Spring 2021

Decoding Gut Microbial Metabolites through G-Protein Coupled Receptor (GPCR) Activation

Phu Khat Nwe
Yale University Graduate School of Arts and Sciences, phukhat.nwe@gmail.com

Abstract

Decoding Gut Microbial Metabolites through G-Protein Coupled Receptor (GPCR) Activation

Phu Khat Nwe

2021

The microbiome encodes for a complex web of metabolites of which scientists are just starting to deconvolute. While a lot of focus has been on investigating the implications of the microbial metabolome on health and disease physiologies, we have merely uncovered the tip of the interactome of microbes and host G-Protein Coupled Receptors (GPCRs). Early literature has reported a plethora of short chain fatty acids fermented by dietary fibers acting as GPCR agonists. A few other studies have showcased that gut microbes produced N-acyl amides and secondary bile acids mimicking host ligands and therefore interacting with these GPCRs. Chapter 2 and 3 showcases the different strategies to mine GPCR agonists from the commensal microbiota.

Chapter 1 describes a collective review of the current state of microbiome research, particularly on the chemistry and biosynthesis of the microbiome. Despite advances with metagenomic, culturomic, transcriptomic and metabolomic approaches, about three-quarter of the microbiome space, thus termed "dark matter" still needs to be deciphered. This chapter highlights a few of the major classes of microbial metabolites and their interactions in the microbe-microbe and host-microbe axes including implications in GPCR physiology and health and disease dynamics.

Chapter 2 details an orthogonal platform bridging microbial metabolites and host GPCRs. Here, screening 144 bacterial supernatants against 316 host GPCRs utilizing a high throughput GPCR assay called the PRESTO-Tango led to a gamut of host-microbe, microbe-microbe and diet-microbe-host interactions. Through combinatorial screening, quantitative metabolomics and *in vivo* studies, we established a new

approach for parsing the microbiota metabolome and uncovered multiple biologically relevant host-microbiota metabolome interactions.

Chapter 3 illustrates *E. coli* functionalization of indole metabolites under acidic stress response. Employing metabolomic and small molecule discovery techniques led to the structural characterization of indole analogues which form spontaneously under acidic environment. A high throughput GPCR screen of these metabolites showcased potentially mild microbe-host GPCR interactions. A virulence assay alludes the potential inhibitory effects of indole oligomerization towards pathogenic microbes.

Decoding Gut Microbial Metabolites through G-Protein Coupled Receptor (GPCR) Activation

A Dissertation

Presented to the Faculty of the Graduate School

Of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Phu Khat Nwe

Dissertation Director: Jason Michael Crawford

June 2021

Decoding Gut Microbial Metabolites through

G-Protein Coupled Receptor (GPCR) Activation

List of Figures, Tables, and Schemes

Chapter 1

Chapter 2

Acknowledgements

The GPCR collaboration started off as a side project, but has slowly evolved into my thesis. In parallel, the collaboration between the Crawford lab and the Palm lab grown tremendously to a wonderful collaboration. I thank Jason Crawford for being my advisor and allowing me to work on this project. I am especially thankful to Noah Palm for his openness and reception to different ideas and I really admire his eloquence in writing and the ability to elevate the scientific storytelling. I am also very grateful that I get to work with Haiwei Chen for his humbleness, brilliance and scientific rigor. I always learn something new from my frequent conversations with Haiwei. Lastly, I thank Sarah Slavoff for always being very encouraging and engaging with me for thought-provoking conversations.

My PhD journey would've been very different if I didn't get support from my lab members and the really special people who I get to call my friends and my partner. There have been a lot of ups and downs along my graduate school journey, and my partner Pol Arranz-Gibert has been my strongest support, and for that, I am eternally grateful. Special thanks to Christina Cho and Koen Vanderschuren who lifted me up and empowered me with kindness and compassion when I was really down. I am grateful to Hyun Bong Park and Joonseok Oh for being like my second mentors. Thank you Jaymin Patel for always being very selflessly helping with anything that I need. I thank the rest of the lab members for always being a lighthearted bunch. Lastly, I would like to thank my McDougal partners, Meaghan Sullivan, Courtney Smith and Becky Lacroix for being strong female peers. We went through various different hurdles and you empowered me with your creativity and attitude. Because of you, graduate school is a little more bearable, and I will never forget the little kindness that each and every one of you has shown towards me.

Chapter 1

Chemistry and Enzymology Encoded by the Human Microbiome

The following chapter is adapted from my published work as a second author (Shine, E.E.; **Nwe, P.-K**.; Crawford, J.M. Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, Elsevier, **2019**, ISBN 9780124095472

1.1 Introduction

The Russian scientist Élie Metchnikoff first proposed the role of gut bacteria on host physiology, predating the modern era of microbiome research by more than 100 years.[1] His general hypotheses that gut bacteria are modulators of host physiology, and that dysregulation of gut homeostasis maintained by certain bacterial species could lead to a disease state mostly stand the test of time, even now as we are firmly rooted within the 21st century.[1] We now understand the collection of bacterial, archaeal, fungal, and viral species residing within eukaryotes (termed "microbiota," "microbiome" referring to all the genes they contain) – to be immensely dense, vastly diverse among individuals, and associated with human diseases.[2-4] Collections of microbes at every surface within the body – skin, respiratory, urogenital, and gastrointestinal tracts – can influence pathogen infection rates, stimulate immune system development, and contribute to inflammation.[5-6] Studies have shown that individuals can be >90% different in terms of gut microbiome species composition, suggesting diversity in chemistry and enzymology and the biological functions that they regulate.[7-8]

The implications of humans carrying within each of us a symbiotic collection of organisms consisting of billions of different cells, each collective different from our neighbor, is enormous (Figure 1). For one, it is hypothesized that we are more microbe than *Homo sapiens:* in terms of gene and cell numbers, although the microbial cells are much smaller than the human cells. Estimates of bacteria-to-human cell ratios have ranged in the past from 10:1 to more recently 1:1.[9] But more importantly, no matter what the

ratio, the genetic capacity of information encoded in the human genome is vastly outweighed by the contribution from the microbiome. Indeed, the combined human microbiome encodes roughly 150-fold more genes than our own genome.[10] Furthermore, the genetic diversity within bacteria is astonishing. In the human gut alone, nearly 10 million non-redundant reference genes have been identified.[11] Bacteria are metabolic and chemical factories capable of utilizing a far richer and more diverse array of substrates than our own cells. Most of the details behind these chemical transformations remain hidden among thousands of uncharacterized species and genes, further complicated by the vast, mysterious network of microbe-

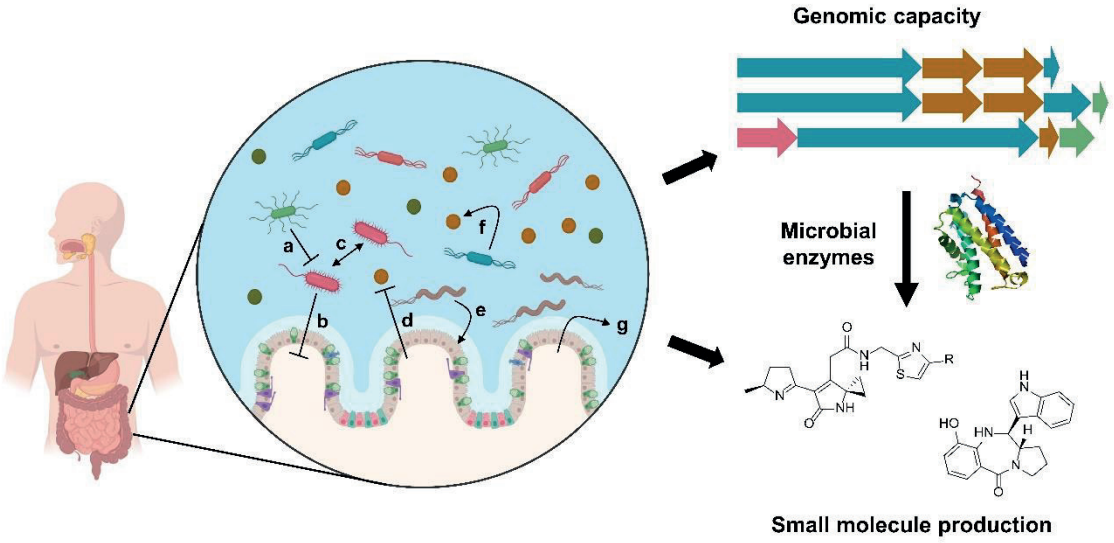


Figure 1. General overview of enzyme-mediated host-microbial interactions discussed in this chapter. First, bacterial inter- and intra-species competition (a) exists in many forms, such as the secretion of specialized metabolites, allowing for fitness advantages and niche expansions. Second, pathogenic or parasitic factors exhibiting negative effects on the host (b) are of much interest for their immunomodulatory and/or carcinogenic roles. Next, bacterial communication (c) mediated by signaling molecules can coordinate group behaviors among bacteria within species or across taxonomic lines. Pathogen exclusion by the host (d) may occur through the production of antimicrobial factors, ensuring that pathogens are prevented from penetrating the mucosal layer. Commensals and symbionts use enzymatic machinery to confer benefits to the host (e), *i.e.*, the synthesis of vitamins and processing of host bile acids. Evolved cooperation between bacterial species (f) that depend on nearby synthesis of "public goods" may explain why many bacterial species within the microbiome cannot be cultured independently outside of the host. Finally, bacterial species can metabolize foreign chemical scaffolds consumed by the host (g). Importantly, we highlight in the context of the microbiome, how interactions (a) and (c) may affect the degree to which interactions (b) and (d) present disease.

microbe and microbe-host interactions in a dense and dynamic community. We are just beginning to explore the tip of the iceberg in regard to how genetic diversity within the microbiome leads to metabolic and chemical diversity, and how this diversity contributes to host phenotypic consequences. Detailed mechanisms often remain obscure, and ironically, for many illnesses we find ourselves not too far from where Metchnikoff himself stood in 1890 in terms of *how* the microbiome precedes or predicates disease at the molecular level.

After years of sequencing work to establish the presence or absence of bacterial community members between "healthy" and "diseased" patients, microbiome researchers have only more recently begun to focus their attention on the hunt for molecular mechanisms that might underlie differences in presented pathologies. Our framework for the understanding of disease still largely lies with Koch's postulates, which are based on the pathogenesis of typically one microbe causing disease and that microbe conferring disease when inoculated in other healthy subjects.[12-13] Yet Koch's postulates, for example, have failed to identify one particular bacterial species involved for all patients with obesity or inflammatory bowel disease (IBD) – two conditions which multitudes of studies show are heavily influenced by the composition of the microbiome.[14-17] Molecular Koch's postulates first put forward by Stanley Falkow in 1988, which focuses on the underlying molecules of disease including cytotoxins, provides a framework for examining pathogenesis at the molecular level beyond species-level classifications, but these revised postulates still fail to take into account the complex interactions among different microbiome members that can regulate the pathogenic molecules causing disease and how the function of those molecules can also vary with differing host susceptibilities — *i.e.*, other environmental exposures and host genetics.[12-13] Consistent with these challenges, microbiome research has shown that mere presence or absence of a bacterial species or a bacterial metabolic product is not enough to account for disease incidence alone, and seemingly healthy patients can have high carriage rates of these markers.

The field's partial focus on pathogens is well-warranted. Even for complex IBDs, it was shown that bacteria penetrating the mucosa – analogous to pathogen invasion – were coated with immunoglobulin A

(IgA) in this space, and that human IBD-derived bacteria identified to have an IgA coating could cause disease when transplanted into healthy mice.[18-19] Like other immune mechanisms intended to protect the host, IgA can be co-opted by select microbiome members to promote mucosal colonization instead.[20] That said, pathogenic interactions still only account for a fraction of microbe-associated impacts on human physiology. The challenge is that many bacteria can possess these pathogenic qualities which may depend on the microbial composition, diet, and the host. Our understanding of diseases that have such high interpersonal variation as cancer, autoimmune disease, and metabolic disorders can no longer focus on just pathogenesis, but must also take into account broader host-microbe interactions (Figure 1).[13]

Host-microbe symbiotic interactions can be classified as mutualistic (benefiting both organisms), commensalistic (benefiting one symbiont), or parasitic (benefiting one organism and harming the other).[21-22] Microbes can exist as obligate partners, depending on each other for nutrients, or can exist as enemies in combat.[23-26] The chemical and molecular underpinnings of all these interactions, both individual and in complex metabolic exchange markets, are critical for understanding the mysteries of the microbiome and its impact on human health. In this chapter, it is the goal to introduce the reader to the chemistry and function of biosynthetic pathways found in representative microbiota members that provide an entry point for understanding human disease within the context of bacteria-mammalian symbioses that have evolved for millennia.

1.2 Natural product biosynthesis encoded in the microbiome

This chapter will focus on the biosynthesis and function of representative specialized metabolites encoded in the microbiome. Molecules associated with pathogenesis, such as siderophores and (lipo)polysaccharides, have been extensively studied for many decades and have been reviewed elsewhere.[27-30] We also do not cover very low-molecular weight molecules of major functional importance in the microbiome, such as primary metabolites including short chain fatty acids,[31-33] reactive nitrogen species,[34] and reactive oxygen species.[34-35] Instead, this section will introduce and discuss five major, classical classes of natural products produced by bacterial members of the microbiome: ribosomally-

encoded and posttranslationally-modified peptides (RiPPs), nonribosomal peptide synthetase (NRPS) products, polyketide synthase (PKS) products, hybrids of the latter, and *N*-acyl-amides. We highlight selected examples in recent literature. Particular attention will be paid toward identifying these types of chemistries from functional metagenomics or bioinformatic genome mining studies, as well as their interactions with other bacterial species that could have implications for community dynamics and/or disease severity.

1.2.1 RiPPs

The term RIPPs (or ribosomally-encoded and posttranslationally-modified peptides) is a relatively recent nomenclature designation given to a diverse set of modified peptides that are derived from ribosomal protein synthesis.[36] Ribosomal precursor peptides are subjected to diverse posttranslational modifications, including proteolytic cleavage, alkylation (*e.g.*, radical SAM-dependent methylation), prenylation, dehydration, heterocycle formation, redox modifications, and diverse macrocyclization reactions among others. The resulting structures are endowed with better target recognition, decreased susceptibility to protease degradation, and increased structural stability relative to standard peptides.

Almost all RiPPs are synthesized from a structural precursor peptide, typically 20-110 residues in length that includes a signal peptide sequence and leader peptide sequence flanking a core peptide at either the N- or C-terminus (Figure 2). The leader peptide facilitates recognition of a fairly common set of RiPP processing enzymes and sometimes directly participates in catalysis.[36] Signaling sequences are more or less conserved, while the processing enzymes can tolerate mutations leading to considerable structural variation within the core peptide. Together, these features highlight the evolutionary advantage of accessing high chemical diversity at a low genetic cost.[36]

As with many metabolic pathways and natural products, in prokaryotes, the biosynthetic genes for RiPPs are typically clustered together at a single genetic locus, enabling identification of complete RiPP pathways from genomic data using bioinformatic approaches.[37-40] Analysis of ~65,000 prokaryotic genomes identified approximately 30,000 candidate RiPP biosynthetic gene clusters, which included open reading

frames for the precursor peptide as well as associated tailoring enzymes.[38] RiPPs are nearly universally distributed throughout the prokaryote domain of life, especially in the human microbiome. Indeed, in human-associated microbiomes, RiPPs are slightly enriched over environmental microbiomes, and they are broadly distributed across every sampled body site.[37]

Traditionally, RiPPs have been grouped into distinct subclasses based on shared biosynthetic or structural features. Two subclasses that were particularly abundant within human microbiome samples were the thioazole/oxazole-modified microcins (TOMMs) and thiopeptides.[37] Notably, these RiPP families have potent antibacterial activities against Gram-negative and Gram-positive isolates, respectively. Mounting evidence supports that RiPPs are utilized by microbes *in vivo* during inter- and intra-species competition for resources and establishment of niche colonization.[41] Thus, RiPPs are essential for understanding microbe-microbe interactions and how community structure and dynamics are established. For instance, an isolated thiopeptide, lactocillin, from a vaginal isolate of *Lactobacillus gasseri* was found to harbor antibiotic activity against a wide-range of Gram-positive organisms, but it was inactive against vaginal commensals *Lactobacillus jensenii* and *Lactobacillus crispatus*.[37] Correspondingly, production of lactocillin likely helps to maintain the vaginal microbiome community structure that is dominated by just four *Lactobacillus* species.[42] In contrast to the almost uniform vaginal microbiome, a healthy steady-state gut microbiome is rich in species diversity, which is sustained by microcins that have a narrow-spectrum of activity toward closely related competitors. As an illustration, the gut commensal *E. coli* Nissle 1917 is able to produce microcins M and H47 under inflammatory conditions to limit the expansion of closely related Enterobacteriaceae and exert select probiotic effects.[43] To illustrate the diversity within this family of microbiome-encoded small molecules, we will focus on the structural and enzymatic features of the two related RiPPs, thiocillin and microcin B17/P1.

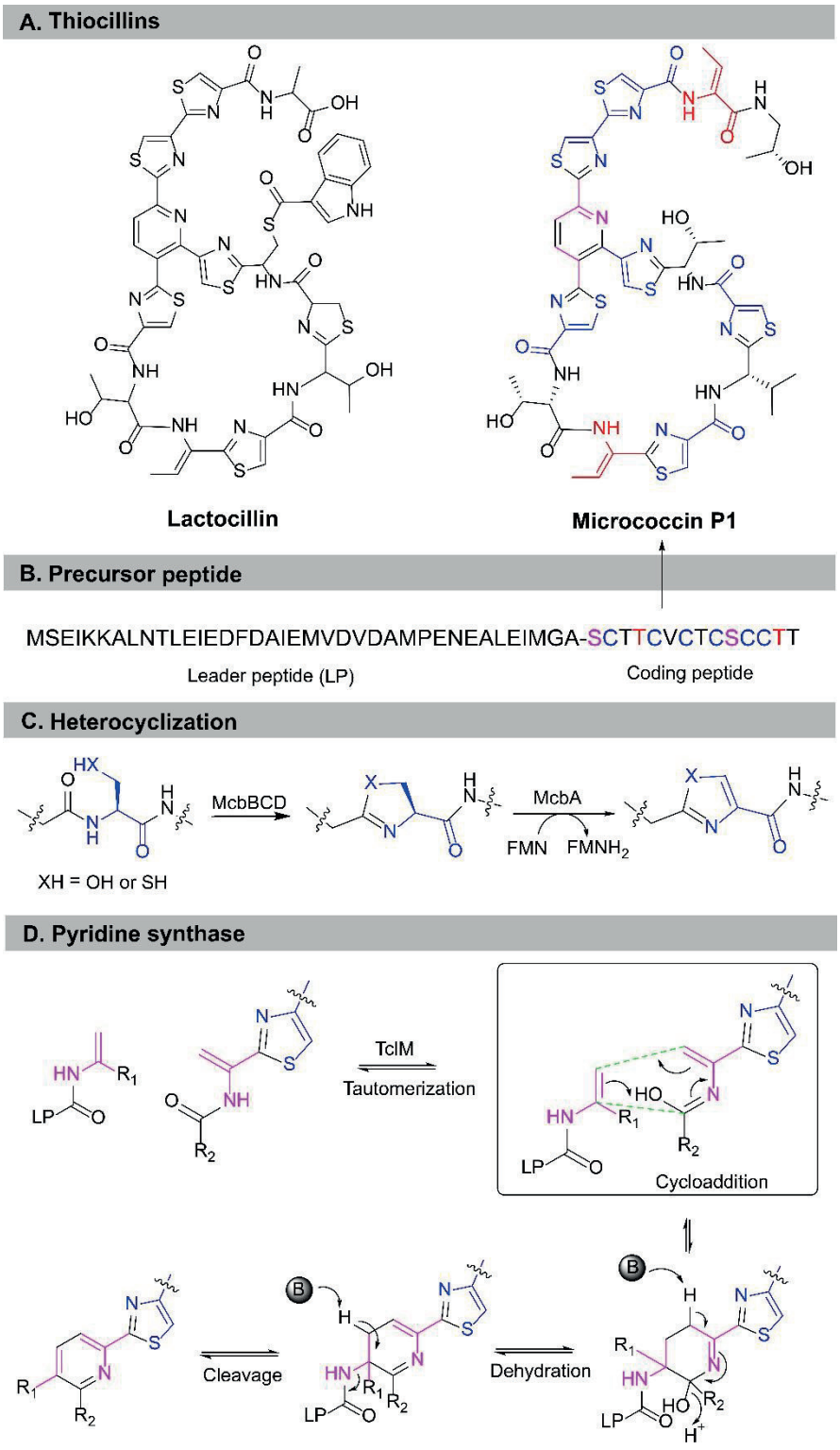


Figure 2. Proposed biosynthesis of thiocillins as a representative RiPP pathway. (A) Structures of lactocillin and micrococcin P_1 (B) Precursor peptide for micrococcin P_1 (C) Proposed heterocyclization reaction of RiPPs, focusing on microcin B17 as an example (D) Proposed pyridine synthase reaction of thiocillins proceeds through a Diels Alder reaction.

Microcins

Microcins have traditionally been the name for ribosomally-derived peptides from Gram-negative bacterial species. Chemically, microcins are a part of the linear azol(in)e-containing peptide family defined by the presence of thiazole and (methyl)oxazole heterocycles in an otherwise linear core peptide.[38] Microcins can be further subdivided into class I, IIa, and IIb according to size and genetic characteristics[44]. Genetic loci that encode microcins typically include the precursor peptide, associated modification enzymes, a self-immunity gene to protect against self-poisoning, and the genes needed for export out of the cell.

Class I microcins generally have a mass below 5 kDa and their self-resistance genes are often located distally to the biosynthetic gene cluster, an exception to the gene "clustering" phenomenon in bacteria and fungi.[44] An example of a class I microcin is that of microcin B17 (MccB17) encoded by strains of *E. coli* bearing a 70 kb conjugative pMccB17 plasmid.[45-46] A three-gene system consisting of *mcbB*, *mcbC*, and *mcbD* associate to form the microcin B17 synthetase which transforms 14 residues within the backbone into heterocylic moieties, thus activating the molecule for antibiotic activity.[47-48] McbD, as part of the McbBCD complex (Figure 2C), initiates the cyclodehydration of serine and cysteine residues by ATP phosphorylation.[49] The zinc-dependent McbB cooperates with McbD to catalyze the formation of oxazolines and thiazolines.[49-50] These heterocycles are further oxidized/aromatized by the flavin-dependent dehydrogenase McbA. The leader peptide is then cleaved off by chromosomally encoded proteases (TldD/E) prior to exportation.[51-53] The resulting, fully processed microcin B17 is a narrow-spectrum DNA gyrase inhibitor against several members of Enterobacteriaceae.[54-55]

Class II microcins contain processed peptides that are within 5-10 kDa. Class IIa refers to microcins containing disulfide bonds, and class IIb microcins may contain additional C-terminal modifications, such as conjugation to a siderophore.[44] As mentioned above, *E. coli* Nissle 1917 encodes microcin H47, which has been speculated to be attached to a siderophore (an iron-scavenging molecule).[56-57] Indeed, microcin H47 has been shown to be synthesized and possess antibacterial activity during iron starvation conditions.[43] Microcin H47 is thought to target recipient bacteria through their cognate siderophore receptors.[56, 58] This

"Trojan-horse" mechanism of action leads to species selectivity, facilitating competition with closely related Enterobacteriaceae that share similar resources and locale.[58-59]

Thiocillin

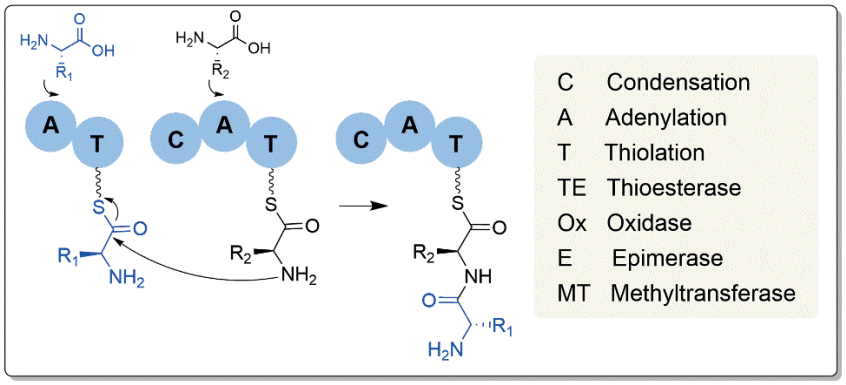
Thiocillin belongs to the thiopeptide class of RiPPs due to the presence of multiple thiazoles connected to a central six-membered nitrogenous ring – in this case, a pyridine. Thiocillin targets the bacterial 50S ribosome and exhibits potent antibacterial activity against a number of drug resistant pathogens.[60-61] Its biosynthesis in the native host *Bacillus cereus* begins with ribosomal translation of a 52-amino acid precursor peptide that contains a 38-residue leader sequence and a 14-amino acid core peptide (Figure 2A,B).[62] In addition to the four copies of the pre-processed open reading frame (*tclE-H*), the gene cluster additionally contains two lantibiotic-type dehydratases, cyclodehydratase, dehydrogenase, four proposed resistance elements, and other tailoring enzymes that modify individual residues in a stochastic manner.[63] Combinations of some of these modifications result in a panel of structural variants produced by *B. cereus*: thiocillin I, II, III[64], and IV[62], micrococcin P_1 and P_2[65], and YM-266183 and 266184[66].

Isolation of peptide precursors from individual gene deletion strains within the thiocillin gene cluster suggested that key posttranslational modifications such as thiazole formation, serine and threonine dehydration to form dehydroalanine-type residues, and C-terminal oxidation/decarboxylation precede the final maturation step in pyridine ring formation.[63, 67] The enzyme TclM was demonstrated to catalyze the formal [4+2] cycloaddition between two dehydroalanine moieties (Figure 2D)[63]. Following cycloaddition, subsequent dehydration and elimination of the 38-amino acid leader peptide results in the aromatic pyridine product. TclM represents a rare class of Diels-Alderases that recognize a heavily decorated substrate in order to catalyze a non-spontaneous reaction. Base catalysis was demonstrated to enhance enzyme substrate promiscuity, displaying value for its use as a biocatalyst to make thiocillin analogs.[63, 67]

1.2.2 Nonribosomal peptides

The microbiome also encodes peptides derived from a parallel and complementary biosynthetic route to the RiPPs through the use of nonribosomal peptide synthetases (NRPSs). Whereas RiPPs use ribosome tRNA binding sites (E, P, A) and a mRNA to construct peptides,[68] NRPS enzymes instead divide the chemical reactions across multiple catalytic proteins or domains often within multidomain enzymes. While this process requires much more cellular energy and coding capacity within the genome relative to ribosomal protein synthesis, it has a much larger structural and functional repertoire.[69]

The majority of NRPSs represent multi-modular enzymes that act in an assembly line fashion. Individual amino acids are activated as AMP esters by an adenylation (A) domain and tethered as thioester linkages to the flexible phosphopantetheinyl arm of a thiolation (T) domain, a peptidyl carrier protein (PCP) (Box 1/Figure B1).[70-71]



Box 1/Figure B1. General schematic of amino acid activation and condensation in NRPS biosynthesis. Additional tailoring enzymes modify the growing chain (domains in grey).

In addition to standard proteinogenic substrates, NRPS A-domains are able to select from hundreds of nonproteinogenic amino acids, dramatically expanding their substrate pool relative to aminoacyl-tRNA synthetases used for ribosomal peptide synthesis.[72] The condensation (C) domain catalyzes peptide bond formation between the thioester of the aminoacyl-T domain of an upstream module with the amine of the aminoacyl-T domain of a downstream module, extending the peptide one unit at a time and shuttling the growing peptide downstream.[70-71] The growing peptide chain can be further modified with tailoring

domains, such as epimerases (E), cyclases (Cy), oxidases (Ox), reductases (R), and methyltransferases (MT). The terminal module typically invokes a thioesterase (TE) domain to release the product via hydrolysis or macrocyclization. However, reductive release also represents a common termination strategy, which we will highlight below. Finally, released products can be subjected to additional post-assembly line transformations, including glycosylation, redox modification, and proteolytic cleavage.[69]

Some of the most well-studied NRPS-derived products include siderophores, which are known virulence factors that chelate iron with exceptional affinity and promote iron acquisition.[73-74] Antibiotics, such as penicillin, are also well-studied NRPS-derived products.[75-76] While the majority of antibiotic and other specialized metabolic pathways have been characterized in soil- and sediment-dwelling microbes, where they are used for nutrient competition and cell-to-cell signaling,[77-80] examples of NRPS products encoded within human-associated microbiota species have only more recently been elucidated.

Bioinformatic analysis of biosynthetic gene clusters within the human microbiome suggests that the abundance of NRPS-derived products within human-associated bacterial species are lower than those in other prokaryotes, especially soil and aquatic bacteria.[29, 37, 81] However, abundance should not be conflated with biological importance. Of those microbiota harboring NRPS-containing pathways, many have undergone genome reduction and harbor chromosomes as small as 2-3 Mb, suggesting that the often large NPRS pathways maintained are crucial for survival and colonization within the gut. With one notable exception (pyrazinones), most NRPS-containing biosynthetic gene clusters were found in <5% of 752 healthy, human microbiome samples.[37] Nearly all of the bioinformatically-identified NRPS gene clusters have never been described before, meaning that these pathways represent a rich source of novel molecules that may have been selected to function in the human environment.[29, 37, 81]

Lugdunin

Using a classical bioassay-guided metabolite characterization approach, Zipperer *et al.* identified a nonribosomal peptide encoded within a *Staphylococcus* isolate from a collection of bacteria isolated from the nasal passages of healthy volunteers[82]. They found that *S. lugdunensis* IVK28 harbored antibiotic

activity against methicillin-resistant *Staphylococcus aureus* (MRSA) under iron-limiting conditions. Transposon mutagenesis of strain IVK28 led to the identification of an approximately 30 kb operon that, when mutated, led to a complete loss of antibiotic activity. This operon contained four NRPS genes, a 4′-phosphopantetheinyl transferase (PPTase) that posttranslationally activates the carrier proteins (T domains), a monooxygenase, a type-II editing thioesterase that removes stalled intermediates to improve biosynthetic processing, an annotated regulator, and a putative biosynthetic tailoring enzyme of unknown function. Collectively, each enzyme had less than 35% identity to any other described enzyme and was found exclusively in *S. lugdunensis.*[82]

Isolation and characterization of the antibiotic revealed a new NRPS product termed lugdunin, a cyclic peptide containing a thiazolidine heterocycle and three D-form amino acids. The authors hypothesized that the thiazolidine ring is formed upon reductive release of a heptapeptide scaffold, whereby the amine of the *N*-terminal L-Cys nucleophilically attacks the C-terminal L-Val carbonyl to form a macrocyclic imine (Figure 3).

Nucleophilic attack by the cysteine thiol group to this Schiff base results in a five-membered thiazolidine heterocycle. The structure of lugdunin was larger than expected from bioinformatic analysis, as there were only five adenylation (A) domains encoded within the operon. Iterative incorporation of three consecutive valine moieties with alternating L- and D-configurations by an iterative LugC adenylation domain was proposed to account for this discrepancy - resulting in a heptapeptide metabolite.[82]

Lugdunin was demonstrated to possess antibiotic activity against Gram-positive bacteria *in vitro*, and it was also effective in a mouse skin *S. aureus* infection model. Either purified lugdunin or its producer *S. lugdunensis* applied to the infection site led to a strong reduction of viable *S. aureus* within the epidermis. Correspondingly, the authors found a strong inverse relationship between *S. lugdunesis* and *S. aureus* by analyzing nasal swabs of patients, suggesting that lugdunin is indeed expressed within the nasal passages and can lower the carriage rate of *S. aureus.*[82]

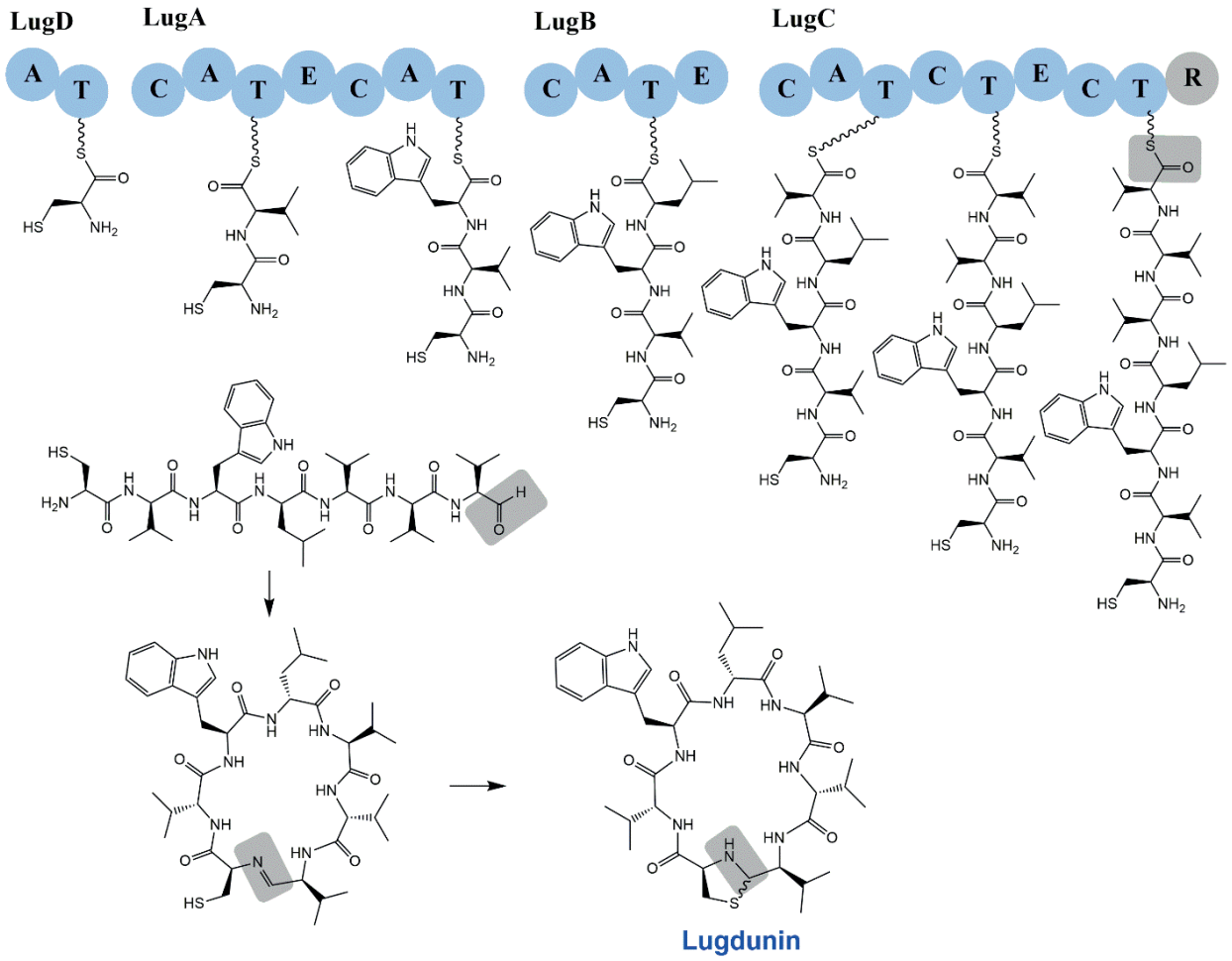


Figure 3. Biosynthesis and structure of lugdunin. The linear precursor is released via reduction and spontaneously cyclized to form lugdunin. Aldehyde chemistry is typical of reductive release reactions catalyzed by reductase (R) domains, shaded in gray

Tilivalline

Antibiotic-associated hemorrhagic colitis (AAHC) is a disease phenotype accompanying perturbed gut microbial homeostasis as a result of antibiotic treatment. In the absence of *Clostridium difficile*, AAHC is found to be associated with overgrowth of a pathobiont strain, *Klebsiella oxytoca*.[83-85] Schneditz *et al.* first described tilivalline, an enterotoxic nonribosomal peptide produced by this pathobiont, which is present in 2-10% of healthy human GI microbiome samples.[86] Following brief penicillin therapy, *K. oxytoca* claims an opened environmental niche and this dysbiotic population shift induces acute colitis, as demonstrated in a mouse model. While tilivalline is thought to be the chief virulence factor of *K. oxytoca* responsible for

disease pathology, tilivalline production provided no advantage for colonization in a rodent model for antibiotic-induced dysbiosis. The genomic context of tilivalline shares >50% sequence similarity to pyrrolobenzodiazepine biosynthesis genes present in Gram-positive soil bacteria, such as actinomycetes.[86] Indeed, tilivalline was characterized as a pentacyclic pyrrolobenzodiazepine (PDB)[86], a member of a broader structural class known to bind and alkylate DNA.[87-89] Tilivalline is the first PBD shown to be formed by the enteric microbiota and thus represents a new class of enterobacterial toxins.[86]

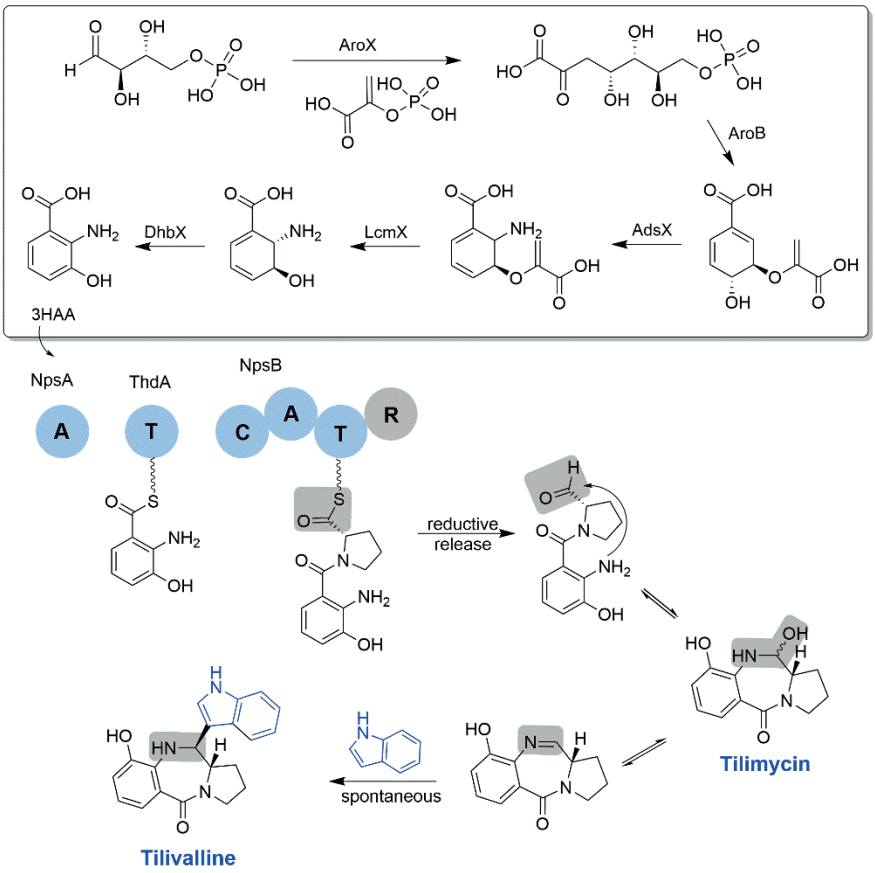


Figure 4. Biosynthesis and structure of tilivalline from 3-hydroxy-anthranillic acid (3HAA). Proposed biosynthesis of precursor 3HAA is illustrated in upper box. NRPS gene cluster containing NpsA, NpsB and ThdA utilizes 3HAA and L-proline as substrates to produce tilivalline. Cytotoxic tilimycin is a stable intermediate which results from cyclization of the aldehyde precursor

Tilivalline is produced from two transcriptional operons: one encoding the synthesis of a hydroxyanthranilic acid substrate[90-91] and one encoding the NRPS operon containing a standalone adenylation domain, thiolation domain, and one NRPS termination module with a C-A-T-R domain architecture (Figure 4).[86] NRPS adenylation domain NpsA loads the precursor 3-hydroxy-anthranillic acid (3HAA) onto thiolation domain ThdA. The condensation of 3HAA and L-Pro by NpsB is followed by an NADPH-dependent reductive release. The linear aldehyde undergoes a spontaneous cyclization event to give rise to tilimycin (or kleboxymycin), which was identified as another enterotoxin from the pathway.[90-91] A tryptophanase gene (*tnaA*) was reported to convert L-Trp to free indole, which then serves as a nucleophile to form the final product tilivalline in a spontaneous Friedel-Crafts-type alkylation reaction (Figure 4).[90, 92] Systematic gene knockout and metabolite profiling experiments established a 4-hydroxyphenylacetate-3-monooxygenase (*hmoX*), 2-amino-2-deoxy-isochorismate synthase (*adsX*), isochorismatase (*icmX*), 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (*dhbX*), 2-keto-3-deoxy-D-arabino-heptolosonate phosphate synthase (*aroX*), and distal gene 3-dehydroquinate synthase (*aroB*) to be responsible for 3HAA substrate biosynthesis.[90-92]

While tilivalline and tilimycin share the same pyrrolobenzodiazepine core, they exert distinct host phenotypes.[90-91, 93] The most recent study by Unterhauser *et al.* established a detailed mode of action for both enterotoxins. As predicted from the PBD class of cytotoxins, the intermediate tilimycin binds and stabilizes the minor groove of double-stranded DNA, promoting dysregulation of DNA damage responses and replication stress. This causes DNA double strand breaks and G_1/S phase cell cycle arrest both in cell culture and *in vivo*. Tilivalline, on the other hand, was reported to stabilize microtubules which constitute the cytoskeleton of tubulin. Binding of tilivalline to microtubules inhibits tubulin depolymerization, leading to mitotic arrest. Treatment of tilimycin or tilivalline to colon epithelial cells showed that both cytotoxins can induce cellular apoptosis through independent modes of action. Collectively, both enterotoxins are able to induce apoptotic disease phenotypes which are hallmarks of AAHC.[93] Given that up to 10% of humans harbor tilivalline-producing *K. oxytoca* and are asymptomatic,[85] decoding resistance mechanisms to keep *K. oxytoca* in check or neutralize tilivalline activity represent an interesting future direction.

Pyrazinones

A conserved NRPS pathway in *Staphylococcus aureus* and other skin-associated staphylococci encodes a family of pyrazinone natural products such as phevalin, tyrvalin and leuvalin.[94-95] Phevalin was originally discovered in 1995 from soil actinomycetes and was shown to inhibit calcium-dependent calpain proteases in eukaryotes.[96] A study in human keratinocytes also showed phevalin to be involved in modestly regulating its gene expression.[97] Similar pyrazinones, isolated from *Photorhabdus* species, were also shown to have calpain protease inhibition.[98] As such, it is worth noting that an additional class of synthetic pyrazinones are currently under clinical trials for treatment of hepatitis C virus NS3 protease inhibition.[99-100] Despite these roles, the exact function of pyrazinones is still unknown. No antibiotic activity has been reported by these compounds, and they have been hypothesized to have a role in signaling and pathogenesis of *S. aureus*.[94, 101] Indeed, a structurally related compound 3, 5-dimethylpyrazin-2-ol (DPO) was described to act as an autoinducer in *Vibrio cholerae* quorum sensing.[102]

Similar NRPS pathways with a reductase domain that encode dihydropyrazinones and pyrazinones were reported in several human gut members from Firmicutes and some Gram-negative strains. These pyrazinone-encoding gene clusters were present in >90% of fecal samples from the Human Microbiome Project, showcasing the wide distribution of these genes and the prevalence of dipeptide aldehydes and pyrazinones in the human gut.[103] Select pyrazinones and precursor dipeptide aldehydes are illustrated in Figure 5AB.

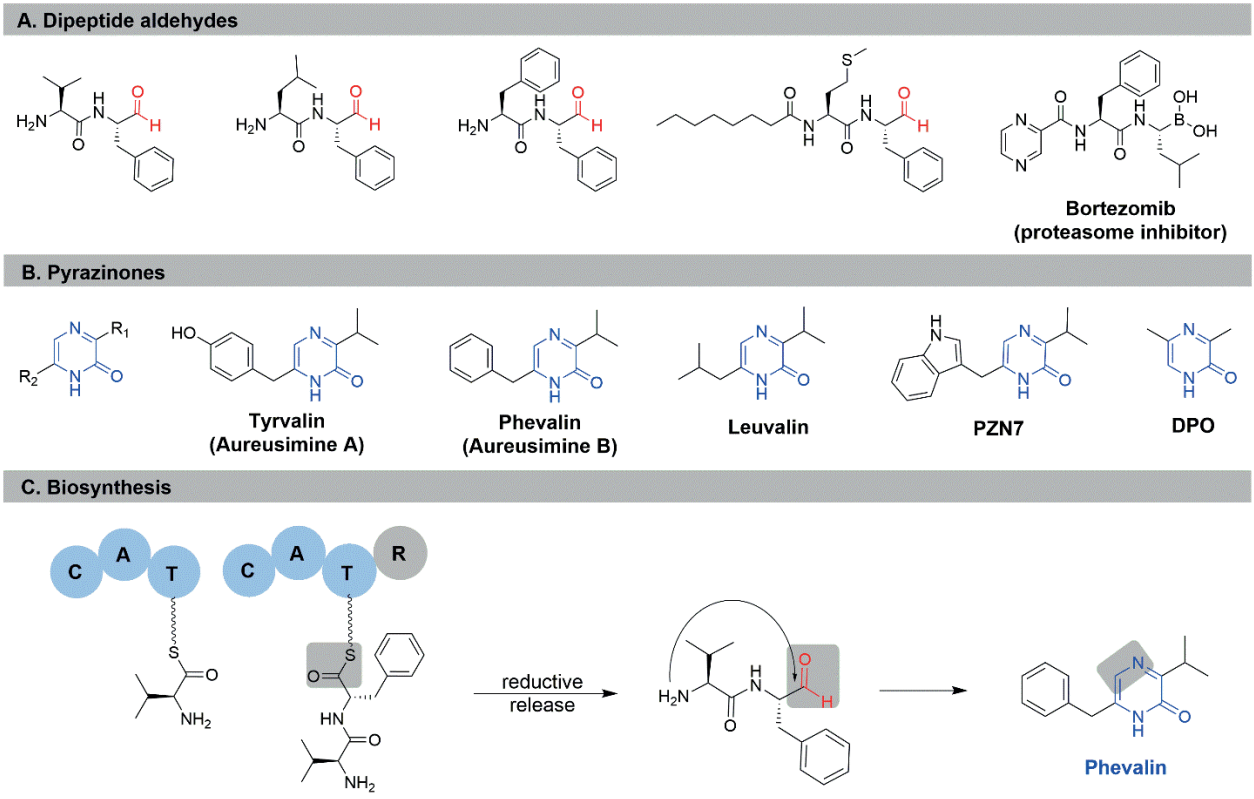


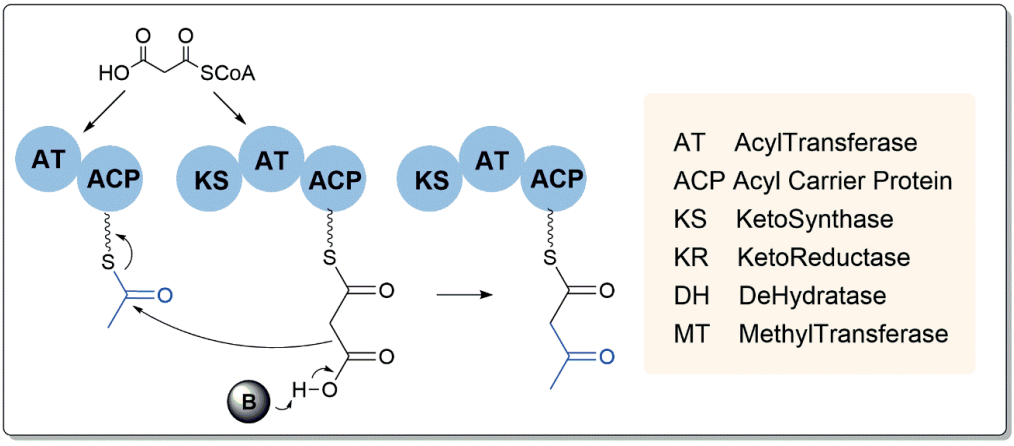
Figure 5. (A) Representative dipeptide aldehydes. The aldehyde moieties highlighted in red represent the pharmacophore for protease inhibition. For example, bortezomib, a clinical proteasome inhibitor, is a structural mimic of the dipeptide aldehyde backbone. (B) Representative pyrazinones. Pyrazinone core is highlighted in blue, and R_1 and R_2 groups are derived from various amino acid side chains. (C) Biosynthesis and structure of phevalin. PznA (AusA) encodes A_1-T_1-C-A_2-T_2-R domains as shown while PznB (AusB) encodes a phosphopantetheinyl transferase.

The pyrazinone precursor dipeptide aldehydes are themselves a noteworthy class of protease inhibitors. Dipeptide aldehydes are usually encoded by dimodular NRPS systems with a conserved reductase domain for reductive release in a NADPH-dependent manner (Figure 5C). The released aldehydes are stable under physiological conditions for hours and undergo cyclization to form dihydropyrazinones. In the presence of oxygen, dihydropyrazinones undergo irreversible oxidation to pyrazinones, and molecular oxygen becomes more available at the epithelial surface.[104-105] However, since the gut environment is largely anaerobic with the exception of inflammatory circumstances, oxygen is generally in low abundance and oxidation is slow. This leads to an accumulation of the dipeptide aldehydes free in solution, available for targeting proteases. The electrophilic aldehydes are known to be targeted by nucleophilic cysteine or

serine residues in proteases, specifically cathepsins involved in modulating intestinal epithelial cells.[95, 103] Ruminopeptin, an *N*-acylated dipeptide aldehyde, was characterized from a gut commensal *Ruminococcus bromii* and was reported to exert inhibitory activities against *S. aureus* endoproteinase GluC.[106]

1.2.3 Polyketides

Polyketides are a structurally and functionally divergent class of small molecules derived from the homologation and functionalization of simple acyl- and malonyl-CoA-derived substrates. The chemical logic of polyketide synthases (PKSs) mirrors that of fatty acid biosynthesis. As with NRPSs, central features of polyketide synthesis involve tethering the elongating acyl chains to a thioester linkage on an acyl-carrier-protein (ACP), the functional equivalent of T domains (PCPs) in NRPSs. A ketosynthase (KS) domain catalyzes C-C bond formation between two acyl-thioester linkages via effective decarboxylative thio-Claisen condensation (Box 2/Figure B2). Additional β-keto processing by tailoring enzymes, such as ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains may or may not be present to reduce the growing chain to a "programmed" oxidation state. Similar to NRPSs, chain termination often ends with a thioesterase (or a reductase) to release the product via hydrolysis or macrocyclization (or reductive release). Additional post-assembly line modifications of the polyketide scaffold can follow,



Box 2/Figure B2. General schematic of effective decarboxylative thio-Claisen condensation mediated by the KS domain in PKS biosynthesis. Common PKS tailoring enzymes/domains are listed in grey.

including acylation, alkylation, glycosylation, prenylation, and redox modifications.[69]

The organization of PKS systems can vary. Well-known products like the virulence factor mycolactone from *Mycobacterium ulcerans*[107] are produced by giant type I modular PKSs, in which the catalytic domains are organized in large multidomain proteins connected by interprotein junctions.[69, 108] Type II PKSs typically produce aromatic polyketides like the antibiotic tetracycline from soil bacteria, in which the catalytic components are expressed as discrete proteins and used in an iterative manner.[69, 109-112] However, type II systems are also invoked as modular building blocks in the construction of other molecular classes, such as the tetrahydropyridine alkaloids widely distributed in *Pseudomonas* species, in which the discrete proteins are thought to be used once in the catalytic cycle.[113-115] Indeed, both type I and type II systems can function modularly or iteratively. Yet another enzymatic design is reflected in type III polyketides, which are homodimers that form a single active site responsible for priming, extension, and cyclization reactions and act in an iterative manner. Type III systems are most often found in plants, many of which are responsible for the production of dietary polyketides relevant to diet and the intestinal microbiota.[116] In this section, we selected an atypical bacterial type II PKS system found in *Photorhabdus* species that produce "plant-like" stilbenes. We believe that this system serves as a model of the structural and functional cross-section among bacteria, dietary plants, and host phenotypic outcomes.

