gut microbiota from adult and school-aged patients with asthma. Although this report was entirely *in silico*, some of the signatures of asthma revealed here could be indicative of gut-lung axis mechanisms. A prominent difference found between the healthy and asthma cohorts was that of fatty acid metabolism pathway abundance. An enrichment of genes encoding enzymes that synthesize long-chain fatty acids (LCFAs), especially mono-unsaturated fatty acids, was an unexpected result associated with the asthma cohort. However, mono-unsaturated LCFAs have been linked to asthma in several studies. Palmitoleate, for example, affects macrophage polarization[212], and is negatively associated with upper airway neutrophil counts[174] and lung function[213] in male adults with asthma. Additionally, oleate is linked to asthma risk in adults via dietary intake[180]. However, whether microbial LCFA production contributes to or modifies asthma is not yet determined. Host lipid profiles are known to be affected by the gut microbiota and are important in metabolic syndrome, which includes asthma under its umbrella, as well as inflammation in general[214]. The finding in Chapter 2 could be indicative of diet-related differences between the asthma and healthy cohorts that were not tested by the MARS study. In this case the fatty acid metabolism signature may be an artifact of the substrates being made available to gut microbes via the diet. But if these differences are not due to dietary intake, they

could be accumulated in response to systemic mucosal inflammation thought to occur in those with asthma. Perhaps organisms of the microbiota have evolved to produce LCFA end products that create protective niches and/or prevent adverse effects from immune cells. Regardless, in light of existing literature, the metagenomic signature found here suggests that the gut microbiota makes LCFA in moderate-to-severe asthma that can then alter lung inflammation. A clinical asthma study with stool and serum lipidomics alongside gut metagenomic sequencing would further describe the implications of this finding.

References

1. Olin, A. *et al.* Stereotypic Immune System Development in Newborn Children. *Cell* **174**, 1277-1292.e14 (2018).

2. Donaldson, G. P. *et al.* Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science (80-.).* **360**, 795–800 (2018).

3. Brüssow, H. Turning the inside out: the microbiology of atopic dermatitis. *Environ. Microbiol.* **18**, 2089–2102 (2016).

4. Baviera, G. *et al.* Microbiota in healthy skin and in atopic eczema. *Biomed Res. Int.* **2014**, (2014).

5. Nakamura, Y. *et al.* Staphylococcus δ-toxin induces allergic skin disease by activating mast cells. *Nature* **503**, 397–401 (2013).

6. Allakhverdi, Z. *et al.* Thymic stromal lymphopoietin is released by human epithelial cells in response to microbes, trauma, or inflammation and potently activates mast cells. *J. Exp. Med.* **204**, 253–258 (2007).

7. Ong, P. Y. *et al.* Endogenous Antimicrobial Peptides and Skin Infections in Atopic Dermatitis. *N. Engl. J. Med.* **347**, 1151–1160 (2002).

8. Trompette, A. *et al.* Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat. Med.* **20**, 159–166 (2014).

9. Huang, Y. J. *et al.* The airway microbiome in patients with severe asthma: Associations with disease features and severity. *J. Allergy Clin. Immunol.* **136**, 874–884 (2015).

10. Taylor, S. L. *et al.* Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology. *J Allergy Clin Immunol* **141**, 94–103 (2018).

11. Alnahas, S. *et al.* IL-17 and TNF-α are key mediators of Moraxella catarrhalis triggered exacerbation of allergic airway inflammation. *Front. Immunol.* **8**, 1–11 (2017).

12. Cait, A. *et al.* Microbiome-driven allergic lung inflammation is ameliorated by short-chain fatty acids. *Mucosal Immunol.* **11**, 785–795 (2018).

13. Fujimura, K. E. *et al.* Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat. Med.* **22**, 1187–1191 (2016).

14. Smith, P. M. *et al.* The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science (80-.).* **341**, 569–574 (2013).

15. van Meel, E. R., Jaddoe, V. W. V., Bønnelykke, K., de Jongste, J. C. & Duijts, L. The role of respiratory tract infections and the microbiome in the development of asthma: A narrative review. *Pediatr. Pulmonol.* **52**, 1363–1370 (2017).

16. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32 (2015).

17. Dickson, R. P., Erb-Downward, J. R., Martinez, F. J. & Huffnagle, G. B. The Microbiome and the Respiratory Tract. *Annu. Rev. Physiol.* **78**, 481–504 (2016).

18. Bisgaard, H. *et al.* Childhood Asthma after Bacterial Colonization of the Airway in Neonates. *N Engl J Med* **357**, 1487–1495 (2007).

19. Biesbroek, G. *et al.* Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. *Am. J. Respir. Crit. Care Med.* **190**, 1283–1292 (2014).

20. Teo, S. M. *et al.* The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* **17**, 704–715 (2015).

21. Teo, S. M. *et al.* Airway Microbiota Dynamics Uncover a Critical Window for Interplay of Pathogenic Bacteria and Allergy in Childhood Respiratory Disease. *Cell Host Microbe* **24**, 341-352.e5 (2018).

22. Charlson, E. S. *et al.* Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am. J. Respir. Crit. Care Med.* **184**, 957–963 (2011).

23. Durack, J. *et al.* Bacterial biogeography of adult airways in atopic asthma. *Microbiome* **6**, 1–16 (2018).

24. Hilty, M. *et al.* Disordered microbial communities in asthmatic airways. *PLoS One* **5**, (2010).

25. Einarsson, G. G. *et al.* Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax* **71**, 795–803 (2016).

26. Durack, J. *et al.* Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. *J. Allergy Clin. Immunol.* **140**, 63–75 (2017).

27. Wenzel, S. E. Asthma phenotypes: The evolution from clinical to molecular approaches. *Nat. Med.* **18**, 716–725 (2012).

28. Mccann, J. R., Mason, S. N., Auten, R. L., Geme, W. S. & Seed, C. Early-Life Intranasal Colonization with Nontypeable Haemophilus influenzae Exacerbates Juvenile Airway Disease in Mice. *Infect. Immun.* **84**, 2022–2030 (2016).

29. Gollwitzer, E. S. *et al.* Lung microbiota promotes tolerance to allergens in neonates via PD-L1. *Nat. Med.* **20**, 642–647 (2014).

30. Berera, K. *et al.* Gut microbiota from multiple sclerosis patients enables spontaneous autoimmune encephalomyelitis in mice. *PNAS* **114**, 10719–10724 (2017).

31. Cekanaviciute, E. *et al.* Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. *Proc. Natl. Acad. Sci.* **114**, 201711235 (2017).

32. Wu, H. J. *et al.* Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* **32**, 815–827 (2010).

33. Bradley, C. P. *et al.* Segmented Filamentous Bacteria Provoke Lung Autoimmunity by Inducing Gut-Lung Axis Th17 Cells Expressing Dual TCRs. *Cell Host Microbe* **22**, 697-704.e4 (2017).

34. Olszak, T. *et al.* Microbial Exposure During Early Life Has Persistent Effects on Natural Killer T Cell Function. *Science (80-.).* **336**, 489–494 (2012).

35. Arrieta, M.-C. *et al.* Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med* **7**, 307–152 (2015).

36. Wilson, N. G., Hernandez-Leyva, A. & Kau, A. L. The ABCs of wheeze: Asthma and bacterial communities. *PLoS Pathog.* **15**, (2019).

37. Kuruvilla, M. E., Lee, F. E. H. & Lee, G. B. Understanding Asthma Phenotypes, Endotypes, and Mechanisms of Disease. *Clin. Rev. Allergy Immunol.* **56**, 219 (2019).

38. Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* **562**, 583–588 (2018).

39. Hufnagl, K., Pali-Schöll, I., Roth-Walter, F. & Jensen-Jarolim, E. Dysbiosis of the gut and lung microbiome has a role in asthma. *Semin. Immunopathol.* **42**, 75–93 (2020).

40. Michalovich, D. *et al.* Obesity and disease severity magnify disturbed microbiome-immune interactions in asthma patients. *Nat. Commun.* **10**, (2019).

41. Wang, Q. *et al.* A metagenome-wide association study of gut microbiota in asthma in UK adults. *BMC Microbiol.* **18**, 1–7 (2018).

42. Chiu, C. Y. *et al.* Gut microbial dysbiosis is associated with allergen-specific IgE responses in young children with airway allergies. *World Allergy Organ. J.* **12**, 100021 (2019).

43. Liu, A. *et al.* Adjunctive Probiotics Alleviates Asthmatic Symptoms via Modulating the Gut Microbiome and Serum Metabolome. *Microbiol. Spectr.* **9**, 1–17 (2021).

44. Barcik, W. *et al.* Histamine-secreting microbes are increased in the gut of adult asthma patients. *J. Allergy Clin. Immunol.* **138**, 1491-1494.e7 (2016).

45. Barcik, W. *et al.* Bacterial secretion of histamine within the gut influences immune responses within the lung. *Allergy Eur. J. Allergy Clin. Immunol.* **74**, 899–909 (2019).

46. Bischoff, S. C. *et al.* Intestinal permeability - a new target for disease prevention and therapy. *BMC Gastroenterol.* **14**, 1–25 (2014).

47. Damms-Machado, A. *et al.* Gut permeability is related to body weight, fatty liver disease, and insulin resistance in obese individuals undergoing weight reduction. *Am. J. Clin. Nutr.* **105**, 127–135 (2017).

48. Dupont, C. *et al.* Food Induced Alterations of Intestinal Permeability in Children with Cow's Milk-Sensitive Enteropathy and Atopic Dermatitis. *J. Pediatr. Gastroenterol. Nutr.* 459–465 (1989).

49. Ventura, M. T. *et al.* Intestinal permeability in patients with adverse reactions to food. *Dig. Liver Dis.* **38**, 732–736 (2006).

50. Perrier, C. & Corthésy, B. Gut permeability and food allergies. *Clin. Exp. Allergy* **41**, 20–28 (2011).

51. Chelakkot, C., Ghim, J. & Ryu, S. H. Mechanisms regulating intestinal barrier integrity and its pathological implications. *Exp. Mol. Med.* **50**, (2018).

52. Hayes, C. L. *et al.* Commensal microbiota induces colonic barrier structure and functions that contribute to homeostasis. *Sci. Rep.* **8**, 1–14 (2018).

53. Hiippala, K. *et al.* The potential of gut commensals in reinforcing intestinal barrier function and alleviating inflammation. *Nutrients* **10**, (2018).

54. Yu, L. C.-H. Host-microbial interactions and regulation of intestinal epithelial barrier function: From physiology to pathology. *World J. Gastrointest. Pathophysiol.* **3**, 27 (2012).

55. Yu, L. C. H. Microbiota dysbiosis and barrier dysfunction in inflammatory bowel disease and colorectal cancers: exploring a common ground hypothesis. *J. Biomed. Sci.* **25**, 79 (2018).

56. Thevaranjan, N. *et al.* Age-Associated Microbial Dysbiosis Promotes Intestinal Permeability, Systemic Inflammation, and Macrophage Dysfunction. *Cell Host Microbe* **21**, 455-466.e4 (2017).

57. Benard, A. *et al.* Increased intestinal permeability in bronchial asthma. *J. Allergy Clin. Immunol.* **97**, 1173–1178 (1996).

58. Hijazi, Z. *et al.* Intestinal permeability is increased in bronchial asthma. *Arch. Dis. Child.* **89**, 227–229 (2004).

59. Thio, C. L. P., Chi, P. Y., Lai, A. C. Y. & Chang, Y. J. Regulation of type 2 innate

lymphoid cell–dependent airway hyperreactivity by butyrate. *J. Allergy Clin. Immunol.* **142**, 1867-1883.e12 (2018).

60. Bordin, M., D'Atri, F., Guillemot, L. & Citi, S. Histone deacetylase inhibitors up-regulate the expression of tight junction proteins. *Mol. Cancer Res.* **2**, 692–701 (2004).