Tapinarof

Tapinarof is produced by all known members of the *Photorhabdus* genus, which exist in a tripartite symbiosis with insect larvae and nematodes.[117] In addition to its pathogenic role in insects and its mutualistic role with its host nematode, *Photorhabdus asymbiotica* is known to cause systemic infections and severe soft tissue infections including the skin in humans.[118] Indeed, *P. asymbiotica* is now recognized as an emerging human pathogen.[119] Tapinarof is a topical drug currently in phase 3 clinical trials to treat skin disorders such as atopic dermatitis and the autoimmune disease psoriasis. Its clinical efficacy is thought to be derived from its activation of the xenobiotic arylhydrocarbon receptor (AhR) and nuclear factor erythroid 2-related factor 2 (Nrf2)-antioxidant signaling pathways.[120-121] The AhR is a transcription factor

activated by a diverse set of ligands and contributes to the regulation of immune signaling, disease progression, and skin barrier integrity. In human cells and in a mouse model, AhR activation leads to reduced cytokine expression and inflammation.[121] Additionally, tapinarof activates the antioxidant response to scavenge reactive oxygen species in a Nrf2-dependent manner.[120] Tapinarof has also been reported to have moderate antifungal and Gram-positive antibacterial activities.[122-124]

The tapinarof biosynthetic pathway is semi-clustered among four distinct regions of the genome and encodes several interesting enzymatic features. First, like NRPSs, the starter substrates are amino acids, which are then converted into competent PKS substrates (Figure 6). Phenylalanine is converted to cinnamic acid (StlA)[125] and activated as an acyl-CoA thioester (StlB)[126], whereas leucine is transaminated and then oxidatively decarboxylated to its acyl-CoA thioester (BkdAB) in a typical branched-chain fatty acid biosynthetic process.[126] After one round of canonical polyketide extension, these substrates are condensed in a head-to-tail fashion by an atypical ketosynthase (KS) StlD to form a carboxylated-cyclohexanedione product.[126-127] Intriguingly, in this transformation, the KS catalyzes not only a Claisen reaction characteristic of this family of enzymes, but also a Michael reaction. The product is then recognized by an aromatase StlC to decarboxylate and aromatize the final product.[126-127] *Photorhabdus* species also encode a flavin-dependent epoxidase thought to detoxify accumulating stilbenes.[128] Notably, *Photorhabdus* is one of two known bacterial stilbene producers that convergently evolved to produce "plant-like" compounds typical of the plant type III PKS systems – *e.g.*, resveratrol stilbene synthase.[125-127, 129-130]

Tapinarof is closely related to other dietary stilbenes such as resveratrol from grapes, green tea, and berries.[131] These dietary stilbenes have traditionally been shown to induce anti-inflammatory and anti-oxidant activities and are sometimes utilized as "alternative therapies" for IBDs, including Crohn's disease and ulcerative colitis.[132] Similar to tapinarof, polyphenols from green tea extracts were shown to modulate Nrf2 signaling to induce antioxidant and anti-inflammatory effects in a preclinical IBD mouse model.[133] However, IBD clinical trials for stilbene treatments ultimately failed due to inter-individual variability.[132] As such, our growing understanding of xenobiotic transformations of metabolites by gut microbiota

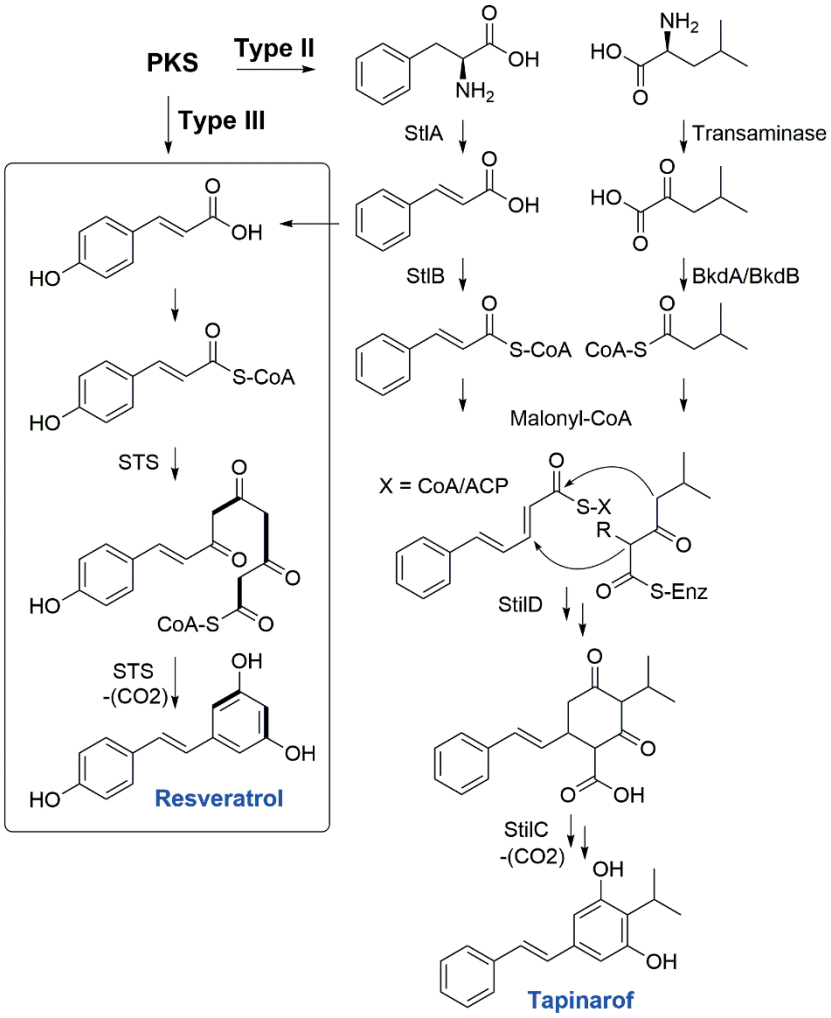


Figure 6. Biosynthesis of tapinarof and resveratrol. The atypical pathway features a novel stilbene synthesis mechanism condensing two units in a head to tail manner by a type II PKS system. Typical type III PKS synthesis of resveratrol as a representative of plant stilbenes is shown in the left box.

suggests that the inter-individual variations in these clinical trials might be due to the vast differences in the gut microbiome among patients.[134] We believe that these studies could be revisited in the context of illuminating a functional tripartite interaction: dietary plant metabolites, the microbiota, and the host.

1.2.4 Hybrid NRPS/PKS

Both NRPS and PKS systems utilize analogous protein domains in which a covalently attached phosphopantetheinyl cofactor provides a flexible thioester linkage to their biosynthetic intermediates. This architecture enables them to interface with each other to create hybrid NRPS-PKS pathways.[69] This collaboration depends on the efficiency of a ketosynthase or condensation domain to recognize a non-cognate thioester substrate to catalyze bond formation. This also requires precise protein-protein domain interactions to accept and usher the growing intermediates along the biosynthetic assembly line.[69] We also acknowledge that these types of systems can sometimes accept "free" substrates derived from other non-PKS/NRPS biosynthetic systems.[135-138]

In an analysis of ~2,500 bacterial genomes, the Sivonen group reported 3,339 NRPS/PKS gene clusters, commonly occurring in the phyla of *Proteobacteria, Actinobacteria,* and *Firmicutes* – all of which have been described as common components of the human microbiota.[139] A third of these described gene clusters encoded hybrid proteins containing NRPS and PKS core domains, and a tenth of which included individual catalytic domains in separate open reading frames – which differs from the dominant "modular" classification of NRPS/PKS enzymes. Hybrid pathways typically tend to be larger and possess more domain types, enabling more chemically diverse scaffolds than either PKSs or NRPSs alone.[139] In addition, stand-alone NRPS/PKS domains such as individual acyl-transferases (AT) can process iterative or modular loading of acyl monomers *in trans*.[140-143] Collectively, this means that hybrid NRPS/PKS gene clusters often violate the "collinearity" principle, in which the final product can be "read out" by the composition and order of biosynthetic domains.[69] Thus, while bioinformatic algorithms can now easily identify the presence or absence of these pathways, it is becoming increasingly difficult to predict a PKS/NRPS product from gene sequence or even targeted biochemical reactions alone. Consequently, hybrid NRPS/PKS pathways

offer an opportunity to uncover interesting, less-explored chemical space within natural products, in addition to the unknown biological roles that they mediate within host-microbe and microbe-microbe interactions.[69, 139]

Colibactin

It has been estimated that *Escherichia coli* resides in up to 90% of humans.[144] Some strains of *E. coli* and *Klebsiella pneumoniae* harbor an ~54 kb genomic island that encodes for the small molecule genotoxin termed colibactin.[145-147] Colibactins are secondary metabolites produced by the colibactin biosynthetic gene cluster (*clb* or *pks*), and their structural characterization have been of intense interest to the natural products, microbiology, and cancer cell biology communities.[148-152] Production of colibactins in live bacteria co-cultured with mammalian cells reveal induced mammalian cell genotoxicity and cytotoxicity *in vitro* and *in vivo*, including DNA double stranded breaks, cell cycle arrest, and genomic instability phenotypes.[145, 153-156] Presence of the pathway promotes carcinogenesis in colitis-susceptible mice treated with azoxymethane, and clinical data show an increased prevalence of *clb*-encoding *E. coli* in inflammatory bowel disease (IBD) and colorectal cancer (CRC) patients over healthy patient controls.[154, 156-157] The risk of colibactin-mediated carcinogenesis could be increased upon interaction with additional strains, as tumor formation in mice was increased in mice co-colonized with toxin-positive *Bacteroides fragilis* and *clb*$^+$ *E. coli* over mice mono-colonized with each respective strain.[158]

While the *clb* pathway is diversity-oriented, producing many pathway-dependent metabolites with variable activities, the key mode of action of colibactin related to disease has been supported as DNA interstrand crosslinking.[159] Indeed, Nougayrède, Oswald, and co-workers demonstrated that *clb*$^+$ bacteria crosslink exogenous DNA and activate DNA crosslink repair machinery of the Fanconi anemia pathway.[159] Crosslinking has been confirmed and a mono-adenine adduct has been characterized by two separate groups.[160-161] The subsequent activation of downstream γH2AX, a marker of DNA double strand breaks, and cellular enlargement (megalocytosis) was found to be dependent on every biosynthetic gene in the

pathway.[145] The presence or absence of uncharacterized peptidase ClbL led to cellular crosslinking versus alkylation, respectively, indicating that ClbL is required for the crosslinking phenotype.[162] Colibactin is an intriguing molecule given that the biosynthetic pathway is found in *E. coli* Nissle 1917 (EcN, 99% similarity)[142], a probiotic used for the treatment of ulcerative colitis and irritable bowel syndrome.[163] Remarkably, inactivation of the pathway also abrogates EcN's beneficial probiotic activity as well as colibactin's genotoxocity, although cross-talk effects of other EcN metabolic pathways have not been ruled out.[163]

While mounting evidence exists for colibactin's link to CRC development[154, 156-158, 164], an open question remains as to the benefits of colibactin to *clb*$^+$ strains. Because both intestinal inflammation and the presence of the *clb* gene cluster are required for colibactin-mediated carcinogenesis in multiple mouse models, one can speculate that colibactin allows a fitness advantage during inflammatory conditions.[154] Activation of the p53 DNA damage response is known to modulate expression of toll-like receptors and interferon signaling[165], and inflammatory conditions provide Enterobacteriaceae access to new nutrients in the intestinal tract leading to blooms.[166-167] Additionally, cross-talk is observed for colibactin and siderophore biosynthesis through sharing of a phosphopantheinyl transferase required for posttranslational carrier protein activation.[168] Since siderophores provide certain bacteria with colonization advantages due to iron acquisition, colibactin synthesis could contribute to niche establishment during inflammatory conditions, where iron is limiting.[27, 43, 147]

The mysterious biology surrounding the pathway has resulted in a surge of investigations by multiple labs over the last few years in characterizing the biosynthesis and structure of colibactin.[148-152] While the fully mature structure of colibactin is currently under peer review, a wealth of publications has become publicly available in recent years describing colibactin's unique biosynthetic, structural, and mechanistic features. The colibactin pathway is regulated by a LuxR-type regulator[169] and encodes hybrid NRPS/PKS megasynthetases, accessory enzymes, a type II editing thioesterase that uniquely hydrolyses multiple pathway intermediates[170-172], a transporter,[173-174] and a resistance protein.[175-176] The unexpected

chemical reactivity, instability, and biosynthetic logic employed by the pathway help to explain, in part, the challenges behind isolating the final product. Here, we will discuss several interesting features of the pathway, including a prodrug activation step, cyclopropane formation from S-adenosyl-methionine, and α-aminomalonyl-carrier protein biosynthesis.

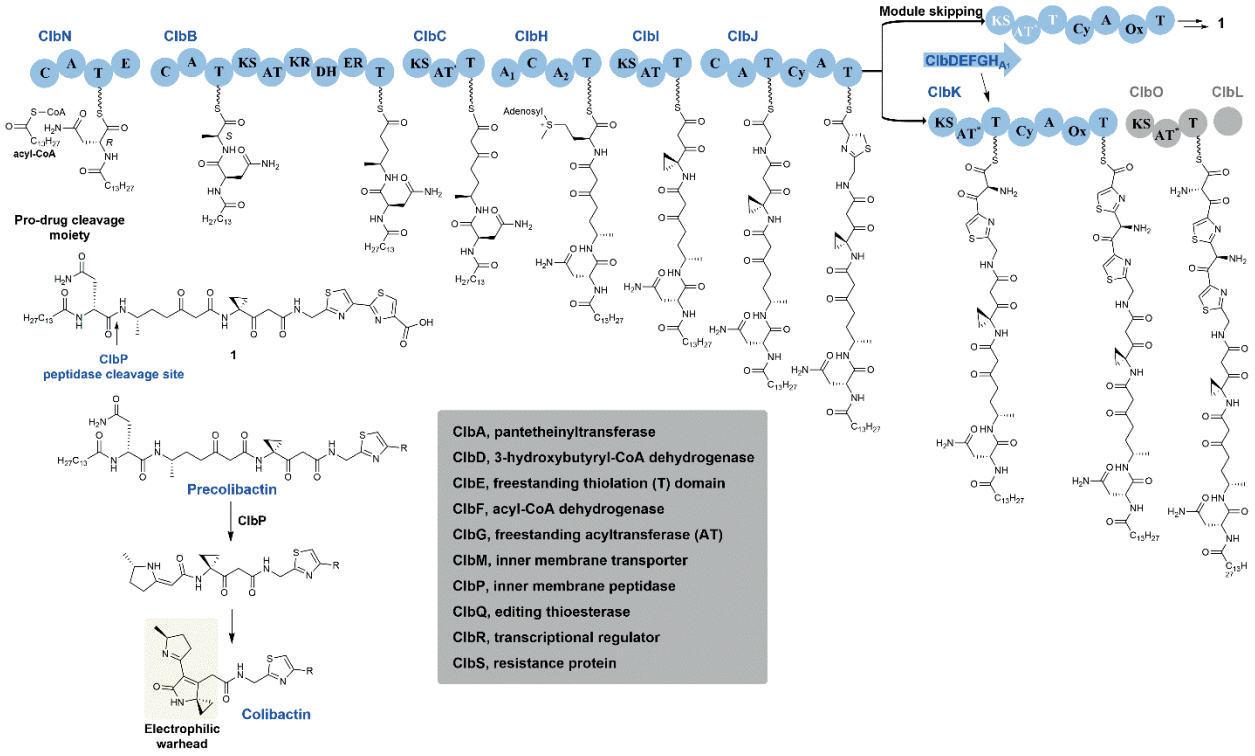


Figure 7. Proposed biosynthesis of colibactin. Precolibactin is synthesized by a hybrid NRPS-PKS pathway. Highlighted features in the chapter include 1) prodrug activation, 2) cyclopropane formation, and 3) α-aminomalonyl-biosynthesis. 1) An N-acyl-D-Asn prodrug motif is cleaved by peptidase ClbP, leading to the free N-acyl-amide and subsequent spontaneous cyclodehydrations reactions that set the electrophilic colibactin warhead. 2) Methionine labeling studies supported that the cyclopropane was methionine-derived. Protein biochemical studies established that the methionine-derived unit was SAM. Interestingly, hydrolytic derailment of the ClbH product leads to an AHL, whereas ClbI is required to shuttle the intermediate downstream and form the cyclopropane motif. 3) ClbH-A_1 activates L-Ser and transfers it onto carrier protein ClbE to initiate α-aminomalonyl-biosynthesis. Dehydrogenases ClbD and ClbF oxidize the L-serinyl-ClbE to α-aminomalonyl-ClbE, and trans-acyltransferase ClbG transfers the unit to the PKS module of ClbK and the PKS module ClbO in the proposed biosynthesis. Peptidase ClbL is required to convert *clb*-pathway-derived DNA alkylators into DNA interstrand crosslinkers. NRPS-PKS enzyme domains are represented as blue circles, final domain steps of precolibactin biosynthesis are shown in grey circles, the spontaneous cylodehyrdrations leading to the active colibactin warhead is shown, and additional proteins are listed in the grey box.

Prodrug cleavage and colibactin activation

Biosynthesis of colibactin (Figure 7) begins with the formation of a prodrug termed precolibactin, which is hydrolyzed by an inner membrane peptidase ClbP.[177-178] Cleavage of precolibactin releases *N*-acyl-D-Asn[179-181] and linear colibactins, the latter of which undergo sequential cyclodehydrations to form an electrophilic 'warhead.'[151, 170, 182] The 'warhead' is reminiscent of the duocarmycin family of DNA alkylators, which alkylate DNA via cyclopropane ring opening.[183] Consistently, the cyclopropane in synthetic colibactins closely resembling characterized colibactin metabolites has been shown to be required for DNA alkylation.[170, 182] Indeed, a *gem*-dimethyl analog of the warhead was inactive.[182] Naturally produced colibactin-adenine nucleobase adducts characterized from DNA exposed to *clb*$^+$ *E. coli* are in agreement with this proposal.[160-161]

Cyclopropane formation

The NRPS module ClbH is a unique NRPS with a noncanonical A_1-C-A_2-T domain architecture. While ClbH A_1 activates L-Ser[170, 184-185] (discussed below), domain targeted metabolomics, genetics, and isotopic labeling studies showed that A_2 was responsible for L-Met isotopic labeling of the cyclopropane moiety.[170, 186-187] The "Met-derived precursor (for example, *S*-adenosyl-Met)"[186] was confirmed to be derived directly from *S*-adenosyl-Met (SAM) via protein biochemical studies.[188] Interestingly, downstream ClbI was shown to be required for cyclopropane formation, and in its absence, an acyl-homoserine lactone (AHL) was formed as a ClbH-derailment product instead.[188] Acyl homoserine lactones (AHLs) are commonly involved in quorum sensing and modulation of gene expression in Gram-negative bacteria.[189] The production of an AHL by the *clb* pathway, the presence of its LuxR-type transcriptional regulator ClbR[169], and the knowledge that the native *E. coli* LuxR-regulator SdiA responds to AHLs[190-191], suggest that *clb*-derived AHLs and related molecules may be intercepted as signaling molecules (see also below). These biological connections remain exciting open questions in the field.

Installation of an α-aminomalonyl-derived unit

A bioinformatic analysis of the colibactin pathway predicted the possible incorporation of an L-Ser-derived α-aminomalonyl extender unit[186] based on homology to known pathways,[192-193] although Ser isotope labels were not detected in the dataset at the time[186]. Unequivocal establishment of the incorporation of this extender unit came from protein biochemical studies[184-185] and characterization of a product incorporating this moiety[171]. *In vitro* studies indicated that ClbH-A1 loads L-Ser onto the freestanding PCP ClbE (represented as T). Dehydrogenases ClbD and ClbF then oxidize seryl-ACP to aminomalonyl-ACP.[184-185] Transfer of the α-aminomalonyl unit from ClbE to a T domain in ClbC, ClbI, ClbK, and ClbO was shown to be mediated by freestanding acyltransferase ClbG *in vitro*.[185] However, some of these reported activities appeared to be the result of promiscuous activity of purified proteins *in vitro,* as corresponding metabolites for loading of this unit onto ClbC and ClbI have not been detected in cell culture. Gene inactivation studies further supported that the PKS domain of ClbK accepted α-aminomalonyl-ACP from the trans-acyltransferase ClbG.[171] Selective inactivation of $ClbK_{PKS}$ by genome editing also demonstrated that PKS extension using an α-aminomalonyl-unit was essential for inducing genotoxicity in cell culture.[162] Biosynthetic divergence points introduced through domain skipping of $ClbK_{PKS}$ can produce molecules with differing DNA-damaging effects: bithiazole scaffolds can alkylate DNA *in vitro*, while α-amino-malonyl-ACP biosynthetic machinery is required for DNA crosslinking.[162, 170] We speculate that the diversity arising from this pathway could provide an evolutionary benefit for the producer.

1.2.5 *N*-acyl-amides

N-acyl-amides represent a broad family of microbial and mammalian signaling molecules, including the well-studied subfamily *N*-acylhomoserine lactones (AHLs), noted above. AHLs serve as autoinducers in Gram-negative bacteria and regulate a signaling phenomenon known as quorum sensing.[194-197] Quorum sensing is a bacterial communication method that links bacterial cell density to collective genetic behaviors; concentration of an autoinducer controls the transcription of programs as diverse as

biofilm formation, virulence factor expression, antibiotic production, bioluminescence, sporulation, and DNA-uptake.[195] Population-dependency and cellular stress[138, 198-199] regulation of these processes are often crucial for fitness and survival within a host, both for pathogens and commensals.[200] AHLs are typically synthesized by autoinducer synthases, which catalyze the condensation of *S*-adenosyl-Met and an acyl carrier protein (ACP)-linked fatty acid.[200] Many bacterial species encode receptors for AHLs but no known biosynthetic genes within their own genomes, suggesting that AHLs are a common chemical language that can be overheard by potentially competing bacterial species.[201]

Metagenomic analysis of DNA from soil microbes shed light on an additional class of enzymes that share limited sequence similarity with autoinducer synthases, yet catalyze the production of chemically similar molecules.[202] *N*-acyl amino acid synthases (NASs) utilize acyl-ACPs and amino acids in lieu of SAM as substrates.[202] Again, the *N*-acyl-amide products of these enzymes are hypothesized to have a role in signaling much like their AHL counterparts. This is supported by evidence that an acyl-transferase class of NASs are genetically linked to two-component signal transduction systems in bacteria.[203]

Other *N*-acyl-amides have been described that arise from mechanistically distinct enzymatic routes. A family of *N*-acyl-L-histidines produced by the "accidental" human pathogen *Legionella pneumophila* was shown to be dependent on an annotated ATP-grasp protein.[204] This ATP-grasp protein likely functions as an ATP-dependent ligase (Figure 8A) that condenses free medium-chain-length fatty acids from primary metabolism with free L-His substrates, which is in contrast to the transferase (Figure 8B) activity of AHL synthases and NASs.[204] Notably, this pathway is highly upregulated during *L. pneumophila* macrophage infection, and suggests a conserved role for *N*-acyl-amides in bacterial- mammalian signaling. Additional unknown enzymatic mechanisms can account for *N*-acyl-amide synthesis. For example, a recent report of an *N*-acyl-amide encoded by diverse *Pseudomonas* species was speculated to arise from an unidentified halogenase reaction (Figure 8C) on a tetrahydropyridine alkaloid precursor,[114-115] mimicking a similar sequence observed in plants.[205] These examples indicate that there are a variety of biosynthetic routes, some

of which have been characterized, to this class of signaling molecules, indicating that putative *N*-acyl-amide pathways are dramatically underrepresented in genome mining analysis of the human microbiome. Ultimately, the various enzymatic pathways converge to create a shared *N*-acyl-amide core structure.

Mammalian *N*-acyl-amides, such as endocannabinoids, are noted to have analgesic and anti-inflammatory properties and many of these molecules target G-protein-coupled-receptors (GPCRs).[206] GPCRs are one of the largest families of mammalian membrane receptors and are implicated in a variety of diseases, notably in obesity, cancer, and IBDs – all conditions which have been widely found to have a connection to the microbiome composition.[207] Given that the symbiosis of mammals and microbes have evolved over millennia, it should not be surprising that human-associated bacteria have developed

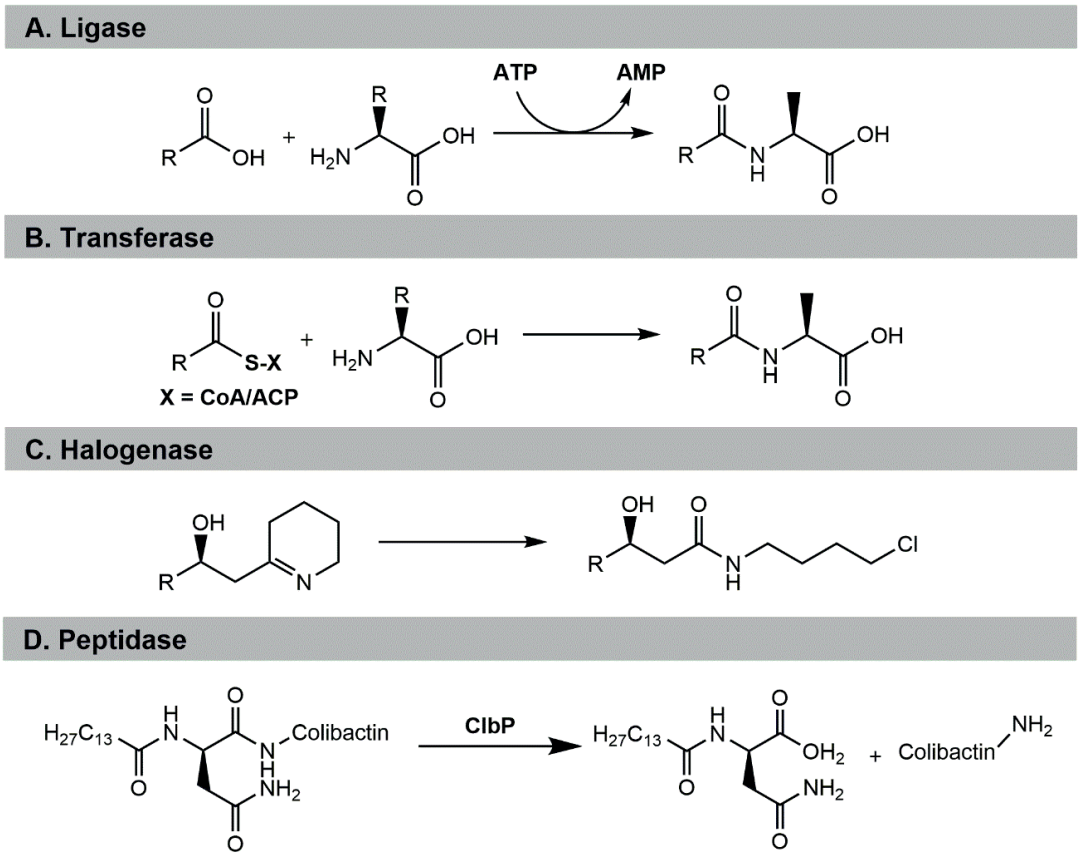


Figure 8. General schematic of enzymes involved in N-acyl amide biosynthesis. (A) ATP-dependent ligase couples a fatty acid with an amino acid. (B) A transferase, representative of a typical N-acyl synthase, condenses a fatty acid thioester with amino acid to form an N-acyl amide scaffold. (C) Proposed halogenase reaction using an alkaloid precursor (D) Inner membrane ClbP cleaves N-myristoyl-D-Asn to activate colibactin.

molecular mimicry that can intercept and selectively activate mammalian transcriptional regulatory networks. The convergence on *N*-acyl-amide signaling between bacteria and humans likely derives from their simple coupling of two abundant primary metabolites, amino acids and fatty acids.

_N_-acyl-D-Asparagine type molecules

The colibactin pathway, as mentioned above, is found to produce both *N*-acyl-amides and AHLs. *N*-myristoyl-D-Asn (NMDA) is one of the most abundant molecules produced by the *clb*$^+$ pathway and is the major species resulting from ClbP peptidase-mediated cleavage of precolibactins (Figure 8D).[179-181] In a panel of central nervous system receptor assays, NMDA was found to harbor modest 10 μM-level antagonistic activity against the dopamine 5 (D_5) transporter and the 5-hydroxytryptamine (5-HT_7) GPCR.[181] As with other *N*-acyl amides derived from the soil environment[208-210], NMDA also exhibited mild growth inhibitory activity against *Bacillus subtilis* in the low μM-range, which might be one of the many contributors to robust *clb*$^+$ *E. coli* colonization within the gut.[181] In suspension cultures, NMDA accumulated to ~27 μM suggesting that these *in vitro* activities may be physiologically relevant.[181]

The *clb* pathway in Nissle 1917 was demonstrated to also produce the *N*-acyl-amides *N*-dodecanoyl-Asn-γ-aminobutyric-acid and *N*-dodecanoyl-Asn-β-aminobutyric-acid (Figure 9A).[211] The production of these molecules was found to be largely dependent on ClbN and ClbB.[211] As ClbB normally activates and incorporates an L-Ala residue prior to polyketide extension, it is unclear if γ-amino-butyric acid (GABA) or β-aminobutyric acid (BABA) incorporation is catalyzed by ClbB or is spontaneous, as a Δ*clbB* strain did not fully abolish the production of these products.[211] However, the incidence of these compounds is notable, as GABA in particular is a primary neurotransmitter throughout the mammalian central nervous system.[211] Cenac and coworkers went on to show that *N*-dodecanoyl-Asn-GABA could cross the epithelial barrier and inhibit calcium signaling via the $GABA_B$ GPCR. The authors speculate that this mechanism could be involved in modulating visceral hypersensitivity, a key mechanism of underlying abdominal pain associated with irritable bowel syndrome. It is unclear if these molecules reach a physiologically relevant

concentration *in vivo*, although synergistic interactions with *Lactobacillus* species, which produce high levels of GABA, could be considered for the potential probiotic enhancement of *E. coli* Nissle 1917 in future studies.[212]

Commendamide

Using a functional metagenomics approach, Brady and coworkers developed a screening assay to identify bacterial effectors from gut-associated bacterial genomes that activated mammalian signaling pathways.[213] Functional metagenomics is a culture-free approach that involves the construction of DNA libraries from whole-genomic sample extractions, allowing expression analysis from both cultured and currently uncultured microbes.[213] In this case, the authors screened cell-free supernatants from an *E. coli* library of ~75,000 cosmid clones to search for small molecules that activated NF-κB – a central transcriptional hub controlling cellular responses to a variety of stressors (*e.g.*, pathogen-associated molecular patterns, or PAMPs). Characterization of one of these hits revealed the structure *N*-acyl-3-hydroxy-palmitoyl glycine, termed commendamide (for *com*mensal *m*imicking *end*ogenous amide) (Figure 9B). Characterization of the elicitor gene showed commensal bacterial effector gene (Cbeg12) of the *N*-acyl-synthase family to be responsible for producing commendamide. Genes highly similar to Cbeg12 appeared to be restricted within the Bacteroidetes phylum, suggesting a non-pathogenic role for commensal-host signaling.[213] Commendamide shares structural similarity with human-produced long-chain endocannabinoids such as palmitoylethanolamide, *N*-acyl-palmitoyl glycine, and anandamide.[213] These endogenous *N*-acyl-amides are traditionally known to induce GPCRs and targets in pain and

inflammation. Commendamide was shown to activate human GPCR GPR132/G2A, which has been linked to autoimmunity and atherosclerosis through immune cell expansion, differentiation, and chemotaxis.[214]

A follow-up study by the same group mined *N*-acyl synthases in the human microbiome and showed that this class of enzymes was enriched within the gut relative to other environments and that the biosynthetic products structurally overlap to known, endogenous ligands.[215] Through these identification efforts, the authors characterized six additional *N*-acyl amides. Two of these compounds were similar to

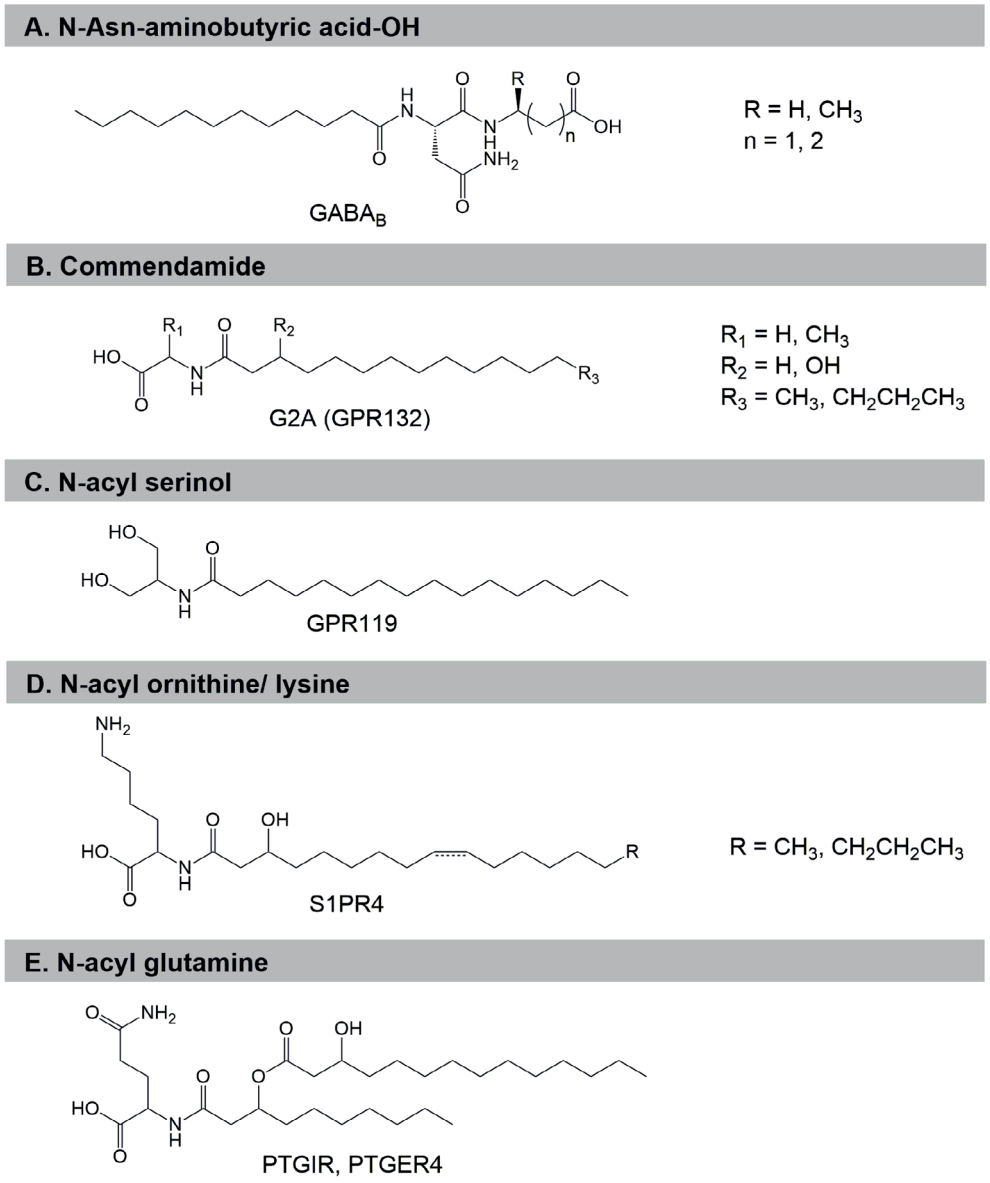


Figure 9. Selected N-acyl amide structures and their respective GPCR targets.

previously characterized commendamide, and not surprisingly, were shown to activate GPR132/G2A. An *N*-acyl serinol (Figure 9C) was found to activate GPR119 due to structural similarities with endogenous ligand oleoylethanolamide. Just like natural ligands, microbiota-encoded *N*-acyl serinol was shown to affect GPR119-dependent glucose homeostasis and gastric emptying through glucagon-like peptide 1 (GLP-1) release. Similarly, *N*-palmitoyl ornithine/lysine derivatives (Figure 9D) act as specific activators for sphingosine-1-phosphate receptor 4 (S1PR4) known to modulate Th17 polarization and chemotaxis. Lastly, *N*-acylglutamine (Figure 9E) was reported as a new antagonist for two prostaglandin receptors, PTGIR and PTGER4.[215] These studies highlight the growing importance of bacterial metabolites as modulators of host GPCRs[216] by mimicking eukaryotic ligands, and the need to identify further enzymatic chemistries responsible for mediating such interactions.[215]

1.3 Metabolic exchange and resistance mechanisms

In this section, we highlight examples of metabolite modification undertaken by the microbiome, rather than *de novo* metabolite synthesis. Microbes can both alter metabolites produced by the host, such as bile acids, or xenobiotics consumed by the host – which often underlie phenotypic variation among individuals.[217] Among bacterial species, both producers and competitors evolve resistance mechanisms to molecularly spar with one another, which can affect community composition. Below, we explore selected examples of toxin resistance and xenobiotic metabolism and look at the enzymatic capacity of the microbiome to detoxify, transform, and ultimately evade the effects of certain types of functional small molecules.

1.3.1 Resistance mechanisms

Biosynthetic gene clusters often contain transporter, regulatory, and resistance genes that confer protection to a producing host.[37] This is a necessary survival adaptation, particularly for producers of natural products with cytotoxic effects. Self-resistance strategies have been studied extensively, as they form the genetic reservoir of antibiotic resistance genes that poses dire future consequences for our health care

system.[218-220] As such, mining for new chemical scaffolds within the human microbiome should take into account resistance elements, as they are often transcribed on mobile genetic elements and can undermine the efficacy of newly discovered natural products with therapeutic potential.[221] By analyzing the structural information of putative resistance genes, modes of action and/or cellular targets sometimes can be inferred prior to structural characterization and biological assays.[222-223]

Self-resistance takes a number of different chemical and molecular forms, including prodrug synthesis, efflux systems, target modifications, and chemical modifications.[222] As mentioned above for colibactin, prodrug synthesis can involve a "protecting group" that is removed during activation; in this case, upon export of the genotoxic product in the periplasm and distal to the host genome. For colibactin, the NMDA prodrug motif is also thought to participate in molecular recognition by the inner membrane cation-coupled multidrug and toxic compound extrusion family (MATE) transporter ClbM[173-174]. One of the most common variants of efflux systems are ATP-binding cassette transporters that efflux natural products to the extracellular environment with remarkable affinity.[224]

Target modification is another common resistance strategy. Here, the binding site of the natural product target is mutated in a manner to prevent self-toxicity. A classic example of target modification includes methylation of the ribosome to confer resistance to aminoglycosides.[225] An alternate strategy recognized in biosynthetic gene clusters is the presence of an additional copy of a target gene with often an amino acid substitution(s) to confer resistance. Thus, one genomic copy could in principle serve as "bait"[226] for the antimicrobial and the other could compensate for its loss of function.[222] Lastly, chemical modifications are often employed to deactivate small molecules through addition of a functional group (*e.g.*, phosphorylation or acetylation are common addition strategies), or degrade them entirely through use of hydrolytic enzymes. For one example, the microcin C producer encodes a resistance element within the

microcin locus that serves as a peptidase, detoxifying this compound and preventing antibacterial activity through the cleavage of the C-terminal aspartate from a nucleotide.[227]

Resistance elements can also uncover new biochemistries that could potentially be exploited for therapeutics. In the case of colibactin, the *clb* pathway-encoded enzyme ClbS protects the producer from colibactin that may reenter the cell under certain conditions via diffusion. ClbS is the only encoded enzyme in the bacterial *clb* pathway not required to mediate host genotoxicity, as it was shown that exponential-

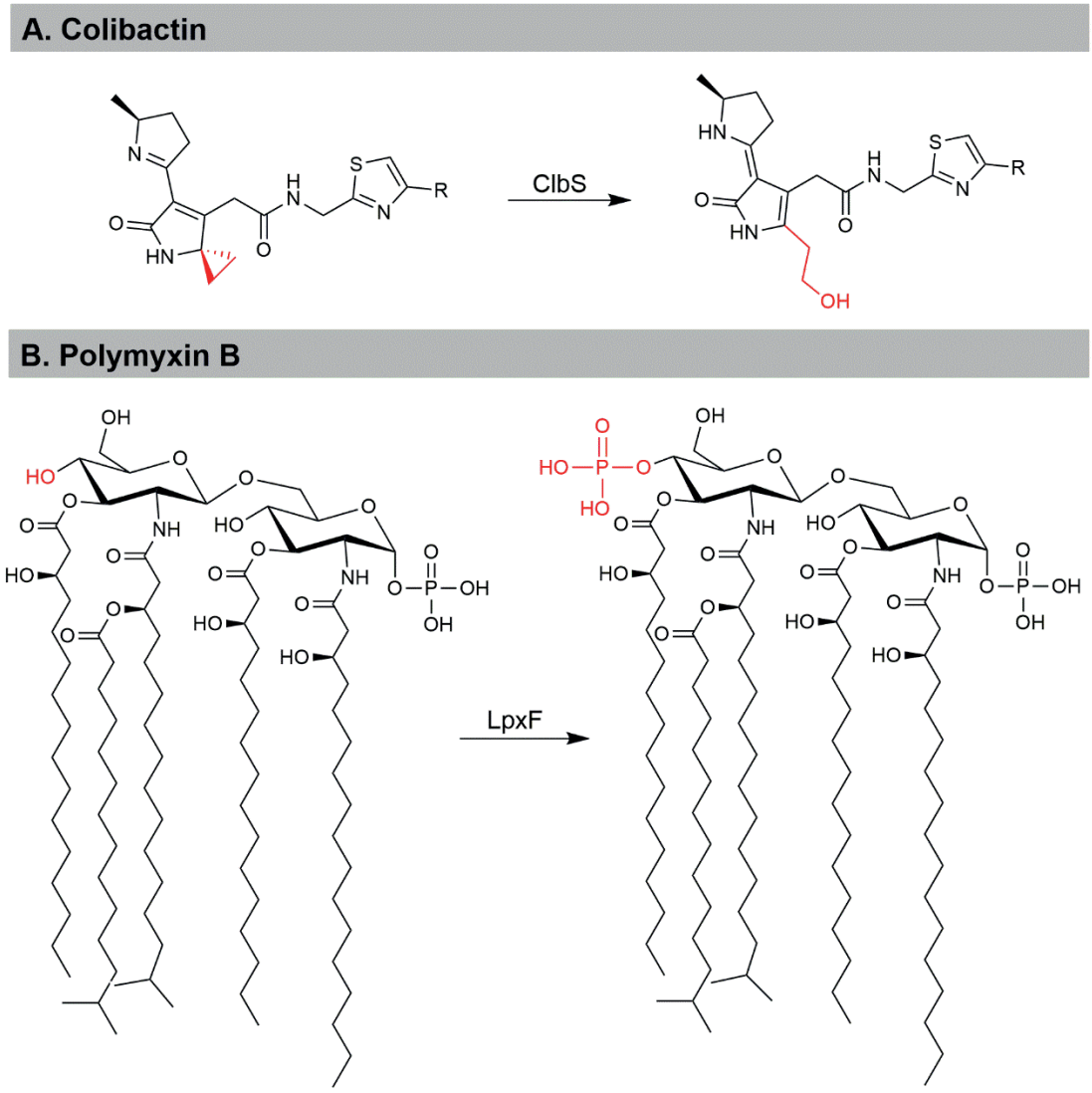


Figure 10. Resistance mechanisms of colibactin and polymyxin B. (A) Cyclopropane hydrolase ClbS hydrolyzes cyclopropane moiety of colibactin, rendering it inactive. (B) LpxF dephosphorylates lipopolysaccharide for polymyxin B resistance.

phase cultures of a *clbS* mutant were capable of inducing the same level of genotoxicity to that of their wildtype counterparts.[175] In contrast, late-phase cultures of the *clbS* mutant induce less DNA damage due to a growth defect caused by the accumulation of genotoxic colibactins and a subsequent induction of the SOS DNA damage response in the producing organism.[175] *In vitro* biochemical studies revealed that ClbS can act as an unprecedented "cyclopropane hydrolase," hydrolytically opening the cyclopropane residue required for genotoxicity and generating an innocuous product (Figure 10A).[176, 182] This is a newly reported enzymatic function, and X-ray crystallography studies place this enzyme within the previously uncharacterized DINB_2 enzyme superfamily.[176] While ClbS is thought to be localized in the cytosol of *clb*$^+$ strains, ClbS is catalytically active and can confer protection to colibactin-mediated genotoxicity when ectopically expressed in mammalian cells[175] or exogenously supplemented in a *clb*$^+$ bacteria-human cell transient infection model[159, 162].

Resistance mechanisms provide a fascinating vantage point for analyzing the complex dynamic between the host and its microbes. For instance, host innate immunity invokes the use of antimicrobial peptides to restrict bacterial growth to the colonic lumen and prevent invasion of the epithelial tissue barrier.[228] Goodman and coworkers described a resistance mechanism toward the antimicrobial peptide polymyxin B (Figure 10B) within the Bacteroidetes phylum.[229] A single 4'-phosphatase gene (*lpxF*) encoded in *Bacteroides thetaiotaomicron*, a common commensal member of the mammalian gut, was responsible for observed under-phosphorylation of the lipid A anchor in the lipopolysaccharide. This, in turn, conferred cationic polymyxin B resistance by four orders of magnitude over the *lpxF* isogenic mutant strain by altering the cell surface charge. The authors demonstrate that this enzymatic activity endows this taxonomic clade with a long term persistence phenotype in mice and resiliency to inflammatory conditions.[229] Alternatively, it was recently shown that β-lactamase-encoding bacteria including drug-resistant *E. coli* can convert the β-lactam amoxicillin into a new immunostimulant that lacks antibacterial activity, a functional transformation process.[230] Antibiotics often cause inflammation in a patient-dependent manner[231-234], and microbiome-mediated structural and functional transformations of antibiotics and other

drugs represents an area with much promise. In a last example, amidohydrolases that degrade *N*-acyl-amides have been described, lending to the hypothesis that certain bacterial strains can intercept and perhaps silence bacterial cell-to-cell signaling, which could further contribute to shaping microbiome composition.[235-236] Like with quorum sensing, certain strains may be able to sense these *N*-acyl-amide molecular signals but may or may not be capable of synthesizing them.[190] Indeed, determining general molecular mechanisms of microbial fitness and function in the gut will be important to understand host inflammation and the progression of dysbiosis.

1.3.2 Xenobiotic Metabolism

Microbes encounter a vast, foreign library of molecules (xenobiotics) ingested or introduced into the host, including dietary compounds, industrial chemicals and pollutants, and pharmaceuticals. Not surprisingly, given the vast genomic potential of the microbiome and the promiscuity of bacterial metabolic enzymes to catalyze non-cognate reactions, microbes have the capacity to directly alter the chemical structures of non-native compounds.[134, 237] The uniqueness of each microbial gene pool housed within each host imparts an individual metabolic signature, which can have major implications for drug metabolism in particular. Outside of host genetics, the inter-individual variability of the microbiome among patients could explain the often wide therapeutic ranges required for drug efficacy and/or variable side effects.[134, 237] A new emphasis is being placed on the microbiome in regards to pharmacology, and it is anticipated that future precision medicines could take into account an individual's microbiome composition.

Many microbial enzymes associated with xenobiotic detoxification and metabolism are some of the most abundant and widely distributed enzyme classes across sequenced microbiomes.[238-240] A common theme involves hydrolytic and reductive chemistries manifesting in a diverse array of enzyme families - from metalloenzymes operating in anoxic conditions, to proteases, hydrolases, lyases, radical enzymes, and transferases.[134, 241-242] Bacteria are able to degrade xenobiotics to use as energy inputs or carbon/nitrogen sources for primary metabolism, as well as use them as terminal electron acceptors during respiration. Due to often relaxed substrate selectivity of microbial enzymes, predicting xenobiotic transformation capacity

from sequence analysis is a major challenge.[134] Systematically predicting and identifying genes that can selectively modify certain pharmacophore scaffolds is an open problem in the field. For these reasons, microbial contributions to drug metabolism is often assessed currently in an individual drug-by-drug basis. Here, we will comment on two case studies: digoxin and gemcitabine.

Digoxin

The cardiac glycoside digoxin, derived from foxglove (*Digitalis purpurea)* plant extracts inhibits Na^+/K^+ ATPases in cardiac myocytes and is used to treat heart failure and arrhythmias.[243] Early studies pointed to a role for the microbiome in affecting patient outcome with digoxin prescription, as co-administering digoxin with antibiotics decreased the levels of dihydrodigoxin, an inactive metabolic derivative, that was excreted.[244] A single bacterial sample isolated from dihydrodigoxin-excreting

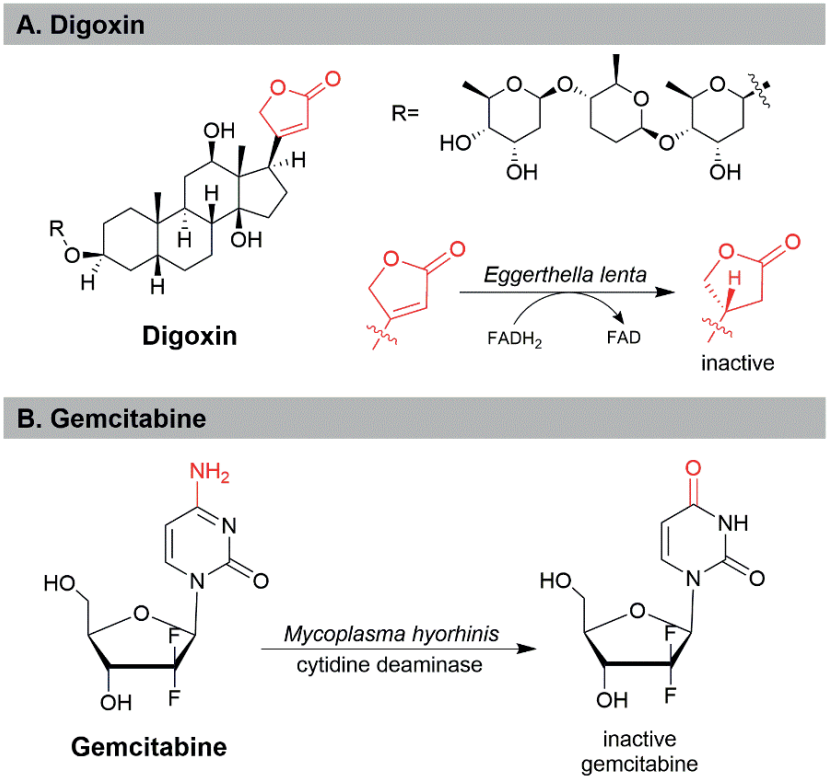


Figure 11. Xenobiotic transformation mechanisms of Digoxin and Gemcitabine (A) Flavin-dependent reductase of *Eggerthella lenta* reduces the lactone ring of digoxin for inactivation (B) *Mycoplasma hyorhinis* uses cytidine deaminase to inactivate gemcitabine

individuals, *Eggerthella lenta*, was found to reduce the lactone ring of digoxin *in vitro* (Figure 11A).[245] A study by Turnbaugh and colleagues used transcriptional profiling and comparative genomics to identify a two-gene operon (cgr1 and cgr2) up-regulated by digoxin.[246] More recently, Cgr1 and Cgr2 were proposed to form a membrane-bound protein complex that shuttles electrons through a series of cytochromes and [4Fe-4S] clusters to generate hydride species through Flavin Adenine Dinucleotide (FAD) cofactors. The terminal hydride then reduces the lactone moiety of digoxin, rendering its therapeutic effect inactive.[247]

Gemcitabine

Gemcitabine is a nucleoside analog (2'2'-difluorodeoxycytidine) used to treat patients with pancreatic, lung, breast, and bladder cancers. Differential patient responses to chemotherapeutics is an established phenomenon observed in the clinic, and studies suggested a secreted factor within isolated human dermal fibroblasts (HDFs) was responsible for gemcitabine drug resistance.[248] Bacterial contamination of HDFs with *Mycoplasma hyorhinis* ended up being identified as the root cause for the observed phenotype, and gemcitabine inactivation (Figure 11B) was found to be dependent on the single expression of a long form of the bacterial enzyme cytidine deaminase (CDD_L).[249] Deep sequencing of pancreatic ductal adenocarcinomas isolated from patients confirmed that intratumor bacteria mainly belonging to Gammaproteobacteria and positive for CDD_L were present at high incidence within these samples. Of those bacteria isolated and cultured in pure form from these samples, 93% were able to render human colon carcinoma cell lines fully resistant to gemcitabine.[250]

1.4 New frontiers in the human microbiome

The structures and enzymology highlighted herein represent the tip of the iceberg of the coding capacity of the microbiome; a majority of possible chemistries and metabolic interactions on all levels remain uncharacterized. At the genetic level, about three-quarters of the prevalent microbiome open reading frames detected by metagenomic sequencing within the gut consist of uncharacterized orthologous

groups and/or completely novel gene families.[250] This means that a majority of enzymes cannot be mapped to any known functions within existing databases such as the NCBI-NR database of non-redundant protein sequences or KEGG (Kyoto Encyclopedia of Genes and Genomes). In the absence of experimental manipulation strategies – *i.e.*, many systems are genetically intractable or strains are currently unculturable – assigning function to unknown orthologous genes detected in sequence alone remains a difficult task.[251] Additionally, the overwhelming majority of unknown ions, or "molecular features," detected within animal-associated metabolomes are said to represent "dark matter".[252] Greater than 90% of metabolites identified in microbiome-derived metabolomic datasets do not have a match in any public databases, and some estimates claim that about 1.8% of spectra in an untargeted metabolomic data collection is associated with a putative chemical assignment.[253-254] Moreover, simply changing the extraction and metabolomic conditions greatly expands on this diversity. While some of these molecules will represent "junk" from the host perspective – we are discussing the intestinal tract after all – others could open completely new areas of signaling biology. Indeed, the host-bacteria interface represents a new frontier in chemical ecology and natural products research.

In light of the major unknowns and the magnitude of data available, how does one continue to pinpoint unique enzymes and metabolic products that can tip the scale toward health or disease? Bioinformatic tools based on known chemical languages of natural product biosynthesis have proven to be useful in predicting possible chemical mediators from the abundance of genetic data we now have available.[81, 255-260] However, biosynthetic enzymes related to RiPPs, NRPSs, PKSs, hybrids, and other fairly well-characterized proteins discussed herein will similarly only account for a small fraction of this chemical space. The grey and often debatable "boundaries" for what constitutes primary versus secondary metabolism will continue to blur. There are great opportunities for the natural products field, especially for identifying novel small molecules in a functional context regardless of biosynthetic origins.

The continued integration of microbiology, chemistry, and cell biology will be essential for parsing apart the functional mysteries of the microbiome – for both novel metabolites and the genes responsible for their production. On the genetics level, mere sequence alone can obscure biological importance: apparent

"redundancies" are common and not well understood. For example, some "duplicated" genes at the annotation level carryout completely different functions.[251] Scalable, unbiased methods for detecting relevant bacterial strains and specific genes that influence physiological phenotypes in mammalian hosts are essential for beginning to identify complex interactions within their biological context.[261-264] Spatial distribution of these bacteria among distinct physical niches within the gut may also have vastly different consequences for the microbial ecosystem.[265] For example, recent research put forth by Elinav and colleagues presented evidence that fecal samples do not accurately capture the microbial composition near the mucosal surface.[266] As this local microenvironment is essential for host-microbe interactions[267], sequencing downstream fecal samples as a measure of species' abundance in the gastrointestinal tract tells only an incomplete story. This area near the mucosal/epithelial layer is also known to be microaerobic, or at least relatively oxygenated in comparison to where most anaerobic microbe members thrive within the gut.[104, 268] Consequently, aerobic and/or spontaneous oxidation reactions likely occur more readily at the host epithelial surface. Along these lines, other spontaneous chemistries regulated by substrate supply or chemical instabilities need to be considered. Advanced detection and analysis methods, such as imaging mass spectrometry, are being employed to begin to understand spatial localizations of microbial chemistries in the gut.[269-271] Together, a concerted effort between multiple fields will help to uncover and define functional molecules at the host-bacteria interface.

1.5 References

1. Lee, W.; Brey, P., How microbiomes influence metazoan development: insights from history and *Drosophila* modeling of gut-microbe interactions. *Annu Rev Cell Dev Biol* **2013,** *29*, 571-592.
2. Consortium, T. H. M. P., Structure, function and diversity of the healthy human microbiome. *Nature* **2012,** *486*, 207-214.
3. Yatsunenko, T.; Rey, F.; Manary, M.; Trehan, I.; Dominguez-Bello, M.; Contreras, M.; Magris, M.; Hidalgo, G.; Baldassano, R.; Anokhin, A.; Heath, A.; Warner, B.; Reeder, J.; Kuczynski, J.; Caporaso, J.; Lozupone, C.; Lauber, C.; Clemente, J.; Knights, D.; Knight, R.; Gordon, J., Human gut microbiome viewed across age and geography. *Nature* **2012,** *486*(7402), 222-227.
4. Yadav, M.; Verma, M.; Chauhan, N., A review of metabolic potential of human gut microbiome in human nutrition. *Arch Microbiol* **2018,** *200* (2), 203-217.
5. Kamada, N.; Seo, S.; Chen, G.; Núñez, G., Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol* **2013,** *13* (5), 321-335.
6. Hooper, L.; Littman, D.; Macpherson, A., Interactions between the microbiota and the immune system. *Science* **2012,** *336* (6086), 1268-1273.
7. Dorrestein, P. C.; Mazmanian, S. K.; Knight, R., Finding the missing links among metabolites, microbes, and the host. *Immunity* **2014,** *40* (6), 824-832.
8. Adnani, N.; Rajski, S.; Bugni, T., Symbiosis-inspired approaches to antibiotic discovery. *Nat ProdRep* **2017,** *34*(7), 784-814.
9. Sender, R.; Fuchs, S.; Milo, R., Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell* **2016,** *164* (3), 337-340.
10. Cani, P.; Delzenne, N., The gut microbiome as therapeutic target. *Pharmacol Ther* **2011,** *130* (2), 202-212.
11. Li, J.; Jia, H.; Cai, X.; Zhong, H.; Feng, Q.; Sunagawa, S.; Arumugam, M.; Kultima, J.; Prifti, E.; Nielsen, T.; Juncker, A.; Manichanh, C.; Chen, B.; Zhang, W.; Levenez, F.; Wang, J.; Xu, X.; Xiao, L.; Liang, S.; Zhang, D.; Zhang, Z.; Chen, W.; Zhao, H.; Al-Aama, J.; Edris, S.; Yang, H.; Wang, J.; Hansen, T.; Nielsen, H.; Brunak, S.; Kristiansen, K.; Guarner, F.; Pedersen, O.; Doré, J.; Ehrlich, S.; Consortium, M.; Bork, P.; Wang, J., An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* **2014,** *32* (8), 834-841.
12. Falkow, S., Molecular Koch's postulates applied to microbial pathogenicity. *Rev Infect Dis* **1988,** *10*, S274-S276.
13. Byrd, A.; Segre, J., Adapting Koch's postulates. *Science* **2016,** *351* (6270), 224-226.
14. Ley, R.; Turnbaugh, P.; Klein, S.; Gordon, J., Microbial ecology: human gut microbes associated with obesity. *Nature* **2006,** *444* (7122), 1022-1023.
15. Turnbaugh, P.; Hamady, M.; Yatsunenko, T.; Cantarel, B.; Duncan, A.; Ley, R.; Sogin, M.; Jones, W.; Roe, B.; Affourtit, J.; Egholm, M.; Henrissat, B.; Heath, A.; Knight, R.; Gordon, J., A core gut microbiome in obese and lean twins. *Nature* **2009,** *457*, 480-484.
16. Seksik, P.; Rigottier-Gois, L.; Gramet, G.; Sutren, M.; Pochart, P.; Marteau, P.; Jian, R.; Doré, J., Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut* **2003,** *52* (2), 237-242.
17. Garrett, W.; Gallini, C.; Yatsunenko, T.; Michaud, M.; DuBois, A.; Delaney, M.; Punit, S.; Karlsson, M.; Bry, L.; Glickman, J.; Gordon, J.; Onderdonk, A.; Glimcher, L.,

Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **2010,** *8* (3), 292-300.
18. Palm, N.; de Zoete, M.; Cullen, T.; Barry, N.; Stefanowski, J.; Hao, L.; Degnan, P.; Hu, J.; Peter, I.; Zhang, W.; Ruggiero, E.; Cho, J.; Goodman, A.; Flavell, R., Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **2014,** *158* (5), 1000-1010.
19. Schofield, W.; Palm, N., Gut microbiota: IgA protects the pioneers. *Curr Biol* **2018,** *28* (18), R1117-R1119.
20. Donaldson, G.; Ladinsky, M.; Yu, K.; Sanders, J.; Yoo, B.; Chou, W.; Conner, M.; Earl, A.; Knight, R.; Bjorkman, P.; Mazmanian, S., Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science* **2018,** *360* (6390), 795-800.
21. Casadevall, A.; Pirofski, L., Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. *Infect Immun* **2000,** *68* (12), 6511-6518.
22. Brinker, P.; Fontaine, M.; Beukeboom, L.; Falcao Salles, J., Host, symbionts, and the microbiome: the missing tripartite interaction. *Trends Microbiol (in press)* **2019**.
23. Sczesnak, A.; Segata, N.; Qin, X.; Gevers, D.; Petrosino, J.; Huttenhower, C.; Littman, D.; Ivanov, I., The genome of Th17 cell-inducing segmented filamentous bacteria reveals extensive auxotrophy and adaptations to the intestinal environment. *Cell Host Microbe* **2011,** *10* (3), 260-272.
24. Rakoff-Nahoum, S.; Foster, K.; Comstock, L., The evolution of cooperation within the gut microbiota. *Nature* **2016,** *533*(7602), 255-259.
25. Strassmann, J.; Gilbert, O.; Queller, D., Kin discrimination and cooperation in microbes. *Annu Rev Microbiol* **2011,** *65* (1), 349-367.
26. Verster, A.; Ross, B.; Radey, M.; Bao, Y.; Goodman, A.; Mougous, J.; Borenstein, E., The landscape of type VI secretion across human gut microbiomes reveals its role in community composition. *Cell Host Microbe* **2017,** *22* (3), 411-419.e4.
27. Saha, R.; Saha, N.; Donofrio, R.; Bestervelt, L., Microbial siderophores: a mini review. *J Basic Microbiol* **2013,** *53* (4), 303-317.
28. Fischbach, M.; Lin, H.; Liu, D.; Walsh, C., How pathogenic bacteria evade mammalian sabotage in the battle for iron. *Nat chem biol* **2006,** *2* (3), 132-138.
29. Donia, M.; Fischbach, M., Small molecules from the human microbiota. *Science* **2015,** *349* (6246), 1254766.
30. Alexander, C.; Rietschel, E., Bacterial lipopolysaccharides and innate immunity. *J Endotoxin Res* **2001,** *7* (3), 167-202.
31. Chang, P.; Hao, L.; Offermanns, S.; Medzhitov, R., The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proc Natl Acad Sci USA* **2014,** *111* (6), 2247-2252.
32. Morrison, D.; Preston, T., Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut microbes* **2016,** *7* (3), 189-200.
33. Ohira, H.; Tsutsui, W.; Fujioka, Y., Are short chain fatty acids in gut microbiota defensive players for inflammation and atherosclerosis? *J Atheroscler Thromb* **2017,** *24* (7), 660-672.
34. Bäumler, A.; Sperandio, V., Interactions between the microbiota and pathogenic bacteria in the gut. *Nature* **2016,** *535* (7610), 85-93.
35. Jones, R.; Mercante, J.; Neish, A., Reactive oxygen production induced by the gut microbiota: pharmacotherapeutic implications. *Curr Med Chem* **2012,** *19* (10), 1519-1529.

36. Arnison, P.; Bibb, M.; Bierbaum, G.; Bowers, A.; Bugni, T.; Bulaj, G.; Camarero, J.; Campopiano, D.; Challis, G.; Clardy, J.; Cotter, P.; Craik, D.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P.; Entian, K.; Fischbach, M.; Garavelli, J.; Göransson, U.; Gruber, C.; Haft, D.; Hemscheidt, T.; Hertweck, C.; Hill, C.; Horswill, A.; Jaspars, M.; Kelly, W.; Klinman, J.; Kuipers, O.; Link, A.; Liu, W.; Marahiel, M.; Mitchell, D.; Moll, G.; Moore, B.; Müller, R.; Nair, S.; Nes, I.; Norris, G.; Olivera, B.; Onaka, H.; Patchett, M.; Piel, J.; Reaney, M.; Rebuffat, S.; Ross, R.; Sahl, H.; Schmidt, E.; Selsted, M.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R.; Tagg, J.; Tang, G.; Truman, A.; Vederas, J.; Walsh, C.; Walton, J.; Wenzel, S.; Willey, J.; van der Donk, W., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for universal nomenclature. *Nat Prod Rep* **2013,** *30*, 108-160.
37. Donia, M.; Cimermancic, P.; Schulze, C.; Wieland Brown, L.; Martin, J.; Mitreva, M.; Clardy, J.; Linington, R.; Fischbach, M., A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* **2014,** *158* (6), 1402-1414.
38. Skinnider, M.; Johnston, C.; Edgar, R.; Dejong, C.; Merwin, N.; Rees, P.; Magarvey, N., Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proc Natl Acad Sci USA* **2016,** *113* (42), E6343-6351.
39. Tietz, J.; Schwalen, C.; Patel, P.; Maxson, T.; Blair, P.; Tai, H.; Zakai, U.; Mitchell, D., A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol* **2017,** *13*(5), 470-478.
40. van Heel, A.; de Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O., BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res* **2013,** *41* (W1), W448-W453.
41. Sassone-Corsi, M.; Raffatellu, M., No vacancy: how beneficial microbes cooperate with immunity to provide colonization resistance to pathogens. *J Immunol* **2015,** *194* (9), 4081-4087.
42. Ravel, J.; Gajer, P.; Abdo, Z.; Schneider, G.; Koenig, S.; McCulle, S.; Karlebach, S.; Gorle, R.; Russell, J.; Tacket, C.; Brotman, R.; Davis, C.; Ault, K.; Peralta, L.; Forney, L., Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci USA* **2011,** *108*, 4680-4687.
43. Sassone-Corsi, M.; Nuccio, S.; Liu, H.; Hernandez, D.; Vu, C.; Takahashi, A.; Edwards, R.; Raffatellu, M., Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* **2016,** *540* (7632), 280-283.
44. Duquesne, S.; Destoumieux-Garzón, D.; Peduzzi, J.; Rebuffat, S., Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep* **2007,** *24* (4), 708-734.
45. Baquero, F.; Bouanchaud, D.; Martinez-Perez, M.; Fernandez, C., Microcin plasmids: a group of extrachromosomal elements coding for low-molecular-weight antibiotics in *Escherichia coli*. *J Bacteriol* **1978,** *135* (2), 342-347.
46. San Millan, J.; Hernandez-Chico, C.; Pereda, P.; Moreno, F., Cloning and mapping of the genetic determinants for microcin B17 production and immunity. *J Bacteriol* **1985,** *163* (1), 275-281.
47. Genilloud, O.; Moreno, F.; Kolter, R., DNA sequence, products and transcriptional pattern of the genes involved in production of the DNA replication inhibitor microcin B17. *J Bacteriol* **1989,** *171* (2), 1126-1135.
48. San Millán, J.; Kolter, R.; Moreno, F., Plasmid genes required for microcin B17 production. *J Bacteriol* **1985,** *163* (3), 1016-1020.

49. Li, Y.; Milne, J.; Madison, L.; Kolter, R.; Walsh, C., From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* **1996,** *274* (5290), 1188-93.
50. Dunbar, K.; Melby, J.; Mitchell, D., YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *Nat Chem Biol* **2012,** *8* (6), 569-575.
51. Madison, L.; Vivas, E.; Li, Y.; Walsh, C.; Kolter, R., The leader peptide is essential for the post-translational modification of the DNA-gyrase inhibitor microcin B17. *Mol. Microbiol.* **1997,** *23*(1), 161-168.
52. Allali, N.; Afif, H.; Couturier, M.; Van Melderen, L., The highly conserved TldD and TldE proteins of *Escherichia coli* are involved in microcin B17 processing and in CcdA degradation. *J Bacteriol* **2002,** *184*(12), 3224-3231.
53. Garrido, M.; Herrero, M.; Kolter, R.; Moreno, F., The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell. *EMBO J* **1988,** *7* (6), 1853-1862.
54. Asensio, C.; Pérez-Díaz, J., A new family of low molecular weight antibiotics from enterobacteria. *Biochem Biophys Res Commun* **1976,** *69*(1), 7-14.
55. Heddle, J.; Blance, S.; Zamble, D.; Hollfelder, F.; Miller, D.; Wentzell, L.; Walsh, C.; Maxwell, A., The antibiotic microcin B17 is a DNA gyrase poison: characterisation of the mode of inhibition. *J Mol Biol* **2001,** *307* (5), 1223-1234.
56. Patzer, S.; Baquero, M.; Bravo, D.; Moreno, F.; Hantke, K., The colicin G, H and X determinants encode microcins M and H47, which might utilize the catecholate siderophore receptors FepA, Cir, Fiu and IroN. *Microbiology* **2003,** *149*(9), 2557-2570.
57. Vassiliadis, G.; Destoumieux-Garzón, D.; Lombard, C.; Rebuffat, S.; Peduzzi, J., Isolation and characterization of two members of the siderophore-microcin family, microcins M and H47. *Antimicrob Agents Chemother* **2010,** *54* (1), 288-297.
58. Nolan, E.; Walsh, C., Investigations of the MceIJ-catalyzed posttranslational modification of the microcin E492 C-terminus: linkage of ribosomal and nonribosomal peptides to form "trojan horse" antibiotics. *Biochemistry* **2008,** *47* (35), 9289-9299.
59. Rebuffat, S., Microcins in action: amazing defence strategies of Enterobacteria. *Biochem Soc Trans* **2012,** *40* (6), 1456-1462.
60. Bagley, M.; Dale, J.; Merritt, E.; Xiong, X., Thiopeptide antibiotics. *Chem Rev* **2005,** *105* (2), 685-714.
61. Hughes, R.; Moody, C., From amino acids to heteroarenes - Thiopeptide antibiotics as heterocyclic peptides from Nature. *Applied Chemistry* **2007,** *119* (42).
62. Wieland Brown, L.; Acker, M.; Clardy, J.; Walsh, C.; Fischbach, M., Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. *Proc Natl Acad Sci USA* **2009,** *106* (8), 2549-2553.
63. Bowers, A.; Walsh, C.; Acker, M., Genetic interception and structural characterization of thiopeptide cyclization precursors form *Bacillus cereus*. *J Am Chem Soc* **2010,** *132* (35), 12182-12184.
64. Shoji, J.; Hinoo, H.; Wakisaka, Y.; Koizumi, K.; Mayama, M.; Matsuura, S.; Matsumoto, K., Isolation of three new antibiotics, thiocillins I, II and III, related to micrococcin P. Studies on antibiotics from the genus *Bacillus*. VIII. *J antibiot (Tokyo)* **1976,** *29* (4), 366-374.
65. Bycroft, B.; Gowland, M., The structures of the highly modified peptide antibiotics Micrococcin P_1 and P_2. *JCS Chem Comm* **1978**, (6), 256-258.

66. Suzumura, K.; Yokoi, T.; Funatsu, M.; Nagai, K.; Tanaka, K.; Zhang, H.; Suzuki, K., YM-266183 and YM-266184, novel thiopeptide antibiotics produced by *Bacillus cereus* isolated from a marine sponge II. Structure elucidation. *J antibiot (Tokyo)* **2003,** *56* (2), 129-134.
67. Bogart, J.; Bowers, A., Thiopeptide pyridine synthase TbtD catalyzes an intermolecular formal aza-Diels-Alder reactions. *J Am Chem Soc* **2019,** *141* (5), 1842-1846.
68. Ogle, J.; Brodersen, D.; Clemons Jr, W.; Tarry, M.; Carter, A.; Ramakrishnan, V., Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **2001,** *292* (5518), 897-902.
69. Fischbach, M.; Walsh, C., Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **2006,** *106* (8), 3468-3496.
70. Marahiel, M.; Stachelhaus, T.; Mootz, H., Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem Rev* **1997,** *97* (7), 2651-2673.
71. von Döhren, H.; Keller, U.; Vater, J.; Zocher, R., Multifunctional peptide synthetases. *Chem Rev* **1997,** *97*(7), 2675-2705.
72. Schwarzer, D.; Finking, R.; Marahiel, M., Nonribosomal peptides: from genes to products. *Nat Prod Rep* **2003,** *20* (3), 275-287.
73. Cox, C.; Rinehart Jr, K.; Moore, M.; Cook Jr, J., Pyochelin: novel structure of an iron-chelating growth promoter for *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* **1981,** *78* (7), 4256-4260.
74. Crosa, J.; Walsh, C., Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev* **2002,** *66* (2), 223-249.
75. Baldwin, J.; Sir Edward Abraham, The biosynthesis of penicillins and cephalosporins. *Nat Prod Rep* **1988,** *5* (2), 129-145.
76. Aharonowitz, Y.; Cohen, G., Penicillin and cephalosporin biosynthetic genes: structure, organization, regulation and evolution. *Annu Rev Microbiol* **1992,** *46*, 461-495.
77. Müller, C.; Oberauner-Wappis, L.; Peyman, A.; Amos, G.; Wellington, E.; Berg, G., Mining for nonribosomal peptide synthetase and polyketide synthase genes revealed a high level of diversity in the *Sphagnum* bog metagenome. *Appl Environ Microbiol* **2015,** *81* (15), 5064-5072.
78. Valiante, V., The cell wall integrity signaling pathway and its involvement in secondary metabolite production. *J Fungi (Basel)* **2017,** *3*(4), 68.
79. Yim, G.; Wang, H.; Davies FRS, J., Antibiotics as signalling molecules. *Philos Trans R Soc Lond B Biol Sci* **2007,** *362* (1483), 1195-1200.
80. Linares, J.; Gustafsson, I.; Baquero, F.; Martinez, J., Antibiotics as intermicrobial signaling agents instead of weapons. *Proc Natl Acad Sci USA* **2006,** *103* (51), 19484-19489.
81. Cimermancic, P.; Medema, M.; Claesen, J.; Kurita, K.; Wieland Brown, L.; Mavrommatis, K.; Pati, A.; Godfrey, P.; Koehrsen, M.; Clardy, J.; Birren, B.; Takano, E.; Sali, A.; Linington, R.; Fischbach, M., Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **2014,** *158* (2), 412-421.
82. Zipperer, A.; Konnerth, M.; Laux, C.; Berscheid, A.; Janek, D.; Weidenmaier, C.; Burian, M.; Schilling, N.; Slavetinsky, C.; Marschal, M.; Willmann, M.; Kalbacher, H.; Schittek, B.; Brötz-Oesterhelt, H.; Grond, S.; Peschel, A.; Krismer, B., Human commensals producing a novel antibiotic impair pathogen colonization. *Nature* **2016,** *535* (7613), 511-516.
83. Beaugerie, L.; Metz, M.; Barbut, F.; Bellaiche, G.; Bouhnik, Y.; Raskine, L.; Nicolas, J.; Chatelet, F.; Lehn, N.; Petit, J.; Group, T. I. C. S., *Klebsiella oxytoca* as an agent of antibiotic-associated hemorrhagic colitis. *Clin Gastroenterol Hepatol* **2003,** *1* (5), 370-376.

84. Högenauer, C.; Langner, C.; Beubler, E.; Lippe, I.; Schicho, R.; Gorkiewicz, G.; Krause, R.; Gerstgrasser, N.; Krejs, G.; Hinterleitner, T., *Klebsiella oxytoca* as a causative organism of antibiotic-associated hemorrhagic colitis. *N Engl J Med* **2006,** *355*(23), 2418-2426.
85. Zollner-Schwetz, I.; Högenauer, C.; Joainig, M.; Weberhofer, P.; Gorkiewicz, G.; Valentin, T.; Hinterleitner, T.; Krause, R., Role of *Klebsiella oxytoca* in antibiotic-associated diarrhea. *Clin Infect Dis* **2008,** *47* (9), e74-e78.
86. Schneditz, G.; Rentner, J.; Roier, S.; Pletz, J.; Herzog, K.; Bücker, R.; Troeger, H.; Schild, S.; Weber, H.; Breinbauer, R.; Gorkiewicz, G.; Högenauer, C.; Zechner, E., Enterotoxicity of a nonribosomal peptide causes antibiotic-associated colitis. *Proc Natl Acad Sci USA* **2014,** *111* (36), 13181-13186.
87. Hurley, L., Pyrrolo(1,4)benzodiazepine antitumor antibiotics. Comparative aspects of anthramycin, tomaymycin and sibiromycin. *J antibiot (Tokyo)* **1977,** *30* (5), 349-370.
88. Hurley, L.; Reck, T.; Thurston, D.; Langley, D.; Holden, K.; Hertzberg, R.; Hoover, J.; Gallagher Jr, G. J.; Faucette, L.; Mong, S.; Johnson, R., Pyrrolo[1,4]benzodiazepine antitumor antibiotics: relationship of DNA alkylation and sequence specificity to the biological activity of natural and synthetic compounds. *Chem Res Toxicol* **1988,** *1*(5), 258-268.
89. Hartley, J., The development of pyrrolobenzodiazepines as antitumour agents. *Expert Opin Investig Drugs* **2011,** *20* (6), 733-744.
90. Dornisch, E.; Pletz, J.; Glabonjat, R.; Martin, F.; Lembacher-Fadum, C.; Neger, M.; Högenauer, C.; Francesconi, K.; Kroutil, W.; Zangger, K.; Breinbauer, R.; Zechner, E., Biosynthesis of the enterotoxic pyrrolobenzodiazepine natural product tilivalline. *Angew Chem Int Ed* **2017,** *56* (46), 14753-14757.
91. Tse, H.; Gu, Q.; Sze, K.; Chu, I.; Kao, R.; Lee, K.; Lam, C.; Yang, D.; Shing-Chiu Tai, S.; Ke, Y.; Chan, E.; Chan, W.; Dai, J.; Leung, S.; Leung, S.; Yuen, K., A tricyclic pyrrolobenzodiazepine produced by *Klebsiella oxytoca* is associated with cytotoxicity in antibiotic-associated hemorrhagic colitis. *J Biol Chem* **2017,** *292* (47), 19503-19520.
92. von Tesmar, A.; Hoffmann, M.; Fayad, A.; Hüttel, S.; Schmitt, V.; Herrmann, J.; Müller, R., Biosynthesis of the *Klebsiella oxytoca* pathogenicity factor tilivalline: heterologous expression, *in vitro* biosynthesis, and inhibitor development. *ACS Chem Biol* **2018,** *13*, 812-819.
93. Unterhauser, K.; Pöltl, L.; Schneditz, G.; Kienesberger, S.; Glabonjat, R.; Kitsera, M.; Pletz, J.; Josa-Prado, F.; Dornisch, E.; Lembacher-Fadum, C.; Roier, S.; Gorkiewicz, G.; Lucena, D.; Barasoain, I.; Kroutil, W.; Wiedner, M.; Loizou, J.; Breinbauer, R.; Díaz, J.; Schild, S.; Högenauer, C.; Zechner, E., *Klebsiella oxytoca* enterotoxins tilimycin and tilivalline have distinct host DNA-damaging and microtubule-stabilizing activities. *Proc Natl Acad Sci USA* **2019,** *116*(9), 3774-3783.
94. Wyatt, M.; Wang, W.; Roux, C.; Beasley, F.; Heinrichs, D.; Dunman, P.; Magarvey, N., *Staphylococcus aureus* nonribosomal peptide secondary metabolites regulate virulence. *Science* **2010,** *329*(5989), 294-296.
95. Zimmermann, M.; Fischbach, M., A family of pyrazinone natural products form a conserved nonribosomal peptide synthetase in *Staphylococcus aureus*. *Chem Biol* **2010,** *17* (9), 925-930.
96. Alvarez, M.; White, C.; Gregory, J.; Kydd, G.; Harris, A.; Sun, H.; Gillum, A.; Cooper, R., Phevalin, a new calpain inhibitor, from a *Streptomyces* sp. *J Antibiot* **1995,** *48* (10), 1165-1167.
97. Secor, P.; Jennings, L.; James, G.; Kirker, K.; deLancey Pulcini, E.; Mclnnerney, K.; Gerlach, R.; Livinghouse, T.; Hilmer, J.; Bothner, B.; Fleckman, P.; Olerud, J.; Stewart, P., Phevalin

(aureusimine B) production by *Staphylococcus aureus* biofilm and impacts on human keratinocyte gene expression. *PLoS One* **2012,** *7* (7), e40973.
98. Park, H.; Crawford, J., Pyrazinone protease inhibitor metabolites from *Photorhabdus luminescens*. *J antibiot (Tokyo)* **2016,** *69* (8), 616-621.
99. Belfrage, A.; Abdurakhmanov, E.; Åkerblom, E.; Brandt, P.; Alogheli, H.; Neyts, J.; Danielson, U.; Sandström, A., Pan-NS3 protease inhibitors of hepatitis C virus based on an R^3-elongated pyrazinone scaffold. *Eur J Med Chem* **2018,** *148*, 453-464.
100. Gising, J.; Belfrage, A.; Alogheli, H.; Ehrenberg, A.; Åkerblom, E.; Svensson, R.; Artursson, P.; Karlén, A.; Danielson, U.; Larhed, M.; Sandström, A., Achiral pyrazinone-based inhibitors of the hepatitis C virus NS3 protease and drug-resistant variants with elongated substituents directed toward the S2 pocket. *J Med Chem* **2014,** *57*(5), 1790-1801.
101. Kyeremeh, K.; Acquah, K.; Camas, M.; Tabudravu, J.; Houssen, W.; Deng, H.; Jaspars, M., Butrepyrazinone, a new pyrazinone with an unusual methylation pattern from a Ghanaian *Verrucosispora* sp. K51G. *Mar. Drugs* **2014,** *12* (10), 5197-5208.
102. Papenfort, K.; Silpe, J.; Schramma, K.; Cong, J.; Seyedsayamdost, M.; Bassler, B., A *Vibrio cholerae* autoinducer–receptor pair that controls biofilm formation. *Nat Chem Biol* **2017,** *13*(5), 551-557.
103. Guo, C.; Chang, F.; Wyche, T.; Backus, K.; Acker, T.; Funabashi, M.; Taketani, M.; Donia, M.; Nayfach, S.; Pollard, K.; Craik, C.; Cravatt, B.; Clardy, J.; Voigt, C.; Fischbach, M., Discovery of reactive microbiota-derived metabolites that inhibit host proteases. *Cell* **2017,** *168* (3), 517-526.
104. Carlson-Banning, K.; Sperandio, V., Enterohemorrhagic *Escherichia coli* outwits hosts through sensing small molecules. *Curr Opin Microbiol* **2018,** *41*, 83-88.
105. Lustri, B.; Sperandio, V.; Moreira, C., Bacterial chat: intestinal metabolites and signals in host-microbiota-pathogen interactions. *Infect Immun* **2017,** *85* (12), e00476-17.
106. Schneider, B.; Balskus, E., Discovery of small molecule protease inhibitors by investigating a widespread human gut bacterial biosynthetic pathway. *Tetrahedron* **2018,** *74* (26), 3215-3230.
107. George, K.; Chatterjee, D.; Gunawardana, G.; Welty, D.; Hayman, J.; Lee, R.; Small, P., Mycolactone: a polyketide toxin from *Mycobacterium ulcerans* required for virulence. *Science* **1999,** *283* (5403), 854-857.
108. Staunton, J.; Weissman, K., Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **2001,** *18* (4), 380-416.
109. White, S.; Zheng, J.; Zhang, Y.; Rock, C., The structural biology of Type II fatty acid biosynthesis. *Annu Rev Biochem* **2005,** *74*, 791-831.
110. Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **2003,** *7*(2), 285-295.
111. Ridley, C.; Lee, H.; Khosla, C., Evolution of polyketide synthases in bacteria. *Proc Natl Acad Sci USA* **2008,** *105* (12), 4595-4600.
112. Zhou, H.; Li, Y.; Tang, Y., Cyclization of aromatic polyketides from bacteria and fungi. *Nat Prod Rep* **2010,** *27* (6), 839-868.
113. Sandmann, A.; Dickschat, J.; Jenke-Kodama, H.; Kunze, B.; Dittmann, E.; Müller, R., A type II polyketide synthase from the gram-negative bacterium *Stigmatella aurantiaca* is involved in Aurachin alkaloid biosynthesis. *Angew Chem Int Ed Engl* **2007,** *46* (15), 2712-2716.

114. Lozano, G.; Bravo, J.; Garavito Diago, M.; Park, H.; Hurley, A.; Peterson, S.; Stabb, E.; Crawford, J.; Broderick, N.; Handelsman, J., Introducing THOR, a Model Microbiome for Genetic Dissection of Community Behavior. *mBio* **2019,** *10* (2), e02846-18.
115. Lozano, G.; Park, H.; Bravo, J.; Armstrong, E.; Denu, J.; Stabb, E.; Broderick, N.; Crawford, J.; Handelsman, J., Bacterial analogs of plant piperidine alkaloids mediate microbial interactions in a rhizosphere model system. *bioRxiv* **2018**, 499731.
116. Shimizu, Y.; Ogata, H.; Goto, S., Type III polyketide synthases: functional classification and phylogenomics. *ChemBioChem* **2017,** *18* (1), 50-65.
117. Paul, V.; Frautschy, S.; Fenical, W.; Nealson, K., Antibiotics in microbial ecology: isolation and structure assignment of several new antibacterial compounds from the insect-symbiotic bacteria *Xenorhabdus* spp. *J. Chem. Ecol.* **1981,** *7*(3), 589-597.
118. Weissfeld, A.; Halliday, R.; Simmons, D.; Trevino, E.; Vance, P.; O'Hara, C.; Sowers, E.; Kern, R.; Koy, R.; Hodde, K.; Bing, M.; Lo, C.; Gerrard, J.; Vohra, R.; Harper, J., *Photorhabdus asymbiotica*, a pathogen emerging on two continents that proves that there is no substitute for a well-trained clinical microbiologist. *J Clin Microbiol* **2005,** *43* (8), 4152-4155.
119. Gerrard, J.; Waterfield, N.; Vohra, R.; ffrench-Constant, R., Human infection with Photorhabdus asymbiotica: an emerging bacterial pathogen. *Microbes Infect* **2004,** *6* (2), 229-237.
120. Kobayashi, A.; Kang, M.; Watai, Y.; Tong, K.; Shibata, T.; Uchida, K.; Yamamoto, M., Oxidative and electrophilic stresses activate Nrf2 through inhibition of ubiquitination activity of Keap1. *Mol Cell Biol* **2006,** *26* (1), 221-229.
121. Smith, S.; Jayawickreme, C.; Rickard, D.; Nicodeme, E.; Bui, T.; Simmons, C.; Coquery, C.; Neil, J.; Pryor, W.; Mayhew, D.; Rajpal, D.; Creech, K.; Furst, S.; Lee, J.; Wu, D.; Rastinejad, F.; Willson, T.; Viviani, F.; Morris, D.; Moore, J.; Cote-Sierra, J., Tapinarof is a natural AhR agonist that resolves skin inflammation in mice and humans. *J Invest Dermatol* **2017,** *137*(10), 2110-2119.
122. Shi, D.; An, R.; Zhang, W.; Zhang, G.; Yu, Z., Stilbene derivatives from *Photorhabdus temperata* SN259 and their antifungal activities against phytopathogenic fungi. *J Agric Food Chem* **2017,** *65* (1), 60-65.
123. Hu, K.; Li, J.; Li, B.; Webster, J.; Chen, G., A novel antimicrobial epoxide isolated from larval *Galleria mellonella* infected by the nematode symbiont, *Photorhabdus luminescens* (Enterobacteriaceae). *Bioorg Med Chem* **2006,** *14* (13), 4677-4681.
124. Park, H.; Crawford, J., Lumiquinone A, an □-aminomalonate-derived aminobenzoquinone from *Photorhabdus luminescens*. *J Nat Prod* **2015,** *78*(6), 1437-1441.
125. Williams, J.; Thomas, M.; Clarke, D., The gene *stlA* encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology* **2005,** *151* (8), 2543-2550.
126. Joyce, S.; Brachmann, A.; Glazer, I.; Lango, L.; Schwär, G.; Clarke, D.; Bode, H., Bacterial biosynthesis of a multipotent stilbene. *Angew Chem Int Ed Engl* **2008,** *47*(10), 1942-1945.
127. Mori, T.; Awakawa, T.; Shimomura, K.; Saito, Y.; Yang, D.; Morita, H.; Abe, I., Structural insight into the enzymatic formation of bacterial stilbene. *Cell Chem Biol* **2016,** *23* (12), 1468-1479.
128. Park, H.; Sampathkumar, P.; Perez, C.; Lee, J.; Tran, J.; Bonanno, J.; Hallem, E.; Almo, S.; Crawford, J., Stilbene epoxidation and detoxification in a *Photorhabdus luminescens*-nematode symbiosis. *J Biol Chem* **2017,** *292*(16), 6680-6694.

129. Kontnik, R.; Crawford, J.; Clardy, J., Exploiting a Global Regulator for Small Molecule Discovery in *Photorhabdus luminescens*. *ACS Chem Biol* **2010,** *5* (7), 659-665.
130. Shi, Y.; Bode, H., Chemical language and warfare of bacterial natural products in bacteria–nematode–insect interactions. *Nat Prod Rep* **2018,** *35* (4), 309-335.
131. Shen, T.; Wang, X.; Lou, H., Natural stilbenes: an overview. *Nat Prod Rep* **2009,** *26* (7), 916-935.
132. Martin, D.; Bolling, B., A review of the efficacy of dietary polyphenols in experimental models of inflammatory bowel diseases. *Food Funct* **2015,** *6*(6), 1773-1786.
133. Barnett, M.; Cooney, J.; Dommels, Y.; Nones, K.; Brewster, D.; Park, Z.; Butts, C.; McNabb, W.; Laing, W.; Roy, N., Modulation of colonic inflammation in Mdr1a$^{-/-}$ mice by green tea polyphenols and their effects on the colon transcriptome and proteome. *J Nut Biochem* **2013,** *24*(10), 1678-1690.
134. Koppel, N.; Rekdal, V.; Balskus, E., Chemical transformation of xenobiotics by the human gut microbiota. *Science* **2017,** *356* (6344), eaag2770.
135. Perez, C.; Crawford, J., Characterization of a hybrid nonribosomal peptide–carbohydrate biosynthetic pathway in *Photorhabdus luminescens*. *Biochemistry* **2019,** *58* (8), 1131-1140.
136. Guo, X.; Crawford, J., An atypical orphan carbohydrate-NRPS genomic island encodes a novel lytic transglycosylase. *Chem Biol* **2014,** *21* (10), 1271-1277.
137. Perez, C.; Park, H.; Crawford, J., Functional characterization of a condensation domain that links nonribosomal peptide and pteridine biosynthetic machineries in *Photorhabdus luminescens*. *Biochemistry* **2018,** *57* (3), 354-361.
138. Park, H.; Perez, C.; Barber, K.; Rinehart, J.; Crawford, J., Genome mining unearths a hybrid nonribosomal peptide synthetase-like-pteridine synthase biosynthetic gene cluster. *eLife* **2017,** *6*, e25229.
139. Wang, H.; Fewer, D.; Holm, L.; Rouhiainen, L.; Sivonen, K., Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc Natl Acad Sci USA* **2014,** *111* (25), 9259-9264.
140. Cheng, Y.; Tang, G.; Shen, B., Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci USA* **2003,** *100* (6), 3149-3154.
141. Lohman, J.; Ma, M.; Osipiuk, J.; Nocek, B.; Kim, Y.; Chang, C.; Cuff, M.; Mack, J.; Bigelow, L.; Li, H.; Endres, M.; Babnigg, G.; Joachimiak, A.; Phillips Jr., G.; Shen, B., Structural and evolutionary relationships of "AT-less" type I polyketide synthase ketosynthases. *Proc Natl Acad Sci USA* **2015,** *112* (41), 12693-12698.
142. Engel, P.; Vizcaino, M.; Crawford, J., Gut symbionts from distinct hosts exhibit genotoxic activity via divergent colibactin biosynthesis pathways. *Appl Environ Microbiol* **2015,** *81*(4), 1502-1512.
143. Helfrich, E.; Piel, J., Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* **2016,** *33* (2), 231-316.
144. Secher, T.; Brehin, C.; Oswald, E., Early settlers: which *E.coli* strains do you not want at birth? *Am J Physiol Gastrointest Liver Physiol* **2016,** *311* (1), G123-G129.
145. Nougayrède, J.; Homburg, S.; Taieb, F.; Boury, M.; Brzuszkiewicz, E.; Gottschalk, G.; Buchrieser, C.; Hacker, J.; Dobrindt, U.; Oswald, E., *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **2006,** *313* (5788), 848-851.

146. Johnson, J.; Johnston, B.; Kuskowski, M.; Nougayrede, J.; Oswald, E., Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* pks genomic island. *J Clin Microbiol* **2008,** *46* (12), 3906-3911.
147. Putze, J.; Hennequin, C.; Nougayrède, J.; Zhang, W.; Homburg, S.; Karch, H.; Bringer, M.; Fayolle, C.; Carniel, E.; Rabsch, W.; Oelschlaeger, T.; Oswald, E.; Forestier, C.; Hacker, J.; Dobrindt, U., Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect Immun* **2009,** *77* (11), 4696-4703.
148. Trautman, E.; Crawford, J., Linking biosynthetic gene clusters to their metabolites via pathway-targeted molecular networking. *Curr Top Med Chem* **2016,** *16* (15), 1705-1716.
149. Balskus, E., Colibactin: understanding an elusive gut bacterial genotoxin. *Nat Prod Rep* **2015,** *32* (11), 1534-1540.
150. Taieb, F.; Petit, C.; Nougayrède, J.; Oswald, E., The enterobacterial genotoxins: cytolethal distending toxin and colibactin. *EcoSal Plus* **2016,** *7*(1), doi: 10.1128/ecosalplus.ESP-0008-2016.
151. Healy, A.; Herzon, S., Molecular basis of gut microbiome-associated colorectal cancer: a synthetic perspective. *J Am Chem Soc* **2017,** *139* (42), 14817-14824.
152. Faïs, T.; Delmas, J.; Barnich, N.; Bonnet, R.; Dalmasso, G., Colibactin: more than a new bacterial toxin. *Toxins (Basel)* **2018,** *10* (4), 151.
153. Cuevas-Ramos, G.; Petit, C.; Marcq, I.; Boury, M.; Oswald, E.; Nougayrède, J., *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci USA* **2010,** *107* (25), 11537-11542.
154. Arthur, J.; Perez-Chanona, E.; Mühlbauer, M.; Tomkovich, S.; Uronis, J.; Fan, T.; Campbell, B.; Abujamel, T.; Dogan, B.; Rogers, A.; Rhodes, J.; Stintzi, A.; Simpson, K.; Hansen, J.; Keku, T.; Fodor, A.; Jobin, C., Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **2012,** *338* (6103), 120-123.
155. Secher, T.; Samba-Louaka, A.; Oswald, E.; Nougayrède, J., *Escherichia coli* producing colibactin triggers premature and transmissible senescence in mammalian cells. *PLoS One* **2013,** *8* (10), e77157.
156. Tomkovich, S.; Yang, Y.; Winglee, K.; Gauthier, J.; Mühlbauer, M.; Sun, X.; Mohamadzadeh, M.; Liu, X.; Martin, P.; Wang, G.; Oswald, E.; Fodor, A.; Jobin, C., Locoregional effects of microbiota in a preclinical model of colon carcinogenesis. *Cancer Res* **2017,** *77* (10), 2620-2632.
157. Bonnet, M.; Buc, E.; Sauvanet, P.; Darcha, C.; Dubois, D.; Pereira, B.; Déchelotte, P.; Bonnet, R.; Pezet, D.; Darfeuille-Michaud, A., Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin Cancer Res* **2014,** *20* (4), 859-867.
158. Dejea, C.; Fathi, P.; Craig, J.; Boleij, A.; Taddese, R.; Geis, A.; Wu, X.; DeStefano Shields,1C.; Hechenbleikner, E.; Huso, D.; Anders,R.; Giardiello, F.; Wick, E.; Wang, H.; Wu, S.; Pardoll, D.; Housseau, F.; Sears, C., Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **2018,** *359* (6375), 592-597.
159. Bossuet-Greif, N.; Vignard, J.; Taieb, F.; Mirey, G.; Dubois, D.; Petit, C.; Oswald, E.; Nougayrède, J., The colibactin genotoxin generates DNA interstrand crosslinks in infected cells. *mBio* **2018,** *9* (2), e02393-17.
160. Xue, M.; Shine, E.; Wang, W.; Crawford, J.; Herzon, S., Characterization of natural colibactin-nucleobase adducts by tandem mass spectrometry and isotopic labeling. Support for DNA alkylation by cyclopropane ring opening. *Biochemistry* **2018,** *57* (45), 6391-6394.

161. Wilson, M.; Jiang, Y.; Willalta, P.; Stornetta, A.; Boudreau, P.; Carrá, A.; Brennan, .; Chun, E.; Ngo, L.; Samson, L.; Engelward, B.; Garrett, W.; Balbo, S.; Balskus, E., The human gut bacterial genotoxin colibactin alkylates DNA. *Science* **2019,** *363* (6428), eaar7785.
162. Shine, E.; Xue, M.; Patel, J.; Healy, A.; Surovtseva, Y.; Herzon, S.; Crawford, J., Model colibactins exhibit human cell genotoxicity in the absence of host bacteria. *ACS Chem Biol* **2018,** *13* (12), 3286-3293.
163. Olier, M.; Marcq, I.; Salvador-Cartier, C.; Secher, T.; Dobrindt, U.; Boury, M.; Bacquié, V.; Penary, M.; Gaultier, E.; Nougayrède, J.; Fioramonti, J.; Oswald, E., Genotoxicity of *Escherichia coli* Nissle 1917 strain cannot be dissociated from its probiotic activity. *Gut Microbes* **2012,** *3*(6), 501-509.
164. Buc, E.; Dubois, D.; Sauvanet, P.; Raisch, J.; Delmas, J.; Darfeuille-Michaud, A.; Pezet, D.; Bonnet, R., High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PLoS One* **2013,** *8* (2), e56964.
165. Muñoz-Fontela, C.; Mandinova, A.; Aaronson, S.; Lee, S., Emerging roles of p53 and other tumour-suppressor genes in immune regulation. *Nat Rev Immunol* **2016,** *16*(12), 741-750.
166. Lupp, C.; Robertson, M.; Wickham, M.; Sekirov, I.; Champion, O.; Gaynor, E.; Finlay, B., Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe* **2007,** *2* (2), 119-129.
167. Meador, J.; Caldwell, M.; Cohen, P.; Conway, T., *Escherichia coli* pathotypes occupy distinct niches in the mouse intestine. *Infect Immu* **2014,** *82* (5), 1931-1938.
168. Martin, P.; Marcq, I.; Magistro, G.; Penary, M.; Garcie, C.; Payros, D.; Boury, M.; Olier, M.; Nougayrède, J.; Audebert, M.; Chalut, C.; Schubert, S.; Oswald, E., Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLOS Pathogens* **2013,** *9* (7), e1003437.
169. Homburg, S.; Oswald, E.; Hacker, J.; Dobrindt, U., Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. *FEMS Microbiol Lett* **2007,** *275* (2), 255-262.
170. Trautman, E.; Healy, A.; Shine, E.; Herzon, S.; Crawford, J., Domain-targeted metabolomics delineates the heterocycle assembly steps of colibactin biosynthesis. *J Am Chem Soc* **2017,** *139* (11), 4195-4201.
171. Li, Z.; Li, J.; Gu, J.; Lai, J.; Duggan, B.; Zhang, W.; Li, Z.; Li, Y.; Tong, R.; Xu, Y.; Lin, D.; Moore, B.; Qian, P., Divergent biosynthesis yields a cytotoxic aminomalonate-containing precolibactin. *Nat Chem Biol* **2016,** *12* (10), 773-775.
172. Guntaka, N.; Healy, A.; Crawford, J.; Herzon, S.; Bruner, S., Structure and functional analysis of ClbQ, an unusual intermediate-releasing thioesterase from the colibactin biosynthetic pathway. *ACS Chem Biol* **2017,** *12* (10), 2598-2608.
173. Mousa, J.; Newsome, R.; Yang, Y.; Jobin, C.; Bruner, S., ClbM is a versatile, cation-promiscuous MATE transporter found in the colibactin biosynthetic gene cluster. *Biochem Biophys Res Commun* **2017,** *482* (4), 1233-1239.
174. Mousa, J.; Yang, Y.; Tomkovich, S.; Shima, A.; Newsome, R.; Tripathi, P.; Oswald, E.; Bruner, S.; Jobin, C., MATE transport of the *E. coli*-derived genotoxin colibactin. *Nat Microbiol* **2016,** *1*, 15009.