61. Peng, L., Li, Z.-R., Green, R. S., Holzman, I. R. & Lin, J. Butyrate Enhances the Intestinal Barrier by Facilitating Tight Junction Assembly via Activation of AMP-Activated Protein Kinase in Caco-2 Cell Monolayers. *J. Nutr.* **139**, 1619–1625 (2009).

62. Levan, S. R. *et al.* Elevated faecal 12,13-diHOME concentration in neonates at high risk for asthma is produced by gut bacteria and impedes immune tolerance. *Nat. Microbiol.* **4**, 1851–1861 (2019).

63. Arrieta, M. C. *et al.* A humanized microbiota mouse model of ovalbumin-induced lung inflammation. *Gut Microbes* **7**, 1–11 (2016).

64. Clemente, J. C. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).

65. Anthony, W. E. *et al.* Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults. *Cell Rep.* **39**, 110649 (2022).

66. Palleja, A. *et al.* Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat. Microbiol.* **3**, 1255–1265 (2018).

67. Russell, S. L. *et al.* Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO Rep.* **13**, 440–447 (2012).

68. Johnson, C. C. *et al.* Antibiotic exposure in early infancy and risk for childhood atopy. *J. Allergy Clin. Immunol.* **115**, 1218–1224 (2005).

69. Li, X. *et al.* The infant gut resistome associates with E. coli, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition. *Cell Host Microbe* 1–13 (2021) doi:10.1016/j.chom.2021.03.017.

70. Juhn, Y. J. Risks for infection in patients with asthma (or other atopic conditions): Is asthma more than a chronic airway disease? *J. Allergy Clin. Immunol.* **134**, 247-257.e3 (2014).

71. Gibson, P. G. *et al.* Efficacy of azithromycin in severe asthma from the AMAZES randomised trial. *ERJ Open Res.* **5**, 00056–02019 (2019).

72. Gibson, P. G. *et al.* Effect of azithromycin on asthma exacerbations and quality of life in adults with persistent uncontrolled asthma (AMAZES): a randomised, double-blind, placebo-controlled trial. *Lancet* **390**, 659–668 (2017).

73. Johnston, S. L. *et al.* Azithromycin for Acute Exacerbations of Asthma: The AZALEA

Randomized Clinical Trial. *JAMA Intern. Med.* **176**, 1630–1637 (2016).

74. Brusselle, G. G. *et al.* Azithromycin for prevention of exacerbations in severe asthma (AZISAST): a multicentre randomised double-blind placebo-controlled trial. *Thorax* **68**, 322–329 (2013).

75. Taylor, S. L. *et al.* Long-Term Azithromycin Reduces Haemophilus influenzae and Increases Antibiotic Resistance in Severe Asthma. *Am. J. Respir. Crit. Care Med.* **200**, 309–317 (2019).

76. Feng, Y. *et al.* Antibiotics induced intestinal tight junction barrier dysfunction is associated with microbiota dysbiosis, activated NLRP3 inflammasome and autophagy. *PLoS One* **14**, 1–19 (2019).

77. Kaur, R. & Chupp, G. Phenotypes and endotypes of adult asthma: Moving toward precision medicine. *J. Allergy Clin. Immunol.* **144**, 1–12 (2019).

78. Nadif, R. *et al.* Endotypes identified by cluster analysis in asthmatics and non-Asthmatics and their clinical characteristics at follow-up: The case-control EGEA study. *BMJ Open Respir. Res.* **7**, 1–12 (2020).

79. Cho, Y. S. & Moon, H. B. The role of oxidative stress in the pathogenesis of asthma. *Allergy. Asthma Immunol. Res.* **2**, 183–187 (2010).

80. Camargo, L. do N. *et al.* Effects of Anti-IL-17 on Inflammation, Remodeling, and Oxidative Stress in an Experimental Model of Asthma Exacerbated by LPS. *Front. Immunol.* **8**, (2018).

81. Garcia-Nuñez, M. *et al.* Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *J. Clin. Microbiol.* **52**, 4217–4223 (2014).

82. McAleer, J. P. *et al.* Pulmonary Th17 Antifungal Immunity Is Regulated by the Gut Microbiome. *J. Immunol.* **197**, 97–107 (2016).

83. Fagundes, C. T. *et al.* Transient TLR Activation Restores Inflammatory Response and Ability To Control Pulmonary Bacterial Infection in Germfree Mice. *J. Immunol.* **188**, 1411–1420 (2012).

84. Gauguet, S. *et al.* Intestinal microbiota of mice influences resistance to Staphylococcus aureus pneumonia. *Infect. Immun.* **83**, 4003–4014 (2015).

85. Enaud, R. *et al.* The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks. *Frontiers in Cellular and Infection Microbiology* vol. 10 at https://doi.org/10.3389/fcimb.2020.00009 (2020).

86. Mossad, O. *et al.* Gut microbiota drives age-related oxidative stress and mitochondrial damage in microglia via the metabolite N6-carboxymethyllysine. *Nat. Neurosci.* **25**, 295–305 (2022).

87. Zhou, W. *et al.* The gut microbe Bacteroides fragilis ameliorates renal fibrosis in mice. *Nat. Commun.* **13**, (2022).

88. Stiemsma, L. T. & Michels, K. B. The Role of the microbiome in the developmental origins of health and disease. *Pediatrics* **141**, (2018).

89. Abrahamsson, T. R. *et al.* Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin. Exp. Allergy* **44**, 842–850 (2014).

90. Depner, M. *et al.* Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. *Nat. Med.* **26**, 1766–1775 (2020).

91. Begley, L. *et al.* Gut microbiota relationships to lung function and adult asthma phenotype: A pilot study. *BMJ Open Respir. Res.* **5**, 1–7 (2018).

92. Chiu, C. Y. *et al.* Gut microbial-derived butyrate is inversely associated with IgE responses to allergens in childhood asthma. *Pediatr. Allergy Immunol.* **30**, 689–697 (2019).

93. Jaeger, N. *et al.* Airway Microbiota-Host Interactions Regulate Secretory Leukocyte Protease Inhibitor Levels and Influence Allergic Airway Inflammation. *Cell Rep.* **33**, (2020).

94. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

95. Laigaard, A. *et al.* Dietary prebiotics promote intestinal Prevotella in association with a low-responding phenotype in a murine oxazolone-induced model of atopic dermatitis. *Sci. Rep.* **10**, 1–11 (2020).

96. Larsen, J. M. The immune response to Prevotella bacteria in chronic inflammatory disease. *Immunology* **151**, 363–374 (2017).

97. Chua, H. H. *et al.* Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants. *Gastroenterology* **154**, 154–167 (2018).

98. Walter, J., Armet, A. M., Finlay, B. B. & Shanahan, F. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. *Cell* **180**, 221–232 (2020).

99. Dzidic, M. *et al.* Aberrant IgA responses to the gut microbiota during infancy precede asthma and allergy development. *J. Allergy Clin. Immunol.* **139**, 1017-1025.e14 (2017).

100. Vael, C., Nelen, V., Verhulst, S. L., Goossens, H. & Desager, K. N. Early intestinal Bacteroides fragilis colonisation and development of asthma. *BMC Pulm. Med.* **8**, 1–6 (2008).

101. Nishi, H. *et al.* Hemoglobin is expressed by mesangial cells and reduces oxidant stress. *J.*

Am. Soc. Nephrol. **19**, 1500–1508 (2008).

102. Liu, W., Baker, S. S., Baker, R. D., Nowak, N. J. & Zhu, L. Upregulation of hemoglobin expression by oxidative stress in hepatocytes and its implication in nonalcoholic steatohepatitis. *PLoS One* **6**, (2011).

103. Dougan, J. *et al.* Proteomics-metabolomics combined approach identifies peroxidasin as a protector against metabolic and oxidative stress in prostate cancer. *Int. J. Mol. Sci.* **20**, (2019).

104. Chan, T. K. *et al.* House dust mite–induced asthma causes oxidative damage and DNA double-strand breaks in the lungs. *J. Allergy Clin. Immunol.* **138**, 84-96.e1 (2016).

105. Wang, Y., Lin, J., Shu, J., Li, H. & Ren, Z. Oxidative damage and DNA damage in lungs of an ovalbumin-induced asthmatic murine model. *J. Thorac. Dis.* **10**, 4819–4830 (2018).

106. Ba, X., Aguilera-Aguirre, L., Sur, S. & Boldogh, I. 8-Oxoguanine DNA glycosylase-1-driven DNA base excision repair: Role in asthma pathogenesis. *Current Opinion in Allergy and Clinical Immunology* vol. 15 89–97 at https://doi.org/10.1097/ACI.0000000000000135 (2015).

107. Kau, A. L. *et al.* Functional characterization of IgA-targeted bacterial taxa from undernourished Malawian children that produce diet-dependent enteropathy. *Sci. Transl. Med.* **7**, 1–15 (2015).

108. Obiso, R. J., Lyerly, D. M., Van Tassell, R. L. & Wilkins, T. D. Proteolytic activity of the Bacteroides fragilis enterotoxin causes fluid secretion and intestinal damage in vivo. *Infect. Immun.* **63**, 3820–3826 (1995).

109. Obiso, R. J., Azghani, A. O. & Wilkins, T. D. The Bacteroides fragilis toxin fragilysin disrupts the paracellular barrier of epithelial cells. *Infect. Immun.* **65**, 1431–1439 (1997).

110. Wu, S., Lim, K. C., Huang, J., Saidi, R. F. & Sears, C. L. Bacteroides fragilis enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14979–14984 (1998).

111. Cani, P. D. *et al.* Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* **57**, 1470–1481 (2008).

112. Ueda, J. *et al.* Decreased pulmonary extracellular superoxide dismutase during systemic inflammation. *Free Radic. Biol. Med.* **45**, 897–904 (2008).

113. Cerdeño-Tárraga, A. M. *et al.* Extensive DNA inversions in the B. fragilis genome control variable gene expression. *Science (80-.).* **307**, 1463–1465 (2005).

114. Gueders, M. M. *et al.* Mouse models of asthma: a comparison between C57BL/6 and BALB/c strains regarding bronchial responsiveness, inflammation, and cytokine

production. *Inflamm. Res.* **58**, 845–854 (2009).

115. Fang, L. *et al.* A mouse allergic asthma model induced by shrimp tropomyosin. *Int. Immunopharmacol.* **91**, (2021).

116. Kato, N. *et al.* Prevalence of enterotoxigenic Bacteroides fragilis in children with diarrhea in Japan. *J. Clin. Microbiol.* **37**, 801–803 (1999).

117. Merino, V. R. C. *et al.* Quantitative detection of enterotoxigenic Bacteroides fragilis subtypes isolated from children with and without diarrhea. *J. Clin. Microbiol.* **49**, 416–418 (2011).

118. Wagner, V. E. *et al.* Effects of a gut pathobiont in a gnotobiotic mouse model of childhood undernutrition. *Sci. Transl. Med.* **8**, (2016).

119. Schatz, M. *et al.* Asthma Control Test: Reliability, validity, and responsiveness in patients not previously followed by asthma specialists. *J. Allergy Clin. Immunol.* **117**, 549–556 (2006).

120. Trivedi, M. & Denton, E. Asthma in children and adults—what are the differences and what can they tell us about asthma? *Frontiers in Pediatrics* vol. 7 at https://doi.org/10.3389/fped.2019.00256 (2019).

121. Fuseini, H. & Newcomb, D. C. Mechanisms Driving Gender Differences in Asthma. *Current Allergy and Asthma Reports* vol. 17 at https://doi.org/10.1007/s11882-017-0686-1 (2017).

122. Polosa, R. & Thomson, N. C. Smoking and asthma: Dangerous liaisons. *European Respiratory Journal* vol. 41 716–725 at https://doi.org/10.1183/09031936.00073312 (2013).