175. Bossuet-Greif, N.; Dubois, D.; Petit, C.; Tronnet, S.; Martin, P.; Bonnet, R.; Oswald, E.; Nougayrède, J., *Escherichia coli* ClbS is a colibactin resistance protein. *Mol Microbiol* **2016,** *99* (5), 897-908.
176. Tripathi, P.; Shine, E.; Healy, A.; Kim, C.; Herzon, S.; Bruner, S.; Crawford, J., ClbS is a cyclopropane hydrolase that confers colibactin resistance. *J Am Chem Soc* **2017,** *139* (49), 17719-17722.
177. Dubois, D.; Baron, O.; Cougnoux, A.; Delmas, J.; Pradel, N.; Boury, M.; Bouchon, B.; Bringer, M.; Nougayrède, J.; Oswald, E.; Bonnet, R., ClbP is a prototype of a peptidase subgroup involved in biosynthesis of nonribosomal peptides. *J Biol Chem* **2011,** *286* (41), 35562-35570.
178. Cougnoux, A.; Gibold, L.; Robin, F.; Dubois, D.; Pradel, N.; Darfeuille-Michaud, A.; Dalmasso, G.; Delmas, J.; Bonnet, R., Analysis of structure-function relationships in the colibactin-maturating enzyme ClbP. *J Mol Biol* **2012,** *424* (3-4), 203-214.
179. Brotherton, C.; Balskus, E., A prodrug resistance mechanism is involved in colibactin biosynthesis and cytotoxicity. *J Am Chem Soc* **2013,** *135* (9), 3359-3362.
180. Bian, X.; Fu, J.; Plaza, A.; Herrmann, J.; Pistorius, D.; Stewart, A.; Zhang, Y.; Müller, R., In vivo evidence for a prodrug activation mechanism during colibactin maturation. *ChemBioChem* **2013,** *14*, 1194-1197.
181. Vizcaino, M.; Engel, P.; Trautman, E.; Crawford, J., Comparative metabolomics and structural characterizations illuminate colibactin pathway-dependent small molecules. *J Am Chem Soc* **2014,** *136* (26), 9244-9247.
182. Healy, A.; Nikolayevskiy, H.; Patel, J.; Crawford, J.; Herzon, S., A mechanistic model for colibactin-induced genotoxicity. *J Am Chem Soc* **2016,** *138* (48), 15563-15570.
183. MacMillan, K.; Boger, D., Fundamental relationships between structure, reactivity and biological activity for the Duocarmycins and CC-1065. *J Med Chem* **2009,** *52* (19), 5771-5780.
184. Brachmann, A.; Garcie, C.; Wu, V.; Martin, P.; Ueoka, R.; Oswald, E.; Piel, J., Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor. *Chem Comm* **2015,** *51* (66), 13138-41.
185. Zha, L.; Wilson, M.; Brotherton, C.; Balskus, E., Characterization of polyketide synthase machinery from the pks island facilitates isolation of a candidate precolibactin. *ACS Chem Biol* **2016,** *11* (5), 1287-1295.
186. Vizcaino, M.; Crawford, J., The colibactin warhead crosslinks DNA. *Nat Chemi* **2015,** *7*, 411-417.
187. Bian, X.; Plaza, A.; Zhang, Y.; Müller, R., Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety. *Chem Sci* **2015,** *6* (5), 3154-3160.
188. Zha, L.; Jiang, Y.; Henke, M.; Wilson, M.; Wang, J.; Kelleher, N.; Balskus, E., Colibactin assembly line enzymes use *S*-adenosylmethionine to build a cyclopropane ring. *Nat Chem Biol* **2017,** *13* (10), 1063-1065.
189. Parsek, M.; Val, D.; Hanzelka, B.; Cronan Jr, J.; Greenberg, E., Acyl homoserine-lactone quorum-sensing signal generation. *Proc Natl Acad Sci USA* **1999,** *96* (8), 4360-4365.
190. Michael, B.; Smith, J.; Swift, S.; Heffron, F.; Ahmer, B., SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. *J Bacteriol* **2001,** *183* (19), 5733-5742.

191. Yao, Y.; Martinez-Yamout, M.; Dickerson, T.; Brogan, A.; Wright, P.; Dyson, H., Structure of the *Escherichia coli* quorum sensing protein SdiA: activation of the folding switch by acyl homoserine lactones. *J Mol Biol* **2006,** *355* (2), 262-273.
192. Chan, Y.; Boyne, M.; Podevels, A.; Klimowicz, A.; Handelsman, J.; Kelleher, N.; Thomas, M., Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. *Proc Natl Acad Sci USA* **2006,** *103* (39), 14349-14354.
193. Reimer, D.; Pos, K.; Thines, M.; Grün, P.; Bode, H., A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat Chem Biol* **2011,** *7*, 888-890.
194. Fuqua, W.; Winans, S.; Greenberg, E., Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol* **1994,** *176* (2), 269-275.
195. Papenfort, K.; Bassler, B., Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol* **2016,** *14*(9), 576-588.
196. Kendall, M.; Sperandio, V., What a dinner party! Mechanisms and functions of interkingdom signaling in host-pathogen associations. *mBio* **2016,** *7* (2), e01748-15.
197. Hawver, L.; Jung, S.; Ng, W., Specificity and complexity in bacterial quorum-sensing systems. *FEMS Microbiol Rev* **2016,** *40* (5), 738-752.
198. Bazire, A.; Dheilly, A.; Diab, F.; Morin, D.; Jebbar, M.; Haras, D.; Dufour, A., Osmotic stress and phosphate limitation alter production of cell-to-cell signal molecules and rhamnolipid biosurfactant by *Pseudomonas aeruginosa*. *FEMS Microbiol Lett* **2005,** *253* (1), 125-131.
199. Lee, J.; Wu, J.; Deng, Y.; Wang, J.; Wang, C.; Wang, J.; Chang, C.; Dong, Y.; Williams, P.; Zhang, L., A cell-cell communication signal integrates quorum sensing and stress response. *Nat Chem Biol* **2013,** *9*(5), 339-343.
200. Williams, P., Quorum sensing, communication and cross-kingdom signalling in the bacterial world. *Microbiology* **2007,** *153*(12), 3923-3938.
201. Patankar, A.; González, J., Orphan LuxR regulators of quorum sensing. *FEMS Microbiol Rev* **2009,** *33* (4), 739-756.
202. Brady, S.; Chao, C.; Clardy, J., Long-chain *N*-acyltyrosine synthases from environmental DNA. *Appl Environ Microbiol* **2004,** *70* (11), 6865-6870.
203. Craig, J.; Cherry, M.; Brady, S., Long-chain *N*-acyl amino acid synthases are linked to the putative PEP-CTERM/ Exosortase protein-sorting system in Gram-negative bacteria. *J bacteriol* **2011,** *193* (20), 5707-5715.
204. Tørring, T.; Shames, S.; Cho, W.; Roy, C.; Crawford, J., Acyl histidines: new N-acyl amides from *Legionella pneumophila*. *ChemBioChem* **2017,** *18* (7), 638-646.
205. Blitzke, T.; Porzel, A.; Masaoud, M.; Schmidt, J., A chlorinated amide and piperidine alkaloids from *Aloe sabaea*. *Phytochemistry* **2000,** *55* (8), 979-982.
206. Burstein, S., *N*-acyl amino acids (elmiric acids): endogenous signaling molecules with therapeutic potential. *Mol Pharmacol* **2018,** *93*(3), 228-238.
207. Cani, P.; Plovier, H.; Van Hul, M.; Geurts, L.; Delzenne, N.; Druart, C.; Everard, A., Endocannabinoids - at the crossroads between the gut microbiota and host metabolism. *Nat Rev Endocrinol* **2016,** *12*(3), 133-143.
208. Hai, Y.; Tang, Y., Biosynthesis of long-chain *N*-acyl amide by a truncated polyketide synthase–nonribosomal peptide synthetase hybrid megasynthase in fungi. *J Am Chem Soc* **2018,** *140* (4), 1271-1274.

209. Brady, S.; Clardy, J., Long-chain *N*-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. *Journal Am Chem Soc* **2000,** *122* (51), 12903-12904.
210. Craig, J.; Cherry, M.; Brady, S., Long-chain *N*-acyl amino acid synthases are linked to the putative PEP-CTERM/exosortase protein-sorting system in Gram-negative bacteria. *J bacteriol* **2011,** *193* (20), 5707-5715.
211. Pérez-Berezo, T.; Pujo, J.; Martin, P.; Le Faouder, P.; Galano, J.; Guy, A.; Knauf, C.; Tabet, J.; Tronnet, S.; Barreau, F.; Heuillet, M.; Dietrich, G.; Bertrand-Michel, J.; Durand, T.; Oswald, E.; Cenac, N., Identification of an analgesic lipopeptide produced by the probiotic *Escherichia coli* strain Nissle 1917. *Nat Commun* **2017,** *8*, 1314.
212. Dhakal, R.; Bajpai, V.; Baek, K., Production of GABA (□-aminobutyric acid) by microorganisms: a review. *Braz J Microbiol* **2012,** *43* (4), 1230-1241.
213. Cohen, L.; Kang, H.; Chu, J.; Huang, Y.; Gordon, E.; Reddy, B.; Ternei, M.; Craig, J.; Brady, S., Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist. *Proc Natl Acad Sci USA* **2015,** *112* (35), E4825-E4834.
214. Milshteyn, A.; Colosimo, D.; Brady, S., Accessing bioactive natural products from the human microbiome. *Cell Host Microbe* **2018,** *23* (6), 725-736.
215. Cohen, L.; Esterhazy, D.; Kim, S.; Lemetre, C.; Aguilar, R.; Gordon, E.; Pickard, A.; Cross, J.; Emiliano, A.; Han, S.; Chu, J.; Vila-Farrese, X.; Kaplitt, J.; Rogoz, A.; Calle, P.; Hunter, C.; Bitok, J.; Brady, S., Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature* **2017,** *549*(7670), 48-53.
216. Chen, H.; Nwe, P.; Yang, Y.; Rosen, C.; Bielecka, A.; Kuchroo, M.; Cline, G.; Kruse, A.; Ring, A.; Crawford, J.; Palm, N., A forward chemical genetic screen reveals gut microbiota metabolites that modulate host physiology. *Cell* **2019**, *177*, 1-15.
217. Ridlon, J.; Kang, D.; Hylemon, P.; Bajaj, J., Bile acids and the gut microbiome. *Curr Opin Gastroenterol* **2014,** *30* (3), 332-338.
218. D'Costa, V.; McGrann, K.; Hughes, D.; Wright, G., Sampling the antibiotic resistome. *Science* **2006,** *311* (5759), 374-377.
219. Sommer, M.; Dantas, G.; Church, G., Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **2009,** *325* (5944), 1128-1131.
220. Crofts, T.; Gasparrini, A.; Dantas, G., Next-generation approaches to understand and combat the antibiotic resistome. *Nat Rev Microbiol* **2017,** *15*(7), 422-434.
221. Ochman, H.; Lawrence, J.; Groisman, E., Lateral gene transfer and the nature of bacterial innovation. *Nature* **2000,** *405* (6784), 299-304.
222. Almabruk, K.; Dinh, L.; Philmus, B., Self-resistance of natural product producers: past, present, and future focusing on self-resistant protein variants. *ACS Chem Biol* **2018,** *13* (6), 1426-1437.
223. Tang, X.; Li, J.; Millán-Aguiñaga, N.; Zhang, J.; O'Neill, E.; Ugalde, J.; Jensen, P.; Mantovani, S.; Moore, B., Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chem Biol* **2015,** *10* (12), 2841-2849.
224. Mousa, J.; Bruner, S., Structural and mechanistic diversity of multidrug transporters. *Nat Prod Rep* **2016,** *33* (11), 1255-1267.
225. Wilson, D., Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol* **2014,** *12* (1), 35-48.

226. Kling, A.; Lukat, P.; Almeida, D.; Bauer, A.; Fontaine, E.; Sordello, S.; Zaburannyi, N.; Herrmann, J.; Wenzel, S.; König, C.; Ammerman, N.; Barrio, M.; Borchers, K.; Bordon-Pallier, F.; Brönstrup, M.; Courtemanche, G.; Gerlitz, M.; Geslin, M.; Hammann, P.; Heinz, D.; Hoffmann, H.; Klieber, S.; Kohlmann, M.; Kurz, M.; Lair, C.; Matter, H.; Nuermberger, E.; Tyagi, S.; Fraisse, L.; Grosset, J.; Lagrange, S.; Müller, R., Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science* **2015,** *348* (6239), 1106-1112.
227. Tikhonov, A.; Kazakov, T.; Semenova, E.; Serebryakova, M.; Vondenhoff, G.; Van Aerschot, A.; Reader, J.; Govorun, V.; Severinov, K., The mechanism of microcin C resistance provided by the MccF peptidase. *J Biol Chem* **2010,** *285* (49), 37944-37952.
228. Ganz, T., Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* **2003,** *3* (9), 710-720.
229. Cullen, T.; Schofield, W.; Barry, N.; Putnam, E.; Rundell, E.; Trent, M.; Degnan, P.; Booth, C.; Yu, H.; Goodman, A., Gut microbiota. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* **2015,** *347* (6218), 170-175.
230. Oh, J.; Patel, J.; Park, H.; Crawford, J., β-Lactam biotransformations activate innate immunity. *J Org Chem* **2018,** *83* (13), 7173-7179.
231. Rubin, B.; Tomoaki, J., *Antibiotics as anti-inflammatory and immunomodulatory agents*. Birkhäuser Basel: New York, 2005.
232. Anuforom, O.; Wallace, G.; Piddock, L., The immune response and antibacterial therapy. *Med Microbiol Immunol* **2015,** *204* (2), 151-159.
233. Yang, J.; Bhargava, P.; McCloskey, D.; Mao, N.; Palsson, B.; Collins, J., Antibiotic-induced changes to the host metabolic environment inhibit drug efficacy and alter immune function. *Cell Host Microbe* **2017,** *22* (6), 757-765.e3.
234. Gopinath, S.; Kim, M.; Rakib, T.; Wong, P.; van Zandt, M.; Barry, N.; Kaisho, T.; Goodman, A.; Iwasaki, A., Topical application of aminoglycoside antibiotics enhances host resistance to viral infections in a microbiota-independent manner. *Nat Microbiol* **2018,** *3* (5), 611-621.
235. Matsumoto, J.; Nagai, S., Amidohydrolases for N-short and long chain fatty acyl-L-amino acids from *Mycobacteria*. *J Biochem* **1972,** *72* (2), 269-279.
236. Shintani, Y.; Fukuda, H.; Okamoto, N.; Murata, K.; Kimura, A., Isolation and characterization of N-long chain acyl aminoacylase from *Pseudomonas diminuta*. *J Biochem* **1984,** *96*(3)*,* 637-643.
237. Khersonsky, O.; Tawfik, D., Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem* **2010,** *79*, 471-505.
238. Linhardt, R.; Galliher, P.; Cooney, C., Polysaccharide lyases. *Appl Biochem Biotechnol* **1987,** *12* (2), 135-176.
239. Kaoutari, A.; Armougom, F.; Gordon, J.; Raoult, D.; Henrissat, B., The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* **2013,** *11*(7), 497-504.
240. Levin, B.; Huang, Y.; Peck, S.; Wei, Y.; Martínez-del Campo, A.; Marks, J.; Franzosa, E.; Huttenhower, C.; Balskus, E., A prominent glycyl radical enzyme in human gut microbiomes metabolizes *trans*-4-hydroxy-L-proline. *Science* **2017,** *355* (6325), eaai8386.
241. Sousa, T.; Paterson, R.; Moore, V.; Carlsson, A.; Abrahamsson, B.; Basit, A., The gastrointestinal microbiota as a site for the biotransformation of drugs. *Int J Pharm* **2008,** *363* (1-2), 1-25.
242. Rajakovich, L.; Balskus, E., Metabolic functions of the human gut microbiota: the role of metalloenzymes. *Nat Prod Rep (in press)* **2019**.

243. Brunton, L.; Chabner, B.; Knollmann, B., *Goodman & Gilman's the pharmacological basis of therapeutics, 12e*. New York: McGraw-Hill, 2011: United States, 2011.
244. Lindenbaum, J.; Rund, D.; Butler, V.; Tse-Eng, D.; Saha, J., Inactivation of digoxin by the gut flora: reversal by antibiotic therapy. *N Engl J Med* **1981,** *305* (14), 789-794.
245. Saha, J.; Butler Jr, V.; Neu, H.; Lindenbaum, J., Digoxin-inactivating bacteria: identification in human gut flora. *Science* **1983,** *220* (4594), 325-327.
246. Haiser, H.; Gootenberg, D.; Chatman, K.; Sirasani, G.; Balskus, E.; Turnbaugh, P., Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* **2013,** *341* (6143), 295-298.
247. Koppel, N.; Bisanz, J.; Pandelia, M.-E.; Turnbaugh, P.; Balskus, E., Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *eLife* **2018**, *7*, e33953.
248. Straussman, R.; Morikawa, T.; Shee, K.; Barzily-Rokni, M.; Qian, Z.; Du, J.; Davis, A.; Mongare, M.; Gould, J.; Frederick, D.; Cooper, Z.; Chapman, P.; Solit, D.; Ribas, A.; Lo, R.; Flaherty, K.; Ogino, S.; Wargo, J.; Golub, T., Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **2012,** *487*(7408), 500-504.
249. Geller, L.; Barzily-Rokni, M.; Danino, T.; Jonas, O.; Shental, N.; Nejman, D.; Gavert, N.; Zwang, Y.; Cooper, Z.; Shee, K.; Thaiss, C.; Reuben, A.; Livny, J.; Avraham, R.; Frederick, D.; Ligorio, M.; Chatman, K.; Johnston, S.; Mosher, C.; Brandis, A.; Fuks, G.; Gurbatri, C.; Gopalakrishnan, V.; Kim, M.; Hurd, M.; Katz, M.; Fleming, J.; Maitra, A.; Smith, D.; Skalak, M.; Bu, J.; Michaud, M.; Trauger, S.; Barshack, I.; Golan, T.; Sandbank, J.; Flaherty, K.; Mandinova, A.; Garrett, W.; Thayer, S.; Ferrone, C.; Huttenhower, C.; Bhatia, S.; Gevers, D.; Wargo, J.; Golub, T.; Straussman, R., Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science* **2017,** *357* (6356), 1156-1160.
250. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; Mende, D.; Li, J.; Xu, J.; Li, S.; Li, D.; Cao, J.; Wang, B.; Liang, H.; Zheng, H.; Xie, Y.; Tap, J.; Lepage, P.; Bertalan, M.; Batto, J.; Hansen, T.; Le Paslier, D.; Linneberg, A.; Nielsen, H.; Pelletier, E.; Renault, P.; Sicheritz-Ponten, T.; Turner, K.; Zhu, H.; Yu, C.; Li, S.; Jian, M.; Zhou, Y.; Li, Y.; Zhang, X.; Li, S.; Qin, N.; Yang, H.; Wang, J.; Brunak, S.; Doré, J.; Guarner, F.; Kristiansen, K.; Pedersen, O.; Parkhill, J.; Weissenbach, J.; Meta, H. I. T. C.; Bork, P.; Ehrlich, S.; Wang, J., A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010,** *464*, 59-65.
251. Degnan, P.; Barry, N.; Mok, K.; Taga, M.; Goodman, A., Human gut microbes use multiple transporters to distinguish vitamin B_{12} analogs and compete in the gut. *Cell Host Microbe* **2014,** *15* (1), 47-57.
252. Peisl, B.; Schymanski, E.; Wilmes, P., Dark matter in host-microbiome metabolomics: tackling the unknowns–a review. *Anal Chim Acta* **2018,** *1037*, 13-27.
253. Roume, H.; El Muller, E.; Cordes, T.; Renaut, J.; Hiller, K.; Wilmes, P., A biomolecular isolation framework for eco-systems biology. *ISME J* **2013,** *7*(1), 110-121.
254. da Silva, R.; Dorrestein, P.; Quinn, R., Illuminating the dark matter in metabolomics. *Proc Natl Acad Sci USA* **2015,** *112* (41), 12549-12550.
255. Medema, M.; Blin, K.; ; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M.; Weber, T.; Takano, E.; Breitling, R., antiSMASH: rapid identification, annotation and analysis of

secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **2011,** *39* (suppl_2), W339-W346.

256. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P.; Cullum, J.; Hranueli, D., *ClustScan*: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and *in silico* prediction of novel chemical structures. *Nucleic acids res* **2008,** *36* (21), 6882-6892.
257. Weber, T.; Rausch, C.; Lopez, P.; Hoof, I.; Gaykova, V.; Huson, D.; Wohlleben, W., CLUSEAN: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J Biotechnol* **2009,** *140* (1-2), 13-17.
258. Skinnider, M.; Dejong, C.; Rees, P.; Johnston, C.; Li, H.; Webster, A.; Wyatt, M.; Magarvey, N., Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res* **2015,** *43* (20), 9645-9662.
259. Blin, K.; Wolf, T.; Chevrette, M.; Lu, X.; Schwalen, C.; Kautsar, S.; Suarez Duran, H.; de los Santos, E.; Kim, H.; Nave, M.; Dickschat, J.; Mitchell, D.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S.; Weber, T.; Medema, M., antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **2017,** *45* (W1), W36-W41.
260. Skinnider, M.; Merwin, N.; Johnston, C.; Magarvey, N., PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res* **2017,** *45* (W1), W49-W54.
261. Faith, J.; Ahern, P.; Ridaura, V.; Cheng, J.; Gordon, J., Identifying gut microbe–host phenotype relationships using combinatorial communities in gnotobiotic mice. *Sci Transl Med* **2014,** *6* (220), 220ra11.
262. Goodman, A.; Wu, M.; Gordon, J., Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. *Nat Protoc* **2011,** *6*(12), 1969-1980.
263. Martens, E.; Chiang, H.; Gordon, J., Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **2008,** *4* (5), 447-457.
264. Dantas, G.; Sommer, M.; Degnan, P.; Goodman, A., Experimental approaches for defining functional roles of microbes in the human gut. *Annu Rev of Microbiol* **2013,** *67* (1), 459-475.
265. Donaldson, G.; Lee, S.; Mazmanian, S., Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* **2015,** *14*(1), 20-32.
266. Zmora, N.; Zilberman-Schapira, G.; Suez, J.; Mor, U.; Dori-Bachash, M.; Bashiardes, S.; Kotler, E.; Zur, M.; Regev-Lehavi, D.; Brik, R.; Federici, S.; Cohen, Y.; Linevsky, R.; Rothschild, D.; Moor, A.; Ben-Moshe, S.; Harmelin, A.; Itzkovitz, S.; Maharshak, N.; Shibolet, O.; Shapiro, H.; Pevsner-Fischer, M.; Sharon, I.; Halpern, Z.; Segal, E.; Elinav, E., Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. *Cell* **2018,** *174* (6), 1388-1405.e21.
267. Mowat, A.; Agace, W., Regional specialization within the intestinal immune system. *Nat Rev Immunol* **2014,** *14*(10), 667-685.
268. Marteyn, B.; West, N.; Browning, D.; Cole, J.; Shaw, J.; Palm, F.; Mounier, J.; Prévost, M.; Sansonetti, P.; Tang, C., Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* **2010,** *465*(7296), 355-358.
269. Rath, C.; Dorrestein, P., The bacterial chemical repertoire mediates metabolic exchange within gut microbiomes. *Curr Opin Microbiol* **2012,** *15* (2), 147-154.

270. Traxler, M.; Watrous, J.; Alexandrov, T.; Dorrestein, P.; Kolter, R., Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. *mBio* **2013,** *4* (4), e00459-13.
271. Yang, Y.; Xu, Y.; Straight, P.; Dorrestein, P., Translating metabolic exchange with imaging mass spectrometry. *Nat Chem Biol* **2009,** *5*(12), 885-887.

Chapter 2

A Forward Chemical Genetic Screen Reveals Gut Microbiota Metabolites that Modulate Host Physiology

This chapter is adapted from the published work as a second author. (Chen, H.; **Nwe, P.-K.**; Yang, Y.; Rosen, C. E.; Bielecka, A. A.; Kuchroo, M.; Cline, G.W.; Kruse, A. C.; Ring, A. M.; Crawford, J.M.; Palm, N.W. *Cell*, **2019**, *177* (5): 1217-1231.

2.1 INTRODUCTION

The human gut microbiota produces thousands of unique small molecules that can potentially affect nearly all aspects of human physiology, from regulating immunity in the gut to shaping mood and behavior [1-5]. These metabolites can act locally in the intestine or can accumulate up to millimolar concentrations in the serum and potentially act systemically [5-8]. Recent studies have employed state-of-the-art genomic and metabolomic approaches to evaluate the microbiota metabolome under various conditions [9-13]. These efforts have begun to reveal the enormously complex intra- and inter-species microbial chemistries that potentially impinge on host physiology, as well as the impact of gut microbes on the processing of dietary small molecules and medical drugs. In addition, they underscore the importance of continuing to develop new approaches to explore the bioactive microbiota metabolome [8].

G-protein coupled receptors (GPCRs) are the largest family of membrane proteins encoded in the human genome (including over 350 conventional non-olfactory GPCRs) and are critical sensors of a variety of small molecules. GPCRs exert pivotal roles in regulating diverse aspects of host physiology, such as vision, mood, pain, and immunity [14]. Specific GPCRs are also known to sense microbial metabolites. Microbiota-derived short-chain fatty acids can activate GPR41, GPR43, and GPR109A and contribute to immune-microbiota homeostasis in the intestine [3, 4]. Succinate produced by commensal protists can induce expansion of tuft cells in the small intestine through activation of the succinate receptor [15-17]. Finally, recent

studies have continued to reveal novel microbiota-derived GPCR ligands that shape host physiology [18]. Thus, the microbiota metabolome is a rich source of potential GPCR ligands.

Here, we use host-sensing of microbiota metabolites by conventional GPCRs as a lens to reveal microbial production of bioactive small molecules that may impact host physiology. Building on recent developments in high-throughput screening of the complete GPCRome and activity-guided microbial metabolite identification approaches, we develop a high-throughput pipeline to screen human gut microbes for the ability to produce ligands that can activate human GPCRs[5, 8, 13, 19]. We show that dozens of diverse human gut microbes produce agonists for GPCRs from a variety of families, including both well-characterized GPCRs and orphan GPCRs with no known natural small molecule ligands. We demonstrate that specific gut microbe-derived GPCR ligands, including small molecules resulting from biotransformations of dietary compounds, can potentially influence both local and systemic physiology. Using bioassay-guided chemical characterization techniques, we identify bacterial-derived L-Phe as an agonist for two orphan GPCRs. Finally, we reveal a unique pathway of metabolite exchange between bacterial strains that can contribute to the production of the potent psychoactive trace amine phenethylamine, which can cross the blood-brain barrier and trigger lethal phenethylamine poisoning after administration of an FDA-approved monoamine oxidase inhibitor. Thus, we have established an orthogonal approach for elucidating physiologically-relevant microbiota metabolite-host interactions and uncovered multiple diet-microbe-host and microbe-microbe-host metabolic axes.

2.2 RESULTS

2.2. 1 A forward chemical genetic screen to identify bioactive microbiota metabolites

We set out to establish a high-throughput screening system to identify specific human gut microbes that produce agonists or antagonists of conventional GPCRs. We developed a pipeline for parsing the microbiota metabolome based on the transformative GPCR screening technology Parallel Receptor-ome Expression and Screening via Transcriptional Output-Tango (PRESTO-Tango) [19]. This technology leverages the Tango β-arrestin recruitment assay to simultaneously measure the activation of nearly all non-olfactory GPCRs (Figure 1, S1A, B) [19, 20]. We thus proceeded to exploit this assay to perform a broad-ranging screen of bioactive metabolites produced by diverse members of the human gut microbiota.

We previously assembled personalized gut microbiota culture collections from eleven inflammatory bowel disease patients through high-throughput anaerobic culture methods and next-generation sequencing [21]. This collection yielded 144 unique bacterial isolates from five phyla, nine classes, eleven orders, and twenty families, as well as multiple strains that were assigned to the same species. Thus, this culture collection enables us to examine the effects of phylogenetically diverse taxa while also providing insights into potential strain-specific differences between members of the same species (Table S1). We cultured all members of our collection individually in a medium specialized for the cultivation of diverse gut commensals (gut microbiota medium; GMM) [22] and screened their supernatants for activation or inhibition of nearly all conventional GPCRs using PRESTO-Tango (Figure 1, 2). PRESTO-Tango screening was performed essentially as described by Kroeze et al. [19] and GPCR activation or inhibition was determined by comparing stimulation with a given bacterial supernatant to stimulation of the same GPCR with bacterial media alone (details described in Methods). No stimulation and bacterial media alone controls were included as separate conditions in each experiment to correct for day to day and experiment to experiment variability across the screen.

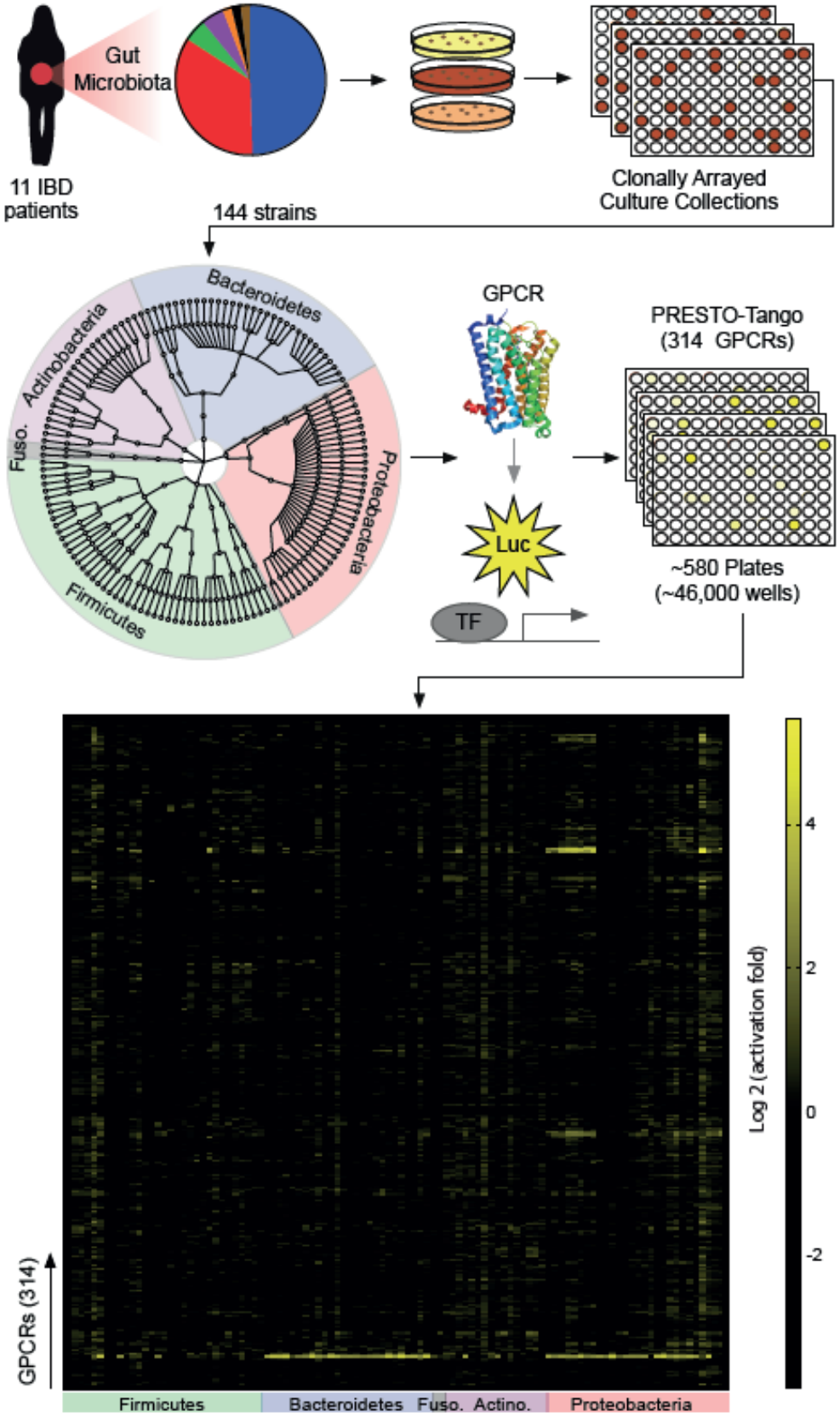

Gut Microbiota
11 IBD patients
Clonally Arrayed Culture Collections
144 strains
Bacteroidetes
Actinobacteria
Fuso.
Proteobacteria
Firmicutes
GPCR
PRESTO-Tango (314 GPCRs)
Luc
TF
~580 Plates (~46,000 wells)
GPCRs (314)
Log 2 (activation fold)
4
2
0
-2
Firmicutes
Bacteroidetes
Fuso.
Actino.
Proteobacteria

Figure 1. A forward chemical genetic screen identifies human gut microbes that activate GPCRs.

We isolated 144 unique human gut bacteria spanning five phyla, nine classes, eleven orders, and twenty families from 11 inflammatory bowel disease patients via high-throughput anaerobic culturomics and massively barcoded 16S rRNA gene sequencing. Bacterial isolates were grown in monoculture in a medium specialized for the cultivation of human gut microbes (gut microbiota medium) and supernatants from individual bacterial monocultures were screened against the near-complete non-olfactory GPCRome (314 conventional GPCRs) using the high-throughput assay Parallel Receptor-ome Expression and Screening via Transcriptional Output-Tango (PRESTO-Tango).

2.2.2 Human gut microbes produce compounds that activate both well-characterized and orphan GPCRs

PRESTO-Tango screening revealed a diverse array of hits, including bacterial-derived metabolite mixtures that activated well-characterized GPCRs as well as mixtures that activated orphan receptors. Across the complete data set, we observed activation of at least one GPCR from almost every class by at least one metabolite mixture (Figure 2). One specific pattern of GPCR activation tracked closely with gross phylogeny—strains belonging to the phyla Bacteroidetes and Proteobacteria potently activated the succinate receptor (Sucr1), while strains belonging to the phyla Firmicutes, Fusobacteria and Actinobacteria largely failed to activate this receptor (Table S2). However, most activation patterns did not correlate with phylogeny, and there were multiple examples of bacterial strains that exhibited unique GPCR agonist activities despite being assigned to the same species (Table S2). We also noted that stimulation with GMM alone activated specific GPCRs when compared to PBS and that media harvested from specific microbes sometimes reversed these effects, either due to bacterial consumption of GPCR ligands in the media or bacterial production of GPCR antagonists (Table S3).

Given the high sensitivity of β-arrestin recruitment-based assays like PRESTO-Tango, the veracity of individual hits from this screen will need to be confirmed using alternative methods. Nonetheless, these data demonstrate that human gut microbes produce a remarkable array of GPCR ligands.

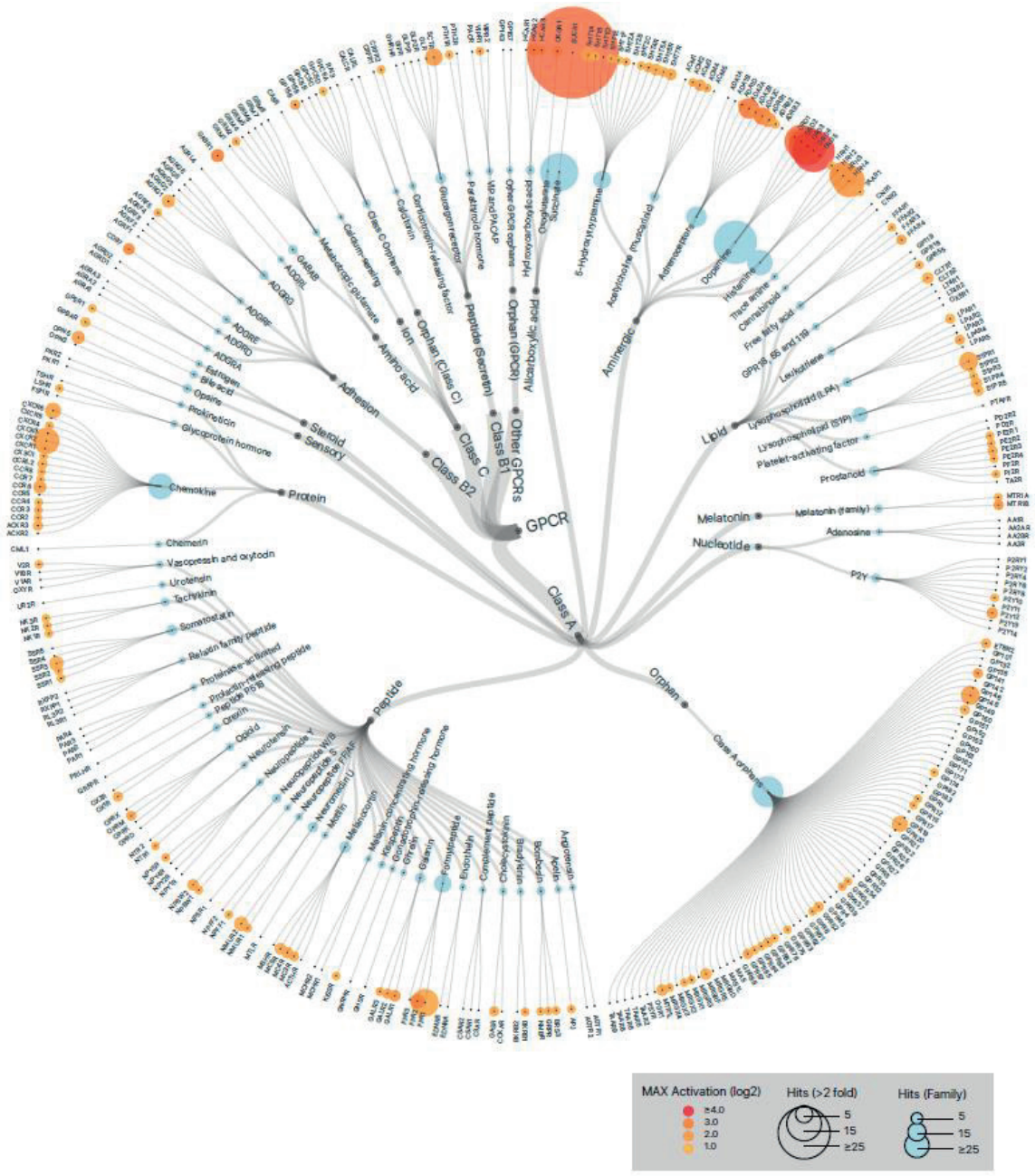


Figure 2. Members of the human gut microbiota produce metabolites that activate diverse human GPCRs.

GPCR activation by metabolomes from a human gut microbiota culture collection consisting of 144 strains isolated from 11 IBD patients. Data is displayed on a hierarchical tree of GPCRs organized by class, ligand type, and receptor family. Color intensity represents the maximum magnitude of activation (log 2) over background (gut microbiota medium alone) across the entire data set—*i.e.*, the maximum activation of a given GPCR by any microbial metabolome in our collection. Radii of the circles at each tip are scaled based on the number of strains that activated a given receptor or receptor family (*i.e.*, number of hits across the complete data set). Hits are defined as activation of a given receptor more than two-fold over background. An interactive html graphic representing these data is included in the supplemental information. Graphics were generated in collaboration with visavisllc using d3.js.

2.2.3 Human gut microbes produce compounds that activate aminergic receptors

Besides the succinate receptor, the next most prevalent class of GPCRs activated by gut commensals was the aminergic receptors, which are expressed in diverse tissues and cell types and regulate a wide variety of core physiological processes ranging from neurotransmission to immunity (Figure 2) [23-25]. These hits included bacterial-derived activators of the dopamine (DRDs), histamine (HRHs) and adrenergic receptor families. We observed that more than a dozen commensal supernatants activated DRD2-4 or HRH2-4 (Figure 3A, S2). For example, ten strains from the phylum Proteobacteria activated both DRDs and HRHs, including eight strains from the species *Morganella morganii* (Figure 3B, C, S2). This suggests that production of DRD and HRH agonists is a conserved feature of *M. morganii*. In contrast, we observed that two strains of *L. reuteri* in our collection activated HRHs, while two distinct strains of *L.reuteri* failed to activate HRHs despite displaying similar growth kinetics (Figure 3C, S2 and data not shown). One strain of *Streptococcus*, but not two related isolates, activated DRD2-4, and two unclassified strains of Enterobacteriaceae activated HRH1-4 and DRD2 but failed to activate other DRDs (Figure 3C, S2).

M. morganii was previously reported to produce various biogenic amines, including dopamine, tyramine, and phenethylamine (PEA) [26, 27]. We noted that all *M. morganii* supernatants potently activated DRD2-4, but not DRD1 and 5 (Figure 3A, B, S2). In contrast, dopamine itself efficiently activated all five dopamine receptors (Figure S1A). Therefore, we suspected that *M. morganii* might primarily produce a metabolite that is structurally related to dopamine and can act as a selective ligand for DRD2-4 but not DRD1 or 5 (Figure S3A). We examined the ability of all possible upstream and downstream metabolites in the mammalian dopamine pathway to activate DRD1-5 via PRESTO-Tango (Figure S3A, B). We found that a variety of compounds in this pathway activated various dopamine receptors; in particular, we noted that PEA and tyramine showed identical activation patterns to *M. morganii* supernatant (Figure S3B and C). Therefore, we examined the concentrations of dopamine, PEA, and tyramine in *M. morganii* supernatants. In line with a previous report [27], we found that *M. morganii* produced only trace amounts of

dopamine and no detectable tyramine, but instead secreted significant quantities of the potent trace amine PEA which, unlike dopamine and tyramine, can readily cross the blood-brain barrier (Figure 3D, S3D, E) [28].

Previous reports have also suggested that *M. morganii* produces histamine [26]. We confirmed that *M. morganii* strains indeed secreted significant amounts of histamine by ELISA and that our *M. morganii* strains encode a previously described histidine decarboxylase; in addition, we found that 48 of 49 previously deposited *M. morganii* strains encoded this histidine decarboxylase (Figure 3E and Table S4, S5). Similarly, two strains of *L. reuteri* and two strains from the Enterobacteriaceae family that activated the histamine receptors also secreted histamine (Figure 3E). However, based on whole genome sequencing, both the histamine-producing and non-histamine-producing strains of *L. reuteri* encoded an identical histidine decarboxylase proenzyme (Table S4, S5). Together, these data reveal that *M. morganii* secretes high levels of PEA, which acts as a potent agonist of the dopamine receptors, and that *M. morganii* and select strains of *L. reuteri* secrete histamine.

In mammals, PEA, dopamine, and tyramine are produced via the decarboxylation of L-Phe, L-DOPA, and L-Tyr, respectively, by the aromatic L-amino acid decarboxylase (AADC; Figure S3A) [29]. Thus, we tested whether *M. morganii* would similarly process these three amino acids into their respective biogenic amines. We used a defined minimal medium (MM) lacking L-Phe, L-DOPA, L-Tyr, and L-His to culture *M. morganii.* Despite normal *M. morganii* growth, we could not detect any production of PEA, tyramine, dopamine, or histamine by liquid chromatography-mass spectrometry (LC-MS) (Figure 3F). However, supplementation with L-Phe or L-His led to the production of high levels of PEA or histamine, and activation of DRD2-4 or HRH2-4 (Figure 3F, G). In contrast, supplementation with L-DOPA or L-Tyr failed to lead to the production of dopamine or tyramine, or activation of DRDs by *M. morganii* supernatants despite similar bacterial growth in all conditions (Figure S4C).

Figure 3

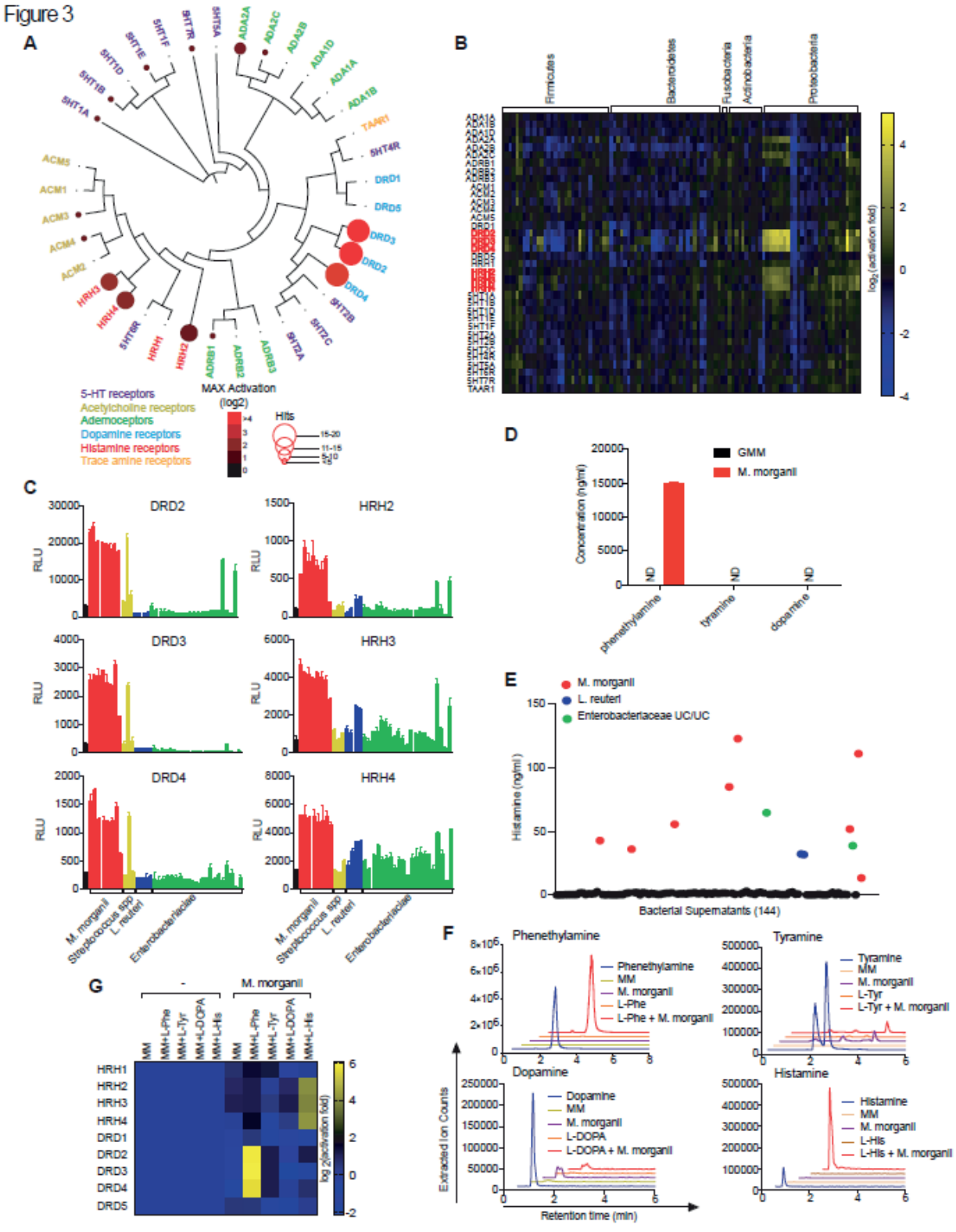

A
5HT1A
5HT1B
5HT1D
5HT1E
5HT1F
5HT1R
5HT5A
ADA2A
ADA2C
ADA2B
ADA1D
ADA1A
ADA1B
TAAR1
5HT4R
DRD1
DRD5
DRD3
DRD2
DRD4
5HT2B
5HT2C
5HT2A
ADRB3
ADRB2
ADRB1
HRH2
HRH1
5HT6R
HRH4
HRH3
ACM2
ACM4
ACM3
ACM1
ACM5
5-HT receptors
Acetylcholine receptors
Adenoceptors
Dopamine receptors
Histamine receptors
Trace amine receptors
MAX Activation (log2)
Hits
B
Firmicutes
Bacteroidetes
Fusobacteria
Actinobacteria
Proteobacteria
log2 (activation fold)
C
DRD2
HRH2
DRD3
HRH3
DRD4
HRH4
RLU
M. morganii
Streptococcus spp
L. reuteri
Enterobacteriaceae
D
GMM
M. morganii
Concentration (ng/ml)
ND
phenethylamine
tyramine
dopamine
E
M. morganii
L. reuteri
Enterobacteriaceae UC/UC
Histamine (ng/ml)
Bacterial Supernatants (144)
F
Phenethylamine
MM
M. morganii
L-Phe
L-Phe + M. morganii
Tyramine
L-Tyr
L-Tyr + M. morganii
Dopamine
L-DOPA
L-DOPA + M. morganii
Histamine
L-His
L-His + M. morganii
Extracted Ion Counts
Retention time (min)
G
-
M. morganii
MM
MM+L-Phe
MM+L-Tyr
MM+L-DOPA
MM+L-His
HRH1
HRH2
HRH3
HRH4
DRD1
DRD2
DRD3
DRD4
DRD5
log 2(activation fold)

Figure 3. Diverse human gut bacteria activate aminergic GPCRs.

(A) Activation of aminergic GPCRs by metabolomes from a human gut microbiota culture collection consisting of 144 strains isolated from 11 IBD patients. GPCR activation was measured by PRESTO-Tango. Screening results are displayed on a phylogenetic tree of aminergic GPCRs. Color intensity represents magnitude of activation over media alone and radii of the circles represents the number of bacteria that activated a given GPCR by more than two-fold over media alone.

(B) Heatmap depicting the activation of aminergic GPCRs by metabolites from 144 human gut bacteria as measured by PRESTO-Tango. Fold induction over stimulation with bacterial media alone is depicted on a log2 scale.

(C) Activation of DRD2-4 and HRH2-4 by select species and strains as measured by PRESTO-Tango. n=3 replicates per isolate.

(D) Quantification of dopamine, phenethylamine and tyramine production by *M. morganii*. Supernatants from 24-hour cultures of *M. morganii* C135 in gut microbiota medium were analyzed by Triple Quadrupole-Mass Spectrometry (QQQ-MS/MS) and compared to those of media controls. n=3 replicates per group.

(E) Quantification of histamine production by 144 isolates of human gut bacteria. All bacteria were grown in gut microbiota medium for 48 hours and then supernatants were probed for histamine production via ELISA.

(F) Mass spectrometric traces of metabolite production by *M. morganii* C135. *M. morganii* can directly convert L-Phe and L-His into phenethylamine and histamine, respectively. However, no conversion of L-Tyr to tyramine or L-DOPA to dopamine was observed. *M. morganii* was cultured in minimal medium (MM) with or without additional L-Phe, L-His, L-Tyr or L-DOPA for 48 hours. Metabolite production was analyzed by Liquid Chromatography-Mass Spectrometry (LC-MS).

(G) *M. morganii*-derived phenethylamine and histamine activate DRD2-4 and HRH2-4, respectively. *M. morganii* C135 were cultured as described in F and supernatants were screened for activity against DRDs and HRHs by PRESTO-Tango. n=3 replicates per group.

Data for all panels other than A and B are representative of at least three independent experiments. Data are presented as mean ± SEM.

This suggests that, unlike mammalian AADC, *M. morganii* selectively converts L-Phe into PEA and cannot efficiently convert L-DOPA or L-Tyr into dopamine or tyramine. While it is currently unclear which specific genes are involved in production of PEA by *M. morganii*, whole genome sequencing revealed the presence of at least 17 decarboxylases that are shared between the two strains of *M. morganii* that we sequenced (Table S4, S5).

2.2.4 Microbiota-derived histamine promotes increased colonic motility through activation of the histamine receptors

Histamine is generated via decarboxylation of L-His [30]. We found that eight strains of *M. morganii* and two strains of *L. reuteri* generated histamine *in vitro* and that supplementation with L-His significantly increased histamine production by these strains; in contrast, two distinct strains of *L. reuteri* failed to produce histamine regardless of supplementation with L-His (Figure 4A). We also cultured *M. morganii* C135 in multiple media either aerobically or anaerobically and found that *M. morganii* supernatants activated DRDs and HRHs regardless of culture conditions (data not shown). To test whether *M. morganii* can also produce histamine *in vivo,* we colonized germ-free mice with two distinct mock communities containing 9 or 10 diverse human gut microbes or with *M. morganii* C135 alone with or without supplementation of 1% L-His in the drinking water *ad libitum* to approximate the effect of an L-His-rich diet (*e.g.*, a meat-heavy diet) (Figure 4B). In addition, we monocolonized mice with two strains of *L. reuteri* with divergent histamine production capabilities: *L. reuteri* C93, which produced significant histamine *in vitro*, and *L. reuteri* C88, which failed to produce histamine *in vitro*. Mice colonized with *M. morganii* C135 or *L. reuteri* C93 exhibited high levels of intestinal histamine production, while mice colonized with the two mock communities or *L. reuteri* C88 showed nearly undetectable intestinal histamine (Figure 4C). In addition, histamine production in *M. morganii* monocolonized mice was significantly increased upon supplementation with dietary L-His (Figure 4C). Finally, we also detected increased histamine in the serum of mice colonized with *M. morganii* (Figure S5A).

We next determined the location of *M. morganii in vivo.* We used modified Niven's agar to enumerate *M. morganii* CFUs in gnotobiotic mice colonized with two mock communities of 9 or 10 diverse human gut microbes plus *M. morganii* C135 [31]. We found that *M. morganii* constitutes approximately 5% of the overall microbial community in this context, primarily inhabits the cecum and colon, and is nearly absent in the small intestine (Figure S5C and Table S3). Previous studies in humans also indicated that *M. morganii* preferentially localizes in tissue- or mucus-associated niches in the colon [32].

Figure 4

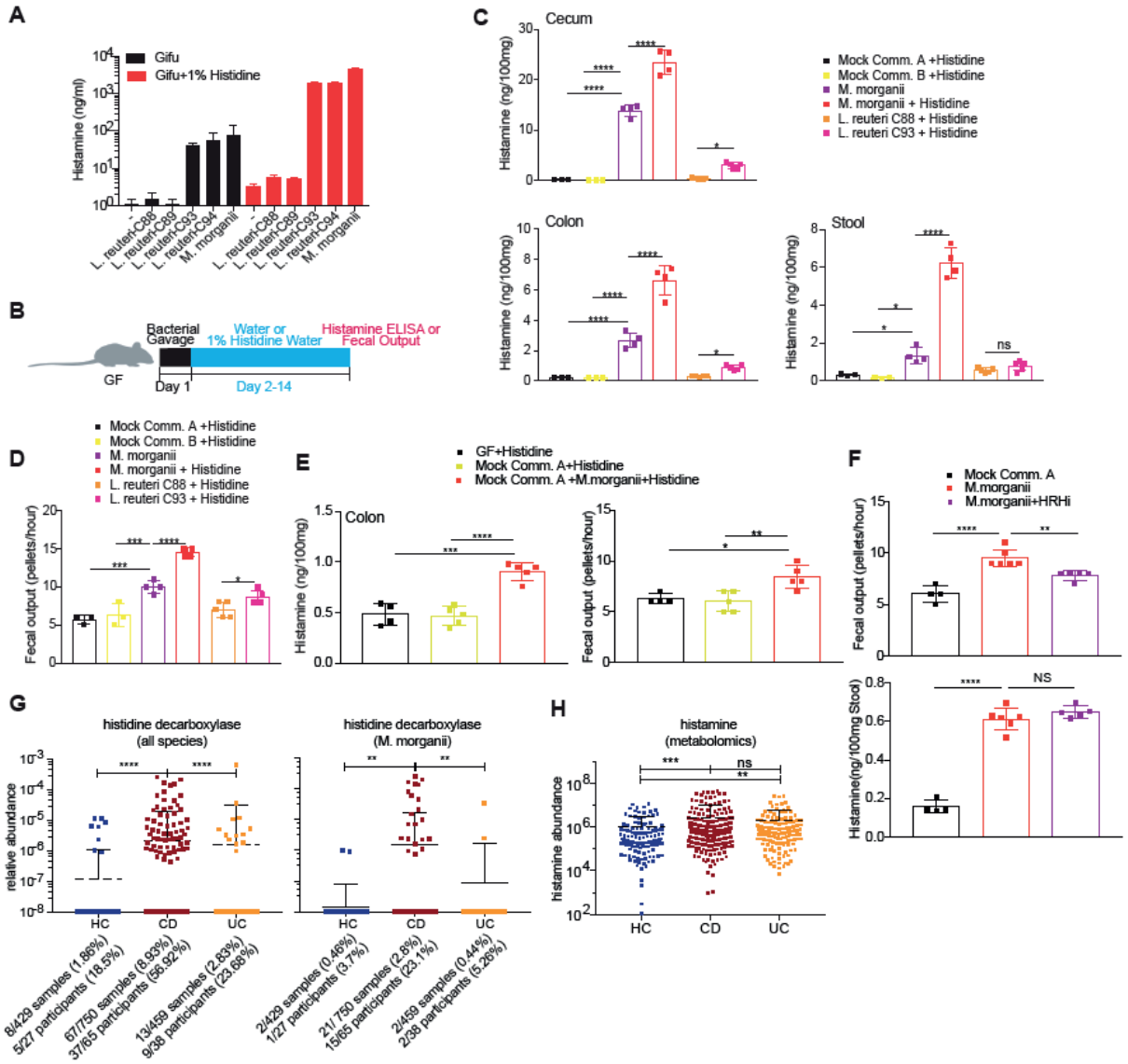


Figure 4. Commensal-derived histamine promotes colon motility.

(A) Production of histamine by *M. morganii* and *L. reuteri* is enhanced by additional L-His. Four strains of *L. reuteri* and one strain of *M. morganii* (C135) were cultured in Gifu medium with or without supplemental L-His and histamine concentrations in the supernatants were measured by ELISA after 48 hours. The addition of L-His increased histamine production by *L. reuteri* C93, C94 and *M. morganii* C135 (all histamine producers), except *L. reuteri* C88 and C89 (background levels in negative controls containing supplemental histidine are due to slight cross-reactivity of the histamine ELISA for histidine). n=3 replicates per isolate.

(B) Experimental design to test *in vivo* histamine production and the effects of histamine-producing bacteria on colon motility.

(C) *M. morganii*- and *L. reuteri*-derived histamine accumulates *in vivo* in the intestines of monocolonized mice. Groups of female germ-free C57Bl/6 mice were colonized with mock communities of 9 or 10 phylogenetically diverse human gut bacteria (Mock Community A or B) or monocolonized with *M. morganii* C135*, L. reuteri* C88 or C93. Mice were fed a conventional diet with or without administration of 1% L-His *ad libitum* in the drinking water. Histamine concentrations in cecal and colonic extracts and feces were measured via ELISA. n=3-5 mice per group.

(D) *M. morganii* C135- and *L. reuteri* C93-derived histamine enhances colon motility. Fecal output for mice treated as described in B were measured by counting the number of fecal pellets produced by a single mouse in one hour. n=3-5 mice per group.

(E) *M. morganii* increases colon motility in the context of a mock gut microbial community. Groups of female germ-free C57Bl/6 mice were colonized with a mock community of 9 phylogenetically diverse human gut bacteria (Mock Community A) with or without *M. morganii* C135 and administered 1% L-His *ad libitum* in the drinking water. Histamine concentrations in colonic extracts were measured via ELISA and fecal output was measured by counting the number of fecal pellets produced by a single mouse in one hour. n=4-5 mice per group.

(F) Histamine receptor inhibition partially reverses the impact of *M. morganii* on colon motility. Groups of female germ-free C57Bl/6 mice were colonized with a mock community of nine phylogenetically diverse human gut bacteria (Mock Community A) or monocolonized with *M. morganii* C135 for two weeks. Mice were then treated with or without a cocktail of four histamine receptor inhibitors (targeting HRH1-4) in the drinking water for one week. Histamine concentrations in feces were measured via ELISA and fecal output was measured by counting the number of fecal pellets produced by a single mouse in one hour. n=4-6 mice per group.

(G and H) Relative abundance of genes encoding histidine decarboxylases (from all bacteria) and histidine decarboxylase (from *M. morganii*) are increased in the microbiomes of patients with Crohn's disease as compared to healthy controls (G). Relative abundance of histamine is increased in IBD patients as compared to healthy controls (H). Metagenomic and metabolomic data from longitudinal stool samples from IBD patients (publicly available from the Human Microbiome Project 2; iHMP) were analyzed for the presence and relative abundance of histidine decarboxylase genes (from all bacteria or from *M. morganii*), or for the relative abundance of histamine. Data shown are a compilation of all data across multiple collection timepoints. Total number of samples or subjects with detectable *M. morganii* are denoted below each plot; a subject was considered positive if *M. morganii* was detectable in one or more samples from that patient across the complete dataset.

Data in all panels are representative of at least two independent experiments. Data are presented as mean ± SEM. One-way ANOVA with Tukey's post-hoc test (C-F) or Kruskall-Wallis with Dunn's multiple comparisons (G-H), *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, NS not significant ($p > 0.05$).

Oral gavage with histamine has been reported to increase colon motility in rodents [33, 34]. We thus hypothesized that gut microbe-derived histamine might also increase intestinal motility. We monitored intestinal motility in mice colonized with two mock communities or monocolonized with *M. morganii* C135 with or without administration of 1% L-His in the water. We found that colonization with *M. morganii* led to a significant increase in fecal output, which was further increased upon supplementation with L-His (Figure 4D). Similarly, we observed that mice colonized with *L. reuteri* C93 showed increased fecal output as compared to mice colonized with *L. reuteri* C88 (Figure 4D).

While monocolonizations are useful for reductionist proof-of-principle experiments, they are inherently artificial. To test whether *M. morganii* impacts host physiology in the context of a more diverse microbial community, we colonized germ-free mice with a mock community consisting of nine phylogenetically diverse human gut microbes with or without *M. morganii* C135 and examined histamine accumulation in the gut and serum as well as fecal output. Although the effects were less profound than those observed with monocolonizations, the addition of *M. morganii* to a mock community also led to an increased accumulation of histamine in the colon and serum (Figure 4E and S5D) as well as increased fecal output (Figure 4E).

To test whether the effects of *M. morganii* on fecal output are dependent on the histamine receptors, we colonized germ-free mice with *M. morganii* C135 and fed these mice with a cocktail of histamine receptor inhibitors *ad libitum* in the drinking water. We found that histamine receptor inhibition significantly inhibited the effects of *M. morganii* on fecal output despite the accumulation of similar levels of histamine in the gut (Figure 4F).

To examine the potential importance of histamine production by *M. morganii* (or other microbes) in human physiology, we mined publicly available metagenomic and metabolomic data from the integrative Human Microbiome Project (iHMP; see methods for details on data accession) to determine the relative abundance of histamine producing enzymes or histamine itself in patients with IBD (n=65 Crohn's disease; CD, n=38 ulcerative colitis; UC) versus healthy controls (n=27; HC) [35]. We found that CD patients exhibited

significantly increased prevalence and abundance of histidine decarboxylase genes as compared to healthy controls or UC patients (Figure 4G). Furthermore, *M. morganii* histidine decarboxylase was the most prevalent contributor to this increase (Figure 4G and Figure S5E). Finally, the abundance of histamine was increased in fecal samples from both CD patients and UC patients as compared to healthy controls as measured by metabolomics (Figure 4H). This observation is in line with previous studies demonstrating increased histamine in intestinal tissues and contents from patients with IBD (though this is often attributed to host-derived histamine) [36].

Together, these data demonstrate that *M. morganii* impacts intestinal motility through the secretion of histamine and activation of histamine receptors, that dietary histidine can enhance these effects, and that bacterial histidine decarboxylases (both generally and from *M. morganii*) are enriched in patients with CD.

2.2.5 *M. morganii* can trigger 'phenethylamine poisoning' when combined with monoamine oxidase inhibition

We detected abundant histamine production by *M. morganii* both *in vitro* and *in vivo*; in contrast, we could only detect low levels of PEA in the colons of *M. morganii* colonized mice (Figure 5A). A possible explanation for this observation is that biogenic amines, such as dopamine, tyramine, and PEA, are rapidly degraded in the intestine by mammalian monoamine oxidases (MAOs) [37]. To reveal the potential production of PEA *in vivo*, we treated germ-free mice or mice monocolonized with *M. morganii* C135 or a *Bacteroides thetaiotaomicron* strain (*B. theta* C34) that does not produce DRD agonists with the irreversible MAO inhibitor (MAOI) phenelzine. We observed increased PEA in the colons of *M. morganii* colonized mice by triple quadrupole MS/MS (QQQ-MS/MS) (Figure 5A); in contrast, colonic PEA was undetectable in germ-free mice and mice colonized with *B. theta* and treated with MAOI (Figure 5A). Unlike many other irreversible MAOIs, phenelzine is still used clinically for the treatment of major depressive disorder, as well as a variety of other psychological disorders including panic, social anxiety, and post-traumatic stress

Figure 5

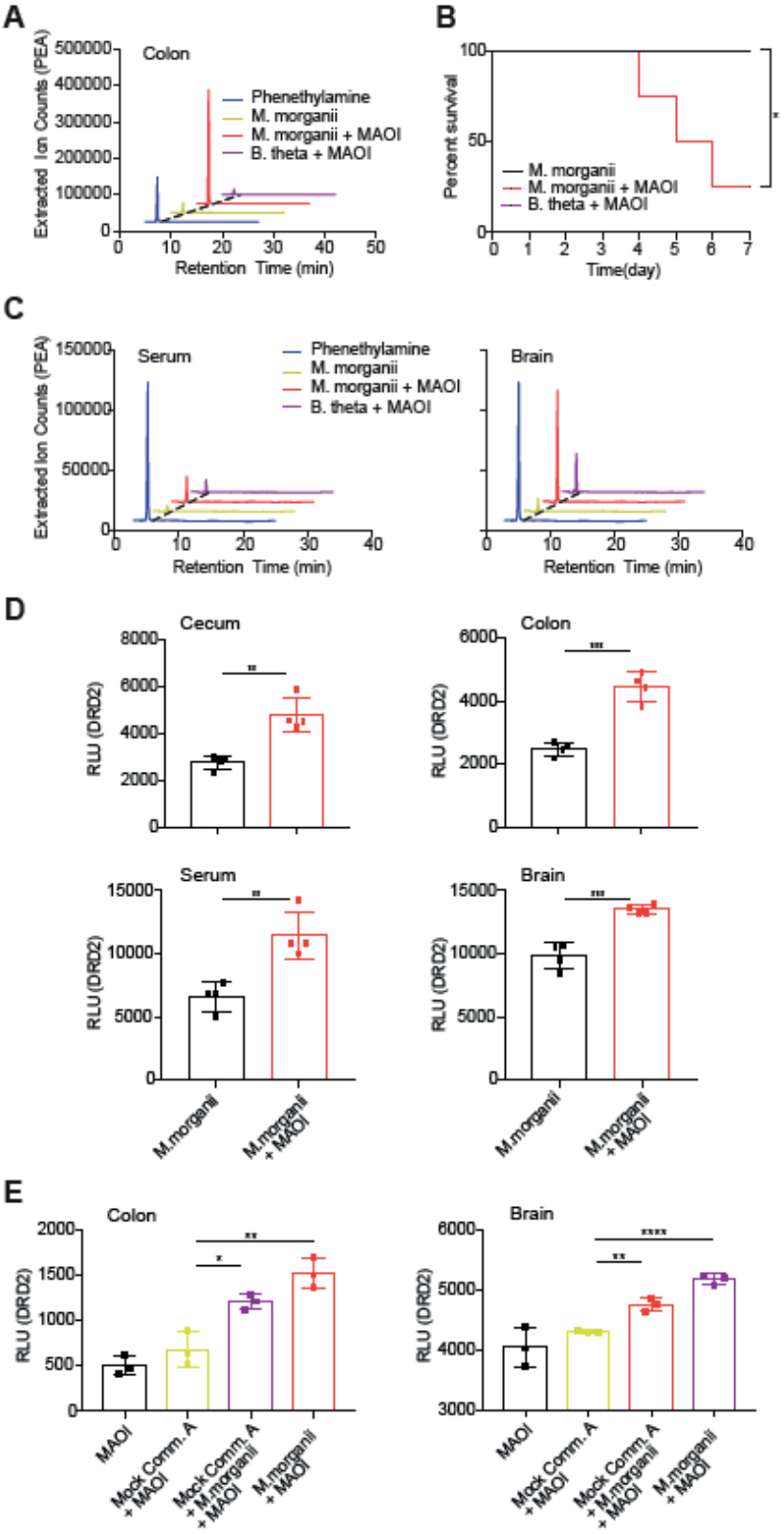

A
Colon
Extracted Ion Counts (PEA)
Retention Time (min)
Phenethylamine
M. morganii
M. morganii + MAOI
B. theta + MAOI
B
Percent survival
Time(day)
M. morganii
M. morganii + MAOI
B. theta + MAOI
C
Serum
Brain
Extracted Ion Counts (PEA)
Retention Time (min)
Phenethylamine
M. morganii
M. morganii + MAOI
B. theta + MAOI
D
Cecum
Colon
Serum
Brain
RLU (DRD2)
M.morganii
M.morganii + MAOI
E
Colon
Brain
RLU (DRD2)
MAOI
Mock Comm. A + MAOI
Mock Comm. A + M.morganii + MAOI
M.morganii + MAOI

Figure 5. *M. morganii*-derived phenethylamine combined with MAOI triggers lethal phenethylamine poisoning.

(A) *M. morganii* produces phenethylamine *in vivo*. Groups of female germ-free C57Bl/6 mice were colonized with *M. morganii* C135 and treated with or without the MAOI phenelzine. Phenethylamine concentration in colonic extracts was examined using QQQ-MS/MS.

(B) Mice colonized with *M. morganii* exhibit lethal phenethylamine poisoning after treatment with the MAOI phenelzine. Groups of female germ-free C57Bl/6 mice were monocolonized with *M. morganii* C135 for one week before treatment with phenelzine in the drinking water. Survival is depicted on a Kaplan-Meier curve. n=4 mice per group.

(C and D) *M. morganii*-colonized mice treated with phenelzine accumulated phenethylamine in the cecum, colon, serum and brain. *M. morganii* C135 and *B. theta* C34 monocolonized female C57Bl/6 mice were treated with or without the MAOI phenelzine in the drinking water. Phenethylamine was detected in the cecum, colon, serum or brain via QQQ-MS/MS (C) or DRD2 PRESTO-Tango (D). n=4 mice per group.

(E) *M. morganii*-derived phenethylamine accumulates in the sera and brains of mice colonized with a mock community of nine phylogenetically diverse human gut microbes (Mock Community A) plus *M. morganii*. Germ-free female C57Bl/6 mice were colonized with a mock community of nine phylogenetically diverse human gut microbes (Mock Community A) with or without *M. morganii* C135, or monocolonized with *M. morganii* C135. All mice were treated with the MAOI phenelzine in the drinking water for one week and then phenethylamine accumulation in the colon and brain was detected using DRD2-Tango as a proxy.

Data in all panels are representative of at least two independent experiments. Data are presented as mean ± SEM. One-way ANOVA with Tukey's post-hoc test (D-E), Kaplan meier and Log rank analysis (B), *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

disorders [38]. We found that mice colonized with *M. morganii* became lethargic within days after treatment with MAOI, and more than half of all mice colonized with *M. morganii* died before the seventh day of treatment (Figure 5B). In contrast, mice monocolonized with *B. theta* C34 and treated with MAOI appeared healthy. Morbidity and mortality after MAOI treatment correlated with elevated levels of PEA in the colon, serum and brains of *M. morganii* monocolonized mice treated with phenelzine, as measured by QQQ-MS/MS (Figure 5C and S5F, G). Finally, we found that cecal and colonic contents, as well as serum and brain extracts from MAOI-treated *M. morganii* colonized mice activated DRD2 (Figure 5D).

To test whether *M. morganii*-derived PEA would also accumulate systemically under more physiological conditions, we colonized germ-free mice with *M. morganii* C135 in the context of a mock community of nine diverse gut microbes and treated these mice with phenelzine. We found that, even in

the presence of a mock community, *M. morganii* colonization led to an increase in PEA in the colon and brain, as measured by DRD2 activation (Figure 5E). Together, these data show that *M. morganii*-derived phenethylamine can accumulate systemically in mice treated with MAOIs.

2.2.6 A unique *Bacteroides* isolate activates GPR56/AGRG1

We observed that specific bacterial supernatants could activate select orphan GPCRs (Figure 6A). To confirm these hits, we repeated our PRESTO-Tango screening procedure using a richer culture medium (Gifu) that supports more robust growth of most of the human gut microbes in our culture collection. This modified procedure significantly expanded the number of positive hits against orphan GPCRs: 17 orphan GPCRs showed greater than four-fold activation compared to media only controls in response to at least one bacterial supernatant (Figure 6B). We noted that metabolites from a strain assigned to the species *B. theta* (*B. theta* C34) activated GPR56/AGRG1 under both culture conditions (Figure 6C). In contrast, other strains of *B. theta* in our collections, including commercially available strains of *B. theta* and multiple other *Bacteroides* species, failed to activate GPR56/AGRG1 (Figure 6D) despite similar bacterial growth (Figure S6A).

2.2.7 The essential amino acid L-Phe activates GPR56/AGRG1 and GPR97/AGRG3

Since there is no known endogenous small molecule ligand for GPR56/AGRG1 [39], we next attempted to identify the specific metabolite produced by *B. theta* C34 that activated GPR56/AGRG1. *B. theta* C34 supernatants were lyophilized, extracted with methanol, and subjected to fractionation by reversed-phase HPLC. All resulting fractions were analyzed for activity via GPR56-Tango and fraction 11 was identified as the active fraction (Figure 6E). High resolution mass spectrometry, NMR and coinjection analyses of F11 showed that the essential amino acid phenylalanine (Phe) is the primary constituent of F11

Figure 6

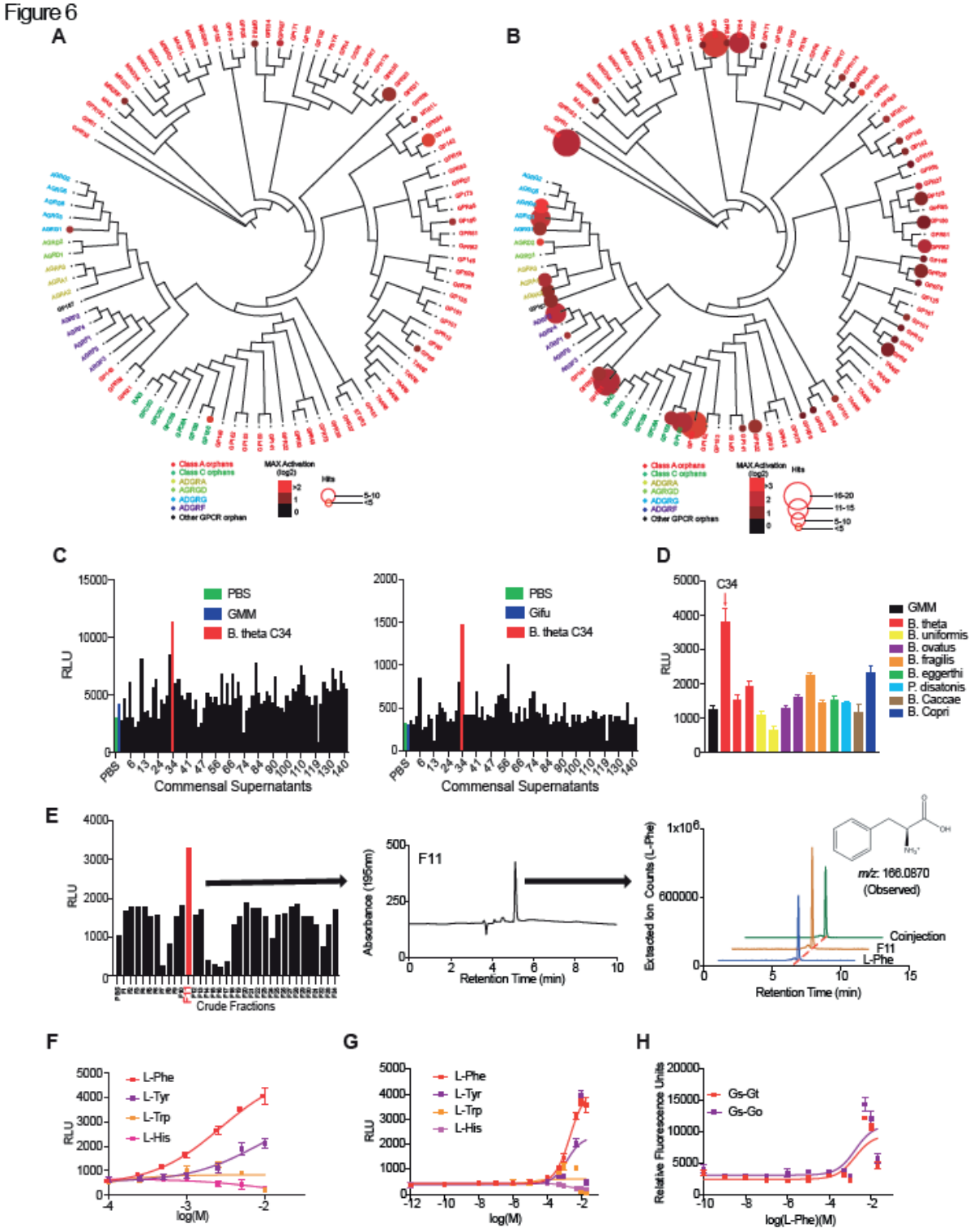

A
B
C
D
E
F
G
H
PBS
GMM
B. theta C34
Gifu
Commensal Supernatants
RLU
C34
B. theta
B. uniformis
B. ovatus
B. fragilis
B. eggerthi
P. disatonis
B. Caccae
B. Copri
Crude Fractions
F11
Absorbance (195nm)
Retention Time (min)
Extracted Ion Counts (L-Phe)
m/z: 166.0870
(Observed)
Coinjection
L-Phe
L-Tyr
L-Trp
L-His
log(M)
Gs-Gt
Gs-Go
Relative Fluorescence Units
log(L-Phe)(M)

Figure 6. A unique strain of *B. thetaiotaomicron* C34 is a prolific producer of L-Phe and activates GPR56/AGRG1.

(A and B) Activation of orphan GPCRs by metabolomes from a human gut microbiota culture collection consisting of 144 strains isolated from 11 IBD patients and grown in either gut microbiota medium (A) or Gifu (B) as measured by PRESTO-Tango. Screening results are displayed on a phylogenetic tree of orphan GPCRs that was constructed and visualized with equal branch lengths using gpcrdb.org, PHYLIP and jsPhyloSVG. Color intensities represent the magnitude of activation over media and radii of circles represent the number of bacteria that activated a given GPCR by more than two-fold.

(C) A single isolate C34 assigned to the species *Bacteroides thetaiotaomicron* activates GPR56/AGRG1 when cultured in gut microbiota medium (GMM: top panel) or Gifu medium (bottom panel). Activation of GPR56/AGRG1 by supernatants from 144 human gut isolates was measured via GPR56 PRESTO-Tango.

(D) *B. theta* strain C34 uniquely activates GPR56/AGRG1. Activation of GPR56/AGRG1 by supernatants from diverse species and strains from the genera *Bacteroides* and *Parabacteroides* culture in GMM was measured via GPR56 PRESTO-Tango. n=3 replicates per isolate.

(E) *B. theta* C34-produced L-Phe activates GPR56/AGRG1. *B. theta* C34 supernatants were fractionated via reversed-phase HPLC and fractions were evaluated for activation of GPR56/AGRG1 via GPR56 PRESTO-Tango. The active fraction (F11) contained a primary constituent that was identified via LC-MS, HRMS-ESI-QTOF, NMR, and advanced Marfey's analyses as L-Phe.

(F and G) L-Phe preferentially activates the orphan receptor GPR56/AGRG1. Activation of GPR56/AGRG1 by titrating doses of pure L-Phe, L-Tyr, L-Trp, and L-His was measured via GPR56 PRESTO-Tango using RPMI 1640 medium containing endogenous L-Phe and L-Tyr (F) or custom medium lacking L-Phe and L-Tyr (G). n=3 replicates per sample.

(H) L-Phe activates G protein-dependent signaling downstream of GPR56/AGRG1. Activation of G proteins downstream of GPR56/AGRG1 by L-Phe as measured by the CRE-SEAP assay. $G\alpha_s$-$G\alpha_t$ and $G\alpha_s$-$G\alpha_o$ chimeras were used to redirect GPR56/AGRG1 signaling to $G\alpha_s$ and enable use of the CRE-SEAP assay. n=3 replicates per sample.

Data in all panels except for A, B, and E are representative of at least three independent experiments. Data are presented as mean ± SEM.

(Figure 6E and S6B). Finally, structural characterization using advanced Marfey's analysis confirmed that L-Phe is the likely bioactive ligand (Figure S6C) [40]. We next tested whether pure L-Phe or structurally related amino acids could activate GPR56/AGRG1 using GPR56-Tango. We found that L-Phe and, to a lesser extent, L-Tyr stereoselectively activated GPR56/AGRG1, while L-Trp and L-His, D-Phe, D-Trp, D-His and D-Tyr showed no activity (Figure 6F and S6D). We hypothesized that L-Phe and L-Tyr in the cell culture medium used for the Tango assay might obscure the full extent of GPR56/AGRG1 activation by L-Phe. Thus, we formulated a custom medium lacking L-Phe and L-Tyr and examined GPR56/AGRG1

activation under these conditions. We found that removal of endogenous L-Phe and L-Tyr greatly increased the sensitivity and magnitude of GPR56 activation by L-Phe and L-Tyr as measured by the Tango assay (Figure 6G); furthermore, GPR56 was essential for this response as cells without GPR56 failed to respond to L-Phe (Figure S6E). We also examined the genomes of *B. theta* C34 and two other strains of *Bacteroides* that failed to activate GPR56/AGRG1 and found that, despite their differential secretion of L-Phe, all four strains showed equivalent presence of enzymes in the shikimate pathway that are involved in the synthesis of L-Phe (Table S4, S5).

To test whether L-Phe can also activate G protein-based signaling downstream of GPR56/AGRG1, we took advantage of promiscuous $G\alpha_s$-$G\alpha_t$ and $G\alpha_s$-$G\alpha_o$ chimeras (a kind gift from Stephen Liberles), which reroute most GPCRs through $G\alpha_s$ and thus enable use of the CRE-SEAP assay to read out G protein activation by many different receptors [41]. Using this system, we found that L-Phe can activate G protein-dependent signaling downstream of GPR56/AGRG1 (Figure 6H and Figure S6F); however, because activation of G protein-dependent signaling required high concentraions of L-Phe (>1mM), it remains unclear whether physiological concentrations of L-Phe will engage this pathway *in vivo.*

GPR56/AGRG1 is a member of the adhesion GPCR family. Adhesion GPCRs characteristically possess large extracellular domains that mediate interactions with a variety of protein ligands, such as components of the extracellular matrix [39]. We thus tested whether the extracellular domain of GPR56/AGRG1 was also required for activation by the small molecule L-Phe by constructing a truncation mutant of GPR56/AGRG1. Although this mutant is expressed normally [42], it failed to respond to L-Phe (Figure S6E). Together, these data show that a unique strain of *B. theta* secretes high levels of L-Phe and that L-Phe is a novel agonist of the adhesion GPCR GPR56/AGRG1.

We next examined whether other orphan GPCRs might also respond to L-Phe. We performed PRESTO-Tango screening of all adhesion and orphan GPCRs stimulated with L-Phe and found that GPR97/AGRG3 was also activated by L-Phe (Figure S7A). Upon further analysis, GPR97/AGRG3 showed greater selectivity toward L-Phe than GPR56/AGRG1—L-Phe, but not L-Tyr, L-Trp, or L-His, activated

GPR97/AGRG3 (Figure S7B). Like GPR56/AGRG1, the extracellular domain of GPR97/AGRG3 was required for its ability to respond to L-Phe (Figure S7C), and removal of L-Phe and L-Tyr from the medium increased the responsiveness of GPR97/AGRG3 to L-Phe (Figure S7B and S7D). Furthermore, L-Phe also activated G protein-dependent signaling downstream of GPR97/AGRG3 (Figure S7E). Notably, both GPR56/AGRG1 and GPR97/AGRG3 belong to the G family of adhesion GPCRs and are closely related evolutionarily (Figure S7F), which may explain their shared ability to detect the essential amino acid L-Phe.

2.2.8 Bacterial metabolic exchange can contribute to *in vivo* production of phenethylamine

Thus far, our reductionist studies revealed that *B. theta* C34 produces large amounts of L-Phe while *M. morganii* C135 can process L-Phe into the trace amine phenethylamine. Thus, we wished to address whether these two bacteria might participate in an active metabolic exchange *in vivo*. The first step in investigating this hypothesis was to determine whether *B. theta* C34 can directly synthesize L-Phe. Using a defined minimal medium that lacks L-Phe (standard amino acid complete medium or SACC; specific formulation described in methods; [9, 43], we observed that *B. theta* C34 could directly synthesize significant amounts of L-Phe *in vitro* (Figure 7A, B). We thus monocolonized mice fed an L-Phe deficient diet with *B. theta* C34 and evaluated the *in vivo* production of L-Phe via QQQ-MS/MS. GF mice fed with an L-Phe deficient diet exhibited reduced concentrations of L-Phe in the feces as compared to GF mice fed a conventional diet (Figure 7C). In contrast, mice colonized with *B. theta* C34 and fed with an L-Phe-deficient diet exhibited significantly increased levels of L-Phe as compared to GF mice fed with an L-Phe-deficient diet (Figure 7C).

We next examined whether *M. morganii* would directly process *B. theta* C34-derived L-Phe into phenethylamine. We cultured *B. theta* C34 in SACC medium and then transferred *B. theta* supernatant to a culture of *M. morganii* C135. We observed that *B. theta* C34-derived L-Phe was efficiently converted into phenethylamine by *M. morganii* C135 (Figure 7D). To test whether metabolic exchange between C34 and *M. morganii* C135 could contribute to *in vivo* production of PEA, we colonized GF mice with either *M. morganii* C135 alone or both *B. theta* C34 and *M. morganii* C135 and fed these mice with a simplified diet We found that mice colonized with *M. morganii* C135 alone and fed an L-Phe deficient diet remained

Figure 7

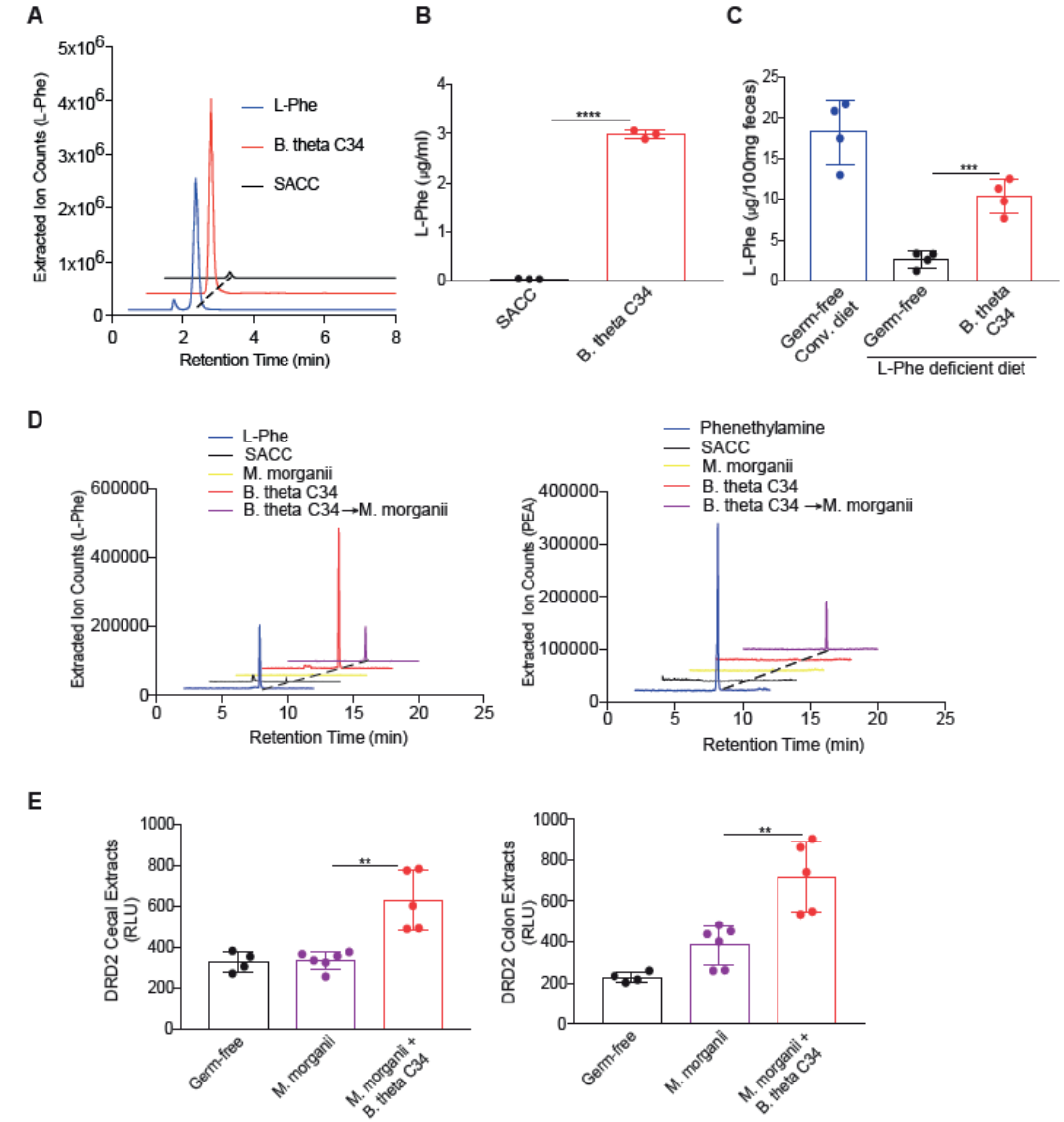


Figure 7. Active metabolic exchange between two commensals supports production of phenethylamine.

(A and B) *B. theta* C34 can directly synthesize L-Phe. L-Phe concentrations in supernatants from C34 grown in a minimal medium (SACC) lacking L-Phe were evaluated by LC-MS (A) and quantitated by QQQ-MS/MS (B). n=3 replicates per sample.

(C) *B. theta* C34 produces L-Phe *in vivo.* Groups of germ-free female C57Bl/6 mice fed a conventional diet or a defined diet lacking L-Phe were colonized with or without *B. theta* C34. Fecal L-Phe concentrations were measured by QQQ-MS/MS one week after colonization. n=4 mice per group.

(D) *M. morganii* C135 consumes *B. theta* C34-derived L-Phe to produce phenethylamine *in vitro*. *B. theta* C34 cultures were grown in SACC medium lacking L-Phe. Supernatants of C34 cultures were later incubated with *M. morganii* C135. QQQ-MS/MS traces of L-Phe and phenethylamine (PEA) levels in these cultures are depicted here.

(E) *B. theta* C34 and *M. morganii* C135 can participate in active metabolic exchange to produce phenethylamine *in vivo*. Groups of female germ-free C57Bl/6 mice were monocolonized with *M. morganii* C135 or co-colonized with *B. theta* C34 and *M. morganii* C135, fed a diet lacking L-Phe and treated with the MAOI phenelzine. Activation of DRD2 by phenethylamine in cecal and colonic extracts was measured by PRESTO-Tango. n=4-6 mice per group.

Data in all panels are representative of at least two independent experiments. Data are presented as mean ± SEM. One-way ANOVA with Tukey's post-hoc test (B-C and E-F), **p < 0.01, ***p < 0.001, ****p < 0.0001.

healthy lacking L-Phe. We then treated these mice with phenelzine to reveal the potential production of PEA. and produced minimal PEA (as measured by DRD2 activation by cecal and colonic extracts) despite MAOI treatment (Figure 7E). In contrast, mice that were bi-colonized with C34 and *M. morganii* C135, fed an L-Phe deficient diet, and treated with MAOI became lethargic by day 4 and produced significant levels of PEA as measured by DRD2 activation by cecal and colonic extracts (Figure 7E). This demonstrates that *B. theta* C34 and *M. morganii* C135 can participate in an active metabolic exchange *in vivo* and that this exchange can potentially contribute to the production of a bioactive trace amine that can have potent effects on host physiology.

2.3 DISCUSSION

Advanced metabolomic, metagenomic and functional genomic approaches have revealed that the human microbiota can produce tens of thousands of unique small molecules [5, 13, 44]. However, the overwhelming complexity of the gut microbiota metabolome obscures facile recognition of chemical relationships between microbes and their hosts [5, 8]. Here, we used host GPCR activation as a lens to detect bioactive metabolites produced by individual gut microbes. This approach revealed a plethora of novel microbiota metabolite-GPCR interactions of potential physiological importance. For example, we uncovered a diet-microbe-host axis that influences intestinal motility through the microbial production of histamine and a tri-partite microbe-microbe-host relationship that results in the production of the potent trace amine phenethylamine. Both of these examples resulted in profound effects on local and systemic host physiology.

We found that dozens of human gut bacteria from diverse phyla, families, species, and strains produced small molecules that activated various GPCRs (as measured by PRESTO-Tango), including both well-characterized GPCRs and orphan GPCRs. We also observed patterns of metabolite production that were largely predictable based on phylogeny, as well as strain-specific differences within a given species. Future studies will be necessary to determine when and why specific pathways are conserved or not in distinct species and strains. We speculate that metabolites resulting from core metabolic processes essential to a given microbe will be highly conserved, while metabolites involved in competitive processes may show considerable variation even between highly related strains. Regardless of the teleological origins of the production of particular microbial metabolites, our results support the notion that human-associated microbes represent a remarkably rich source of small molecules that impact human biology.

Prior studies have employed functional metagenomics screens as well as bioinformatics- and bioassay-guided natural product discovery approaches to uncover novel microbial-derived ligands for host GPCRs, including orphans (*e.g.*, SCFAs and GPR41 and 43, and commendamide and G2A and GPR119); however, these approaches also have notable limitations [5, 13]. For example, while functional metagenomics

screens enable the identification of novel biosynthetic gene clusters and their products from unculturable microorganisms, they are restricted in scope to contiguous biosynthetic gene clusters that can be expressed in heterologous hosts, require large-scale library construction efforts, and necessitate extensive follow up to identify the specific host receptors that recognize novel bioactive compounds. Similarly, while traditional bioassay-guided natural product discovery efforts enable identification of compounds produced by their native sources that can engage a specific receptor or pathway, their utility is largely restricted to cultivatable microorganisms and they typically examine only a single receptor or activity at a time. In contrast, the high-throughput functional profiling approach that we employ here enables simultaneous interrogation of hundreds of receptors and thousands of chemicals and is unconstrained by prior annotations of biosynthetic gene clusters or metabolites (although still dependent on microbial cultivation). We thus anticipate that future expansions of this approach will continue to uncover novel microbial metabolites that impact host physiology and reveal new natural ligands for orphan receptors.

Our initial screens were performed *in vitro*; however, we were particularly interested in examining the possibility that production of GPCR agonists by specific microbes might shape host physiology *in vivo*. We found that histamine production by *M. morganii* or *L. reuteri* promoted increased colon motility (as measured by fecal output), that feeding with exogenous histidine further increased local, *M. morganii*-dependent production of histamine and colonic motility, and that histamine receptor inhibition partially reversed these effects. Since fecal output is an incomplete measure of intestinal motility, future studies will be necessary to determine the mechanistic basis of these change (*e.g.*, effects on fluid secretion versus impacts on the enteric nervous system). We also found that *M. morganii* monocolonized mice fed with histidine exhibited elevated systemic levels of histamine, indicating a potential role for microbiota-derived histamine in shaping systemic immune responses. Indeed, a recent study of the gut microbiome in asthmatic patients found a significant increase in *M. morganii* in asthmatics as compared to healthy controls [45]. Additionally, by mining publicly available metagenomic datasets from the HMP2, we found that histamine decarboxylases (in general and from *M. morganii* specifically) are enriched in patients with Crohn's

disease. This implies that histamine production by the microbiota generally (and by *M. morganii* in particular) may impact human disease [36].

While prior studies had reported production of multiple biogenic amines by *M. morganii* [26, 27], we found that all isolates of *M. morganii* produced the potent trace amine phenethylamine rather than dopamine or tyramine. We also found that treatment of *M. morganii* monocolonized mice with an MAOI led to systemic accumulation of phenethylamine and mortality. Phenethylamine is a potent neuroactive chemical that, unlike dopamine and tyramine, can readily cross the blood-brain barrier [28]. The effects of phenethylamine are thought to be mediated primarily through activation of the trace amine-associated receptors and subsequent release of norepinephrine and dopamine [46-48]. However, our studies suggest that phenethylamine can also act as an agonist for DRD2-4 and may be a biased agonist for DRD1 and 5.

MAOIs were the first FDA approved antidepressants [49]; however, their current usage is limited due to dangerous dietary and drug-drug interactions [38]. Nonetheless, MAOIs remain an important treatment option for patients with refractory depression and other psychiatric disorders [38], as well as Parkinson's disease [50]. Our findings raise an intriguing possibility that interindividual variability in microbial production of phenethylamine could explain the variable efficacy of MAOIs on depression. In particular, phenethylamine enhances mood and is readily able to cross the blood-brain barrier [51], so it is plausible that inhibition of MAO in the gut could enable systemic accumulation of microbially-produced phenethylamine that can act distantly in the brain. Our studies thus underscore the possibility that pharmacological inhibitors of biogenic amine receptors that are thought to act at specific sites (*e.g.*, in the brain) may also exert their pharmacologic activities through modulation of host-microbiota interactions.

Our studies also uncovered a specific *Bacteroides* strain that uniquely produces high levels of the essential amino acid L-Phe and revealed that L-Phe could activate the orphan GPCRs GPR56 and GPR97. Although we did not focus on the physiological roles of these GPCRs here, our finding that L-Phe is a GPR56 and GPR97 agonist raises multiple possibilities. GPR56 is highly expressed in the small intestine and human pancreatic islets [52, 53], and L-Phe concentrations in the jejunum can reach concentrations up to

2 mM after a meal [54]. Thus, GPR56 may act as a nutrient sensor to regulate digestion and satiety. Although L-Phe levels in the serum usually are well below the levels necessary to activate GPR56/97, patients with phenylketonuria (PKU) who cannot degrade L-Phe exhibit serum concentrations of L-Phe higher than 1 mM [55]. Thus, it is theoretically possible that PKU patients may exhibit activation of GPR56 and/or GPR97 in other tissues.

The microbiota metabolome results from a complex web of interactions between diverse microbial species and strains, environmental inputs (*e.g.*, diet), and host factors. Dissecting these metabolite networks will be essential to eventually leverage microbial chemistries for therapeutic benefit. Using a reductionist approach, we discovered two bacterial isolates that traffic in the same small molecule: a unique strain of *B. theta* that is a prolific producer of L-Phe and *M. morganii*, which efficiently converts L-Phe into phenethylamine. Thus, these studies demonstrate the ability of reductionist approaches to reveal metabolic exchanges that may otherwise be missed when examining endpoint microbiota metabolomes produced by complex mixtures of microorganisms. Understanding these metabolic exchange networks will be essential to eventually determine the effect of the microbiota metabolome on host biology under more physiological settings (*i.e.*, in the context of complete gut microbial communities). Towards that end, we examined the effects of *M. morganii* on host physiology in the context of a mock gut microbial community consisting of nine phylogenetically diverse human gut microbes. In this context, we found that *M. morganii* continued to exhibit measurable metabolite-dependent impacts on the host. However, there are almost certainly other gut microbial community conditions where competition for ecological space or metabolic precursors (*e.g.*, L-His or L-Phe), or active degradation of *M. morganii*-derived metabolites may reduce or eliminate the impact of *M. morganii* on the host (or, conversely, may enhance the effects of *M. morganii*).

Our studies underscore the importance of dietary amino-acids (*e.g.*, L-His) in the production of biogenic amines that can shape host physiology. However, they also highlight an alternative source of substrates that are often thought of as primarily derived from diet (*e.g.*, essential amino acids): other members of the microbiota. This leads to the question of when microbial-produced amino acids may

potentially supplement or even replace dietary amino acids in microbial biotransformations. We modeled the possibility that microbe-derived L-Phe may be used as a substrate for biotransformation by *M. morganii* using a simplified and highly-artificial diet that lacks L-Phe. However, bacterial L-Phe may also be important under physiological conditions. For example, dietary amino acids are largely absorbed in the small intestine [54]; thus, colonic microbes such as *M. morganii* have relatively limited access to dietary amino acids as compared to small intestinal organisms. Also, low-protein diets naturally decrease microbial access to dietary amino acids and fasting may reduce intestinal amino acid availability even further [56]. Thus, microbial production of amino acids in the colon may play a critical role in the production of various bioactive microbiota metabolites under a variety of physiologically relevant conditions.

While our reductionist approach using monocultured human gut microbes revealed multiple potentially physiologically important diet-microbe-host and microbe-microbe-host axes, it also suffers from notable limitations that will need to be addressed in future studies to fully understand the *in vivo* bioactive microbiota metabolome. For example, patterns of microbial metabolite production can vary substantially depending on the media used for microbial cultivation (as we observed in our separate screens using GMM versus Gifu), and *in vitro* monoculture conditions fail to capture metabolites that result from interactions with the host organism, biotransformations of compounds absent from the cultivation medium, or interactions with other microbes. Furthermore, the metabolite concentrations produced during *in vitro* cultivation may not reflect *in vivo* metabolite production. Finally, *in vitro* screens fail to reveal the natural tissue distributions of gut microbiota metabolites. Understanding these distributions will be particularly important for metabolites that activate host receptors that are expressed in diverse cell types and tissues. For example, the histamine and dopamine receptors are expressed on cells as diverse

as immune cells, central and peripheral neurons, smooth muscle, epithelial and endothelial cells, and in essentially all tissues including the gut, lung, and brain [25, 36, 57].

In conclusion, while the human gut microbiota metabolome is dauntingly complex and diverse, emerging approaches have begun to reveal the key chemical interactions at the host-microbe interface. We

show here that high-throughput activity-based screening using potential host receptors as a lens can highlight physiologically relevant microbiota metabolites from complex metabolite mixtures. Such host-centric, functional profiling approaches can thus facilitate a mechanistic understanding of how we interact with and are affected by our microbial inhabitants, and have the potential to yield targeted therapeutic interventions aimed at the interface between indigenous microbes and their hosts.

2.4 METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
B.fragilis	ATCC	25285
B.ovatus	ATCC	8483
B.thetaiotaomicron	ATCC	29741
B.uniformis	ATCC	8492
M.morganii	ATCC	25830&49948
Chemicals, peptides, and Recombinant Proteins		
Gifu Anaerobic Broth	VWR	11007-214
Bright-Glo™Luciferase Assay System	Promega	E2620
Histamine Elisa Kit	Enzo Life Sciences	ENZ-KIT140-0001
Desloratadine	Tocris	5958
Tiotidine	Tocris	0826
Iodophenpropit	Tocris	0779
A987306	Tocris	3640
Phenethylamine	Sigma-Aldrich	241008

Tyramine	Sigma-Aldrich	T90344
Dopamine	Sigma-Aldrich	H8502
Histamine	Sigma-Aldrich	H7129
Acetylcholine	Sigma-Aldrich	A2661
L-DOPA	Sigma-Aldrich	D9628
Succinate	Sigma-Aldrich	398055
Serotonin	Sigma-Aldrich	14927
Gastrin	Sigma-Aldrich	G9145
Peptide YY	Anaspec	AS024401
Pancreatic polypeptide	Anaspec	AS-22866
Cholecystokinin	Sigma-Aldrich	C2175
Trace mineral supplement	ATCC	MD-TMS
Vitamin supplement	ATCC	MD-VS
L-glycine	Sigma-Aldrich	G8898
L-valine	Sigma-Aldrich	94619
L-leucine	Sigma-Aldrich	L8000
L-isoleucine	Sigma-Aldrich	I2752
L-methionine	Sigma-Aldrich	64319
L-histidine	Sigma-Aldrich	H8000
L-arginine	Sigma-Aldrich	A5131
L-phenylalanine	Sigma-Aldrich	P2126
L-tyrosine	Sigma-Aldrich	T3754
L-tryptophan	Sigma-Aldrich	T0254
N-methylphenethylamine	Sigma-Aldrich	M68423
Octopamine	Sigma-Aldrich	O0250

Synephrine	Sigma-Aldrich	S0752
Epinephrine	Sigma-Aldrich	E4250
Norepinephrine	Sigma-Aldrich	A7257
3-Methoxytyramine	Sigma-Aldrich	M4251
DMEM	Sigma-Aldrich	D6429
RPMI 1640	Thermo Fisher	21870092
D-phenylalanine	Sigma-Aldrich	673-06-3
FDAA (Marfey's Reagent)	Thermo Fisher	48895

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

6-12 week old germ-free wild-type C57Bl/6 mice were used in all experiments. Both male and female mice were used for these studies and mice were age and sex matched within each experiment (only one sex was used for each independent experiment). We did not observe any obvious sex-specific differences in *in vivo* phenotypes in any of these studies.