123. Howard, E., Orhurhu, V., Huang, L., Guthrie, B. & Phipatanakul, W. The Impact of Ambient Environmental Exposures to Microbial Products on Asthma Outcomes from Birth to Childhood. *Current Allergy and Asthma Reports* vol. 19 at https://doi.org/10.1007/s11882-019-0890-2 (2019).

124. Chung, H. *et al.* Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* **149**, 1578–1593 (2012).

125. Vujkovic-Cvijin, I. *et al.* Host variables confound gut microbiota studies of human disease. *Nature* **587**, 448–454 (2020).

126. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science (80-.).* **326**, 1694–1697 (2009).

127. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).

128. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).

129. Cho, R. L. *et al.* Lipopolysaccharide induces ICAM-1 expression via a c-Src/NADPH oxidase/ROS-dependent NF-κB pathway in human pulmonary alveolar epithelial cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **310**, L639–L657 (2016).

130. Goodwin, A. C. *et al.* Polyamine catabolism contributes to enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15354–15359 (2011).

131. Hwang, S. *et al.* Enterotoxigenic bacteroides fragilis infection exacerbates tumorigenesis in AOM/DSS mouse model. *Int. J. Med. Sci.* **17**, 145–152 (2020).

132. Caffarelli, C. *et al.* Gastrointestinal symptoms in patients with asthma. *Arch. Dis. Child.* **82**, 131–135 (2000).

133. Jesenak, M., Zelieskova, M. & Babusikova, E. Oxidative stress and bronchial asthma in children-causes or consequences? *Front. Pediatr.* **5**, (2017).

134. Akiki, Z. *et al.* High level of fluorescent oxidation products and worsening of asthma control over time. *Respir. Res.* **20**, (2019).

135. Zou, X. L. *et al.* Associations Between Gut Microbiota and Asthma Endotypes: A Cross-Sectional Study in South China Based on Patients with Newly Diagnosed Asthma. *J. Asthma Allergy* **14**, 981–992 (2021).

136. Tajik, N. *et al.* Targeting zonulin and intestinal epithelial barrier function to prevent onset of arthritis. *Nat. Commun.* **11**, (2020).

137. Cardoso-Silva, D. *et al.* Intestinal barrier function in gluten-related disorders. *Nutrients* vol. 11 at https://doi.org/10.3390/nu11102325 (2019).

138. Kuperman, D. A. *et al.* Dissecting asthma using focused transgenic modeling and functional genomics. *J. Allergy Clin. Immunol.* **116**, 305–311 (2005).

139. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522 (2011).

140. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).

141. McMurdie, P. J. & Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput. Biol.* **10**, (2014).

142. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).

143. Nearing, J. T. *et al.* Microbiome differential abundance methods produce different results across 38 datasets. *Nat. Commun.* **13**, 1–16 (2022).

144. Griffith, D. M., Veech, J. A. & Marsh, C. J. Cooccur: Probabilistic species co-occurrence analysis in R. *J. Stat. Softw.* **69**, 1–17 (2016).

145. Planer, J. D. *et al.* Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. *Nature* **534**, 263–266 (2016).

146. Baym, M. *et al.* Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* **10**, 1–15 (2015).

147. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

148. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

149. Yoon, S. H., Ha, S. min, Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **110**, 1281–1286 (2017).

150. Patnode, M. L., Bando, J. K., Krummel, M. F., Locksley, R. M. & Rosen, S. D. Leukotriene B4 amplifies eosinophil accumulation in response to nematodes. *J. Exp. Med.* **211**, 1281–1288 (2014).

151. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

152. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

153. Korotkevich, G. *et al.* Fast gene set enrichment analysis. *bioRxiv* 060012 (2016) doi:10.1101/060012.

154. Knoop, K. A. *et al.* Microbial antigen encounter during a preweaning interval is critical for tolerance to gut bacteria. *Sci. Immunol.* **2**, (2017).

155. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

156. Hu, Y. *et al.* The gut microbiome signatures discriminate healthy from pulmonary tuberculosis patients. *Front. Cell. Infect. Microbiol.* **9**, 1–8 (2019).

157. Wilson, N. G. *et al.* The gut microbiota of people with asthma influences lung inflammation in gnotobiotic mice. *iScience* **26**, 105991 (2023).

158. Asthma-related emergency department visits 2010–2018 | CDC. https://www.cdc.gov/asthma/asthma_stats/asthma-ed-visits_2010-2018.html.

159. Roduit, C. *et al.* High levels of Butyrate and Propionate in early life are associated with protection against atopy. *Allergy* 1–11 (2018) doi:10.1111/all.13660.

160. Zaiss, M. M. *et al.* The Intestinal Microbiota Contributes to the Ability of Helminths to Modulate Allergic Inflammation. *Immunity* **43**, 998–1010 (2015).

161. Ramirez, J. *et al.* Antibiotics as Major Disruptors of Gut Microbiota. *Front. Cell. Infect. Microbiol.* **10**, 1–10 (2020).

162. Kozyrskyj, A. L., Ernst, P. & Becker, A. B. Increased risk of childhood asthma from antibiotic use in early life. *Chest* **131**, 1753–1759 (2007).

163. McKeever, T. M. *et al.* Early exposure to infections and antibiotics and the incidence of allergic disease: A birth cohort study with the West Midlands General Practice Research Database. *J. Allergy Clin. Immunol.* **109**, 43–50 (2002).

164. Hoskin-Parr, L., Teyhan, A., Blocker, A. & Henderson, A. J. W. Antibiotic exposure in the first two years of life and development of asthma and other allergic diseases by 7.5 yr: A dose-dependent relationship. *Pediatr. Allergy Immunol.* **24**, 762–771 (2013).

165. Russell, S. L. *et al.* Perinatal antibiotic treatment affects murine microbiota, immune responses and allergic asthma. *Gut Microbes* **4**, 158–164 (2013).

166. Yang, X. *et al.* Early-life vancomycin treatment promotes airway inflammation and impairs microbiome homeostasis. *Aging (Albany. NY).* **11**, 2071–2081 (2019).

167. Borbet, T. C. *et al.* Influence of the early-life gut microbiota on the immune responses to an inhaled allergen. *Mucosal Immunol.* **15**, 1000–1011 (2022).

168. Rodriguez-R, L. M. & Konstantinidis, K. T. Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics* **30**, 629–635 (2014).

169. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *Elife* **10**, 1–42 (2021).

170. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, (2021).

171. Brayner, B. *et al.* Dietary Patterns Characterized by Fat Type in Association with Obesity and Type 2 Diabetes: A Longitudinal Study of UK Biobank Participants. *J. Nutr.* **151**, 3570–3578 (2021).

172. Wendell, S. G., Baffi, C. & Holguin, F. Fatty acids, inflammation, and asthma. *J. Allergy Clin. Immunol.* **133**, 1255–1264 (2014).

173. Mizuta, K., Matoba, A., Shibata, S., Masaki, E. & Emala, C. W. Obesity-induced asthma: Role of free fatty acid receptors. *Jpn. Dent. Sci. Rev.* **55**, 103–107 (2019).

174. Scott, H. A., Gibson, P. G., Garg, M. L. & Wood, L. G. Airway inflammation is augmented by obesity and fatty acids in asthma. *Eur. Respir. J.* **38**, 594–602 (2011).

175. Snyder, B. M. *et al.* Association between asthma status and prenatal antibiotic prescription fills among women in a Medicaid population. *J. Asthma* **59**, 2100–2107 (2021).

176. Kaminski, J. *et al.* High-Specificity Targeted Funcitonal Profiling in Microbial Communities with ShortBRED. *Nature* **486**, 207–214 (2012).

177. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, 517–525 (2019).

178. Mardia, K., Kent, J. & Bibby, J. Mardia, Kent, Bibby - 1979 - Multivariate Analysis.pdf. 521 (1979).

179. L, C. *et al.* VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* **33**, (2005).

180. Nagel, G. & Linseisen, J. Dietary intake of fatty acids, antioxidants and selected food groups and asthma in adults'. *Eur. J. Clin. Nutr.* **59**, 8–15 (2005).

181. Mizuta, K. *et al.* Novel identification of the free fatty acid receptor FFAR1 that promotes contraction in airway smooth muscle. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **309**, L970–L982 (2015).

182. Felton, E. *et al.* Inflammation in children with cystic fibrosis: contribution of bacterial production of long-chain fatty acids. *Pediatr. Res.* **90**, 99–108 (2021).

183. Niot, I., Poirier, H., Tran, T. T. T. & Besnard, P. Intestinal absorption of long-chain fatty acids: Evidence and uncertainties. *Prog. Lipid Res.* **48**, 101–115 (2009).

184. Vital, M., Howe, A. C. & Tiedje, J. M. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *MBio* **5**, 1–11 (2014).

185. Yassour, M. *et al.* Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Sci. Transl. Med.* **8**, (2016).

186. Schwartz, D. J., Langdon, A. E. & Dantas, G. Understanding the impact of antibiotic perturbation on the human microbiome. *Genome Med.* **12**, 1–12 (2020).

187. Malhotra-Kumar, S., Lammens, C., Coenen, S., Van Herck, K. & Goossens, H. Effect of azithromycin and clarithromycin therapy on pharyngeal carriage of macrolide-resistant streptococci in healthy volunteers: a randomised, double-blind, placebo-controlled study. *Lancet* **369**, 482–490 (2007).

188. Doan, T. *et al.* Gut microbiome alteration in MORDOR I: a community-randomized trial of mass azithromycin distribution. *Nat. Med.* **25**, 1370–1376 (2019).

189. Doan, T. *et al.* Macrolide Resistance in MORDOR I — A Cluster-Randomized Trial in Niger. *N. Engl. J. Med.* **380**, 2271–2273 (2019).

190. Escudeiro, P., Pothier, J., Dionisio, F. & Nogueira, T. Antibiotic Resistance Gene Diversity and Virulence Gene. *mSphere* **4**, 1–13 (2019).

191. Pan, Y. *et al.* Coexistence of Antibiotic Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome Analysis. *mSystems* **5**, (2020).

192. Strati, F. *et al.* Antibiotic-associated dysbiosis affects the ability of the gut microbiota to control intestinal inflammation upon fecal microbiota transplantation in experimental colitis models. *Microbiome* **9**, 1–15 (2021).

193. Casterline, B. W., Hecht, A. L., Choi, V. M. & Bubeck Wardenburg, J. The Bacteroides fragilis pathogenicity island links virulence and strain competition. *Gut Microbes* **0976**, 1–10 (2017).

194. Licari, A., Fuchs, D., Marseglia, G. & Ciprandi, G. Tryptophan metabolic pathway and neopterin in asthmatic children in clinical practice. *Ital. J. Pediatr.* **45**, 1–4 (2019).

195. Van der Leek, A. P., Yanishevsky, Y. & Kozyrskyj, A. L. The kynurenine pathway as a novel link between allergy and the gut microbiome. *Front. Immunol.* **8**, 1374 (2017).

196. De Wolfe, T. J., Arefin, M. R., Benezra, A. & Rebolleda Gómez, M. Chasing Ghosts: Race, Racism, and the Future of Microbiome Research. *mSystems* **6**, 1–8 (2021).

197. Cooper, R. S. Race in Biological and Biomedical Research. *Cold Spring Harb. Perspect. Med.* **3**, (2013).

198. Fitzpatrick, A. M. *et al.* Racial disparities in asthma-related health care use in the National Heart, Lung, and Blood Institute's Severe Asthma Research Program. *J. Allergy Clin. Immunol.* **143**, 2052–2061 (2019).

199. Findley, K., Williams, D. R., Grice, E. A. & Bonham, V. L. Health Disparities and the Microbiome. *Trends Microbiol.* **24**, 847 (2016).

200. Rodriguez-R, L. M., Gunturu, S., Tiedje, J. M., Cole, J. R. & Konstantinidis, K. T. Nonpareil 3: Fast Estimation of Metagenomic Coverage and Sequence Diversity. *mSystems* **3**, (2018).

201. Generate Health. St. Louis Fetal-Infant Mortality Rate 10 Year Report. 47 https://generatehealthstl.org/resources/publications/fimr-10-year-report/ (2017).

202. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

203. Mikheenko, A., Valin, G., Prjibelski, A., Saveliev, V. & Gurevich, A. Icarus: visualizer for de novo assembly evaluation. *Bioinformatics* **32**, 3321–3323 (2016).

204. Mirzayi, C. *et al.* Reporting guidelines for human microbiome research: the STORMS checklist. *Nat. Med.* **27**, 1885–1892 (2021).

205. Wilson, N. G., Hernandez-Leyva, A., Schwartz, D. J., Bacharier, L. B. & Kau, A. L. The gut metagenome harbors metabolic and antibiotic resistance signatures of moderate-to-severe asthma. *bioRxiv* 2023.01.03.522677 (2023) doi:10.1101/2023.01.03.522677.

206. Di Gangi, A., Di Cicco, M. E., Comberiati, P. & Peroni, D. G. Go With Your Gut: The Shaping of T-Cell Response by Gut Microbiota in Allergic Asthma. *Front. Immunol.* **11**, 1–7 (2020).

207. Newcomb, D. C. & Peebles, R. S. Th17-mediated inflammation in asthma. *Curr. Opin. Immunol.* **25**, 755–760 (2013).

208. Bach, J.-F. The Effect of Infections on Susceptibility to Autoimmune and Allergic Diseases. *N Engl J Med* **347**, 911–920 (2002).

209. Fändriks, L. Roles of the gut in the metabolic syndrome: an overview. *J. Intern. Med.* **281**, 319–336 (2017).

210. Nagato, A. C., Bezerra, F. S., Talvani, A., Aarestrup, B. J. & Aarestrup, F. M. Hyperoxia promotes polarization of the immune response in ovalbumin-induced airway inflammation, leading to a th17 cell phenotype. *Immunity, Inflamm. Dis.* **3**, 321–337 (2015).

211. Righetti, R. F. *et al.* Protective effects of anti-IL17 on acute lung injury induced by LPS in mice. *Front. Pharmacol.* **9**, 1021 (2018).

212. Chan, K. L. *et al.* Palmitoleate reverses high fat-induced proinflammatory macrophage polarization via AMP-activated protein kinase (AMPK). *J. Biol. Chem.* **290**, 16979–16988 (2015).

213. Kompauer, I. *et al.* Association of fatty acids in serum phospholipids with lung function and bronchial hyperresponsiveness in adults. *Eur. J. Epidemiol.* **23**, 175–190 (2008).

214. Basak, S., Banerjee, A., Pathak, S. & Duttaroy, A. K. Dietary Fats and the Gut Microbiota: Their impacts on lipid-induced metabolic syndrome. *J. Funct. Foods* **91**, 105026 (2022).

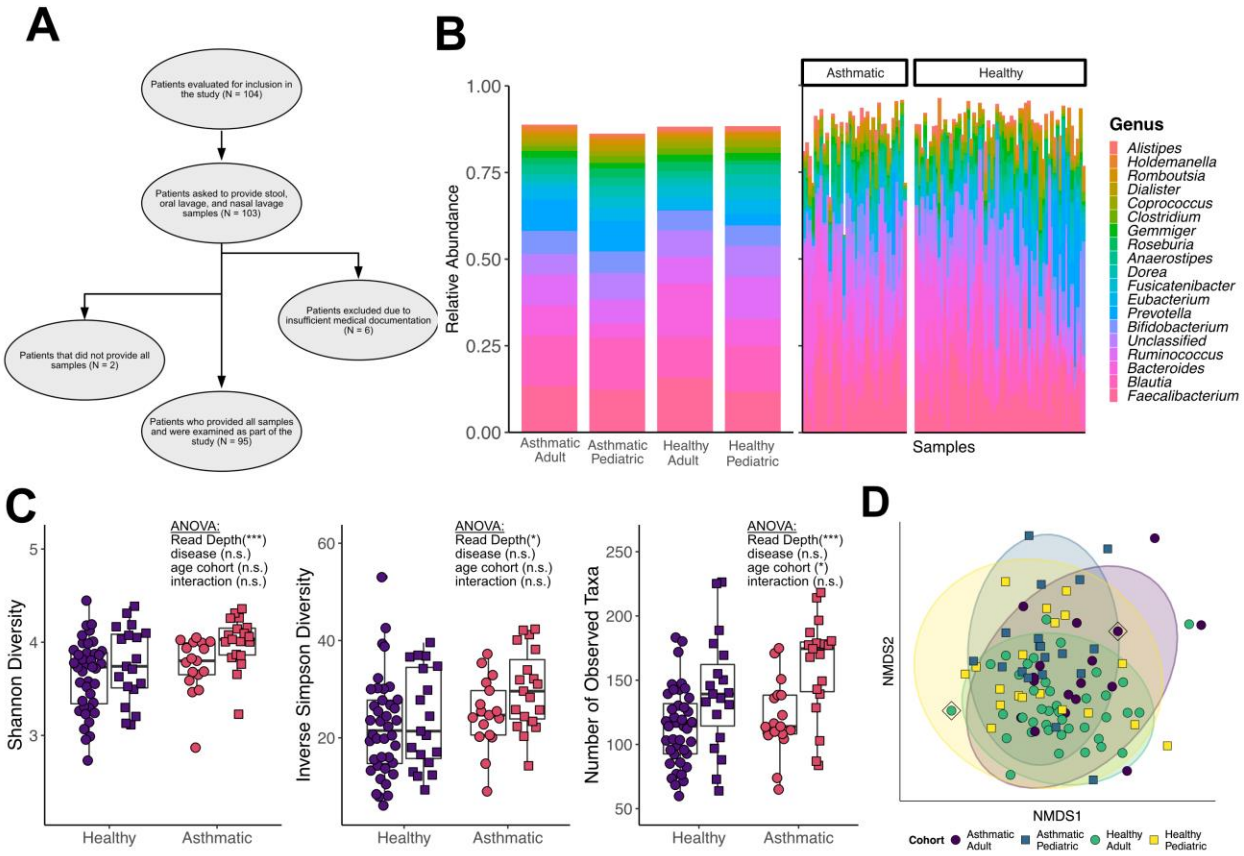

Figure 5.1. Overview of MARS inclusion and 16S rRNA sequencing results summarized to the genus level
A) Flowchart depicting steps for inclusion/exclusion of MARS participants **B)** Relative abundance bar plot at genus level of 95 human fecal samples characterized in this study. Left: mean relative abundance by age group and asthma cohorts; Right: One stacked bar plot per fecal sample. White space represents unclassified taxa. **C)** Shannon, Inverse Simpson, and Observed Number of Taxa alpha diversity based on ASVs in stool samples from the MARS Cohort. Read depth was included as a variable to control for differences in library size. **D)** Non-metric multidimensional scaling (NMDS) on Bray-Curtis dissimilarity of MARS gut microbiomes. Ellipses represent 95% confidence intervals; diamonds indicate donor dyad (MARS0022/MARS0043). For all panels: N=59 Healthy, N=36 Asthmatic.

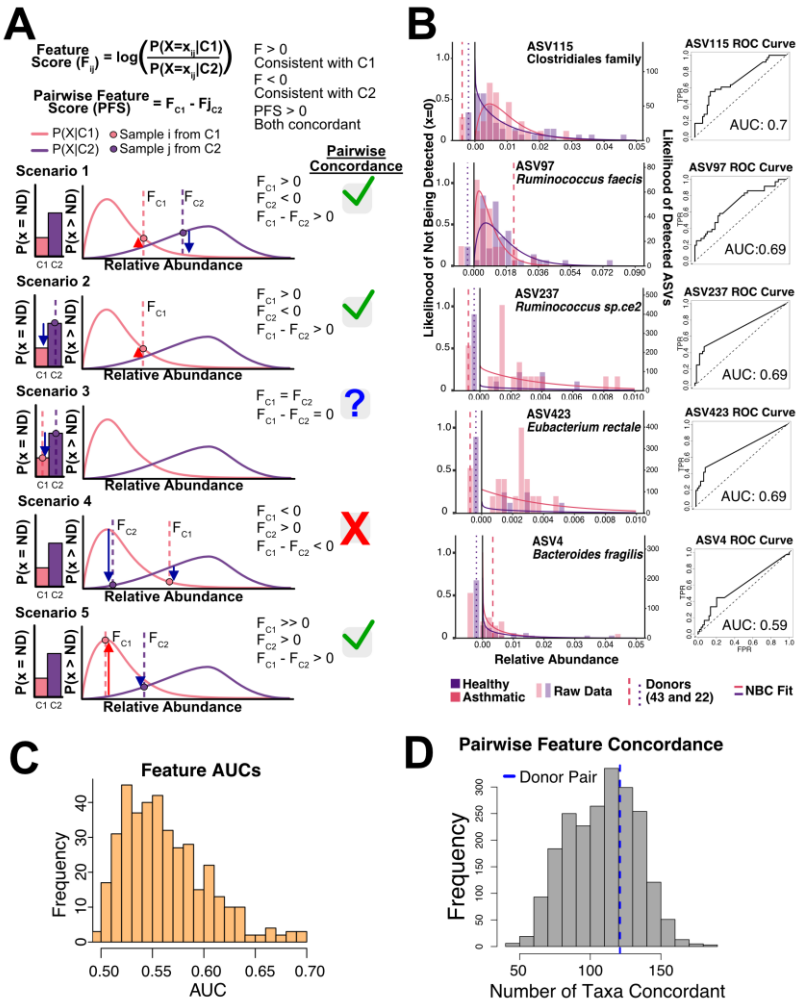

Figure 5.2. Overview of NBC pairwise concordance metric and estimation of taxa importance from AUC scores, related to Figure 2.2 and STAR Methods.

A)The NBC feature score describes the likelihood that a taxon at a given relative abundance would be found in one class over another. A comparison of the feature scores of the same taxon across two different samples provides the pairwise feature score (PFS) that compares the likelihood of these taxa abundances occurring between two samples from different classes. Example scenarios where the pairwise feature score is used to identify concordance with the model are provided. **Scenario 1** shows a case where both the asthma sample and healthy sample are present at relative abundances expected by the model. **Scenario 2** shows a case where a healthy sample is absent, and the asthma sample is present at abundances expected by the model. **Scenario 3** demonstrates a unique situation where the taxon is absent in both samples. In this case, the pair provides no information about the taxon and thus is equal to zero. **Scenario 4** demonstrates the case where the relative abundances of the taxon in both the asthma sample and healthy sample are not consistent with the model. The pairwise feature score in this case is negative, reflecting discordance with the model. **Scenario 5** depicts an edge case, where the relative abundance in the asthma sample is expected, but the relative abundance of the healthy sample is not. The magnitude of the likelihood of the asthma sample is much greater than that of the healthy sample however, and so the pairwise feature score is positive reflecting model concordance in this taxon. **B)** Left: raw data histograms overlaid with NBC curve of best fit and right: ROC curve for highly ranked ASVs and *B. fragilis* (ASV4). Donor samples are represented by dashed lines (pink for MARS0043 and purple for MARS0022). **C)** Distribution of AUC values from all 392 ASVs in NBC training set. **D)** Histogram of raw counts of pairwise concordant features per sample. Donor pair (MARS0022/MARS0043) represented by dashed blue line.