Bacteria

All strains were cultured in gut microbiota medium or Gifu broth at 37 °C under anaerobic conditions and the identities of all strains were confirmed by 16S rRNA gene sequencing.

METHOD DETAILS

Bacterial growth conditions

For PRESTO-Tango screening, all commensals were cultured in gut microbiota medium or Gifu broth for 2 days in an anaerobic chamber (Coy) and commensal supernatants were sterilized by high-speed

centrifugation and sterile filtration (0.22 μm). For *in vitro* studies, *M. morganii* was cultured in minimal medium (MM), or MM with 10 mM L-Phe, 2.5mM L-Tyr, 10 mM L-DOPA, or 10mM L-His for 24 hours. Bacterial supernatants were analyzed by LC-MS.

PRESTO-Tango Assay

HTLA cells, a HEK293 cell line that stably expresses β-arrestin-TEV and tTA-Luciferase (a kind gift from Gilad Barnea, Brown University), were plated in 96-well tissue culture plates (Eppendorf) in DMEM containing 10% FBS and 1% Penicillin/Streptomycin. One day after plating (after reaching approximately 90% confluence), 200 ng per well GPCR-Tango plasmids ([19] in 20 μl DMEM were mixed with 400 ng polyethylenimine (Polysciences) in an equal volume of DMEM and incubated for 20 minutes at room temperature before adding the transfection mixture to the HTLA cells. 16-24 hrs after transfection, medium was replaced with 180μl fresh DMEM containing 1% Penicillin/Streptomycin and 10mM HEPES and 20μl bacterial medium alone or bacterial supernatant. All bacteria were cultured in gut microbiota medium or Gifu under anaerobic conditions for 2 days and bacterial supernatant was isolated via high-speed centrifugation followed by filtration with a 0.22 μm filter. Bacteria that failed to reach an OD of 0.5 after 2 days or caused obvious cell toxicity (*e.g.*, *Clostridum perfringens* isolates) were excluded from further analysis. Supernatants were aspirated 16-24 hr after stimulation and 50 μl per well of Bright-Glo solution (Promega) diluted 20-fold with PBS containing 20mM HEPES was added into each well. After 20 min incubation at room temperature, luminescence was quantified using a Spectramax i3x (Molecular Devices). Activation fold for each sample was calculated by dividing relative luminescence units (RLU) for each condition by RLUs from media alone controls.

Histamine ELISA

All strains were cultured in gut microbiota medium with or without 1% L-His for 24 hours and supernatants were collected via high-speed centrifugation. Cecal and colonic contents and fecal samples were collected and weighed; all samples were suspended in PBS at a ratio of 1:2 (w/v) and were

homogenized by vortexing. Serum and brains were collected, weighed and suspended in PBS at a ratio of 1:2 (w/v). Brains were homogenized by passing through a 21G needle fifty times. All samples were centrifuged and supernatants were collected for histamine ELISA according to the manufacturer's protocol.

Colonization of germ-free mice

Germ-free C57Bl/6 mice were colonized via oral gavage with 200µl of individual bacterial cultures or mixed bacterial consortia. Mock communities A and B consisted of the following taxa: Community A: *Bacteroides spp*; *P. distasonis*; *Peptoniphilus spp*; *B. ovatus*; *Clostridiales UC/UC*; *Lachnospiraceae UC/UC*; *C. stercoris*; *B. uniformis* and *Parabacteroides spp*.; and Community B: *Streptococcus spp*; *C. perfringens*; *B. fragilis*; *Erisipelotrichaceae spp*; *C. aerofaciens*; *Bacteroides UC*; *B. producta*; *Allobaculum spp* and *Oscillospira spp*. All strains were grown to roughly mid-log phase in GMM, mixed in equal ratios based on optical density, and then frozen at -80 °C in GMM containing 20% glycerol in rubber capped 2ml Wheaton vials until use. All gnotobiotic mice were maintained in Techniplast P Isocages and manipulated aseptically for the duration of the experiment. 16S rRNA gene sequencing of the V4 region to confirm colonization and microbial composition was performed essentially as described previously [21] except that data processing and analysis was done using QIIME2-DADA2 [58].

Fecal output assay.

Individual mice were housed in an empty container (1/4 gallon) for 1 hour after which time the fecal pellets were counted and weighed. For mice fed with L-His, mice were given water containing 1% L-His *ad libitum* for 2 weeks before fecal output measurements. Based on an average water intake of 4 ml/ms/day combined with the existing dietary histidine present in conventional mouse chow (5 g/kg; Envigo), feeding of 1% L-His *ad libitum* in the water is equivalent to an overall histidine concentration of ~15g/kg in the diet. For reference, histidine rich foods such as soy protein, egg white, parmesan cheese, cured pork, and beef contain roughly 23, 20.5, 16, 16, and 14 g/kg of histidine, respectively (https://ndb.nal.usda.gov/ndb/nutrients/report/).

CRE-SEAP Assay.

96-well plates were pretreated with 30 µl poly-D-lysine (10ug/ml in water) and incubated at room temperature for 30 minutes. Plates were washed twice with 100 µl PBS and HEK293T were seeded in 100 µl DMEM+10% FBS+1% Penicillin/Streptomycin in each well. When cells were 90% confluent, plasmids were transfected using PEI reagent at a ratio of 1:2. For DRD1, DRD5 and TA1 (which couple to $G\alpha_s$ protein), HEK293T cells were transfected with 50 ng GPCR and 50 ng CRE-SEAP reporter plasmid per well; For DRD2, DRD3, DRD4, GPR56 and GPR97, cells were transfected with 50 ng GPCR, 50 ng CRE-SEAP reporter plasmid (BD Biosciences) and 2.5 ng $G\alpha_s$- $G\alpha_o$ or $G\alpha_s$- $G\alpha_t$ chimeras (a kind gift from Stephen Liberles, Harvard; [41] per well. 6 hours after transfection, medium from the wells was replaced with 180 µl serum-free DMEM and 20 µl DMEM containing putative GPCR ligands was added to each well. After incubating for 48 hours at 37°C, followed by 2 hours of incubation at 70°C, supernatants from each well were mixed with an equal volume of 0.12 mM 4-methylumbelliferyl phosphate in 2 M diethanolamine bicarbonate, pH 10, and incubated at room temperature for 10 minutes. Fluorescence was measured using a SpectraMax plate reader (Molecular Devices).

Histamine Receptor Antagonist Treatment.

For histamine receptor antagonist treatment, 20 uM Desloratadine (HRH1 antagonist), Tiotidine (HRH2 antagonist), Iodophenpropit (HRH3 antagonist) and A987306 (HRH4 antagonist) were added to the drinking water. Colon motility was measured after one week of *ad libitum* treatment.

General Metabolomic Procedures.

NMR spectra were taken using an Agilent 600 MHz NMR system with a cryoprobe. High-resolution MS and tandem MS (MS/MS) data were obtained using an Agilent iFunnel 6550 ESI-HRMS-QTOF (Electron Spray Ionization-High Resolution Mass Spectrometry-Quadrupole Time-of-Flight) instrument on Phenomenex Kinetex 5 µm C18 100Å (4.6 × 250 mm) columns. The Agilent 1260 Infinity system with a Phenomenex Kinetex 5 µm C18 100Å column (4.6 × 250 mm) or an Agilent Poroshell 120

EC-C18 2.7 µm (3.0 x 50 mm) column and a photodiode array (PDA) detector was used for routine sample analysis. An Agilent Prepstar HPLC system with an Agilent Polaris C18-A 5 µm (21.2 × 250 mm) columns were used for sample fractionation and purification.

Metabolite Isolation and Purification.

B. thetaiotaomicron strain C11 was grown in 10 mL of gut microbiota medium under anaerobic conditions at 37°C for 24hr. Supernatant was harvested, lyophilized and extracted with 2mL methanol. The crude extract was then dried and fractionated using a preparative C18 HPLC system. The gradient used was 10-50% acetonitrile in water (with 0.01% trifluoroacetic acid) for 30min, then 100% for 5min. The fractions, which were collected every minute, were dried, resuspended in PBS buffer, and tested for bioactivity using PRESTO-Tango. The active fraction was characterized using ESI-HRMS-QTOF and NMR analyses. Stereochemistry was confirmed by advanced Marfey's analysis (Figure S5D)

Metabolite Quantitation.

Electro Spray Ionization-Triple Quadrupole-Tandem Mass Spectrometry ESI-QQQ-MS/MS was run using Multiple Reaction Monitoring (MRM) mode. An Agilent 6490 ESI-QQQ-MS/MS instrument with a Phenomenex Kinetex 1.7 µm C18 100Å (100 x 2.10mm) column was used for quantitation and calibration. Each standard (L-Phenylalanine, Phenethylamine, Histamine, Tyramine, Dopamine) was optimized using an Agilent Mass Hunter Optimizer. A calibration curve for each standard was established using various concentrations (0 - 25µM range) in triplicate. The gradient constituted 10-100% acetonitrile in ddH_2O (with 0.1% Formic Acid), then a wash step with 100% acetonitrile. The triplicate data was then subjected to linear regression analysis to produce a linear calibration curve. Processing of the experimental samples involved lyophilization and extraction with 100% MeOH (20% volume of original culture volume) before injecting samples. Sample absorbance was subjected to linear calibration to calculate concentrations.

Whole-Genome Sequencing and Annotation, HMP2 Data Mining and Data Deposition.

Whole-genome sequencing. Bacterial DNA was extracted using the DNeasy Ultraclean Microbial Kit (Qiagen) according to the manufacturer's instructions. Sequencing libraries were prepared using the Nextera XT library preparation kit (Illumina) according to the manufacturer's instructions and sequenced on a NovaSeq (2x150; Illumina).

De novo genome assembly. Genome assembly and annotation were performed essentially as described in Dodd et al. [9]. Briefly, all Illumina paired-end reads were filtered and trimmed using Trimmomatic v.0.38 [59] with the following parameters: ILLUMINACLIP: NexteraPE-PE.fa:2:30:12:1:true LEADING:3 TRAILING:3 MAXINFO:40:0.994 MINLEN:36. The four output files after trimming included two (forward and reverse) FASTQ files with paired reads and two FASTQ files with unpaired reads. All four files from each strain were assembled into contigs using SPAdes 3.13.0 [60] with the default parameters for paired-end libraries. Genome coverage was calculated by BBMap (http://sourceforge.net/projects/bbmap/). Summary statistics for each genome assembly and alignment are shown in Table S3.

Genome annotation. For each genome assembly, scaffolds longer than 2000 bp were uploaded to the Rapid Annotation using Subsystem Technology (RAST) server [61] for annotation (using the default RASTtk pipeline). The annotated genomes were downloaded as spreadsheets (Excel XLS format), and the summary of genome annotations as well as the detailed annotations for each strain are shown in Table S3.

HMP2 data acquisition and analysis. Publicly available metagenome and metabolome data from the HMP2 were downloaded from The Inflammatory Bowel Disease Multi'omics Database (IBDMDB), which was funded by the NIH Human Microbiome Project NIDDK U54DE023798 (https://ibdmdb.org). The IBDMDB provides longitudinal meta'ome data on the microbiome of subjects with three clinical diagnoses: nonIBD, CD and UC. Raw metagenomic sequencing data was pre-processed and used to generate taxonomic and functional profiles by the HMP team. The pipeline employed two steps: (1)

MetaPhlAn2 (Truong DT, 2015)-based taxonomic profiling, which uses clade-specific marker genes to identify species-level microbial taxa and their relative abundances using metagenomic data; and (2) HUMAnN2 (Franzosa EA, 2018)-based functional profiling. Briefly, HUMAnN2 implements a tiered search strategy to profile the functional content (gene family, functional pathway, *etc.*) of a meta'ome sample at species-level resolution. In the first tier, based on the known microbial species in a sample (as identified by MetaPhlAn2), HUMAnN2 constructs a custom gene sequence database for the samples by concatenating precomputed, functionally annotated pangenomes of detected species. In the second tier, nucleotide-level mapping of all reads against the sample's pangenome database is performed. In the final search tier, reads that do not align to an identified species' pangenome are then subjected to translated search against a comprehensive, non-redundant protein sequence database (UniRef90 or UniRef50). Per-gene alignment statistics are weighted based on alignment quality, coverage and sequence length to yield gene abundance values. Both taxonomic determinations and functional gene abundances are normalized as relative abundances to facilitate comparisons between samples with different sequencing depths.

The merged tables of taxonomic profiles (https://ibdmdb.org/tunnel/products/HMP2/WGS/1818/taxonomic_profiles.tsv.gz) and functional profiles (https://ibdmdb.org/tunnel/products/HMP2/WGS/1818/ecs.tsv.gz) from metagenomic analyses, the merged table from metabolomics analysis (https://ibdmdb.org/tunnel/products/HMP2/Metabolites/1723/HMP2_metabolomics.csv.gz), and the HMP2 metadata (https://ibdmdb.org/tunnel/products/HMP2/Metadata/hmp2_metadata.csv), were downloaded and analyzed to evaluate the distribution of abundance of histidine decarboxylase and histamine in this cohort. For participants whose clinical diagnosis changed during the course of the study, data points were assigned to the clinical diagnosis at the time of sample collection.

Quantification and Statistical Analysis.

Statistical analysis was performed using Graphpad Prism version 7.0. Data were assessed for normal distribution and plotted in the figures as mean ± SEM. No samples or animals were excluded from the analyses. One-way ANOVA and post hoc analysis with Tukey's test was used to compare the difference between treatment groups. Kaplan-Meier and Log rank analysis was used for survival experiments. Kruskall-Wallis with Dunn's multiple comparisons was used to analyze metagenomic and metabolomic data from the HMP; p-values for Kruskall-Wallis are approximated based on the chi-squared distribution and account for rank ties. Samples sizes are indicated in each figure legend and significant differences are indicated in the figures by *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Media Formulations.

Custom (L-Phe- and L-Tyr-free) Dulbecco's Modified Eagle's Medium (DMEM) formulation

Ingredients	Concentration in Medium (g/L)
Inorganic Salts	
Calcium Chloride	0.2
Ferric Nitrate•$9H_2O$	0.0001
Magnesium Sulfate(anhydrous)	0.09767
Potassium Chloride	0.4
Sodium Bicarbonate	3.7
Sodium Chloride	6.4
Sodium Phosphate Monobasic(anhydrous)	0.109
Amino Acids	
L-Arginine•HCl	0.084
L-Cystine•2HCl	0.0626

L-Glutamine	0.584
Glycine	0.03
L-Histidine•HCl•H_2O	0.042
L-Isoleucine	0.105
L-Leucine	0.105
L-Lysine•HCl	0.146
L-Methionine	0.03
L-Phenylalanine	0
L-Serine	0.042
L-Threonine	0.095
L-Tryptophan	0.016
L-Tyrosine•2Na•2H_2O	0
L-Valine	0.094
Vitamins	
Choline Chloride	0.004
Folic Acid	0.004
myo-Inositol	0.0072
Niacinamide	0.004
D-Pantothenic Acid (hemicalcium)	0.004
Pyridoxine•HCl	0.004
Riboflavin	0.0004
Thiamine•HCl	0.004
Other	
Glucose	4.5
Phenol Red•Na	0.0159
Pyruvic Acid•Na	0.11

Minimal Medium

Ingredients	Concentration in Medium (g/L)
Resazurin	0.0001
$KH2PO_4$	2
$K2HPO_4$	2
$MgCl_2•6H_2O$	0.2
$(NH4)_2SO4$	5
L-Glycine	0.075
L-Valine	0.117
L-Leucine	0.131
L-Isoleucine	0.131
L-Methionine	0.149
L-Histidine	0
L-Arginine	0.174
L-Phenylalanine	0
L-Tyrosine	0
L-Tryptophan	0
$NaHCO_3$	2.5
Cysteine HCl	0.5
Glucose	3.603
Trace Mineral Supplement(g/L)	**10ml**
EDTA	0.5
$MgSO4•7H_2O$	3.0
$MnSO4•H_2O$	0.5
NaCl	1.0
$FeSO4•7H_2O$	0.1
$Co(NO_3)_2•6H_2O$	0.1

$CaCl_2$ (anhydrous)	0.1
$ZnSO_4 \cdot 7H_2O$	0.1
$CuSO_4 \cdot 5H_2O$	0.010
$AlK(SO_4)_2$ (anhydrous)	0.010
H_3BO_3	0.010
$Na_2MoO_4 \cdot 2H_2O$	0.010
Na_2SeO_3 (anhydrous)	0.001
$Na_2WO_4 \cdot 2H_2O$	0.010
$NiCl_2 \cdot 6H_2O$	0.020

Vitamin Supplement (mg/L)	**10ml**
Folic acid	2.0
Pyridoxine hydrochloride	10.0
Riboflavin	5.0
Biotin	2.0
Thiamine	5.0
Nicotinic acid	5.0
Calcium Pantothenate	5.0
Vitamin B12	0.1
p-Aminobenzoic acid	5.0
Thioctic acid	5.0
Monopotassium phosphate	900.0

Standard Amino Acid Complete (SACC) medium

Ingredients	Concentration in Medium (g/L)
Resazurin	0.0001
$KH2PO_4$	2
$K2HPO_4$	2

$MgCl_2•6H_2O$	0.2
$(NH4)_2SO4$	5
L-Glycine	0.075
L-Valine	0.117
L-Leucine	0.131
L-Isoleucine	0.131
L-Methionine	0.149
L-Histidine	0.155
L-Arginine	0.174
L-Phenylalanine	0
L-Tyrosine	0.181
L-Tryptophan	0.204
$NaHCO_3$	2.5
Cysteine HCl	0.5
Glucose	3.603

Trace Mineral Supplement(g/L)	**10ml**
EDTA	0.5
$MgSO4•7H_2O$	3.0
$MnSO4•H_2O$	0.5
NaCl	1.0
$FeSO4•7H_2O$	0.1
$Co(NO_3)_2•6H_2O$	0.1
$CaCl_2$ (anhydrous)	0.1
$ZnSO_4•7H_2O$	0.1
$CuSO_4•5H_2O$	0.010
$AlK(SO_4)_2$ (anhydrous)	0.010
H_3BO_3	0.010
$Na_2MoO_4•2H_2O$	0.010
Na_2SeO_3 (anhydrous)	0.001

$Na_2WO_4 \cdot 2H_2O$	0.010
$NiCl_2 \cdot 6H_2O$	0.020

Vitamin Supplement (mg/L)	**10ml**
Folic acid	2.0
Pyridoxine hydrochloride	10.0
Riboflavin	5.0
Biotin	2.0
Thiamine	5.0
Nicotinic acid	5.0
Calcium Pantothenate	5.0
Vitamin B12	0.1
p-Aminobenzoic acid	5.0
Thioctic acid	5.0
Monopotassium phosphate	900.0

2.5 SUPPLEMARY INFORMATION

Figure S1

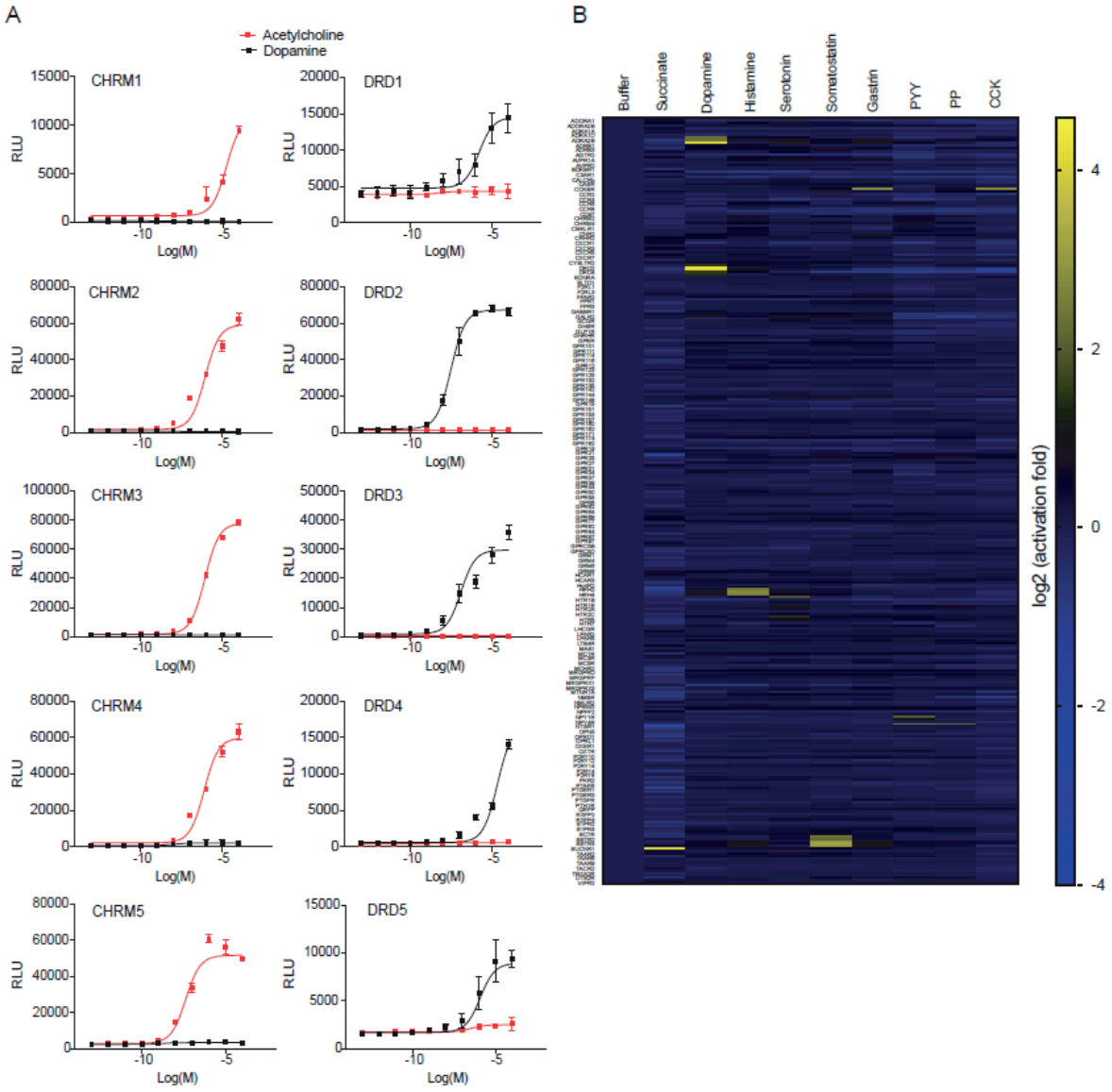


Figure S1. Sensitivity and specificity of the PRESTO-Tango assay, related to Figure 1.

(A) Activation of CHRMs and DRDs by titrating doses of acetylcholine and dopamine as measured by PRESTO-Tango. n=3 replicates per sample.

(B) Activation of GPCRs by defined GPCR ligands as measured by PRESTO-Tango. Activation is depicted on a log 2 scale as a heatmap of 314 GPCRs versus ligands.

Figure S2

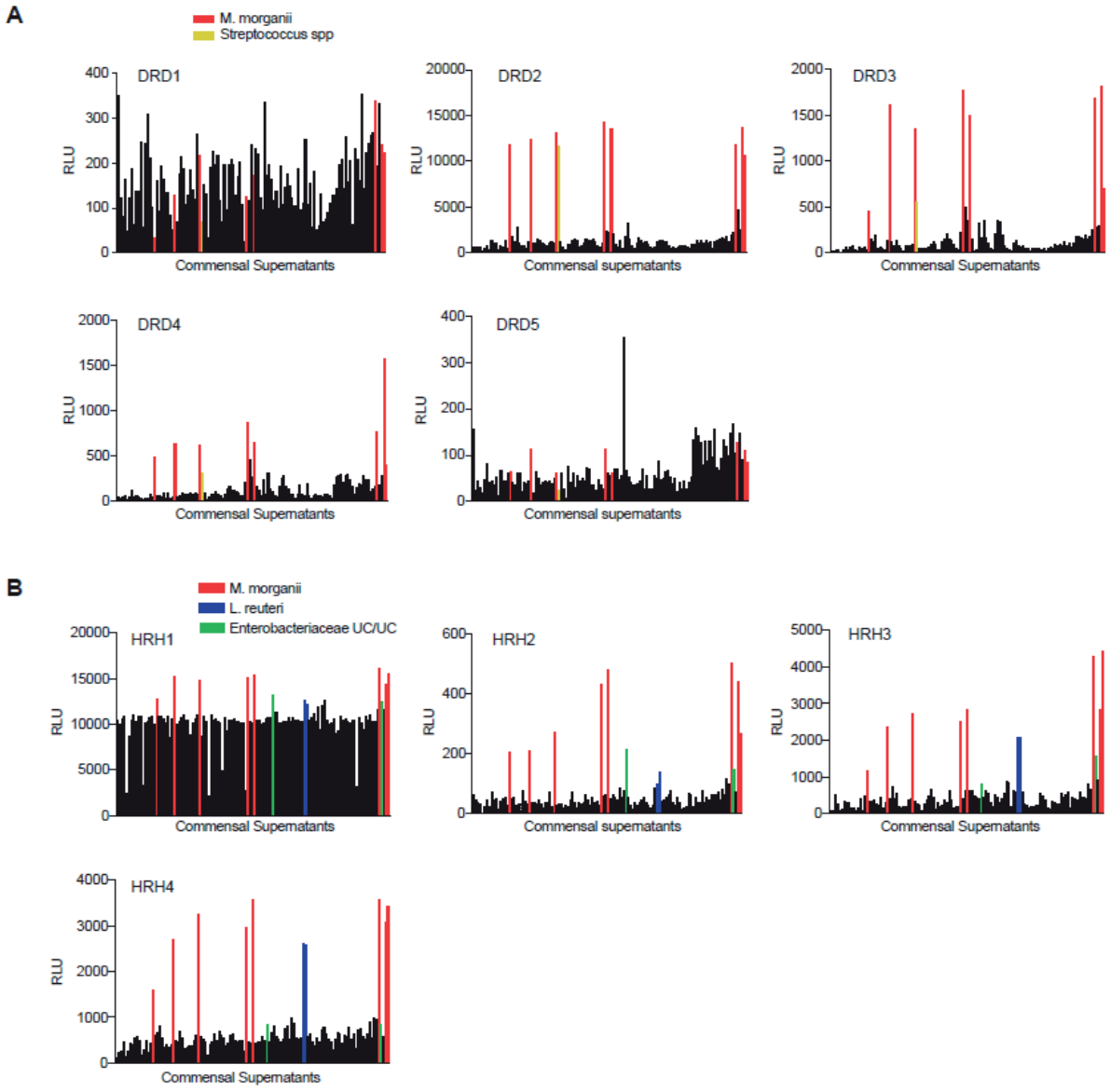


Figure S2. Diverse human gut bacteria activate DRDs and HRHs, related to Figure 3.

Activation of DRD1-5 (A) and HRH1-4 (B) by supernatants from 144 human gut bacteria cultured in gut microbiota medium (GMM) as measured by PRESTO-Tango.

Figure S3

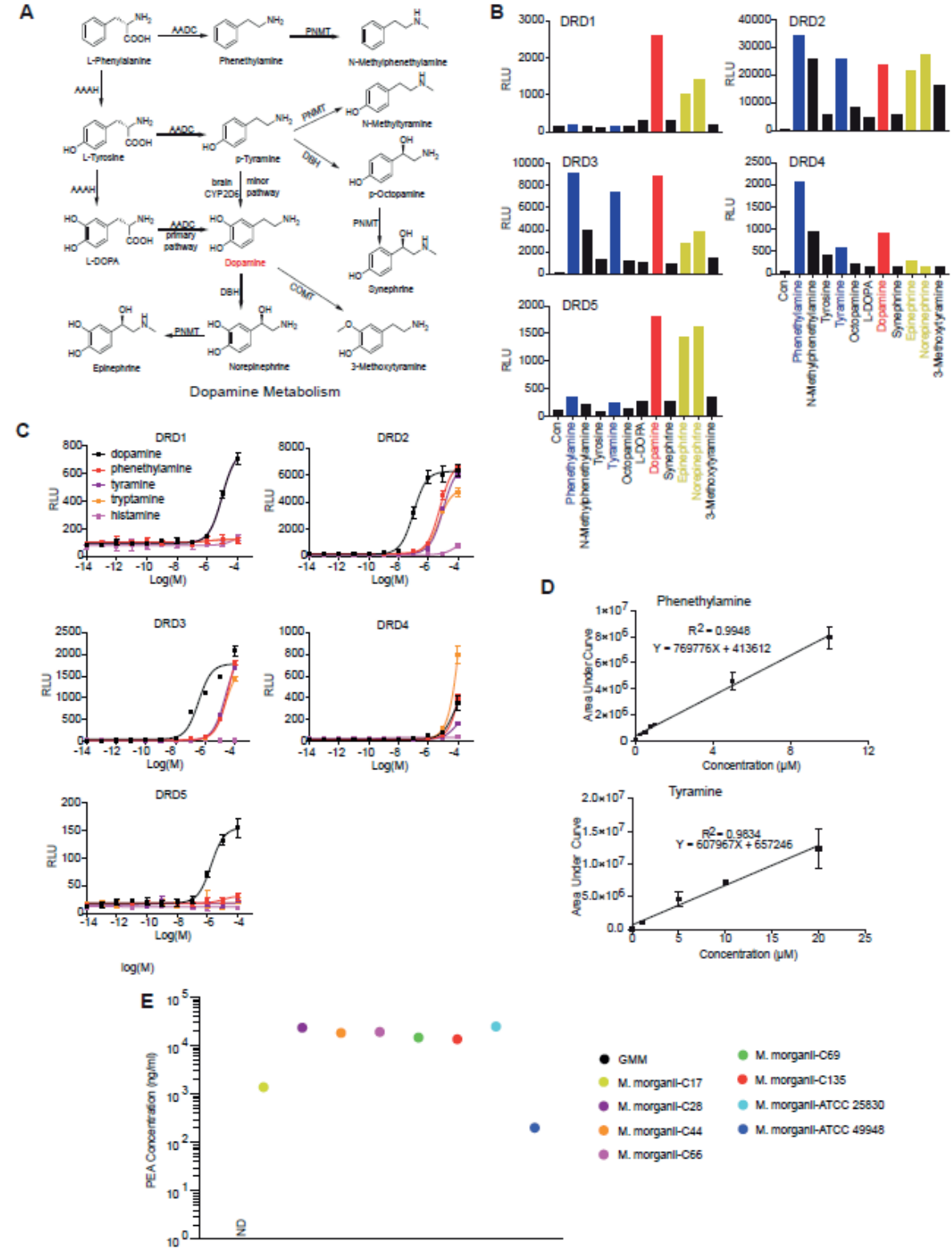

A
L-Phenylalanine
AADC
Phenethylamine
PNMT
N-Methylphenethylamine
AAAH
L-Tyrosine
p-Tyramine
N-Methyltyramine
DBH
p-Octopamine
brain CYP2D6
minor pathway
L-DOPA
primary pathway
Dopamine
Synephrine
COMT
Epinephrine
Norepinephrine
3-Methoxytyramine
Dopamine Metabolism
B
DRD1
DRD2
DRD3
DRD4
DRD5
RLU
Con
Phenethylamine
N-Methylphenethylamine
Tyrosine
Tyramine
Octopamine
L-DOPA
Dopamine
Synephrine
Epinephrine
Norepinephrine
3-Methoxytyramine
C
DRD1
DRD2
DRD3
DRD4
DRD5
dopamine
phenethylamine
tyramine
tryptamine
histamine
RLU
Log(M)
log(M)
D
Phenethylamine
R2 = 0.9948
Y = 769776X + 413612
Area Under Curve
Concentration (μM)
Tyramine
R2 = 0.9834
Y = 607967X + 657246
E
PEA Concentration (ng/ml)
ND
GMM
M. morganII-C17
M. morganII-C28
M. morganII-C44
M. morganII-C66
M. morganII-C69
M. morganII-C135
M. morganII-ATCC 25830
M. morganII-ATCC 49948

Figure S3. Identification of *M. morganii*-derived compounds that activate DRDs and HRHs, related to Figure 3.

(A) Mammalian dopamine metabolism.

(B) Phenethylamine and tyramine serve as selective DRD2/DRD3/DRD4 agonists. Activation of DRD1-5 by metabolites in the mammalian dopamine metabolism pathway was measured via PRESTO-Tango.

(C) DRD1-5 activation by titrating doses of tyramine, dopamine and phenethylamine was measured by PRESTO-Tango (D). n=3 replicates per sample.

(D) Calibration curve for phenethylamine and tyramine on QQQ-MS/MS instrument.

(E) Quantification of phenethylamine production by *M. morganii* strains via QQQ-MS/MS.

Data in all panels are representative of at least two independent experiments.

Figure S4

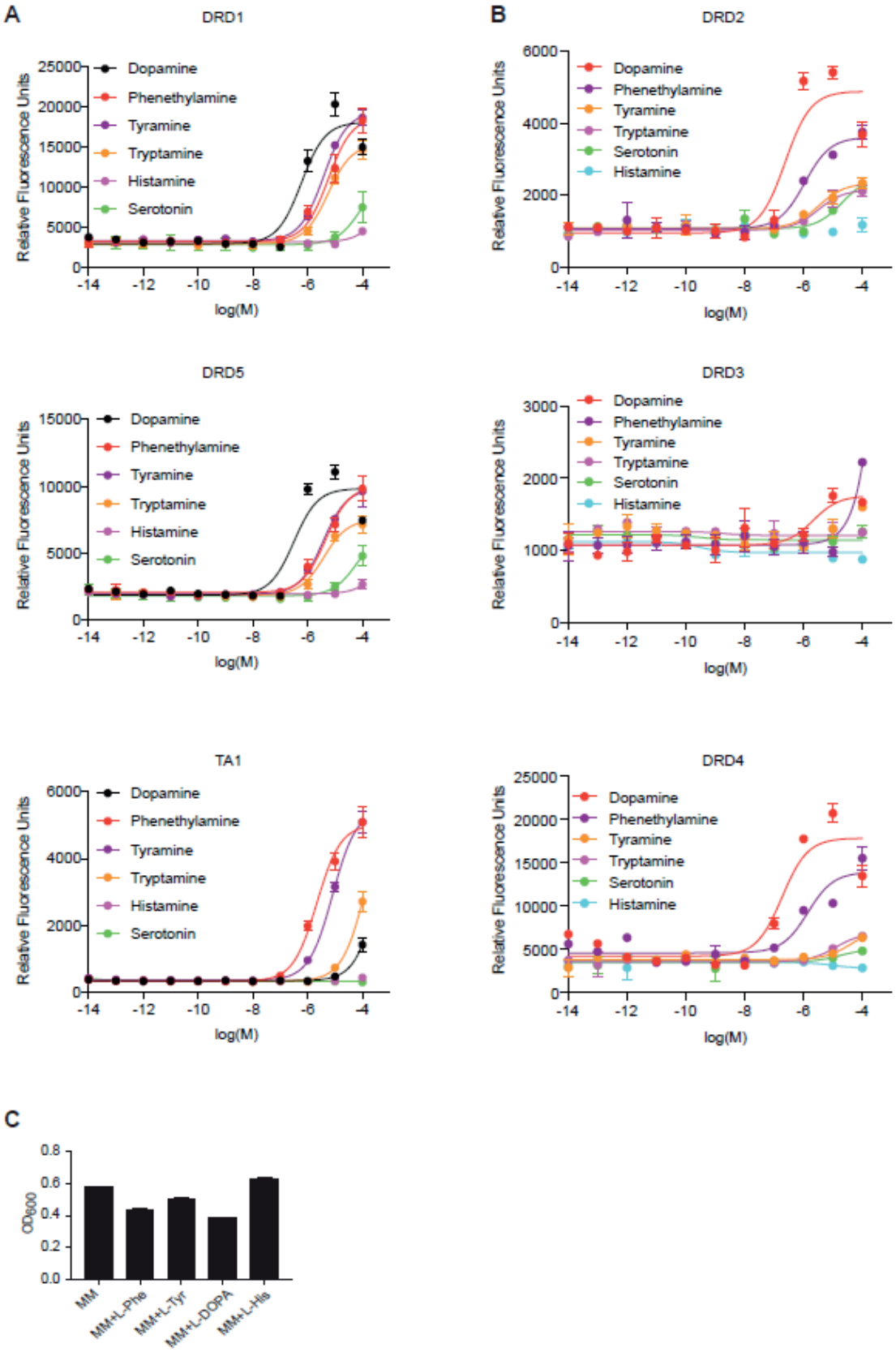

A
DRD1
Relative Fluorescence Units
log(M)
Dopamine
Phenethylamine
Tyramine
Tryptamine
Histamine
Serotonin
B
DRD2
DRD5
DRD3
TA1
DRD4
C
OD600
MM
MM+L-Phe
MM+L-Tyr
MM+L-DOPA
MM+L-His

Figure S4. Activation of G-protein signaling by phenethylamine and related chemicals downstream of the dopamine receptors, related to Figure 3.

(A) Activation of $G\alpha_s$-dependent signaling downstream of DRD1, 5 and TAAR1 by titrating doses of phenethylamine and related chemicals was measured by the CRE-SEAP assay. n=3 replicates per sample.

(B) Activation of G protein-dependent signaling downstream of DRD2-4 by titrating doses phenethylamine and related chemicals was measured by the CRE-SEAP assay. A $G\alpha_s$-$G\alpha_o$ fusion was used to redirect DRD2-4 to $G\alpha_s$ and enable use of the CRE-SEAP assay. n=3 replicates per sample.

(C) OD values for 24 hour cultures of *M. morganii* grown in minimal medium (MM) with or without L-Phe, L-Tyr, L-DOPA or L-His. n=3 replicates per sample.

Data in all panels are representative of at least two independent experiments.

Figure S5

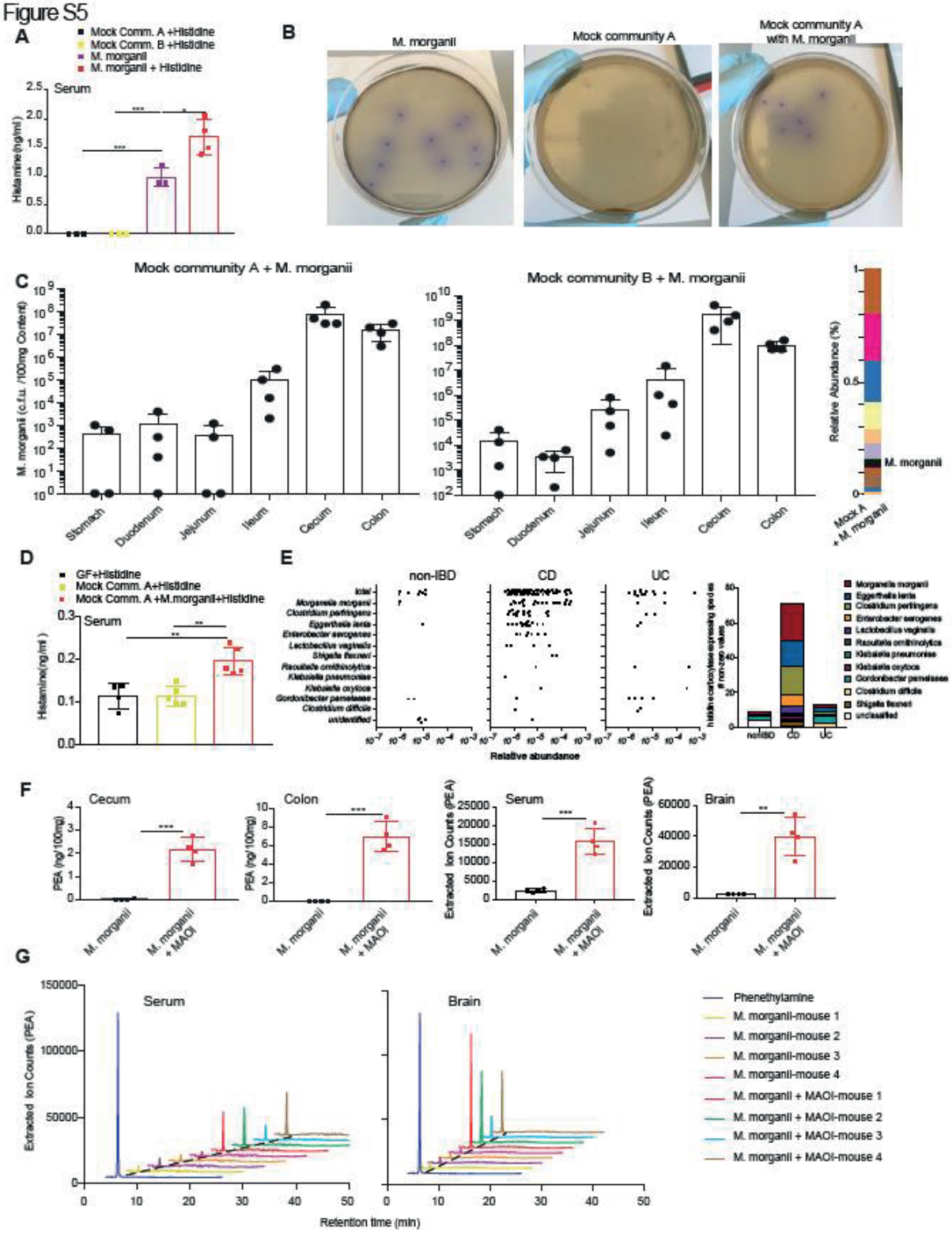

A
Mock Comm. A +Histidine
Mock Comm. B +Histidine
M. morganii
M. morganii + Histidine
Serum
Histamine (ng/ml)
B
M. morganii
Mock community A
Mock community A with M. morganii
C
Mock community A + M. morganii
Mock community B + M. morganii
M. morganii (c.f.u. /100mg Content)
Stomach
Duodenum
Jejunum
Ileum
Cecum
Colon
Relative Abundance (%)
M. morganii
Mock A + M. morganii
D
GF+Histidine
Mock Comm. A+Histidine
Mock Comm. A +M.morganii+Histidine
Serum
Histamine(ng/ml)
E
non-IBD
CD
UC
Relative abundance
nonIBD
CD
UC
F
Cecum
Colon
Serum
Brain
PEA (ng/100mg)
Extracted Ion Counts (PEA)
M. morganii
M. morganii + MAOI
G
Serum
Brain
Extracted Ion Counts (PEA)
Retention time (min)
Phenethylamine
M. morganii-mouse 1
M. morganii-mouse 2
M. morganii-mouse 3
M. morganii-mouse 4
M. morganii + MAOI-mouse 1
M. morganii + MAOI-mouse 2
M. morganii + MAOI-mouse 3
M. morganii + MAOI-mouse 4

Figure S5. *M. morganii* localization and production and accumulation of systemic phenethylamine *in vivo*, related to Figure 4.

(A) Groups of female germ-free C57Bl/6 mice were colonized with mock communities of 9 or 10 phylogenetically diverse human gut bacteria (Mock Community A or B) or monocolonized with *M. morganii* C135. Mice were fed a conventional diet with or without administration of 1% L-His *ad libitum* in the drinking water. Histamine concentrations in serum were measured via ELISA. n=3-5 mice per group.

(B-C) *M. morganii* primarily inhabits the cecum and colon. Groups of female germ-free C57Bl/6 mice were colonized with mock communities of 9 or 10 phylogenetically diverse gut microbes (Mock community A and B, respectively) with or without *M. morganii* C135. *M. morganii* CFUs can be distinguished from other bacteria based on their purple halos when plated on modified Niven's agar. Gastric, small intestinal, cecal and colonic contents from mice colonized with Mock communities A or B and *M. morganii* were plated on Modified Niven's agar to determine *M. morganii* colonization levels at various intestinal loci. Stacked barplot represents relative abundance of bacterial taxa in mice colonized with Mock community A plus *M. morganii* based on 16S rRNA gene sequencing (see also Table S3). n=4 mice per group.

(D) Groups of female germ-free C57Bl/6 mice were colonized with a mock community of 9 phylogenetically diverse human gut bacteria (Mock Community A) with or without *M. morganii* C135. Mice were fed a conventional diet and administered 1% L-His *ad libitum* in the drinking water. Histamine concentrations in serum were measured via ELISA. n=3-5 mice per group.

(E) Contribution of individual species to the relative abundance of histidine decarboxylase genes in the microbiomes of patients with IBD (CD and UC) as compared to healthy controls (non-IBD). Metagenomic data from longitudinal stool samples from IBD patients (publicly available from the Human Microbiome Project 2; iHMP) were analyzed for the presence and relative abundance of histidine decarboxylase genes (see methods for details). Data shown are a compilation of all data across multiple collection timepoints.

(F) Quantification of phenethylamine (PEA) in cecum, colon, serum, and brain from mice monocolonized with *M. morganii* C135 and treated with or without phenelzine (MAOI) via QQQ-MS/MS. n=4 mice per group.

(G) Accumulation of phenethylamine (PEA) in serum and brains of mice monocolonized with *M. morganii* C135 and treated with or without phenelzine (MAOI) as measured via QQQ-MS/MS. n=4 mice per group.

Data in all panels are representative of at least two independent experiments.

Data are presented as mean ± SEM. One-way ANOVA with Tukey's post-hoc test (A and E), *$p < 0.05$, ***$p < 0.001$.

Figure S6

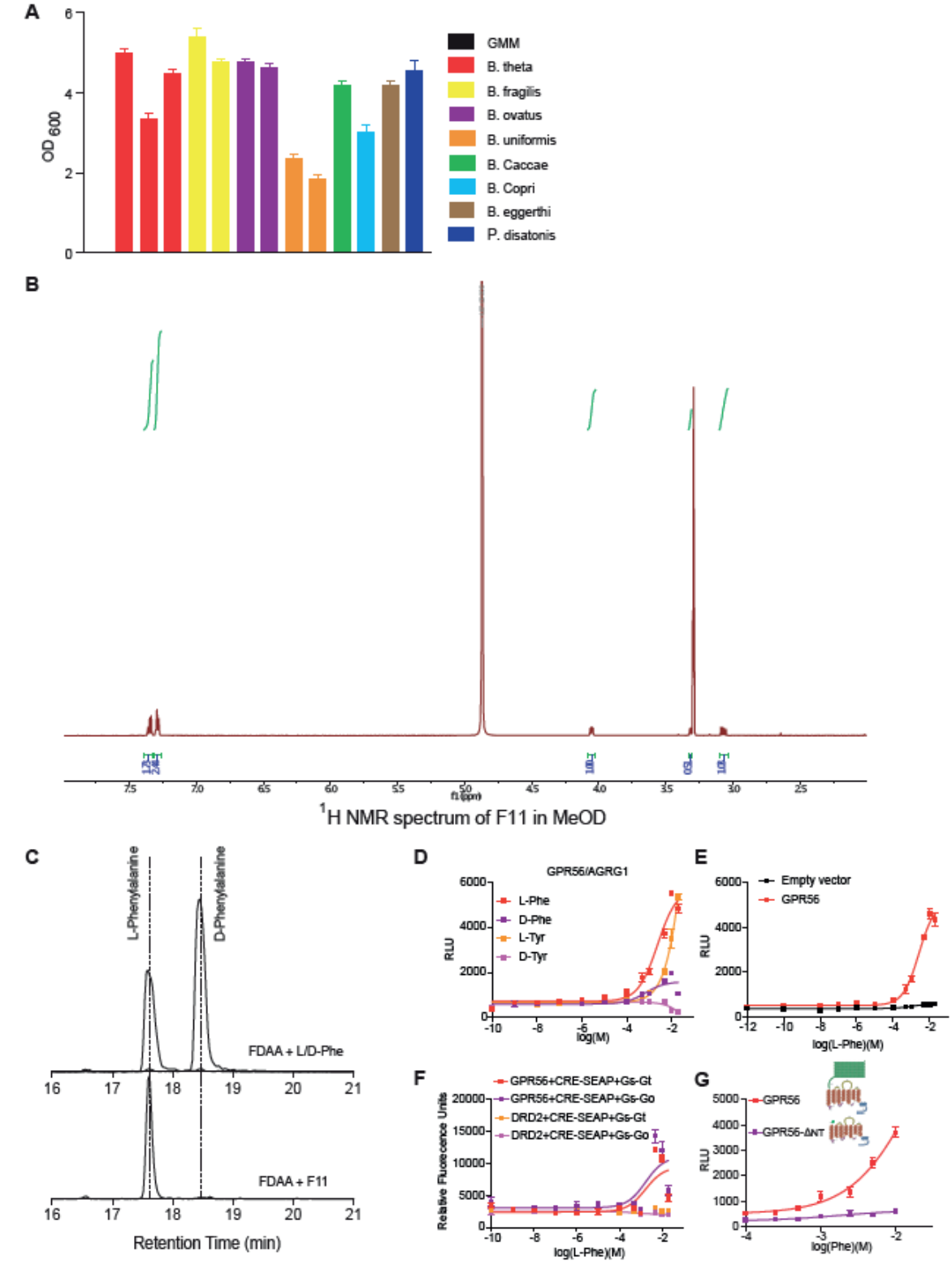

A
OD 600
GMM
B. theta
B. fragilis
B. ovatus
B. uniformis
B. Caccae
B. Copri
B. eggerthi
P. disatonis
B
^{1}H NMR spectrum of F11 in MeOD
C
L-Phenylalanine
D-Phenylalanine
FDAA + L/D-Phe
FDAA + F11
Retention Time (min)
D
GPR56/AGRG1
L-Phe
D-Phe
L-Tyr
D-Tyr
RLU
log(M)
E
Empty vector
GPR56
log(L-Phe)(M)
F
GPR56+CRE-SEAP+Gs-Gt
GPR56+CRE-SEAP+Gs-Go
DRD2+CRE-SEAP+Gs-Gt
DRD2+CRE-SEAP+Gs-Go
Relative Fluorescence Units
log(L-Phe)(M)
G
GPR56
GPR56-ΔNT
RLU
log(Phe)(M)

Figure S6. Effect of different bacterial and culture media on bacterial growth and GPR56/AGRG1 activation, structural characterization of *B. theta* C34 agonist L-Phe, and role of N-terminal domain in GPR56/AGRG1 activation by L-Phe, related to Figure 6.

(A) OD_{600} values of indicated *Bacteroides* and *Parabacteroides* strains cultured in gut microbiota medium (GMM) for 24 hours. n=3 replicates per isolate.

(B) ^{1}H NMR spectrum of active fraction 11 in MeOD revealed Phe as the major component.

(C) Advanced Marfey's analysis verified the stereochemistry of Phe in fraction 11 to be L-Phe. D-Phe in the active fraction was not detected. FDAA is 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (Marfey's Reagent).