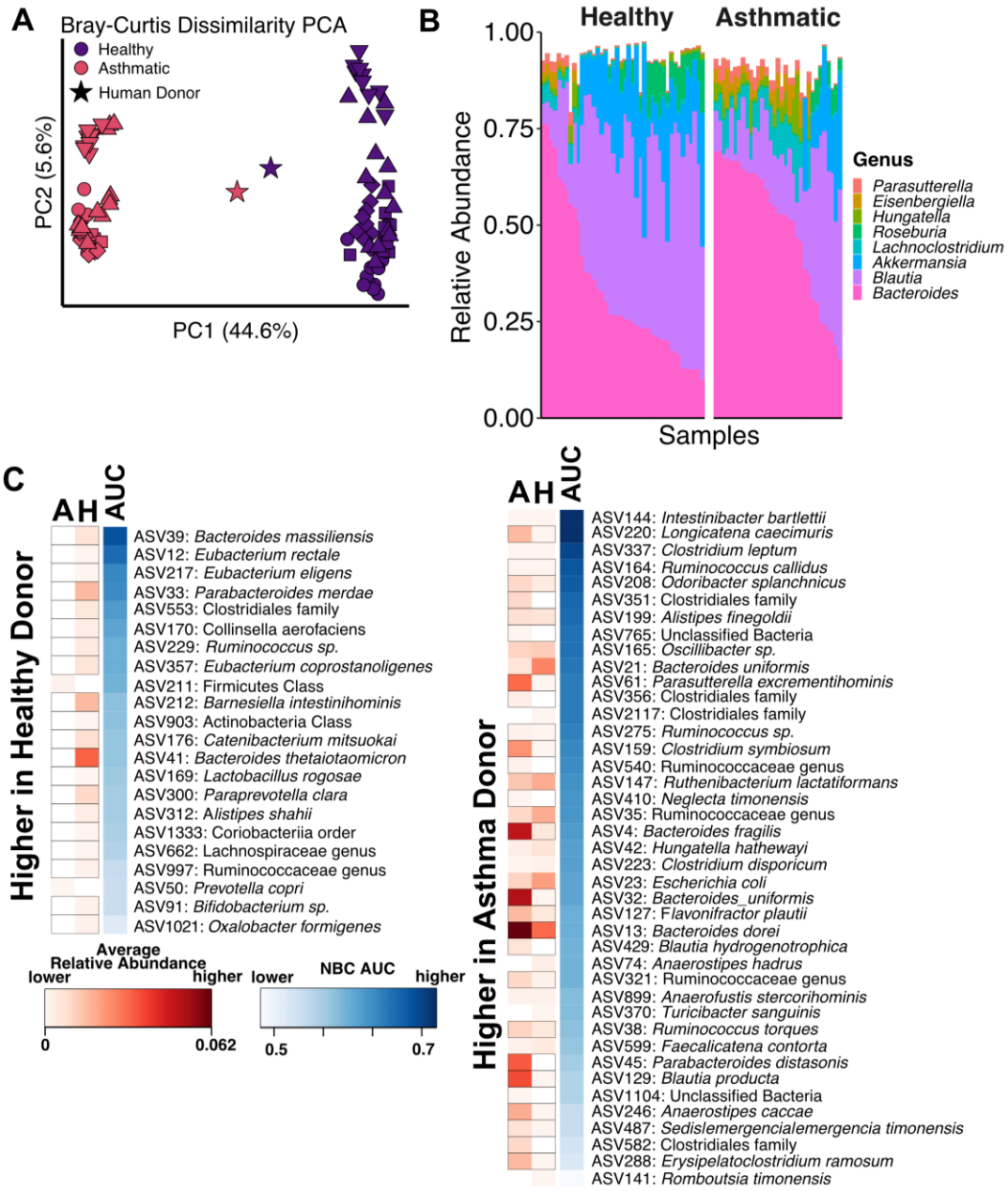

Figure 5.3. 16S rRNA metagenomic characterization of stool from gnotobiotic mice humanized with MARS0022 and MARS0043, related to Figure 2.3 and Figure 2.4.

A) PCoA of the fecal 16S Unifrac distances of humanized gnotobiotic mice colonized for six weeks (n=5-10/group). Shapes correspond to separate experiments and are consistent with previous figure. Star corresponds to human donor sample. **B)** Relative abundance bar plot of all HO and AO mice colonized for six weeks as part of this study (n=5-10/group, 4 experiments). **C)** PCoA of the fecal 16S Unifrac distances of humanized gnotobiotic mice colonized for one week (n=9-10) only along with donor fecal samples (stars). **D)** Heatmap of taxa concordant by NBC and occurring in HO and AO mice colonized with human donors. Statistically significant differences noted by black rectangle outlines (Wilcoxon, two-tailed). All experiments here include 2-5 males and 2-5 females per group.

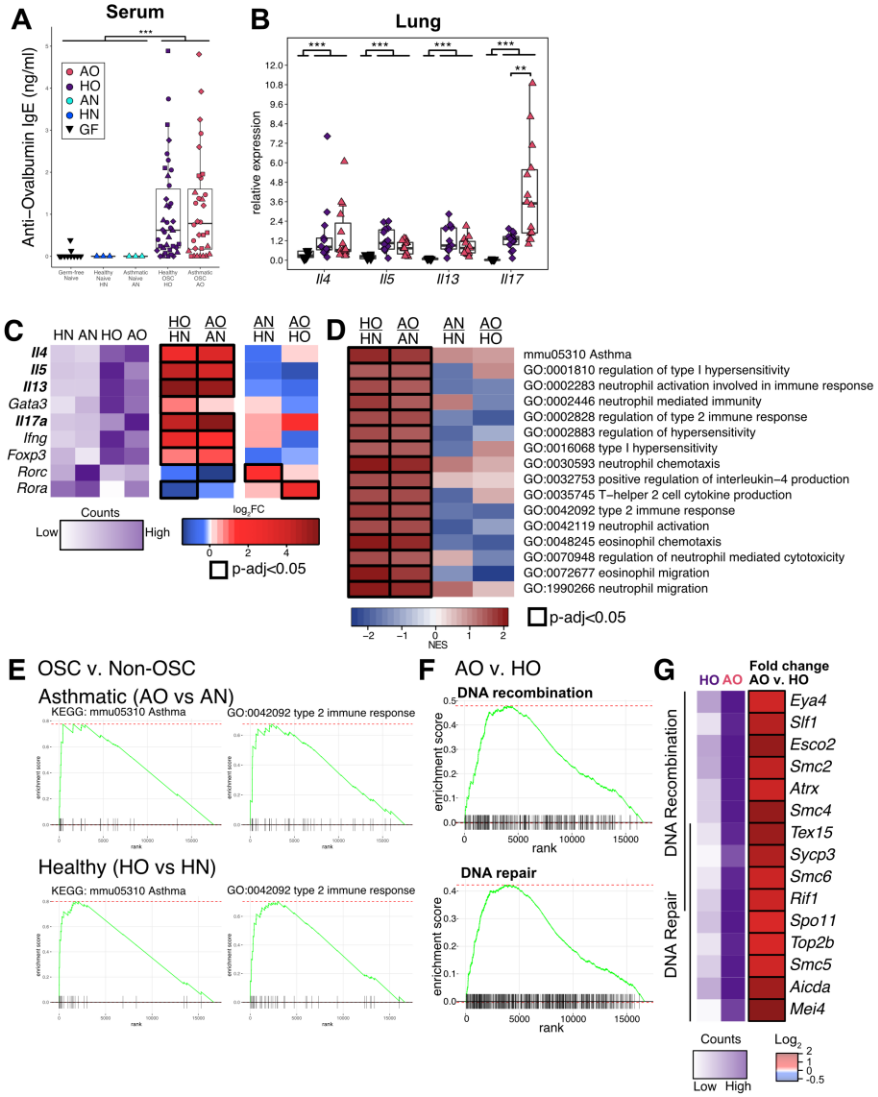

Figure 5.4. Typical AAI markers found in OSC mouse lungs and DNA repair and recombination upregulated in AO compared to HO mice, related to Figure 2.3 and Figure 2.4.

A) Serum anti-ovalbumin IgE measurements from naïve GF, HN, AN, HO and AO mice (n=6-10/group, 4 experiments denoted by shapes). Two-sided Wilcoxon Test result between OSC treated and non-OSC treated mice shown. **B)** Expression of genes encoding Interleukin-4, -5, -13, and -17A measured by RT-qPCR from the lungs of Naïve GF, HO, and AO mice (n=4-10/group, 2 experiments denoted by shapes). Significant two-sided Wilcoxon test results shown. **C)** Heatmap of variance stabilized counts and log2 fold-changes of allergic airway inflammation-related genes. In purple scale: the counts of the genes (z scores scaled by row). In blue-red scale: the log2 fold change. Statistically significant fold changes are outlined in black (n=4-5). **D)** GSEA enrichment heatmap of normalized enrichment scores (NES) of relevant GO and KEGG pathways (n=4-5). **E)** GSEA enrichment plots of the KEGG Asthma pathway and GO Type 2 immune response pathway in OSC vs. Non-OSC mice (n=4-5). **F)** GSEA enrichment plots of GO:0006281 DNA repair and GO:0006310 DNA recombination pathways (n=5/group). **G)** Heatmap of variance stabilized counts of leading edge genes from the GO pathways in (A). In purple scale: the counts of the genes (z scores scaled by row). In blue-red scale: the log2 fold change. Statistically significant fold changes are outlined in black (n=5, p-adjusted<0.05). This experiment includes 2-3 males and 2-3 females per group. The ten naïve germ-free mice are identical to those shown in subsequent figures.

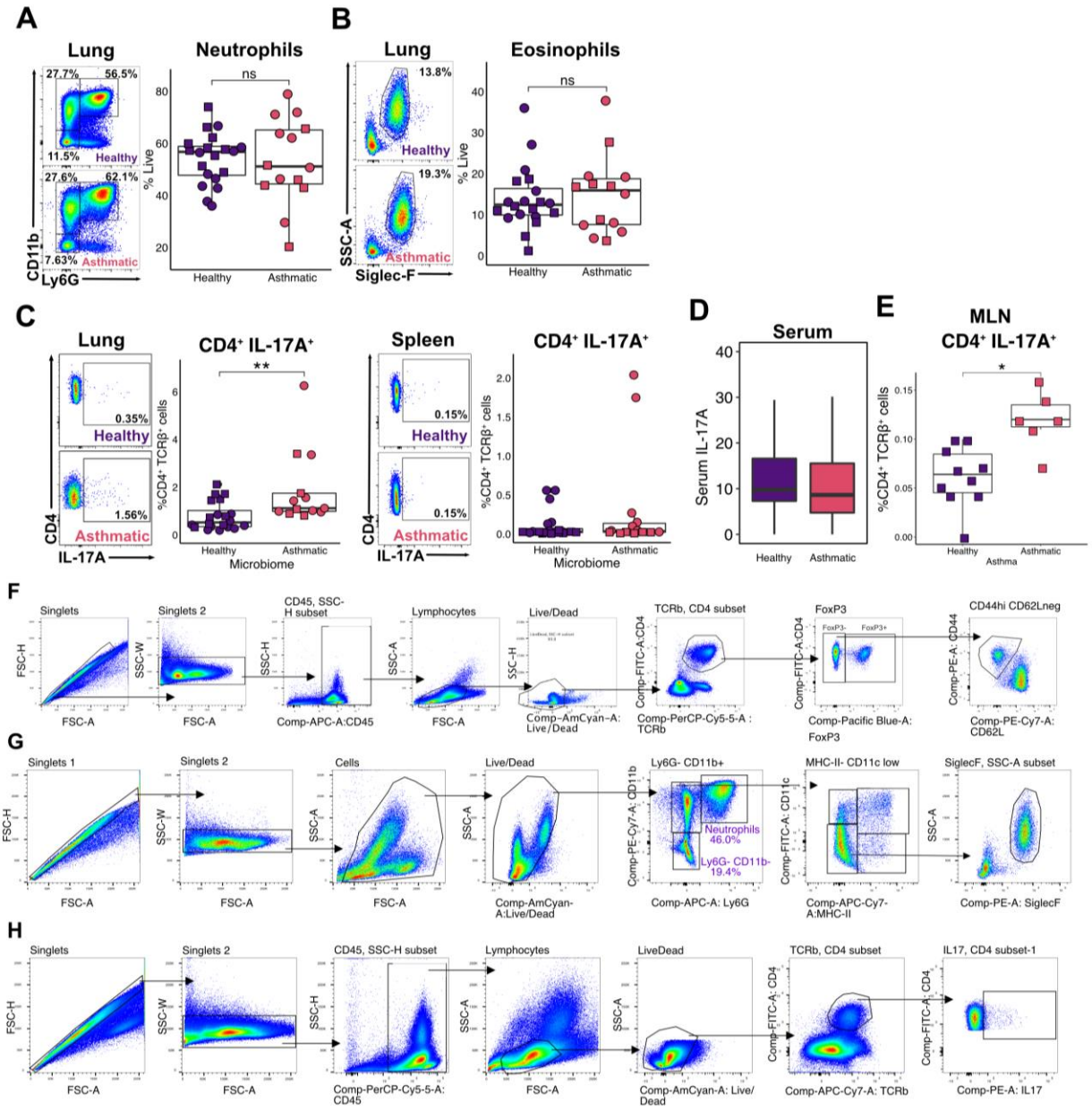

Figure 5.5. Immunophenotyping of humanized gnotobiotic mice, related to Figure 2.3 and STAR Methods.
A - B) Flow cytometry profiling of granulocytes in the lungs of AO and HO mice (n=6-10/group, 2 experiments denoted by shapes). **C)** Intracellular staining of IL-17A following *in vitro* restimulation from lymphocytes isolated from the lungs and spleen of AO and HO mice (n=5-10/group, 2 experiments denoted by shapes). **D)** Serum measurement of IL-17A in AO and HO mice (n=6-10/group, 4 experiments). **E)** Intracellular staining of IL-17A following *in vitro* restimulation from lymphocytes isolated from the mesenteric lymph nodes of AO and HO mice (6-10/group, 1 experiment) **F-H)** Flow cytometry gating strategies for **F)** effector T-cells, **G)** neutrophils and eosinophils, and **H)** IL-17A producing lymphocytes. All experiments here include 2-5 males and 2-5 females per group. Shapes denote separate experiments and are consistent with previous and subsequent figure.

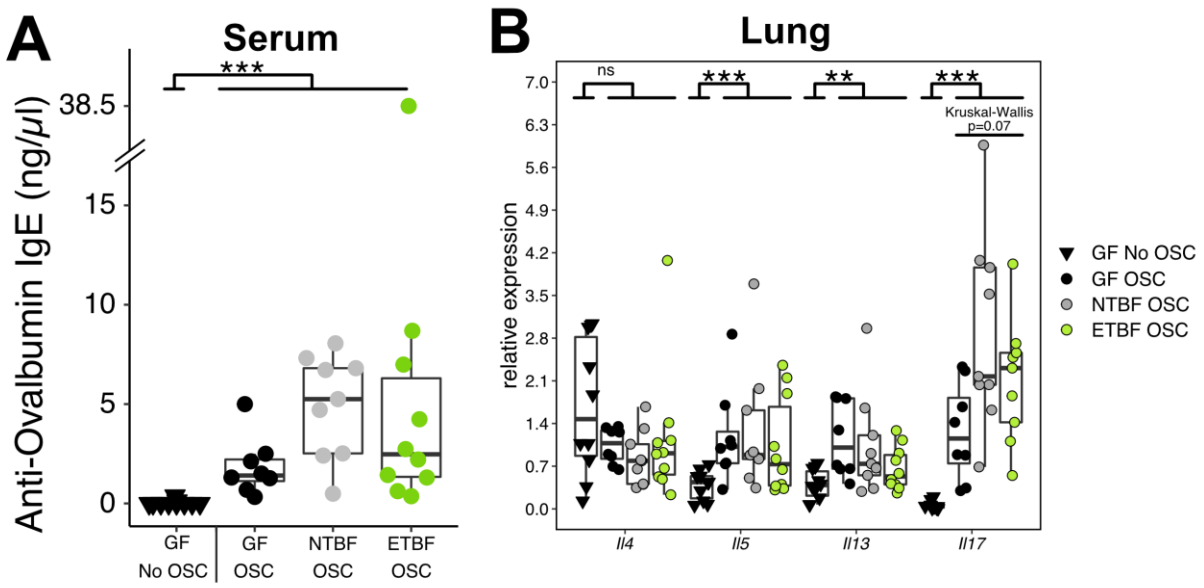

Figure 5.6. Markers of AAI increase after OSC in mice colonized with ETBF, NTBF, or neither but Il17 expression is unchanged between OSC mice, related to Figure 2.5.
A) Serum anti-ovalbumin IgE ELISA results. Two-sided Wilcoxon Test result between OSC treated and non-OSC treated mice shown. Reference naïve GF control OVA-IgE values are the same data shown in Figure 5.3A,B. **B)** Lung cytokine RT-qPCR relative expression. Two-sided Wilcoxon results shown for all lung cytokines between GF No OSC (i.e. naïve) and all OSC mice. Kruskal-Wallis between only OSC groups is shown for *Il17*. All p-values for Kruskal-Wallis between only OSC groups for *Il4, Il5,* and *Il13* were greater than 0.05. A) and B): n=8-10 mice/group; includes 3-4 males and 5-7 females per group.

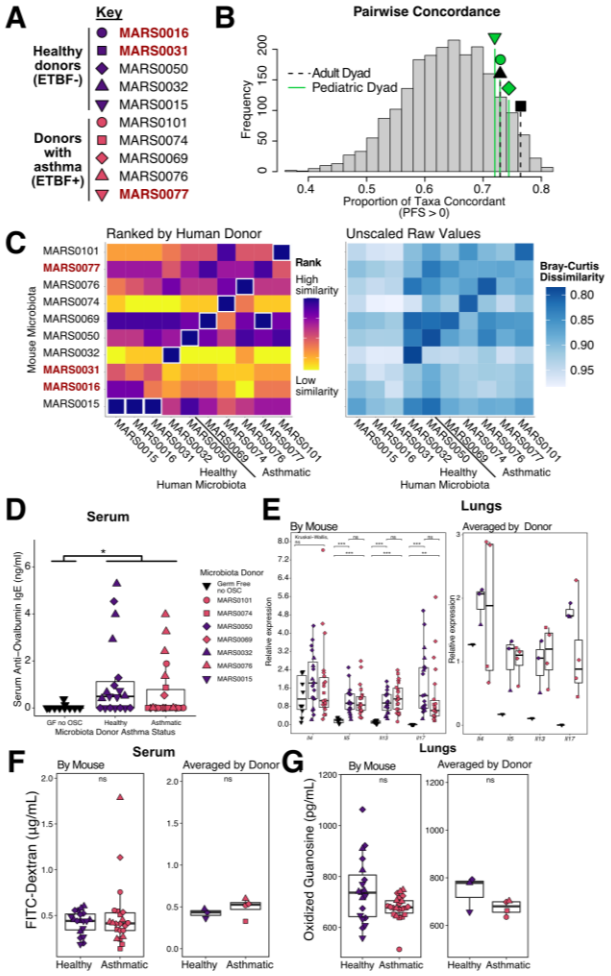

Figure 5.7. Humanization with additional ETBF+ microbiota does not always increase intestinal permeability or markers of AAI compared to humanization with healthy controls, related to Figure 2.6.
A) Legend key of MARS donors used for each humanization. Shapes correspond to matched samples by age group (pediatric vs. adult) and microbiome composition. Red names were excluded from analysis. **B)** Histogram of the proportion of pairwise concordant taxa across all possible healthy-asthma donor dyads (also shown in Figure 2.3C). Vertical dashed line denotes the ETBF+/ETBF- dyads used in this experiment. **C)** Heatmap of Bray Curtis dissimilarity between stool from human donors and recipient mice at the time of sacrifice. The heatmap is presented as ranked across human microbiomes to highlight mouse samples most reflective of their donor (left, top rank outlined in white) and as raw dissimilarity values (right). Mouse recipients in red and bolded were not the most similar to their donor and were excluded from further analysis. **D)** ELISA serum anti-ovalbumin IgE amounts; two-sided Wilcoxon Test result between OSC treated and non-OSC treated mice shown (total Healthy n=19, total Asthmatic n=22). All experimental groups included 1-4 female and 2-4 male mice. Two-sided Wilcoxon test between OSC and non-OSC mice shown. Naïve GF control OVA-IgE values are the same data shown in previous figures. **E)** Lung tissue expression of genes encoding IL-4, -5, -13, and -17A measured by RT-qPCR relative to healthy donors with asterisks and numeric values denoting adjusted p-values from the Kruskal-Wallis post hoc Dunn Test with Benjamini-Hochberg multiple comparisons correction (total Healthy n = 18-19, total Asthmatic n = 21-22). Reference Naïve GF control qPCR are the same data shown in Figure 5.3A and B. **F)** Intestinal permeability of ovalbumin sensitized and challenged (OSC) mice colonized with healthy or asthmatic microbiota following a 2-week colonization (total Healthy n=19, total Asthmatic n = 22). **G)** Oxidized guanosine in lungs of humanized OSC mice (total Healthy n=19, total Asthmatic n = 22).

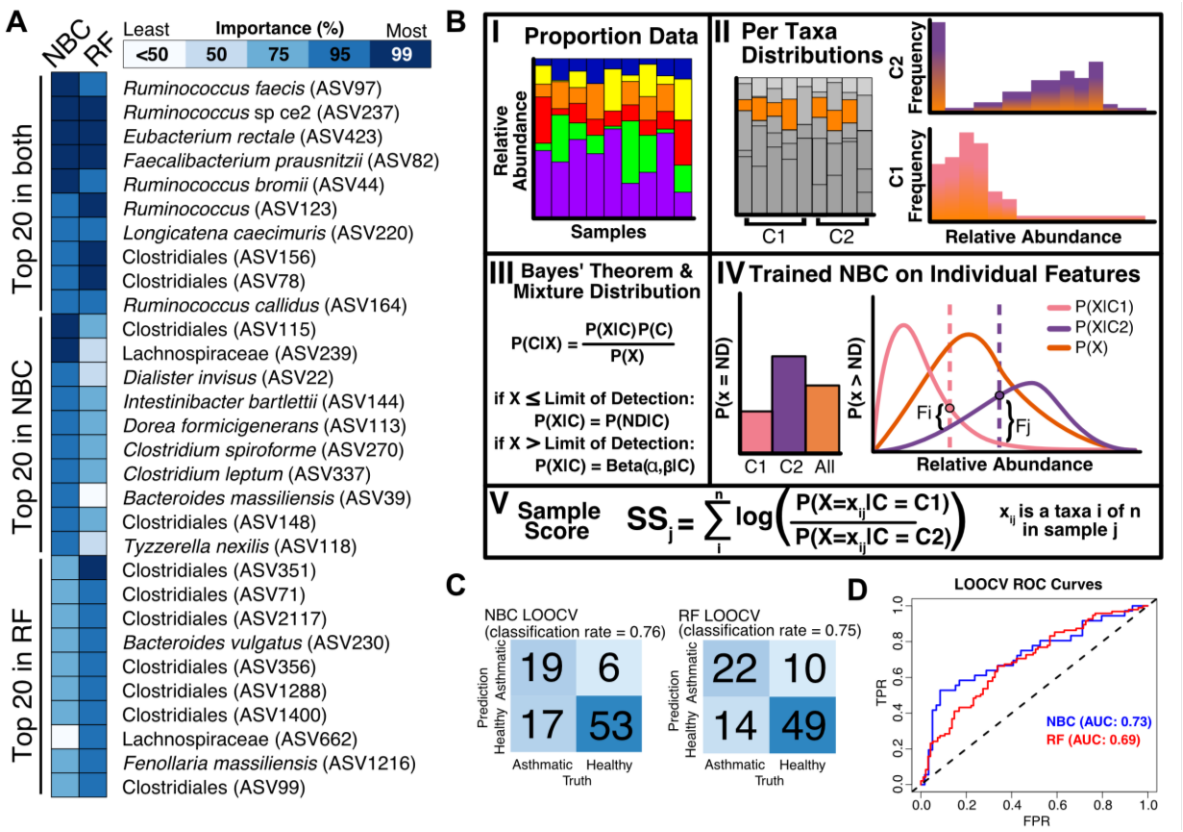

Figure 5.8. A Naïve Bayes' Classifier is an intuitive tool for the exploration of microbiome data, related to Figure 2.2 and STAR Methods.