(D) L-Phe and L-Tyr stereoselectively activate the orphan receptor GPR56/AGRG1. Activation of GPR56/AGRG1 by titrating doses of pure L-Phe, L-Tyr, D-Phe, and D-Tyr (in L-Phe and L-Tyr-free medium) was measured via GPR56-Tango. n=3 replicates per sample.

(E) L-Phe-induced Tango activation is GPR56/AGRG1-dependent. Luciferase expression (RLU) was measured after stimulation of cells transfected with GPR56-Tango or empty vector with titrating doses of L-Phe. n=3 replicates per sample.

(F) L-Phe-induced activation of G protein-dependent signaling in HEK cells is GPR56-dependent. Activation of G proteins downstream of GPR56/AGRG1 by L-Phe as measured by the CRE-SEAP assay. $G\alpha_s$-$G\alpha_t$ and $G\alpha_s$-$G\alpha_o$ chimeras were used to redirect GPR56/AGRG1 signaling to $G\alpha_s$ and enable use of the CRE-SEAP assay. Cells transfected with DRD2-Tango and $G\alpha_s$-$G\alpha_t$ and $G\alpha_s$-$G\alpha_o$ chimeras failed to respond to L-Phe. n=3 replicates per sample.

(G) The extracellular domain of GPR56/AGRG1 is indispensable for GPR56/AGRG1 activation by L-Phe. Activation of GPR56 or GPR56-ΔNT (a mutant lacking the extracellular domain) by titrating doses of L-Phe was measured via PRESTO-Tango. n=3 replicates per sample.

Figure S7

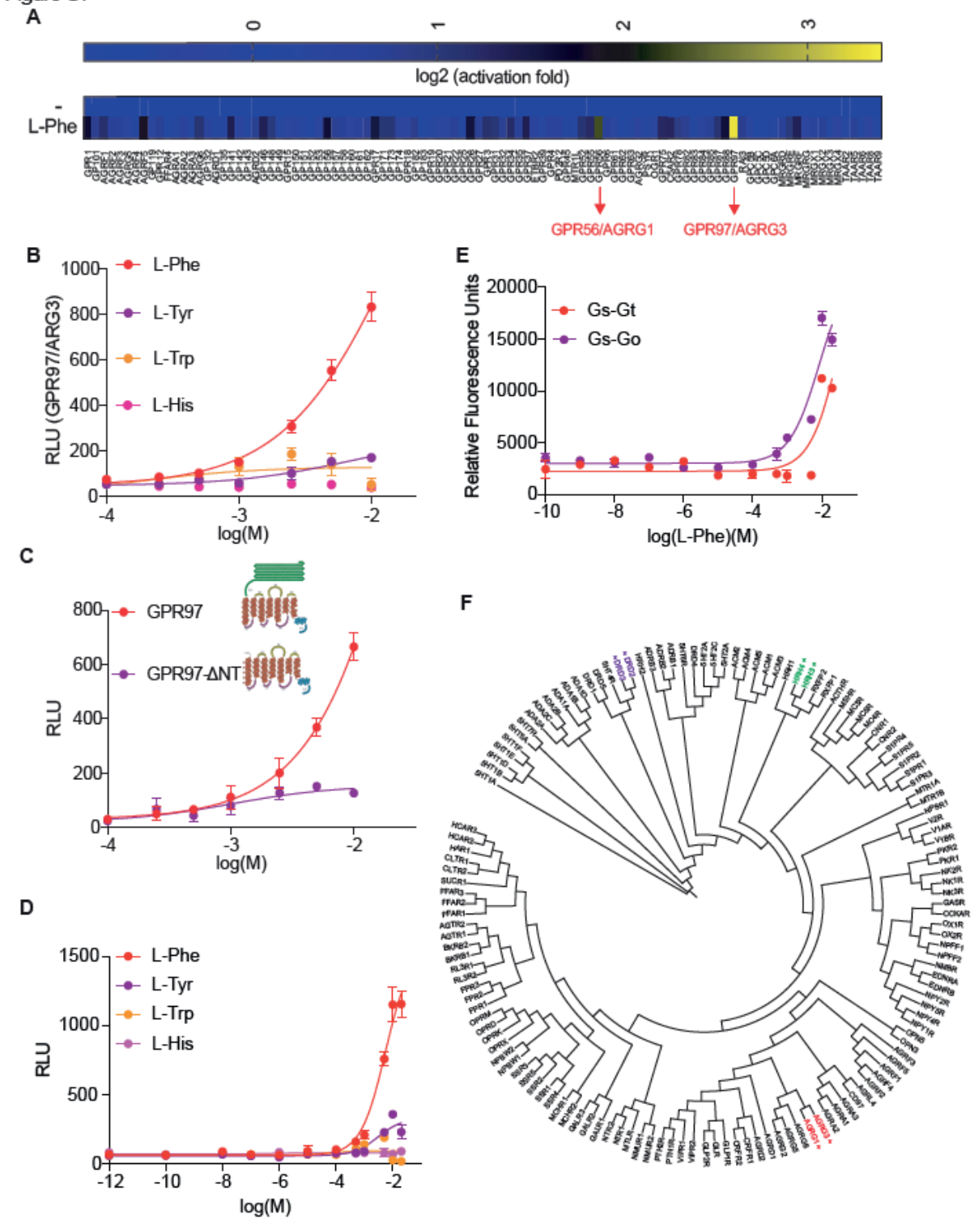

A
log2 (activation fold)
L-Phe
GPR56/AGRG1
GPR97/AGRG3
B
RLU (GPR97/ARG3)
L-Phe
L-Tyr
L-Trp
L-His
log(M)
C
GPR97
GPR97-ΔNT
RLU
log(M)
D
RLU
L-Phe
L-Tyr
L-Trp
L-His
log(M)
E
Relative Fluorescence Units
Gs-Gt
Gs-Go
log(L-Phe)(M)
F

Figure S7. L-Phe activates GPR97/AGRG3, a close relative of GPR56/AGRG1, related to Figure 6.

(A) L-Phe activates GPR56/AGRG1 and GPR97/AGRG3. Activation of all orphan, adhesion and other potential amino acid-sensing GPCRs by L-Phe was evaluated via PRESTO-Tango. n=3 replicates per sample.

(B) L-Phe specifically activates GPR97/AGRG3. Activation of GPR97/AGRG3 by titrating doses of L-Phe, L-Tyr, L-Trp, and L-His was measured via GPR97 PRESTO-Tango. n=3 replicates per sample.

(C) The extracellular domain of GPR97/AGRG3 is indispensable for GPR97/AGRG3 activation by L-Phe. Activation of GPR97 or GPR97-ΔNT (a mutant lacking the extracellular domain) by titrating doses of L-Phe was measured via PRESTO-Tango. n=3 replicates per sample.

(D) L-Phe specifically activates GPR97/AGRG3. Activation of GPR97/AGRG3 by titrating doses of L-Phe, L-Tyr, L-Trp, and L-His was measured via GPR97 PRESTO-Tango in media lacking L-Phe and L-Tyr. n=3 replicates per sample.

(E) L-Phe activates G protein-dependent signaling downstream of GPR97/AGRG3. Activation of G proteins downstream of GPR97/AGRG3 by L-Phe as measured by the CRE-SEAP assay. $G\alpha_s$-$G\alpha_t$ and $G\alpha_s$-$G\alpha_o$ chimeras were used to redirect GPR97/AGRG3 signaling to $G\alpha_s$ and enable use of the CRE-SEAP assay. n=3 replicates per sample.

(F) GPR56/AGRG1 and GPR97/AGRG3 are evolutionarily related. A phylogenetic tree for a subset of GPCRs, including all adhesion GPCRs, was constructed and visualized with equal branch lengths using gpcrdb.org, PHYLIP and jsPhyloSVG.

Data in all panels are representative of at least three independent experiments. Data are presented as mean ± SEM.

2.6 REFERENCES

1. Smith, P. H., MR; Panikov, N; Michaud, M; Gallini, CA; Bohlooly-Y, M; Glickman, JN; Garret, WS. The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science* **2013,** *341*, 569-573.
2. Yano, J. M.; Yu, K.; Donaldson, G. P.; Shastri, G. G.; Ann, P.; Ma, L.; Nagler, C. R.; Ismagilov, R. F.; Mazmanian, S. K.; Hsiao, E. Y. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **2015,** *161* (2), 264-76.
3. Tan, J. K. M., C.; Mariño,E; Macia, L and Mackay, CR. Metabolite-Sensing G Protein-Coupled Receptors-Facilitators of Diet-Related Immune Regulation. *Annu.Rev.Immunol* **2017,** *35*, 371-402.
4. Husted, A. S.; Trauelsen, M.; Rudenko, O.; Hjorth, S. A.; Schwartz, T. W. GPCR-Mediated Signaling of Metabolites. *Cell Metab* **2017,** *25* (4), 777-796.
5. Donia, M. S.; Fischbach, M. A. HUMAN MICROBIOTA. Small molecules from the human microbiota. *Science* **2015,** *349* (6246), 1254766.
6. Pedersen, H. K.; Gudmundsdottir, V.; Nielsen, H. B.; Hyotylainen, T.; Nielsen, T.; Jensen, B. A.; Forslund, K.; Hildebrand, F.; Prifti, E.; Falony, G.; Le Chatelier, E.; Levenez, F.; Dore, J.; Mattila, I.; Plichta, D. R.; Poho, P.; Hellgren, L. I.; Arumugam, M.; Sunagawa, S.; Vieira-Silva, S.; Jorgensen, T.; Holm, J. B.; Trost, K.; Meta, H. I. T. C.; Kristiansen, K.; Brix, S.; Raes, J.; Wang, J.; Hansen, T.; Bork, P.; Brunak, S.; Oresic, M.; Ehrlich, S. D.; Pedersen, O. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **2016,** *535* (7612), 376-81.
7. Perry, R. J.; Peng, L.; Barry, N. A.; Cline, G. W.; Zhang, D.; Cardone, R. L.; Petersen, K. F.; Kibbey, R. G.; Goodman, A. L.; Shulman, G. I. Acetate mediates a microbiome-brain-beta-cell axis to promote metabolic syndrome. *Nature* **2016,** *534* (7606), 213-7.
8. Fischbach, M. A. Microbiome: Focus on Causation and Mechanism. *Cell* **2018,** *174* (4), 785-790.
9. Dodd, D.; Spitzer, M. H.; Van Treuren, W.; Merrill, B. D.; Hryckowian, A. J.; Higginbottom, S. K.; Le, A.; Cowan, T. M.; Nolan, G. P.; Fischbach, M. A.; Sonnenburg, J. L. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* **2017,** *551* (7682), 648-652.
10. Guo, C. J.; Chang, F. Y.; Wyche, T. P.; Backus, K. M.; Acker, T. M.; Funabashi, M.; Taketani, M.; Donia, M. S.; Nayfach, S.; Pollard, K. S.; Craik, C. S.; Cravatt, B. F.; Clardy, J.; Voigt, C. A.; Fischbach, M. A. Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell* **2017,** *168* (3), 517-526 e18.
11. Haiser, H. J.; Gootenberg, D. B.; Chatman, K.; Sirasani, G.; Balskus, E. P.; Turnbaugh, P. J. Predicting and manipulating cardiac drug inactivation by the human gut bacterium Eggerthella lenta. *Science* **2013,** *341* (6143), 295-8.
12. Larsbrink, J.; Rogers, T. E.; Hemsworth, G. R.; McKee, L. S.; Tauzin, A. S.; Spadiut, O.; Klinter, S.; Pudlo, N. A.; Urs, K.; Koropatkin, N. M.; Creagh, A. L.; Haynes, C. A.; Kelly, A. G.; Cederholm, S. N.; Davies, G. J.; Martens, E. C.; Brumer, H. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **2014,** *506* (7489), 498-502.
13. Milshteyn, A.; Colosimo, D. A.; Brady, S. F. Accessing Bioactive Natural Products from the Human Microbiome. *Cell Host Microbe* **2018,** *23* (6), 725-736.
14. Wacker, D.; Stevens, R. C.; Roth, B. L. How Ligands Illuminate GPCR Molecular Pharmacology. *Cell* **2017,** *170* (3), 414-427.

15. Lei, W.; Ren, W.; Ohmoto, M.; Urban, J. F., Jr.; Matsumoto, I.; Margolskee, R. F.; Jiang, P. Activation of intestinal tuft cell-expressed Sucnr1 triggers type 2 immunity in the mouse small intestine. *Proc Natl Acad Sci U S A* **2018,** *115* (21), 5552-5557.
16. Nadjsombati, M. S.; McGinty, J. W.; Lyons-Cohen, M. R.; Jaffe, J. B.; DiPeso, L.; Schneider, C.; Miller, C. N.; Pollack, J. L.; Nagana Gowda, G. A.; Fontana, M. F.; Erle, D. J.; Anderson, M. S.; Locksley, R. M.; Raftery, D.; von Moltke, J. Detection of Succinate by Intestinal Tuft Cells Triggers a Type 2 Innate Immune Circuit. *Immunity* **2018,** *49* (1), 33-41 e7.
17. Schneider, C.; O'Leary, C. E.; von Moltke, J.; Liang, H. E.; Ang, Q. Y.; Turnbaugh, P. J.; Radhakrishnan, S.; Pellizzon, M.; Ma, A.; Locksley, R. M. A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal Remodeling. *Cell* **2018,** *174* (2), 271-284 e14.
18. Cohen, L. J.; Esterhazy, D.; Kim, S. H.; Lemetre, C.; Aguilar, R. R.; Gordon, E. A.; Pickard, A. J.; Cross, J. R.; Emiliano, A. B.; Han, S. M.; Chu, J.; Vila-Farres, X.; Kaplitt, J.; Rogoz, A.; Calle, P. Y.; Hunter, C.; Bitok, J. K.; Brady, S. F. Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature* **2017,** *549* (7670), 48-53.
19. Kroeze, W. K.; Sassano, M. F.; Huang, X. P.; Lansu, K.; McCorvy, J. D.; Giguere, P. M.; Sciaky, N.; Roth, B. L. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat Struct Mol Biol* **2015,** *22* (5), 362-9.
20. Barnea, G.; Strapps, W.; Herrada, G.; Berman, Y.; Ong, J.; Kloss, B.; Axel, R.; Lee, K. J. The genetic design of signaling cascades to record receptor activation. *Proc Natl Acad Sci U S A* **2008,** *105* (1), 64-9.
21. Palm, N. W.; de Zoete, M. R.; Cullen, T. W.; Barry, N. A.; Stefanowski, J.; Hao, L.; Degnan, P. H.; Hu, J.; Peter, I.; Zhang, W.; Ruggiero, E.; Cho, J. H.; Goodman, A. L.; Flavell, R. A. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **2014,** *158* (5), 1000-1010.
22. Goodman, A. L.; Kallstrom, G.; Faith, J. J.; Reyes, A.; Moore, A.; Dantas, G.; Gordon, J. I. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* **2011,** *108* (15), 6252-7.
23. Thurmond, R. L.; Gelfand, E. W.; Dunford, P. J. The role of histamine H1 and H4 receptors in allergic inflammation: the search for new antihistamines. *Nat Rev Drug Discov* **2008,** *7* (1), 41-53.
24. Albuquerque, E. X.; Pereira, E. F.; Alkondon, M.; Rogers, S. W. Mammalian nicotinic acetylcholine receptors: from structure to function. *Physiol Rev* **2009,** *89* (1), 73-120.
25. Beaulieu, J. M.; Gainetdinov, R. R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol Rev* **2011,** *63* (1), 182-217.
26. Özoğul, F. Production of biogenic amines by Morganella morganii, Klebsiella pneumoniae and Hafnia alvei using a rapid HPLC method. *European Food Research and Technology* **2004,** *219* (5), 465-469.
27. Kim, S. H.; Ben-Gigirey, B.; Barros-Velazquez, J.; Price, R. J.; An, H. Histamine and biogenic amine production by Morganella morganii isolated from temperature-abused albacore. *J Food Prot* **2000,** *63* (2), 244-51.
28. Oldendorf, W. Brain uptake of radiolabeled amino acids, amines, and hexoses after arterial injection. *American Journal of Physiology* **1971,** *221*, 1629-1639.
29. Durocher, Y.; Perret, S.; Thibaudeau, E.; Gaumond, M. H.; Kamen, A.; Stocco, R.; Abramovitz, M. A reporter gene assay for high-throughput screening of G-protein-coupled receptors stably or transiently expressed in HEK293 EBNA cells grown in suspension culture. *Anal Biochem* **2000,** *284* (2), 316-26.

30. Liberles, S. D.; Buck, L. B. A second class of chemosensory receptors in the olfactory epithelium. *Nature* **2006,** *442* (7103), 645-50.
31. Lovenberg, W. W., H & Udenfriend, S. Aromatic amino acid decarboxylase. *The Journal of Biological Chemistry* **1962,** *237*, 89-93.
32. Tannase, S. G., BM; Snell, EE. Purification and Properties of a Pyridoxal 5"Phosphate-dependent Histidine Decarboxylase from Morgunellu morganii AM- 15* *The Journal of Biological Chemistry* **1985,** *260*, 6738-6746.
33. Mavromatis, P. Q., PC. ModifiIcation of Niven's Medium for the Enumeration of Histamine-Forming Bacteria and Discussion of the Parameters Associated with Its Use. *Journal of Food Protection* **2002,** *65*, 546-551.
34. Eun, C. S.; Kwak, M. J.; Han, D. S.; Lee, A. R.; Park, D. I.; Yang, S. K.; Kim, Y. S.; Kim, J. F. Does the intestinal microbial community of Korean Crohn's disease patients differ from that of western patients? *BMC Gastroenterol* **2016,** *16*, 28.
35. Tyagi, P.; Mandal, M. B.; Mandal, S.; Patne, S. C.; Gangopadhyay, A. N. Pouch colon associated with anorectal malformations fails to show spontaneous contractions but responds to acetylcholine and histamine in vitro. *J Pediatr Surg* **2009,** *44* (11), 2156-62.
36. Kim, H.; Dwyer, L.; Song, J. H.; Martin-Cano, F. E.; Bahney, J.; Peri, L.; Britton, F. C.; Sanders, K. M.; Koh, S. D. Identification of histamine receptors and effects of histamine on murine and simian colonic excitability. *Neurogastroenterol Motil* **2011,** *23* (10), 949-e409.
37. Integrative, H. M. P. R. N. C. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe* **2014,** *16* (3), 276-89.
38. Smolinska, S.; Jutel, M.; Crameri, R.; O'Mahony, L. Histamine and gut mucosal immune regulation. *Allergy* **2014,** *69* (3), 273-81.
39. Glover, V. S., M; Owen, F; Riley, GJ. Dopamine is a monoamine oxidase B substrate in man. *Nature* **1977,** *265*, 80-81.
40. Fiedorowicz, J. S., KL. The Role of Monoamine Oxidase Inhibitors in Current Psychiatric Practice. *J Psychiatr Pract* **2004,** *10*, 239-248.
41. Purcell, R. H., RA. Adhesion G Protein–Coupled Receptors as Drug Targets. *Annu. Rev. Pharmacol. Toxicol.* **2018,** *58*, 429-49.
42. Bhushan, R.; Bruckner, H. Use of Marfey's reagent and analogs for chiral amino acid analysis: assessment and applications to natural products and biological systems. *J Chromatogr B Analyt Technol Biomed Life Sci* **2011,** *879* (29), 3148-61.
43. Kishore, A.; Purcell, R. H.; Nassiri-Toosi, Z.; Hall, R. A. Stalk-dependent and Stalk-independent Signaling by the Adhesion G Protein-coupled Receptors GPR56 (ADGRG1) and BAI1 (ADGRB1). *J Biol Chem* **2016,** *291* (7), 3385-94.
44. Lovitt, R. W.; Morris, J. G.; Kell, D. B. The growth and nutrition of Clostridium sporogenes NCIB 8053 in defined media. *J Appl Bacteriol* **1987,** *62* (1), 71-80.
45. Nicholson, J. K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S. Host-gut microbiota metabolic interactions. *Science* **2012,** *336* (6086), 1262-7.
46. Barcik, W.; Pugin, B.; Westermann, P.; Perez, N. R.; Ferstl, R.; Wawrzyniak, M.; Smolinska, S.; Jutel, M.; Hessel, E. M.; Michalovich, D.; Akdis, C. A.; Frei, R.; O'Mahony, L. Histamine-secreting microbes are increased in the gut of adult asthma patients. *J Allergy Clin Immunol* **2016,** *138* (5), 1491-1494 e7.

47. Borowsky, B.; Adham, N.; Jones, K. A.; Raddatz, R.; Artymyshyn, R.; Ogozalek, K. L.; Durkin, M. M.; Lakhlani, P. P.; Bonini, J. A.; Pathirana, S.; Boyle, N.; Pu, X.; Kouranova, E.; Lichtblau, H.; Ochoa, F. Y.; Branchek, T. A.; Gerald, C. Trace amines: identification of a family of mammalian G protein-coupled receptors. *Proc Natl Acad Sci U S A* **2001,** *98* (16), 8966-71.
48. Bunzow, J. S., MS; Arttamangkul, S; Harrison, LM; Zhang, G; Quigley, DI; Darland, T; Suchland, KL; Pasumanmula, S; Kennedy, JL; Olson, SB; Magenis, RE; Amara, SG; Grandy, DK. Amphetamine, 3,4-Methylenedioxymethamphetamine, Lysergic Acid Diethylamide, and Metabolites of the Catecholamine Neurotransmitters Are Agonists of a Rat Trace Amine Receptor. *Molecular Pharmacology* **2001,** *60*, 1181-1188.
49. Sotnikova, T. D.; Budygin, E. A.; Jones, S. R.; Dykstra, L. A.; Caron, M. G.; Gainetdinov, R. R. Dopamine transporter-dependent and -independent actions of trace amine beta-phenylethylamine. *J Neurochem* **2004,** *91* (2), 362-73.
50. Ramachandraih, C. S., N; Bar, KJ; Baker, G; Yeragani, VK. Antidepressants: From MAOIs to SSRIs and more. *Indian J Psychiatry* **2011,** *53*, 180-182.
51. Riederer, P.; Laux, G. MAO-inhibitors in Parkinson's Disease. *Exp Neurobiol* **2011,** *20* (1), 1-17.
52. Irsfeld, M.; Spadafore, M.; Pruss, B. M. beta-phenylethylamine, a small molecule with a large impact. *Webmedcentral* **2013,** *4* (9).
53. Amisten, S.; Salehi, A.; Rorsman, P.; Jones, P. M.; Persaud, S. J. An atlas and functional analysis of G-protein coupled receptors in human islets of Langerhans. *Pharmacol Ther* **2013,** *139* (3), 359-91.
54. Duner, P.; Al-Amily, I. M.; Soni, A.; Asplund, O.; Safi, F.; Storm, P.; Groop, L.; Amisten, S.; Salehi, A. Adhesion G Protein-Coupled Receptor G1 (ADGRG1/GPR56) and Pancreatic beta-Cell Function. *J Clin Endocrinol Metab* **2016,** *101* (12), 4637-4645.
55. Adibi, S. M., DW. Protein Digestion in Human Intestine as Reflected in Luminal, Mucosal, and Plasma Amino Acid Concentrations after Meals. *The Journal of Clinical Investigation* **1973,** *52*, 1586-1594.
56. Williams, R. M., CDS; Burnett, JR. Phenylketonuria- An Inborn Error of Phenylalanine Metabolism. *Clin Biochem Rew* **2008,** *29*, 31-41.
57. Pezeshki, A.; Zapata, R. C.; Singh, A.; Yee, N. J.; Chelikani, P. K. Low protein diets produce divergent effects on energy balance. *Scientific Reports* **2016,** *6* (1).
58. Jutel, M.; Akdis, M.; Akdis, C. A. Histamine, histamine receptors and their role in immune pathology. *Clin Exp Allergy* **2009,** *39* (12), 1786-800.
59. Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello, E. K.; Fierer, N.; Pena, A. G.; Goodrich, J. K.; Gordon, J. I.; Huttley, G. A.; Kelley, S. T.; Knights, D.; Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.; Pirrung, M.; Reeder, J.; Sevinsky, J. R.; Turnbaugh, P. J.; Walters, W. A.; Widmann, J.; Yatsunenko, T.; Zaneveld, J.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **2010,** *7* (5), 335-6.
60. Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014,** *30* (15), 2114-20.
61. Nurk, S.; Bankevich, A.; Antipov, D.; Gurevich, A. A.; Korobeynikov, A.; Lapidus, A.; Prjibelski, A. D.; Pyshkin, A.; Sirotkin, A.; Sirotkin, Y.; Stepanauskas, R.; Clingenpeel, S. R.; Woyke, T.; McLean, J. S.; Lasken, R.; Tesler, G.; Alekseyev, M. A.; Pevzner, P. A. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol* **2013,** *20* (10), 714-37.
62. Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R.

A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **2008,** *9*, 75.

Chapter 3

Escherichia coli produce indole-functionalized metabolites under acid stress conditions

3.1 Introduction

Escherichia coli is a prevalent Gram-negative bacterium that resides in greater than 90% of the human population. Though present at less than 1% of the total gut flora, *E. coli* is one of the most studied bacteria for its ubiquitous presence and metabolite production.[62] A common biomarker for the presence of *E. coli* and other microbial strains is the production of indole. Microbial-derived indole acts as an inter- and intra-cellular signal crucial in diverse bacterial physiologies in virulence, biofilm formation, antibiotic stress, acid resistance, and so on. In fact, indole is omnipresent in gram-positive and negative bacteria and can serve as an indicator for growth of certain microbial strains. Indole is hydrolyzed from dietary tryptophan via tryptophanase gene (tnaA) and can accumulate up to millimolar concentrations in stationary phase in *E. coli* and other commensals.

Microbial indole has been shown to inhibit the expression of Locus of Enterocyte Effacement (LEE) pathogenicity gene. LEE gene cluster is a characteristic virulence factor for pathogenic strains such as enterohemorrhagic E. coli (EHEC) and encodes for the protein machinery essential in the assembly of type III secretion systems and injecting bacterial effectors to the host membrane. Kumar *et al.* has demonstrated that a histidine kinase cpxA in non-indole producing pathogens senses indole concentrations in the intestinal epithelium which represses the expression of virulent LEE gene cluster. This example showcases indole as an intercellular signal that is utilized by indole-producing and non-producing gut bacteria to define and colonize their niches.[63]

Indole is also known to regulate antibiotic resistance and consequently persister cell formation in *E. coli*. Upon induction with antibiotics such as rhodamine 6G and SDS, microbial indole upregulates the expression of multidrug efflux pump genes acrD and mdtABC through BaeSR and CpxAR, two component

regulatory sensors involved in stress signaling.[64] Another study by Weatherspoon-Griffin proposed a different multidrug transporting cascade through mar, acrAB and tolC transcription by indole-sensing cpxAR system as a resistance response to antimicrobial peptide protamine.[65] In fact, mutants producing indole were demonstrated to confer population-wide antibiotic shielding to less resistant strains in a continuous *E. coli* cultures treated with gentamicin or norfloxacin.[66]

In response to acidic environments, *E. coli* expresses four inducible acid resistance systems (ARs) as a compensation for survival.[67] The first system AR1 is expressed in the presence of RpoS and cyclic AMP receptor protein (CRP) under mildly acidic pHs (5-6).[68] Its structural components or mechanism are still unclear and need further elucidation. The second system AR2 utilizes glutamate decarboxylase (GadA/B) to decarboxylate glutamate to reduce intracellular proton, while the putative glutamate: GABA antiporter GadC exchanges extracellular glutamate with GABA. In this way, AR2 increases the intracellular pH for neutralization of acid stress and thereby, offer the best protection under pH 3.8 amongst all acid resistance mechanisms.[68-70] The third system, arginine decarboxylase (AdiA) and its antiporter AdiC, is only expressed in the presence of high amount of exogenous arginine and work similarly as AR2 by decarboxylating arginine for neutralization of acidic pH.[68, 71] The fourth system involves acid-inducible lysine decarboxylase (CadAB) in phosphate-starved cells.[72]

While the transcriptional and mechanistic changes under acidic stress and other stress responses have been intensively characterized, the characterization of the stress metabolome is quite lacking. In fact, there have been some studies that show that the production of novel metabolites under stress conditions can invoke distinct host immune activities and mediate bacterial cellular stress signaling. For example, a study by Kim *et al* described indole-derived metabolites called indolokines under cellular redox stress in *E. coli*. These indole stress signals were reported to be derived from indole-3-pyruvate in response to transaminases AspC and TyrB and found to exert persister cell formation, AHR signaling and IL-6 activation. In this study, we utilize acidic stress as an elicitor to uncover previously unknown metabolites that could potentially act as stress response signals under inflammation-mimetic conditions. We report a family of

indole metabolites that oligomerize under acid stress conditions and serve as agonists of the G-protein coupled receptor (GPCR) GPR84, a receptor that is known to regulate cytokine transcription, such as interleukin-12.[73]

3.2 Results and Discussion

3.2.1 Eliciting indole-functionalized Metabolites in *E. coli under* acidic stress

To assess previously uncharacterized stress metabolites in *E. coli*, BW25113 cultures were grown under acidic condition (pH~5) and extracted wildtype strain BW25113 with ethyl acetate. The extracts were then analyzed the organic extracts using single quadrupole liquid chromatography mass spectrometry (LC/MS) coupled with a photodiode array detector (**Figure 1A**). UV analysis indicated a family of indole-related ions that are upregulated in the presence of acid. Further analysis using high resolution mass spectrometry indicated that m/z 118.0651 is indeed indole and that four other ions are indole-derived metabolites (**Supp Fig 1**). Characterization by tandem MS/MS and NMR spectroscopy elucidated the upregulated ions to be indole oligomers **1-5.** Co-injections with commercially available standards (**1,2,4**) and synthesized standards[74, 75] (**3, 5**) validated the structures (**Figure 1B**). To further uncover the potential monomers and oxidized indole analogues, commercial standards **6-10** were purchased and coinjected with the acidified cultures, which led to the additional detection of **6** and **10** in the cultures.

3.2.2 Acid Stress Cultivation and Detection

Commensal bacteria often encounter mildly acidic pH as low as pH 5.3 under inflammatory conditions in the gut[76-78] or as low as pH ~2 when passing through the stomach lumen[78, 79]. We wondered whether a range of mildly acidic pHs can affect the indole polymerization and detection, therefore, we tested the relative production levels of indole oligomers under pH 4-7 conditions. Culture medium containing LB (Luria Broth) was conditioned to mimic acidic conditions in a pH 4-7 range using disodium

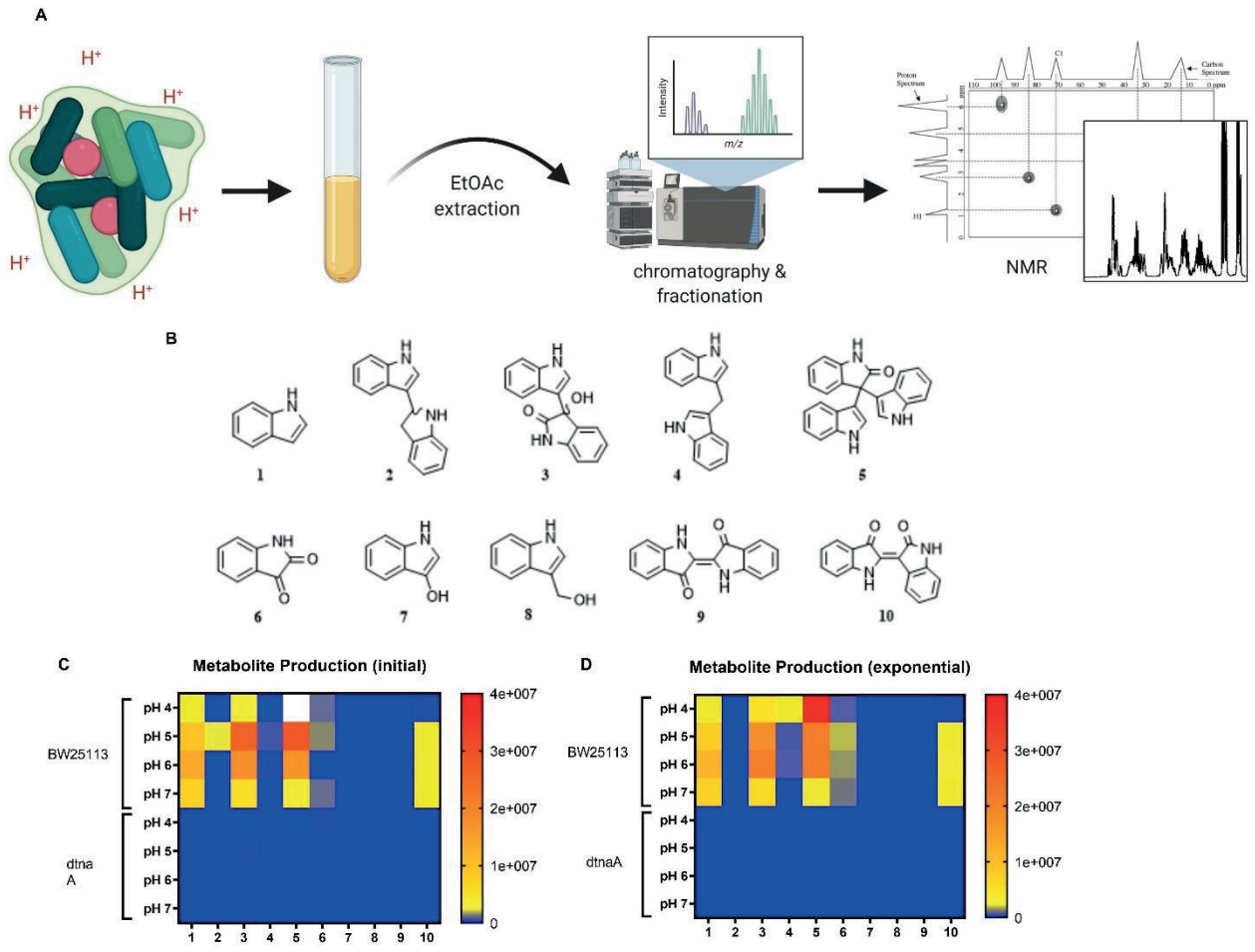


Figure 1. (A) Workflow for characterization of indole metabolites in *E. coli* under acid stress (B) Indole metabolites characterized and/or detected from *E. coli* (C, D) Relative metabolite production levels in E. coli WT BW25113 and Δ*tnaA* strains at pH 4-5 stressed either at initial phase or exponential phase.

phosphate and citric acid with a buffer capacity of >100mM to control for pH (**Table S1**). *E. coli* BW25113 was cultured and stressed with acidic shock pH 4-7 at either initial or exponential phase to capture various acidic stress environments for two days. The cultures were then extracted with ethyl acetate and subjected to pH, growth (OD_{600}) and untargeted metabolomic analyses. All strains grew well under varying pH levels except for pH 4, while the pre- and post- acid induction pH levels remain constant, showing that the pH was well controlled (**Figure S3**). Relative production levels of indole oligomers **1-10** under varying pH conditions are depicted in **Figure 1C, D**. Since free indole is produced in *E. coli* via a tryptophase (*tnaA*),

mutant Δ*tnaA* cultures were prepared and analyzed in parallel under identical conditions. Δ*tnaA* cultures abolished production of **1-10**, supporting that the family is derived from free indole. We also analyzed cell free controls of the pH-conditioned media supplemented 1 mg/mL indole, which supports the spontaneous oligomerization of 1-10 from free indole (**Figure S4**).

3.2.3 Detection of indole polymerization in C10

To establish whether similar products were produced in other indole producers, we analyzed Enterobacteriaceae UC/UC strain C10 isolated from an Inflammatory Bowel Disease (IBD) patient. Using the workflow for *E. coli* BW25113, we observed a similar production profile in the IBD isolate (Figure 2, Supporting Figure 5). These studies suggest that indole oligomerization products are upregulated under acid stress growth conditions.

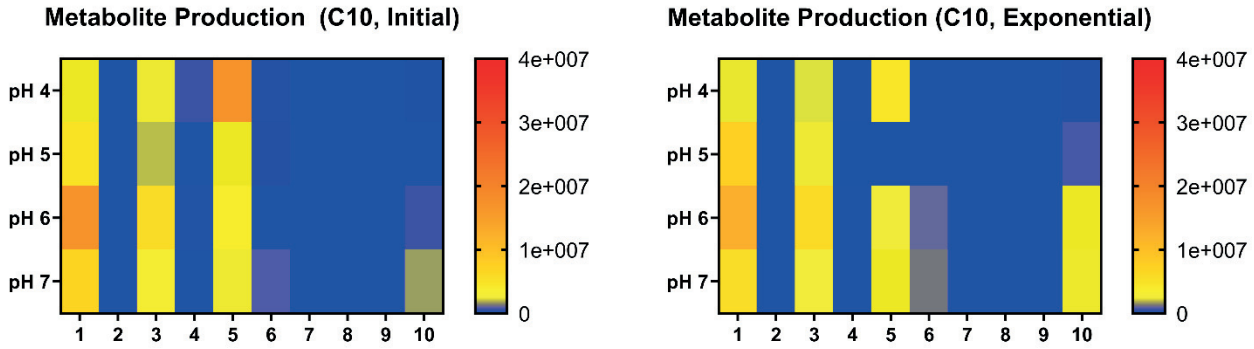


Figure 2. Indole metabolite production levels in Enterobacteriaceae UC/UC strain C10 under pH 4-7 stressed at initial or exponential phase.

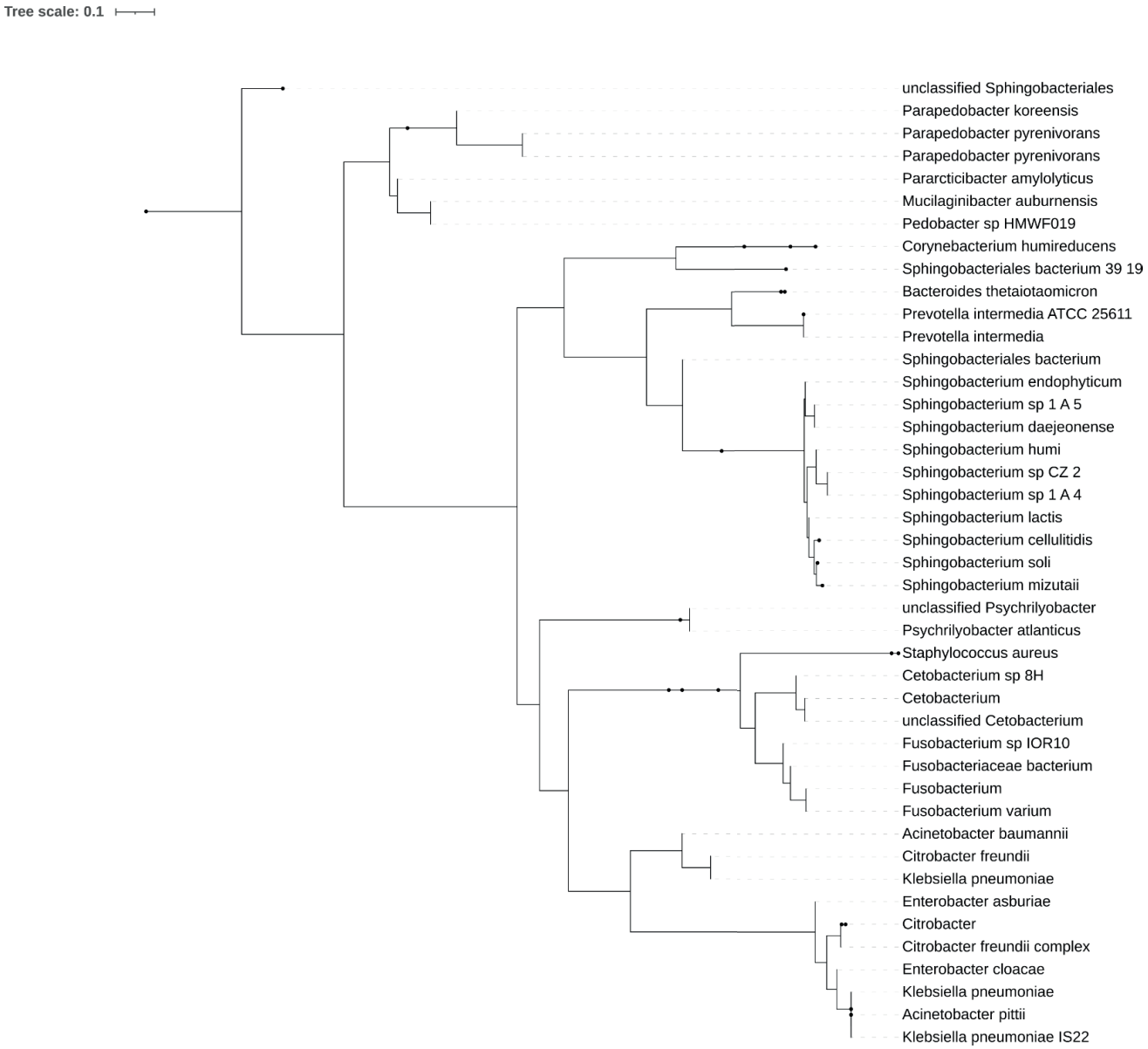


Figure 3. Homologs of *tnaA* gene in intratumor microbiome strains reported in the literature. *E. coli* tryptophanase tnaA protein sequence was blasted against select tumor microbiome strains. The strains with tnaA identity >=40% are constructed in a rooted phylogenetic tree with NGphylogeny and iTOL.

3.2.4 Presence of tryptophanse (*tnaA*) gene in tumor strains

Indole prevalence is an indicator of bacterial growth, in fact, a summary of the prevalence of tnaA gene responsible for indole production in different strains has been reported in this review.[80] Indole functionalization occurs readily at pH 5-7, which is a range representative of mildly acidic tumor microenvironments[81]. Tumor microenvironments are colonized with diverse microbes.[82-85] A comprehensive study on the tumor microbiome from seven solid cancer types was published recently by Nejman and coworkers.[84] Because bacteria can effect tumor progression, we analyzed *tnaA* homologs in the reported strains from this study and in other studies[82, 84, 86]. Tryptophanase is well represented in the tumor microbiome strains (**Figure 3**). This protein alignment data suggest that indole production is relatively ubiquitous in the tumor microbiome and that, under acidic tumor pHs, oligomerization would be expected to occur more readily.

3.2.5 Activation of orphan GPCR84

Diindolylmethane (DIM, **4**) was previously reported to activate G-protein coupled receptor GPR84 through β-arrestin and G_i signaling in an allosteric manner.[87, 88] GPR84 is a medium chain fatty acid (MCFA) receptor that is primarily expressed in the bone marrow and dendritic cells with low expression in gastrointestinal tissues.[89] GPR84 expression is integral for modulation of proinflammatory cytokines and phagocytosis in macrophages. Endogenous ligands (MCFAs) and DIM **4** were reported to upregulate the mRNA expression of interleukin 12 (IL-12) p40 subunit in LPS-induced macrophage cells through GPR84 activation.[73] GPR84 activation by MCFAs was reported to be primarily through G-protein subtype, G_i-G_o signaling.[73] Due to structural similarities to GPR84 agonist **4**, we analyzed metabolites **1-10** and another positive control decanoic acid at 10 μM for GPR84/β-arrestin2 activity using the PRESTO-Tango assays in a dose dependent manner (**Figure 4**) [90]. From these studies, we established carbinol **8** as a new GPR84 agonist through the β-arrestin2 signaling pathway. We also investigated their G-protein signaling activities (**Figure S6**). Indole-3-carbinol and DIM are commonly found in cruciferous vegetables and have been

shown to exert multiple anticarcinogenic properties.[91-94] This study reports the activation of microbial derived DIM **4** as an immunometabolic GPR84 agonist.

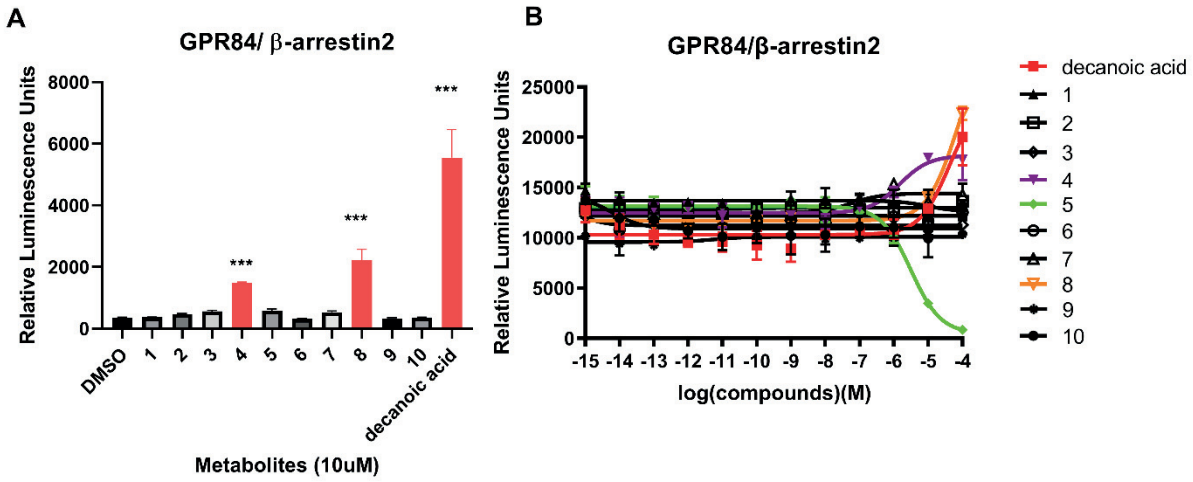


Figure 4. (A) Screening metabolites **1-10** (10 μM) against GPR84/β-arrestin2 activity. Decanoic acid is included as a positive control and DMSO was used as a solvent vehicle negative control. (B) Dose dependent activation of metabolites **1-10** against GPR84/β-arrestin2.

Although *E. coli*'s primary site is in the colon where the pH is neutral, acidic conditions can occur under dysbiotic and inflammatory conditions.[78] Indole derivatives have been reported to exhibit a wealth of biological activities, as summarized in **Table 1**. In fact, several studies have demonstrated antimicrobial activities of metabolites **5**, **6**, and **8**.[95-99] It is also noteworthy that most of these derivatives have anticancer, antioxidant and anti-inflammatory properties through diverse mechanisms of action.[99-101] Collectively, the data suggest that indole producers may take advantage of the spontaneous formation of indole oligomers, especially under acidic conditions, to inhibit other microbes in their microenvironment.

Table 1. Brief overview of reported biological activities of indole metabolites **1-10** in literature

Compounds	Reported Functions	Antimicrobial activity	
1	downregulates bacterial virulence, regulates acid resistance, induces bacterial drug resistance, biofilm formation, quorum sensing signal, AHR ligand		63 64, 80, 102, 103
2	androgen receptor inhibitor		104
3	N/A		
4	GPR84 (IBD, immunostimulatory), radical scavenger, antioxidant, antiangiogenic, androgen receptor antagonist, antitumor, organ transplant rejection, AHR ligand		87, 91-94, 105
5	cytotoxic to Multidrug resistant (MDR) cancer cells	*E.coli, B. subtilis, S. aureus, C.albicans, T. beigelii, S. cerevisiae*	96, 97, 99
6	endogenously produced in CNS; natriuretic peptide receptor type A antagonist	*S. aureus, S. epidermis*	95, 106
7	Antioxidant		100
8	AHR dependent anti-inflammatory response; promote T_{reg} cells while downregulating Th17; antiangiogenic	*MRSA, VRE, VSE, MREC, MRPA*	91, 98, 105, 107
9	AHR ligand, potent induction of IL-10, IL-22		105, 108
10	antiproliferative - inhibition of cell cycle-related kinase, AHR ligand		101, 105, 109

3.3 Experimental Section

General chemical analysis:

Fractionation and isolation of metabolites were performed using an Agilent Prepstar HPLC system with the following columns: Agilent Polaris C18-A 5µm (21.2 x 250mm^2), Agilent Phenyl-Hexyl 5µm (9.4x250mm^2) and Phenomenex Luna C18 (100Å) 10µm (10.0 x 250mm^2). An Agilent 1260 Infinity HPLC with an Agilent 6120 Quadrupole low-resolution Electro Spray Ionization (ESI) mass spectrometer was used to obtain general Liquid Chromatograph Mass Spectrometry (LCMS) data. High-resolution mass spectra were obtained using an Agilent iFunnel 6550 Quadrupole Time-of-Flight (QTOF) instrument coupled to an ESI source.

Characterization of indole metabolites:

E. coli wild type BW25113 (5mL) culture was grown overnight in LB at 37°C. Upon extraction with ethyl acetate and chromatographic analysis of the culture, a family of indole representing peaks were detected at 280nm. To characterize these compounds, 1L culture of BW25113 was incubated overnight in a 37°C in LB medium. The culture was extracted with ethyl acetate, dried and fractionated with Agilent Prepstar HPLC system using Phenomenex Luna C18 (100Å) 10µm (10.0 x 250mm^2) and Agilent Phenyl-Hexyl 5µm (9.4x250mm^2).

Acid stress cultivation conditions:

To condition the acidity of the bacterial cultures, pH 4-7 were adjusted using 0.2M Na_2HPO_4 and 0.1M citric acid buffers accordingly (Supp table 1). Luria Broth (LB) powder was then added to each pH-preadjusted buffer and autoclaved to achieve the desired sterile medium. *E. coli* wild type BW25113 and its Δ*tnaA* knockout strain (from Keio collection at the Yale Coli Genetic Stock Center) were inoculated into pH 4-7 media at initial phase. Another set of BW25113 and Δ*tnaA* strains were grown in normal LB to reach exponential phase, when the cultures were spun down, supernatant decanted and replaced with pH4-7 conditioned LB media for acid stress accordingly. The decanted supernatants were stored and

extracted later with ethyl acetate. All cultures were grown aerobically at 37°C for two days in triplicates. LB and LB containing 1mg/mL indole were also included as cell free controls. pH and OD measurements were taken before and after acid stress. The cultures were then extracted with ethyl acetate, dried and reconstituted in MeOH. The extracts were then analyzed on HR-ESI-QTOF-MS on a Phenomenex Kinetex C18 (100Å) 5µm (4.6 x 250mm^2) using 10-100% acetonitrile in water with 0.1% formic acid for 30min. A clinical isolate C10 strain from UC/UC patient was also pH-conditioned, cultured and analyzed with the same procedure.

Chemical Synthesis of 3 and 5:

Compound **3** was synthesized as previously described.[75] The crude product was purified with Agilent Prepstar HPLC system with a Phenomenex Luna C18 (100Å) 10µm (10.0 x 250mm^2) using 10-100% acetonitrile in water (with no acid) for 30min. Compound **5** was also synthesized as described with minor modifications.[74] Upon reaction completion (~1hr), reaction mixture was extracted with 100mL ethyl acetate: water (1:1 ratio). The organic layer was washed with saturated sodium chloride and dried with anhydrous sodium sulfate. Purification of the crude product was performed with Agilent Prepstar HPLC system with a Phenomenex Luna C18 (100Å) 10µm (10.0 x 250mm^2) using 40-100% acetonitrile in water (with no acid) for 30min.