A) A comparison of taxa identified as important by NBC and RF. Feature importance in the random forest was inferred based on mean decrease in accuracy. Feature importance in the NBC was inferred based on the AUC of an individual taxa's ROC curve in predicting asthma. Colors correspond to percentiles of ranked importance for taxa considered by each of the models. The top 20 predictive features in each model are presented. **B)** Overview of Naïve Bayes' Classifier (NBC) fit with a mixture distribution. **I)** Individual counts of taxa by 16S rRNA sequencing are scaled by total read count per sample. **II)** Relative abundances of each taxon across samples are separated by class. **III)** Bayes' theorem can be used to calculate the probability of a given class given the microbiome composition. We model relative abundance as a mixture distribution such that when a taxon is not detected, it is part of a binary distribution and when detected is part of a beta distribution. **IV)** Using our mixture model, an individual taxon can be modeled overall and as part of a class to calculate the probability of belonging to different classes. **V)** The sample score describes the likelihood that a sample is from one class over another. **C)** Confusion matrices summarizing Leave-One-Out Cross Validation (LOOCV) predictions for NBC and Random Forest (RF), respectively. The NBC performed significantly better than chance (p=8.2e-6 by Fisher's exact test). **D)** Receiver Operating Characteristic (ROC) Curves and area under the ROC curve (AUC) scores from LOOCV NBC and RF models.

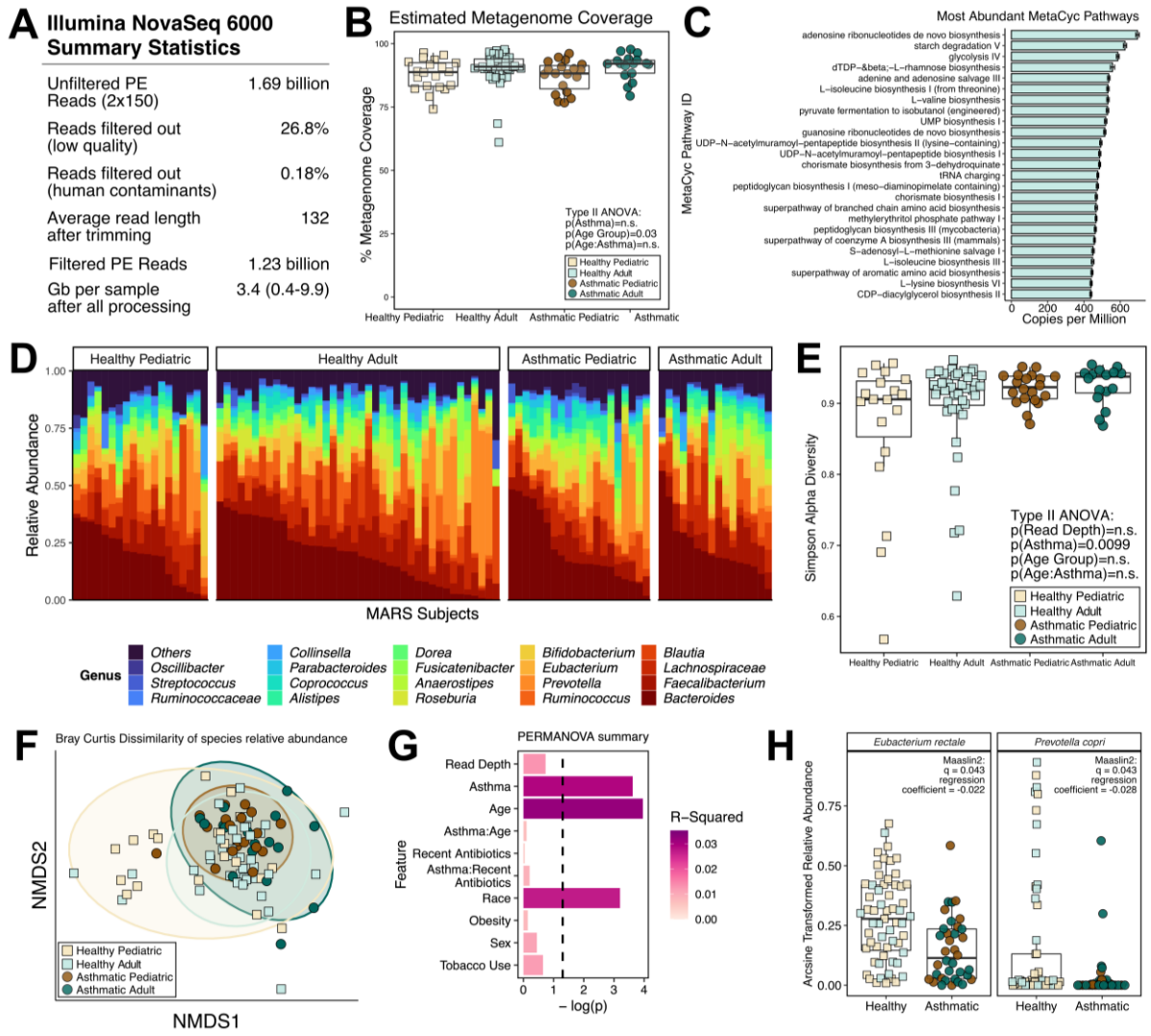

Figure 5.9. MARS whole metagenomic shotgun sequencing captures essential functions and taxonomic shifts of the asthma gut microbiota.

A) Summary of select sequencing statistics from NovaSeq shotgun metagenomic sequencing and subsequent filtering steps. **B)** Boxplot of redundancy-based estimated metagenome coverage (%) as calculated by running the forward reads through the Nonpareil tool. Split into asthma and age group and Two-way Type II ANOVA results shown. **C)** Bar plot of MetaCyc pathway copies per million (CPM) in all MARS samples annotated by HUMANnN pipeline, with horizontal length representing mean and bars the standard error. For all panels: N= 20 healthy children, 39 healthy adults, 19 children with asthma, 17 adults with asthma. **D)** Relative abundance stacked barplots of top abundant bacterial genera split by age group and asthma cohort. **E)** Simpson alpha diversity boxplots split by asthma and age group cohorts (2-Way Type II ANOVA). **F)** NMDS of Bray-Curtis Dissimilarity of species-level relative abundance grouped by age and asthma. **G)** Sequential PERMANOVA to test effect of demographics on beta diversity. Terms were input into the test as ordered from top to bottom of barplot. Dotted vertical line represents a p value of 0.05. Color scale is mapped to the R^2 value. **H)** Arcsine transformed relative abundance boxplots of differentially abundant species as determined by Maaslin2 with age group and race modeled as random effects. For all panels: N= 20 healthy children, 39 healthy adults, 19 children with asthma, 17 adults with asthma.

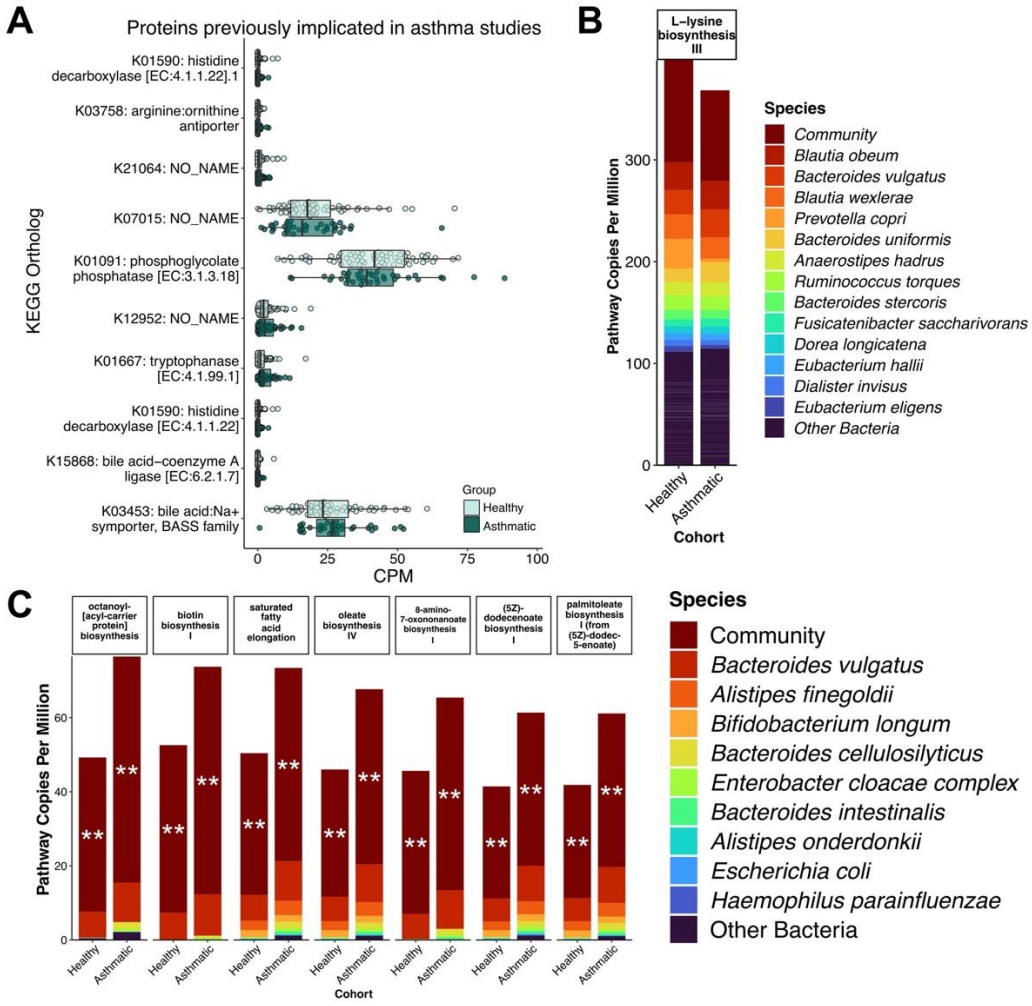

Figure 5.10. KEGG orthologs and differentially abundant MetaCyc fatty acid pathways.
A) Relative abundance of KEGG orthologs previously implicated in asthma. Copies per million (CPM) are counts normalized by gene size and read depth, then total-sum-scaled to one million. **B-C)** Stacked bar plots of differentially abundant pathways mapped to respective taxa including "Community" bin which accounts for the remaining reads that mapped to the pathway but not to any single species by MetaPhlAn3.0/HUMAnN3.0, averaged within asthma or healthy cohorts. **B)** L-lysine biosynthesis III pathway. Only top 13 taxa shown in addition to Community category. **C)** Seven fatty acid metabolism pathways differentially abundant in the asthma cohort. Only top 9 taxa shown in addition to Community category. Stars represent a q value < 0.05 of Wilcoxon tests between the Community pathway richness and *B. vulgatus*-encoded pathway richness. For all panels: N= 59 healthy, 36 individuals with asthma.