G-Protein Coupled Receptor (GPCR) Assay

Ten indole metabolites (10uM) were tested for GPR84/β-arrestin2 activity using the PRESTO-Tango assay.[90] All compounds were dissolved in DMSO and tested in eight replicates. Endogenous ligand decanoic acid was included as a positive control, and DMSO was included as a negative control. Among the ten indole metabolites, carbinol and diindolylmethane (DIM) were shown to have GPR84/β-arrestin2 activity. For dose dependent curves of β-arrestin2 activity, metabolites and the corresponding controls were serially diluted and tested for luciferase activity using the same assay. GPR84 G-protein signaling was investigated using CRE-SEAP plasmids for $G_{\alpha s}$, $G_{\alpha i}$, $G_{\alpha q}$, $G_{\alpha 12}$ response.[110]

3.4 Supplementary Information

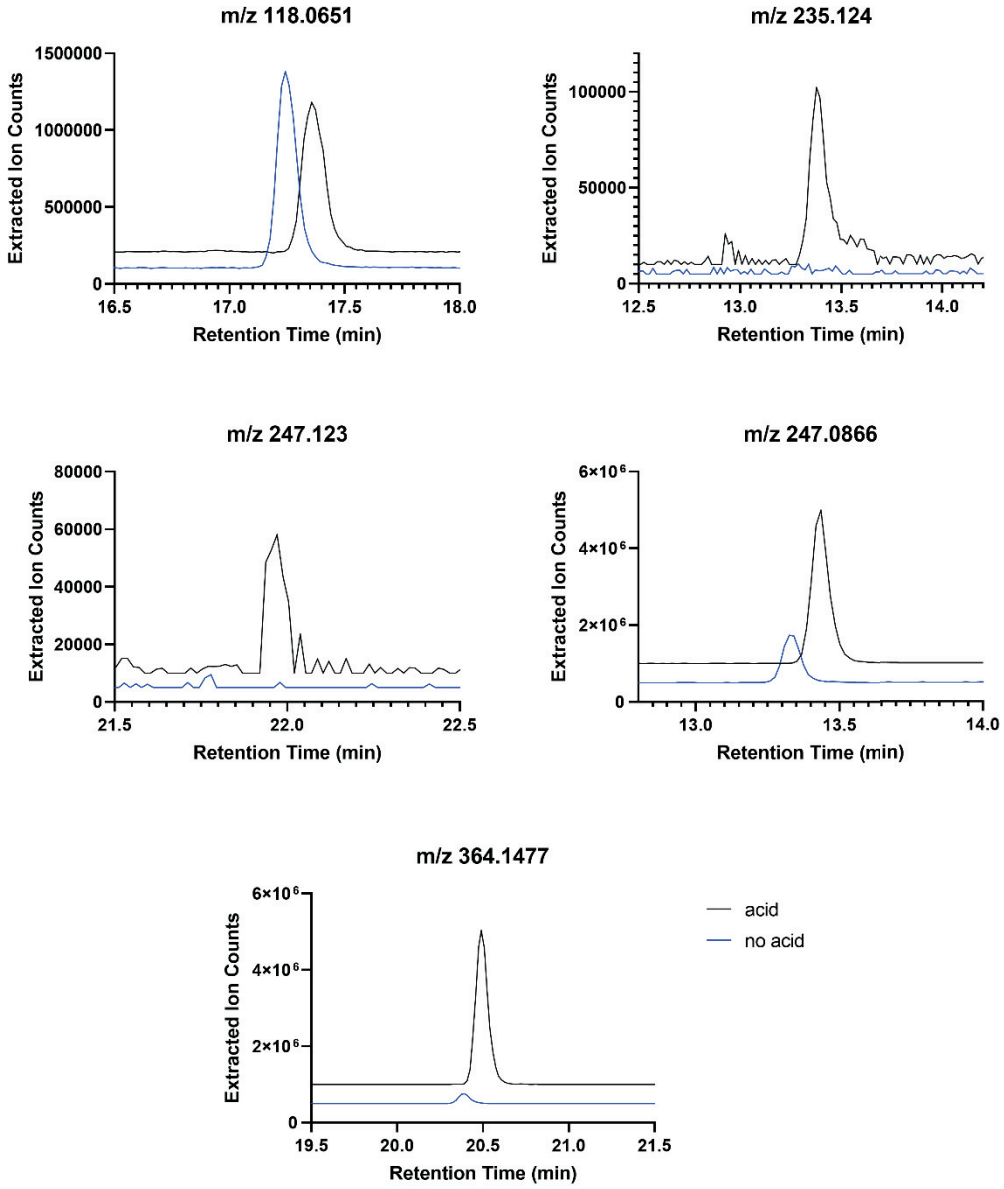


Figure S1. Upregulation of ions in *E. coli* BW25113 cultures as elicited by acidic stress

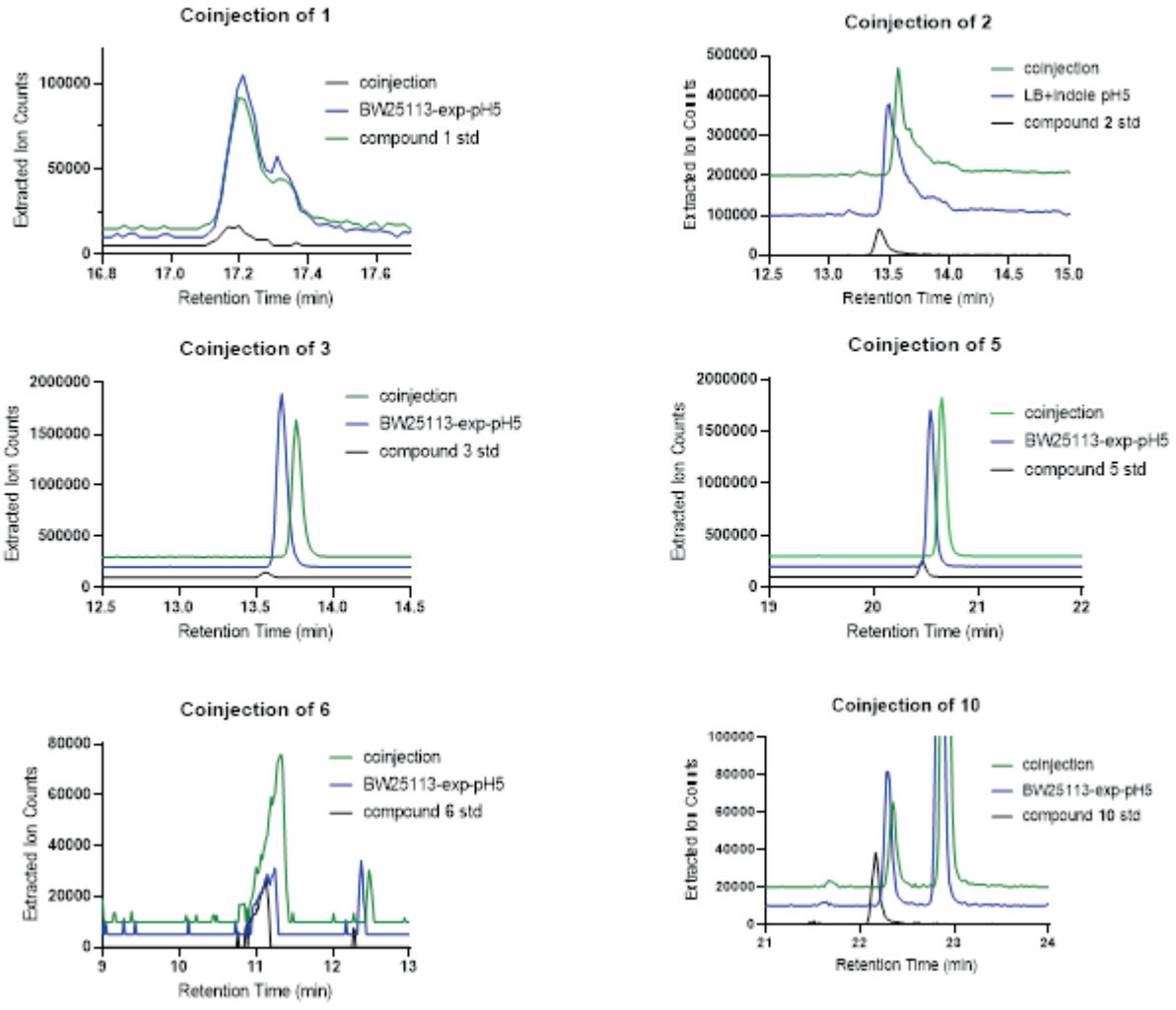


Figure S2. Coinjections of select metabolites **1-10** that are present in *E. coli* extracts under acidic stress

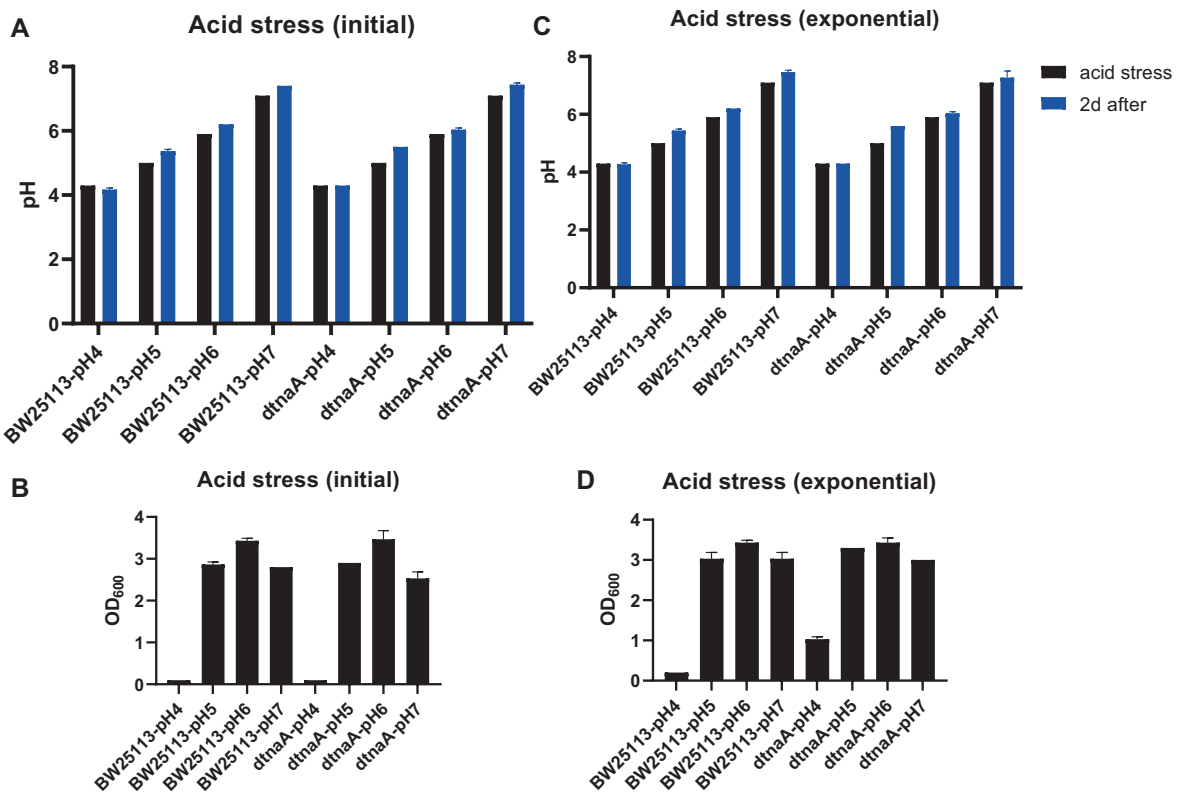


Figure S3. A,C show the pH levels of *BW25113* and *ΔtnaA* cultures before acid stress with conditioned medium (black) and two days after (blue) (A. acid stressed during initial phase; C. acid stressed during exponential phase), B, D show OD_{600} values to monitor the growth of the BW25113 and dtnaA cultures in pH4-7 (B. acid stressed during initial phase; D. acid stressed during exponential phase)

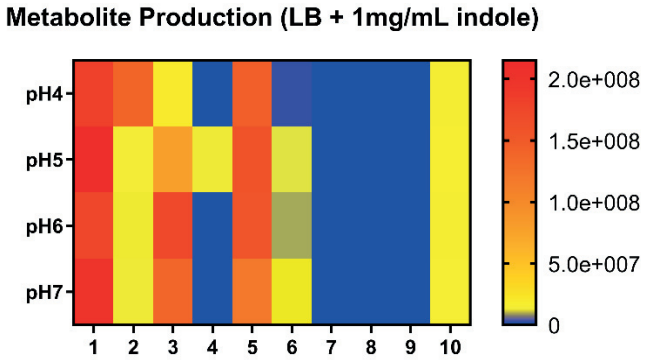


Figure S4. Production of indole metabolites in LB+1mg/mL indole as a cell-free control

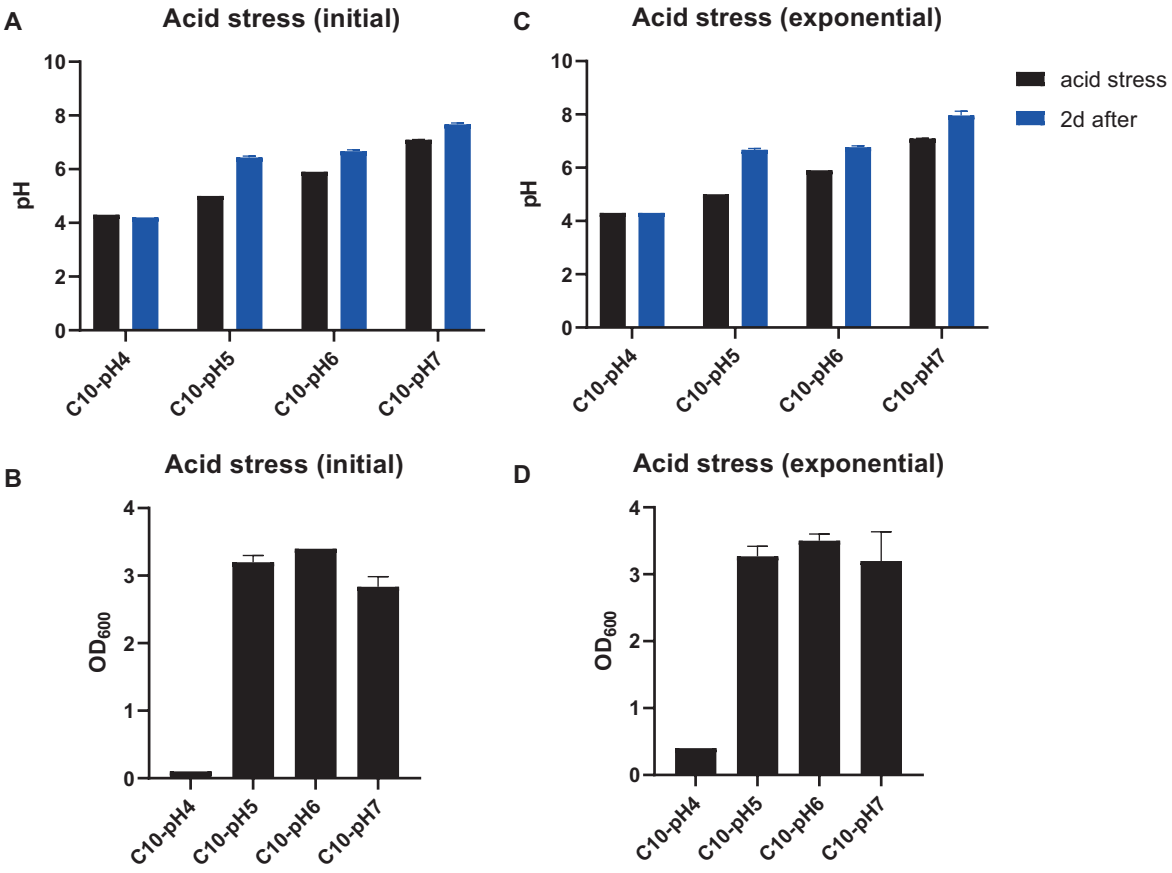


Figure S5. A,C show the pH levels of C10 cultures before acid stress with conditioned medium (black) and two days after (blue) (A. acid stressed during initial phase; C. acid stressed during exponential phase), B, D show OD_{600} values to monitor the growth of the C10 cultures in pH4-7 (B. acid stressed during initial phase; D. acid stressed during exponential phase)

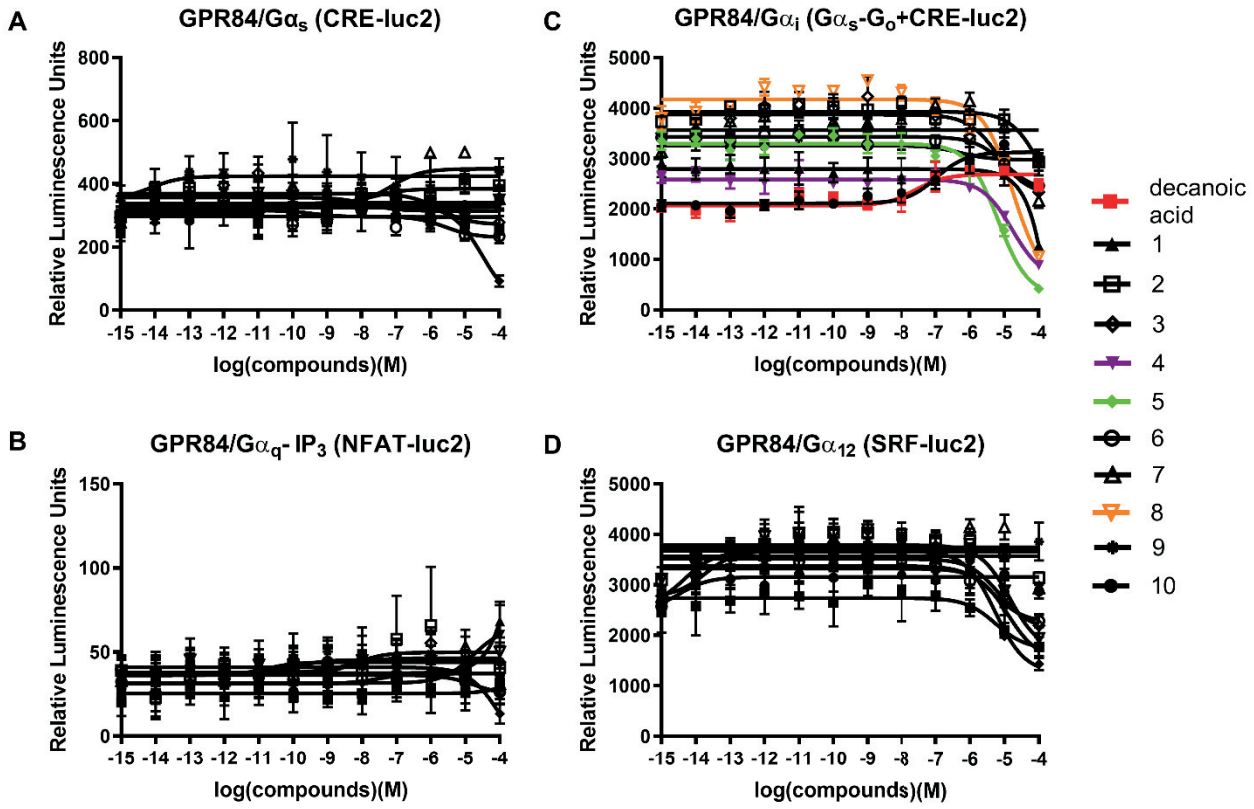


Figure S6. Dose dependent G-protein signaling activities of indole metabolites **1-10**

Table S1. Buffer composition for controlling specific pH (4-7) for bacterial cultures.

0.2M Na2HPO4		0.1M citric acid		pH
mL	mM	mL	mM	
270.2	77	429.8	61	4.3
359.8	103	340.2	49	5
449.4	128	250.6	36	5.9
610.4	174	91	13	7.1

3.5 References

1. Smith, P. H., MR; Panikov, N; Michaud, M; Gallini, CA; Bohlooly-Y, M; Glickman, JN; Garret, WS. The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science* **2013,** *341*, 569-573.
2. Yano, J. M.; Yu, K.; Donaldson, G. P.; Shastri, G. G.; Ann, P.; Ma, L.; Nagler, C. R.; Ismagilov, R. F.; Mazmanian, S. K.; Hsiao, E. Y. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **2015,** *161* (2), 264-76.
3. Tan, J. K. M., C.; Mariño,E; Macia, L and Mackay, CR. Metabolite-Sensing G Protein-Coupled Receptors-Facilitators of Diet-Related Immune Regulation. *Annu.Rev.Immunol* **2017,** *35*, 371-402.
4. Husted, A. S.; Trauelsen, M.; Rudenko, O.; Hjorth, S. A.; Schwartz, T. W. GPCR-Mediated Signaling of Metabolites. *Cell Metab* **2017,** *25* (4), 777-796.
5. Donia, M. S.; Fischbach, M. A. HUMAN MICROBIOTA. Small molecules from the human microbiota. *Science* **2015,** *349* (6246), 1254766.
6. Pedersen, H. K.; Gudmundsdottir, V.; Nielsen, H. B.; Hyotylainen, T.; Nielsen, T.; Jensen, B. A.; Forslund, K.; Hildebrand, F.; Prifti, E.; Falony, G.; Le Chatelier, E.; Levenez, F.; Dore, J.; Mattila, I.; Plichta, D. R.; Poho, P.; Hellgren, L. I.; Arumugam, M.; Sunagawa, S.; Vieira-Silva, S.; Jorgensen, T.; Holm, J. B.; Trost, K.; Meta, H. I. T. C.; Kristiansen, K.; Brix, S.; Raes, J.; Wang, J.; Hansen, T.; Bork, P.; Brunak, S.; Oresic, M.; Ehrlich, S. D.; Pedersen, O. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **2016,** *535* (7612), 376-81.
7. Perry, R. J.; Peng, L.; Barry, N. A.; Cline, G. W.; Zhang, D.; Cardone, R. L.; Petersen, K. F.; Kibbey, R. G.; Goodman, A. L.; Shulman, G. I. Acetate mediates a microbiome-brain-beta-cell axis to promote metabolic syndrome. *Nature* **2016,** *534* (7606), 213-7.
8. Fischbach, M. A. Microbiome: Focus on Causation and Mechanism. *Cell* **2018,** *174* (4), 785-790.
9. Dodd, D.; Spitzer, M. H.; Van Treuren, W.; Merrill, B. D.; Hryckowian, A. J.; Higginbottom, S. K.; Le, A.; Cowan, T. M.; Nolan, G. P.; Fischbach, M. A.; Sonnenburg, J. L. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* **2017,** *551* (7682), 648-652.
10. Guo, C. J.; Chang, F. Y.; Wyche, T. P.; Backus, K. M.; Acker, T. M.; Funabashi, M.; Taketani, M.; Donia, M. S.; Nayfach, S.; Pollard, K. S.; Craik, C. S.; Cravatt, B. F.; Clardy, J.; Voigt, C. A.; Fischbach, M. A. Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell* **2017,** *168* (3), 517-526 e18.
11. Haiser, H. J.; Gootenberg, D. B.; Chatman, K.; Sirasani, G.; Balskus, E. P.; Turnbaugh, P. J. Predicting and manipulating cardiac drug inactivation by the human gut bacterium Eggerthella lenta. *Science* **2013,** *341* (6143), 295-8.
12. Larsbrink, J.; Rogers, T. E.; Hemsworth, G. R.; McKee, L. S.; Tauzin, A. S.; Spadiut, O.; Klinter, S.; Pudlo, N. A.; Urs, K.; Koropatkin, N. M.; Creagh, A. L.; Haynes, C. A.; Kelly, A. G.; Cederholm, S. N.; Davies, G. J.; Martens, E. C.; Brumer, H. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **2014,** *506* (7489), 498-502.
13. Milshteyn, A.; Colosimo, D. A.; Brady, S. F. Accessing Bioactive Natural Products from the Human Microbiome. *Cell Host Microbe* **2018,** *23* (6), 725-736.
14. Wacker, D.; Stevens, R. C.; Roth, B. L. How Ligands Illuminate GPCR Molecular Pharmacology. *Cell* **2017,** *170* (3), 414-427.
15. Lei, W.; Ren, W.; Ohmoto, M.; Urban, J. F., Jr.; Matsumoto, I.; Margolskee, R. F.; Jiang, P. Activation of intestinal tuft cell-expressed Sucnr1 triggers type 2 immunity in the mouse small intestine. *Proc Natl Acad Sci U S A* **2018,** *115* (21), 5552-5557.
16. Nadjsombati, M. S.; McGinty, J. W.; Lyons-Cohen, M. R.; Jaffe, J. B.; DiPeso, L.; Schneider, C.; Miller, C. N.; Pollack, J. L.; Nagana Gowda, G. A.; Fontana, M. F.; Erle, D. J.; Anderson, M. S.; Locksley, R.

M.; Raftery, D.; von Moltke, J. Detection of Succinate by Intestinal Tuft Cells Triggers a Type 2 Innate Immune Circuit. *Immunity* **2018,** *49* (1), 33-41 e7.
17. Schneider, C.; O'Leary, C. E.; von Moltke, J.; Liang, H. E.; Ang, Q. Y.; Turnbaugh, P. J.; Radhakrishnan, S.; Pellizzon, M.; Ma, A.; Locksley, R. M. A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal Remodeling. *Cell* **2018,** *174* (2), 271-284 e14.
18. Cohen, L. J.; Esterhazy, D.; Kim, S. H.; Lemetre, C.; Aguilar, R. R.; Gordon, E. A.; Pickard, A. J.; Cross, J. R.; Emiliano, A. B.; Han, S. M.; Chu, J.; Vila-Farres, X.; Kaplitt, J.; Rogoz, A.; Calle, P. Y.; Hunter, C.; Bitok, J. K.; Brady, S. F. Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature* **2017,** *549* (7670), 48-53.
19. Kroeze, W. K.; Sassano, M. F.; Huang, X. P.; Lansu, K.; McCorvy, J. D.; Giguere, P. M.; Sciaky, N.; Roth, B. L. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat Struct Mol Biol* **2015,** *22* (5), 362-9.
20. Barnea, G.; Strapps, W.; Herrada, G.; Berman, Y.; Ong, J.; Kloss, B.; Axel, R.; Lee, K. J. The genetic design of signaling cascades to record receptor activation. *Proc Natl Acad Sci U S A* **2008,** *105* (1), 64-9.
21. Palm, N. W.; de Zoete, M. R.; Cullen, T. W.; Barry, N. A.; Stefanowski, J.; Hao, L.; Degnan, P. H.; Hu, J.; Peter, I.; Zhang, W.; Ruggiero, E.; Cho, J. H.; Goodman, A. L.; Flavell, R. A. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **2014,** *158* (5), 1000-1010.
22. Goodman, A. L.; Kallstrom, G.; Faith, J. J.; Reyes, A.; Moore, A.; Dantas, G.; Gordon, J. I. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* **2011,** *108* (15), 6252-7.
23. Thurmond, R. L.; Gelfand, E. W.; Dunford, P. J. The role of histamine H1 and H4 receptors in allergic inflammation: the search for new antihistamines. *Nat Rev Drug Discov* **2008,** *7* (1), 41-53.
24. Albuquerque, E. X.; Pereira, E. F.; Alkondon, M.; Rogers, S. W. Mammalian nicotinic acetylcholine receptors: from structure to function. *Physiol Rev* **2009,** *89* (1), 73-120.
25. Beaulieu, J. M.; Gainetdinov, R. R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol Rev* **2011,** *63* (1), 182-217.
26. Özoğul, F. Production of biogenic amines by Morganella morganii, Klebsiella pneumoniae and Hafnia alvei using a rapid HPLC method. *European Food Research and Technology* **2004,** *219* (5), 465-469.
27. Kim, S. H.; Ben-Gigirey, B.; Barros-Velazquez, J.; Price, R. J.; An, H. Histamine and biogenic amine production by Morganella morganii isolated from temperature-abused albacore. *J Food Prot* **2000,** *63* (2), 244-51.
28. Oldendorf, W. Brain uptake of radiolabeled amino acids, amines, and hexoses after arterial injection. *American Journal of Physiology* **1971,** *221*, 1629-1639.
29. Lovenberg, W. W., H & Udenfriend, S. Aromatic amino acid decarboxylase. *The Journal of Biological Chemistry* **1962,** *237*, 89-93.
30. Tannase, S. G., BM; Snell, EE. Purification and Properties of a Pyridoxal 5”Phosphate-dependent Histidine Decarboxylase from Morgunellu morganii AM- 15* *The Journal of Biological Chemistry* **1985,** *260*, 6738-6746.
31. Mavromatis, P. Q., PC. ModifiIcation of Niven’s Medium for the Enumeration of Histamine-Forming Bacteria and Discussion of the Parameters Associated with Its Use. *Journal of Food Protection* **2002,** *65*, 546-551.
32. Eun, C. S.; Kwak, M. J.; Han, D. S.; Lee, A. R.; Park, D. I.; Yang, S. K.; Kim, Y. S.; Kim, J. F. Does the intestinal microbial community of Korean Crohn's disease patients differ from that of western patients? *BMC Gastroenterol* **2016,** *16*, 28.
33. Tyagi, P.; Mandal, M. B.; Mandal, S.; Patne, S. C.; Gangopadhyay, A. N. Pouch colon associated with anorectal malformations fails to show spontaneous contractions but responds to acetylcholine and histamine in vitro. *J Pediatr Surg* **2009,** *44* (11), 2156-62.

34. Kim, H.; Dwyer, L.; Song, J. H.; Martin-Cano, F. E.; Bahney, J.; Peri, L.; Britton, F. C.; Sanders, K. M.; Koh, S. D. Identification of histamine receptors and effects of histamine on murine and simian colonic excitability. *Neurogastroenterol Motil* **2011,** *23* (10), 949-e409.
35. Integrative, H. M. P. R. N. C. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe* **2014,** *16* (3), 276-89.
36. Smolinska, S.; Jutel, M.; Crameri, R.; O'Mahony, L. Histamine and gut mucosal immune regulation. *Allergy* **2014,** *69* (3), 273-81.
37. Glover, V. S., M; Owen, F; Riley, GJ. Dopamine is a monoamine oxidase B substrate in man. *Nature* **1977,** *265*, 80-81.
38. Fiedorowicz, J. S., KL. The Role of Monoamine Oxidase Inhibitors in Current Psychiatric Practice. *J Psychiatr Pract* **2004,** *10*, 239-248.
39. Purcell, R. H., RA. Adhesion G Protein–Coupled Receptors as Drug Targets. *Annu. Rev. Pharmacol. Toxicol.* **2018,** *58*, 429-49.
40. Bhushan, R.; Bruckner, H. Use of Marfey's reagent and analogs for chiral amino acid analysis: assessment and applications to natural products and biological systems. *J Chromatogr B Analyt Technol Biomed Life Sci* **2011,** *879* (29), 3148-61.
41. Liberles, S. D.; Buck, L. B. A second class of chemosensory receptors in the olfactory epithelium. *Nature* **2006,** *442* (7103), 645-50.
42. Kishore, A.; Purcell, R. H.; Nassiri-Toosi, Z.; Hall, R. A. Stalk-dependent and Stalk-independent Signaling by the Adhesion G Protein-coupled Receptors GPR56 (ADGRG1) and BAI1 (ADGRB1). *J Biol Chem* **2016,** *291* (7), 3385-94.
43. Lovitt, R. W.; Morris, J. G.; Kell, D. B. The growth and nutrition of Clostridium sporogenes NCIB 8053 in defined media. *J Appl Bacteriol* **1987,** *62* (1), 71-80.
44. Nicholson, J. K.; Holmes, E.; Kinross, J.; Burcelin, R.; Gibson, G.; Jia, W.; Pettersson, S. Host-gut microbiota metabolic interactions. *Science* **2012,** *336* (6086), 1262-7.
45. Barcik, W.; Pugin, B.; Westermann, P.; Perez, N. R.; Ferstl, R.; Wawrzyniak, M.; Smolinska, S.; Jutel, M.; Hessel, E. M.; Michalovich, D.; Akdis, C. A.; Frei, R.; O'Mahony, L. Histamine-secreting microbes are increased in the gut of adult asthma patients. *J Allergy Clin Immunol* **2016,** *138* (5), 1491-1494 e7.
46. Borowsky, B.; Adham, N.; Jones, K. A.; Raddatz, R.; Artymyshyn, R.; Ogozalek, K. L.; Durkin, M. M.; Lakhlani, P. P.; Bonini, J. A.; Pathirana, S.; Boyle, N.; Pu, X.; Kouranova, E.; Lichtblau, H.; Ochoa, F. Y.; Branchek, T. A.; Gerald, C. Trace amines: identification of a family of mammalian G protein-coupled receptors. *Proc Natl Acad Sci U S A* **2001,** *98* (16), 8966-71.
47. Bunzow, J. S., MS; Arttamangkul, S; Harrison, LM; Zhang, G; Quigley, DI; Darland, T; Suchland, KL; Pasumanmula, S; Kennedy, JL; Olson, SB; Magenis, RE; Amara, SG; Grandy, DK. Amphetamine, 3,4-Methylenedioxymethamphetamine, Lysergic Acid Diethylamide, and Metabolites of the Catecholamine Neurotransmitters Are Agonists of a Rat Trace Amine Receptor. *Molecular Pharmacology* **2001,** *60*, 1181-1188.
48. Sotnikova, T. D.; Budygin, E. A.; Jones, S. R.; Dykstra, L. A.; Caron, M. G.; Gainetdinov, R. R. Dopamine transporter-dependent and -independent actions of trace amine beta-phenylethylamine. *J Neurochem* **2004,** *91* (2), 362-73.
49. Ramachandraih, C. S., N; Bar, KJ; Baker, G; Yeragani, VK. Antidepressants: From MAOIs to SSRIs and more. *Indian J Psychiatry* **2011,** *53*, 180-182.
50. Riederer, P.; Laux, G. MAO-inhibitors in Parkinson's Disease. *Exp Neurobiol* **2011,** *20* (1), 1-17.
51. Irsfeld, M.; Spadafore, M.; Pruss, B. M. beta-phenylethylamine, a small molecule with a large impact. *Webmedcentral* **2013,** *4* (9).
52. Amisten, S.; Salehi, A.; Rorsman, P.; Jones, P. M.; Persaud, S. J. An atlas and functional analysis of G-protein coupled receptors in human islets of Langerhans. *Pharmacol Ther* **2013,** *139* (3), 359-91.

53. Duner, P.; Al-Amily, I. M.; Soni, A.; Asplund, O.; Safi, F.; Storm, P.; Groop, L.; Amisten, S.; Salehi, A. Adhesion G Protein-Coupled Receptor G1 (ADGRG1/GPR56) and Pancreatic beta-Cell Function. *J Clin Endocrinol Metab* **2016,** *101* (12), 4637-4645.
54. Adibi, S. M., DW. Protein Digestion in Human Intestine as Reflected in Luminal, Mucosal, and Plasma Amino Acid Concentrations after Meals. *The Journal of Clinical Investigation* **1973,** *52*, 1586-1594.
55. Williams, R. M., CDS; Burnett, JR. Phenylketonuria- An Inborn Error of Phenylalanine Metabolism. *Clin Biochem Rew* **2008,** *29*, 31-41.
56. Pezeshki, A.; Zapata, R. C.; Singh, A.; Yee, N. J.; Chelikani, P. K. Low protein diets produce divergent effects on energy balance. *Scientific Reports* **2016,** *6* (1).
57. Jutel, M.; Akdis, M.; Akdis, C. A. Histamine, histamine receptors and their role in immune pathology. *Clin Exp Allergy* **2009,** *39* (12), 1786-800.
58. Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello, E. K.; Fierer, N.; Pena, A. G.; Goodrich, J. K.; Gordon, J. I.; Huttley, G. A.; Kelley, S. T.; Knights, D.; Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.; Pirrung, M.; Reeder, J.; Sevinsky, J. R.; Turnbaugh, P. J.; Walters, W. A.; Widmann, J.; Yatsunenko, T.; Zaneveld, J.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **2010,** *7* (5), 335-6.
59. Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014,** *30* (15), 2114-20.
60. Nurk, S.; Bankevich, A.; Antipov, D.; Gurevich, A. A.; Korobeynikov, A.; Lapidus, A.; Prjibelski, A. D.; Pyshkin, A.; Sirotkin, A.; Sirotkin, Y.; Stepanauskas, R.; Clingenpeel, S. R.; Woyke, T.; McLean, J. S.; Lasken, R.; Tesler, G.; Alekseyev, M. A.; Pevzner, P. A. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol* **2013,** *20* (10), 714-37.
61. Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **2008,** *9*, 75.
62. Denamur, E.; Clermont, O.; Bonacorsi, S.; Gordon, D. The population genetics of pathogenic Escherichia coli. *Nature Reviews Microbiology* **2020**.
63. Kumar, A.; Sperandio, V. Indole Signaling at the Host-Microbiota-Pathogen Interface. *mBio* **2019,** *10* (3), e01031-19.
64. Hirakawa, H.; Inazumi, Y.; Masaki, T.; Hirata, T.; Yamaguchi, A. Indole induces the expression of multidrug exporter genes in Escherichia coli. *Molecular Microbiology* **2005,** *55* (4), 1113-1126.
65. Weatherspoon-Griffin, N.; Yang, D.; Kong, W.; Hua, Z.; Shi, Y. The CpxR/CpxA two-component regulatory system up-regulates the multidrug resistance cascade to facilitate Escherichia coli resistance to a model antimicrobial peptide. *J Biol Chem* **2014,** *289* (47), 32571-82.
66. Lee, H. H.; Molla, M. N.; Cantor, C. R.; Collins, J. J. Bacterial charity work leads to population-wide resistance. *Nature* **2010,** *467* (7311), 82-85.
67. Richard, H. T.; Foster, J. W. Acid Resistance in Escherichia coli. In *Advances in Applied Microbiology*, Academic Press: 2003; Vol. 52, pp 167-186.
68. Castanie-Cornet, M. P.; Penfound, T. A.; Smith, D.; Elliott, J. F.; Foster, J. W. Control of acid resistance in Escherichia coli. *J Bacteriol* **1999,** *181* (11), 3525-35.
69. Hersh, B. M.; Farooq, F. T.; Barstad, D. N.; Blankenhorn, D. L.; Slonczewski, J. L. A glutamate-dependent acid resistance gene in Escherichia coli. *J Bacteriol* **1996,** *178* (13), 3978-81.
70. Malashkevich, V. N.; De Biase, D.; Markovic-Housley, Z.; Schlunegger, M. P.; Bossa, F.; Jansonius, J. N. Crystallization and preliminary X-ray analysis of the beta-isoform of glutamate decarboxylase from Escherichia coli. *Acta Crystallogr D Biol Crystallogr* **1998,** *54* (Pt 5), 1020-2.

71. Gong, S.; Richard, H.; Foster, J. W. YjdE (AdiC) is the arginine:agmatine antiporter essential for arginine-dependent acid resistance in Escherichia coli. *J Bacteriol* **2003,** *185* (15), 4402-9.
72. Moreau, P. L. The lysine decarboxylase CadA protects Escherichia coli starved of phosphate against fermentation acids. *J Bacteriol* **2007,** *189* (6), 2249-61.
73. Wang, J.; Wu, X.; Simonavicius, N.; Tian, H.; Ling, L. Medium-chain fatty acids as ligands for orphan G protein-coupled receptor GPR84. *J Biol Chem* **2006,** *281* (45), 34457-64.
74. Karimi-Jaberi, Z.; Fereydoonnezhad, A. One-pot, organocatalytic synthesis of spirooxindoles using citric acid in aqueous media. *Quarterly Journal of Iranian Chemical Communication* **2017,** *5* (Issue 4, pp. 364-493, Serial No. 17), 407-416.
75. You, Z.-H.; Chen, Y.-H.; Tang, Y.; Liu, Y.-K. Organocatalytic Asymmetric Synthesis of Spiro-Bridged and Spiro-Fused Heterocyclic Compounds Containing Chromane, Indole, and Oxindole Moieties. *Organic Letters* **2018,** *20* (21), 6682-6686.
76. Edlow, D. W.; Sheldon, W. H. The pH of Inflammatory Exudates. *Proceedings of the Society for Experimental Biology and Medicine* **1971,** *137* (4), 1328-1332.
77. Dubos, R. J. The Micro-Environment of Inflammation or Metchnikoff Revisited. *Lancet* **1955**, 1-5.
78. Erra Díaz, F.; Dantas, E.; Geffner, J. Unravelling the Interplay between Extracellular Acidosis and Immune Cells. *Mediators of Inflammation* **2018,** *2018*, 1218297.
79. Beasley, D. E.; Koltz, A. M.; Lambert, J. E.; Fierer, N.; Dunn, R. R. The Evolution of Stomach Acidity and Its Relevance to the Human Microbiome. *PLoS One* **2015,** *10* (7), e0134116-e0134116.
80. Lee, J.-H.; Lee, J. Indole as an intercellular signal in microbial communities. *FEMS Microbiology Reviews* **2010,** *34* (4), 426-444.
81. Tannock, I. F.; Rotin, D. Acid pH in Tumors and Its Potential for Therapeutic Exploitation. *Cancer Research* **1989,** *49* (16), 4373.
82. Castellarin, M.; Warren, R. L.; Freeman, J. D.; Dreolini, L.; Krzywinski, M.; Strauss, J.; Barnes, R.; Watson, P.; Allen-Vercoe, E.; Moore, R. A.; Holt, R. A. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome research* **2012,** *22* (2), 299-306.
83. Gao, Z.; Guo, B.; Gao, R.; Zhu, Q.; Qin, H. Microbiota disbiosis is associated with colorectal cancer. *Frontiers in microbiology* **2015,** *6*, 20-20.
84. Nejman, D.; Livyatan, I.; Fuks, G.; Gavert, N.; Zwang, Y.; Geller, L. T.; Rotter-Maskowitz, A.; Weiser, R.; Mallel, G.; Gigi, E.; Meltser, A.; Douglas, G. M.; Kamer, I.; Gopalakrishnan, V.; Dadosh, T.; Levin-Zaidman, S.; Avnet, S.; Atlan, T.; Cooper, Z. A.; Arora, R.; Cogdill, A. P.; Khan, M. A. W.; Ologun, G.; Bussi, Y.; Weinberger, A.; Lotan-Pompan, M.; Golani, O.; Perry, G.; Rokah, M.; Bahar-Shany, K.; Rozeman, E. A.; Blank, C. U.; Ronai, A.; Shaoul, R.; Amit, A.; Dorfman, T.; Kremer, R.; Cohen, Z. R.; Harnof, S.; Siegal, T.; Yehuda-Shnaidman, E.; Gal-Yam, E. N.; Shapira, H.; Baldini, N.; Langille, M. G. I.; Ben-Nun, A.; Kaufman, B.; Nissan, A.; Golan, T.; Dadiani, M.; Levanon, K.; Bar, J.; Yust-Katz, S.; Barshack, I.; Peeper, D. S.; Raz, D. J.; Segal, E.; Wargo, J. A.; Sandbank, J.; Shental, N.; Straussman, R. The human tumor microbiome is composed of tumor type–specific intracellular bacteria. *Science* **2020,** *368* (6494), 973.
85. Thompson, K. J.; Ingle, J. N.; Tang, X.; Chia, N.; Jeraldo, P. R.; Walther-Antonio, M. R.; Kandimalla, K. K.; Johnson, S.; Yao, J. Z.; Harrington, S. C.; Suman, V. J.; Wang, L.; Weinshilboum, R. L.; Boughey, J. C.; Kocher, J.-P.; Nelson, H.; Goetz, M. P.; Kalari, K. R. A comprehensive analysis of breast cancer microbiota and host gene expression. *PLoS One* **2017,** *12* (11), e0188873.
86. Sasaki-Imamura, T.; Yoshida, Y.; Suwabe, K.; Yoshimura, F.; Kato, H. Molecular basis of indole production catalyzed by tryptophanase in the genus Prevotella. *FEMS Microbiology Letters* **2011,** *322* (1), 51-59.
87. Takeda, S.; Yamamoto, A.; Okada, T.; Matsumura, E.; Nose, E.; Kogure, K.; Kojima, S.; Haga, T. Identification of surrogate ligands for orphan G protein-coupled receptors. *Life Sciences* **2003,** *74* (2), 367-377.

88. Pillaiyar, T.; Köse, M.; Sylvester, K.; Weighardt, H.; Thimm, D.; Borges, G.; Förster, I.; von Kügelgen, I.; Müller, C. E. Diindolylmethane Derivatives: Potent Agonists of the Immunostimulatory Orphan G Protein-Coupled Receptor GPR84. *Journal of Medicinal Chemistry* **2017,** *60* (9), 3636-3655.
89. Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.-K.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P.-H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Pontén, F. Tissue-based map of the human proteome. *Science* **2015,** *347* (6220), 1260419.
90. Kroeze, W. K.; Sassano, M. F.; Huang, X.-P.; Lansu, K.; McCorvy, J. D.; Giguère, P. M.; Sciaky, N.; Roth, B. L. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature Structural & Molecular Biology* **2015,** *22* (5), 362-369.
91. Kunimasa, K.; Kobayashi, T.; Kaji, K.; Ohta, T. Antiangiogenic Effects of Indole-3-Carbinol and 3,3′-Diindolylmethane Are Associated with Their Differential Regulation of ERK1/2 and Akt in Tube-Forming HUVEC. *The Journal of Nutrition* **2010,** *140* (1), 1-6.
92. Le, H. T.; Schaldach, C. M.; Firestone, G. L.; Bjeldanes, L. F. Plant-derived 3,3′-Diindolylmethane Is a Strong Androgen Antagonist in Human Prostate Cancer Cells. *Journal of Biological Chemistry* **2003,** *278* (23), 21136-21145.
93. Mattiazzi, J.; Sari, M. H. M.; Lautenchleger, R.; Dal Prá, M.; Braganhol, E.; Cruz, L. Incorporation of 3,3′-Diindolylmethane into Nanocapsules Improves Its Photostability, Radical Scavenging Capacity, and Cytotoxicity Against Glioma Cells. *AAPS PharmSciTech* **2019,** *20* (2), 49.
94. Tomar, S.; Nagarkatti, M.; Nagarkatti, P. S. 3,3'-Diindolylmethane attenuates LPS-mediated acute liver failure by regulating miRNAs to target IRAK4 and suppress Toll-like receptor signalling. *Br J Pharmacol* **2015,** *172* (8), 2133-2147.
95. Chiang, Y.-R.; Li, A.; Leu, Y.-L.; Fang, J.-Y.; Lin, Y.-K. An in vitro study of the antimicrobial effects of indigo naturalis prepared from Strobilanthes formosanus Moore. *Molecules* **2013,** *18* (11), 14381-14396.
96. Fatima, I.; Ahmad, I.; Anis, I.; Malik, A.; Afza, N. Isatinones A and B, new antifungal oxindole alkaloids from Isatis costata. *Molecules* **2007,** *12* (2), 155-162.
97. Kobayashi, M.; Aoki, S.; Gato, K.; Matsunami, K.; Kurosu, M.; Kitagawa, I. Marine natural products. XXXIV. Trisindoline, a new antibiotic indole trimer, produced by a bacterium of Vibrio sp. separated from the marine sponge Hyrtios altum. *Chem Pharm Bull (Tokyo)* **1994,** *42* (12), 2449-2451.
98. Sung, W. S.; Lee, D. G. In Vitro Antimicrobial Activity and the Mode of Action of Indole-3-Carbinol against Human Pathogenic Microorganisms. *Biological and Pharmaceutical Bulletin* **2007,** *30* (10), 1865-1869.
99. Yoo, M.; Choi, S.-U.; Choi, K. Y.; Yon, G. H.; Chae, J.-C.; Kim, D.; Zylstra, G. J.; Kim, E. Trisindoline synthesis and anticancer activity. *Biochemical and Biophysical Research Communications* **2008,** *376* (1), 96-99.
100. Carlos, L. C.; Julio, A.; Maribel, V.-M.; Octavio, P.-L. Antioxidant Activity of an Unusual 3-Hydroxyindole Derivative Isolated from Fruits of Aristotelia chilensis (Molina) Stuntz. *Zeitschrift für Naturforschung C* **2009,** *64* (9-10), 759-762.
101. Knockaert, M.; Blondel, M.; Bach, S.; Leost, M.; Elbi, C.; Hager, G. L.; Nagy, S. R.; Han, D.; Denison, M.; Ffrench, M.; Ryan, X. P.; Magiatis, P.; Polychronopoulos, P.; Greengard, P.; Skaltsounis, L.; Meijer, L. Independent actions on cyclin-dependent kinases and aryl hydrocarbon receptor mediate the antiproliferative effects of indirubins. *Oncogene* **2004,** *23* (25), 4400-4412.
102. Hirakawa, H.; Hayashi-Nishino, M.; Yamaguchi, A.; Nishino, K. Indole enhances acid resistance in Escherichia coli. *Microbial Pathogenesis* **2010,** *49* (3), 90-94.

103. Lee, J.; Bansal, T.; Jayaraman, A.; Bentley, W. E.; Wood, T. K. Enterohemorrhagic Escherichia coli Biofilms Are Inhibited by 7-Hydroxyindole and Stimulated by Isatin. *Applied and Environmental Microbiology* **2007,** *73* (13), 4100.
104. Munuganti, Ravi S. N.; Hassona, Mohamed D. H.; Leblanc, E.; Frewin, K.; Singh, K.; Ma, D.; Ban, F.; Hsing, M.; Adomat, H.; Lallous, N.; Andre, C.; Jonadass, Jon Paul S.; Zoubeidi, A.; Young, Robert N.; Tomlinson Guns, E.; Rennie, Paul S.; Cherkasov, A. Identification of a Potent Antiandrogen that Targets the BF3 Site of the Androgen Receptor and Inhibits Enzalutamide-Resistant Prostate Cancer. *Chemistry & Biology* **2014,** *21* (11), 1476-1485.
105. Hubbard, T. D.; Murray, I. A.; Perdew, G. H. Indole and Tryptophan Metabolism: Endogenous and Dietary Routes to Ah Receptor Activation. *Drug Metab Dispos* **2015,** *43* (10), 1522-1535.
106. Medvedev, A.; Crumeyrolle-Arias, M.; Cardona, A.; Sandler, M.; Glover, V. Natriuretic peptide interaction with [3H]isatin binding sites in rat brain. *Brain Research* **2005,** *1042* (2), 119-124.
107. Busbee, P. B.; Menzel, L.; Alrafas, H. R.; Dopkins, N.; Becker, W.; Miranda, K.; Tang, C.; Chatterjee, S.; Singh, U.; Nagarkatti, M.; Nagarkatti, P. S. Indole-3-carbinol prevents colitis and associated microbial dysbiosis in an IL-22-dependent manner. *JCI Insight* **2020,** *5* (1), e127551.
108. Lin, Y.-H.; Luck, H.; Khan, S.; Schneeberger, P. H. H.; Tsai, S.; Clemente-Casares, X.; Lei, H.; Leu, Y.-L.; Chan, Y. T.; Chen, H.-Y.; Yang, S.-H.; Coburn, B.; Winer, S.; Winer, D. A. Aryl hydrocarbon receptor agonist indigo protects against obesity-related insulin resistance through modulation of intestinal and metabolic tissue immunity. *International Journal of Obesity* **2019,** *43* (12), 2407-2421.
109. Perabo, F. G. E.; Landwehrs, G.; Frössler, C.; Schmidt, D. H.; Mueller, S. C. Antiproliferative and apoptosis inducing effects of indirubin-3′-monoxime in renal cell cancer cells. *Urologic Oncology: Seminars and Original Investigations* **2011,** *29* (6), 815-820.
110. Chen, H.; Nwe, P. K.; Yang, Y.; Rosen, C. E.; Bielecka, A. A.; Kuchroo, M.; Cline, G. W.; Kruse, A. C.; Ring, A. M.; Crawford, J. M.; Palm, N. W. A Forward Chemical Genetic Screen Reveals Gut Microbiota Metabolites That Modulate Host Physiology. *Cell* **2019,** *177* (5), 1217-1231 e18.