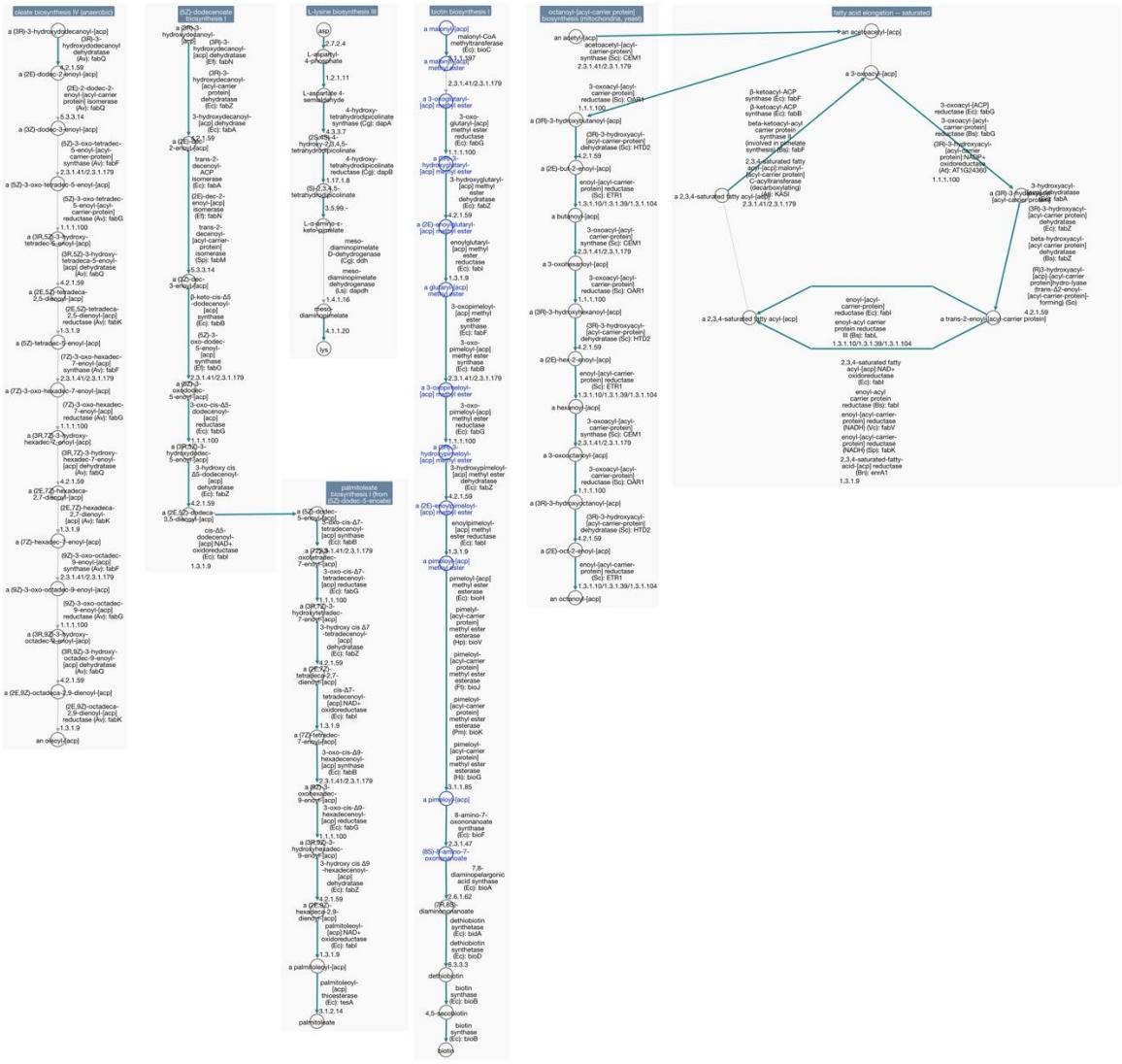

Figure 5.11. Pathway collage for differentially abundant MetaCyc pathways.
"PWY-6519: 8-amino-7-oxononanoate biosynthesis I" is completely overlapping with "BIOTIN-BIOSYNTHESIS-PWY: biotin biosynthesis I" and its steps are highlighted in blue text. Pathway collage made on MetaCyc browser tool.

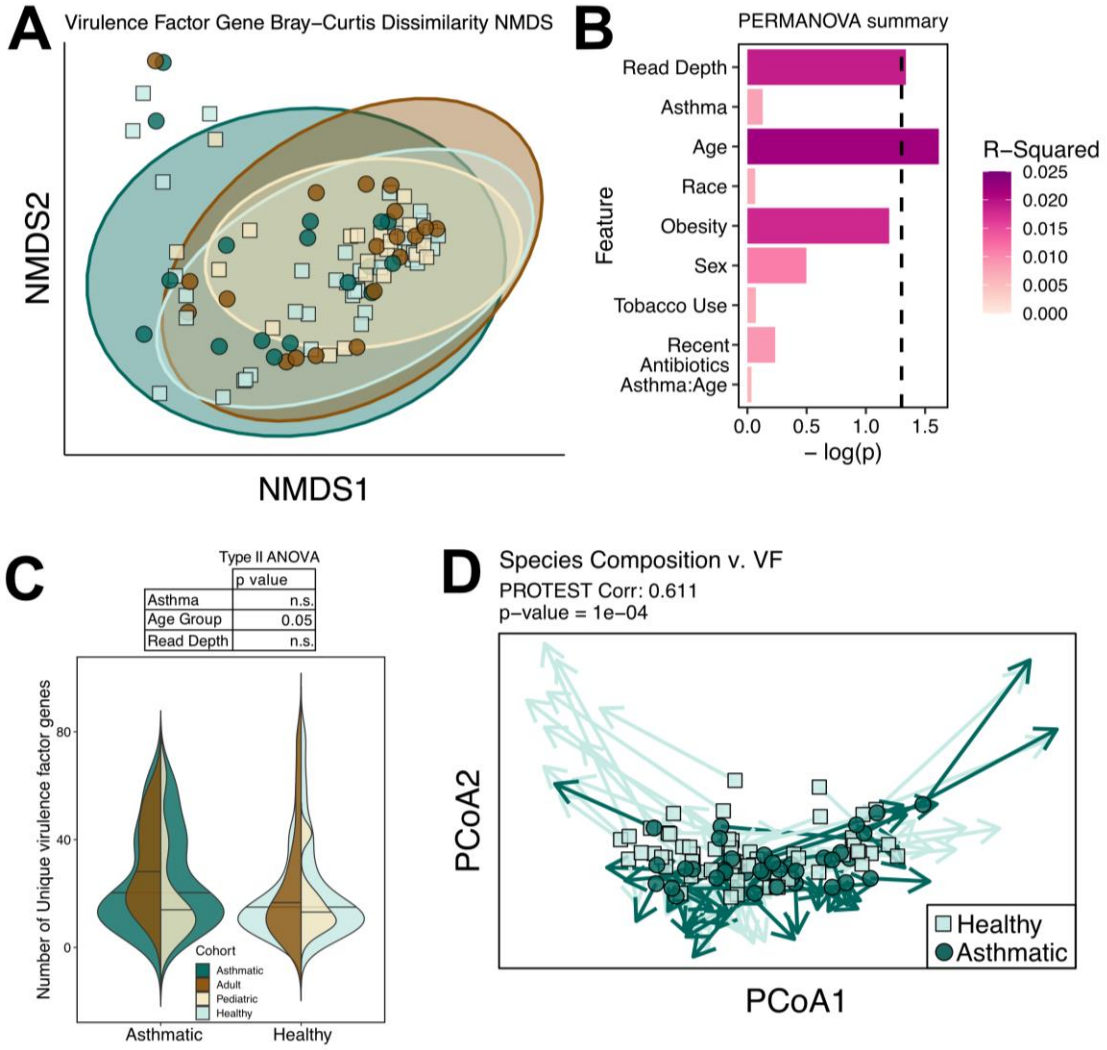

Figure 5.12. Gut virulence factor ecology shifts with age group but not asthma cohort.
A) Total-sum scaled RPKM Bray-Curtis Dissimilarity Non-metric Multidimensional Scaling (NMDS) plot labeled by asthma and age cohorts. Showing two axes out of 5 with stress value=0.09. **B)** Effect of demographic categories on virulence factor profile in A (by sequential PERMANOVA, input terms ordered from top to bottom of barplot). **C)** Stacked violin plots of virulence factor alpha diversity grouped by eithter healthy and asthma cohort (blue green colors in background) or age (brown colors in foreground). Two-Way ANOVA results shown in table above plot. **D)** Procrustes plot and PROTEST analysis between virulence factor profile Bray-Curtis dissimilarity distances and Metaphlan species relative abundance Bray-Curtis dissimilarity distances. Arrows connect the two data points belonging to identical samples. For all panels: N= 20 healthy children, 39 healthy adults, 19 children with asthma, 17 adults with asthma.

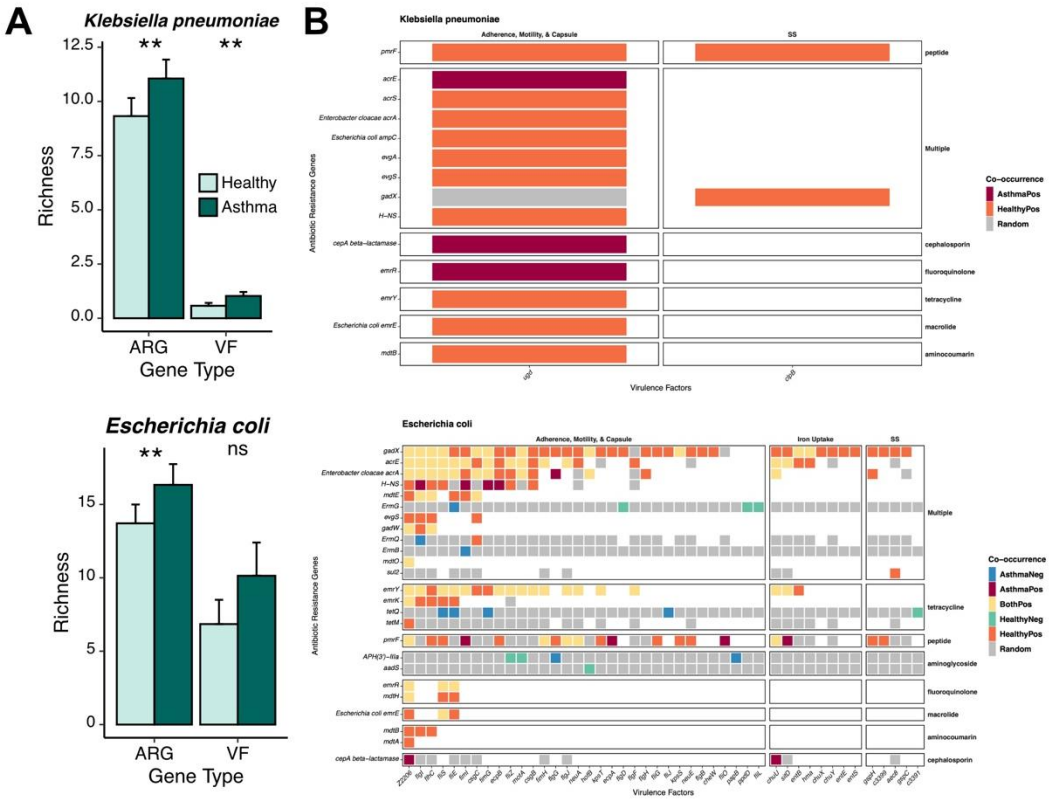

Figure 5.13. Asthma-associated ARG richness and ARG-VF co-occurrence relationships are observed within K. pneumoniae and E. coli.

A) Richness bar plots between antibiotic resistance genes (ARGs) and virulence factors (VFs) grouped by asthma status. **B)** Heatmap of the co-occurrence of each VF/ARG pair colored by the direction in which (positively or negatively co-occurring) and the cohort for which (asthma vs. healthy) the pair had a p-value less than 0.05 via R cooccur function. Blank squares were pairs filtered out due to a lack of observed co-occurrence. SS: secretion systems. (N=36 Healthy, 59 Asthmatic)

Chapter 6: Appendix II – Supplemental Tables

Tables S1-S8 can be found with the associated publication at *iScience*[157].

Tables S8-S13 can be found on bioRxiv[205].