Natural Product Reports

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal

Cite this: DOI: 10.1039/d5np00021a

Mass spectrometry-based metabolomics approaches to interrogate host–microbiome interactions in mammalian systems

Atharva S. Kulkarni, [a] Guilherme M. P. Carrara, †[b] Jiangpeiyun Jin, †[a] Jarrod Laro,†[bc] Thilini Peramuna,[c] Laura-Isobel McCall *[bc] and Neha Garg *[ad]

Covering: 2015 to 2025

Chemical crosstalk is universal to all life, niche-specific, and essential to thrive. This crosstalk is mediated by a large diversity of molecules, including metal ions, small molecules, polysaccharides, nucleic acids, lipids, and proteins. Among these, specialized small molecules referred to as natural products (NPs) play an important role in microbe–drug/environment interactions, microbe–microbe, and microbe–host interactions. Microbial communication using NPs allows microbes to sense quorum, form biofilms, eliminate competition, establish symbiosis, evade immune attack, and respond to stress. In most cases, the elucidation of small molecule mediators and effectors of microbe–host interactions presents a major challenge due to the relatively low abundance of microbial metabolites in a milieu of host, microbe, and environmental metabolites. Advances in analytical instrumentation, such as mass spectrometers, and both experimental as well as computational methods to analyze data, coupled with the use of model organisms, have enabled fundamental discoveries of mechanisms of small molecule-mediated host–microbe interactions. The focus of this review is to detail the approaches applied in the last decade to disentangle microbiome-derived NPs in human and murine model systems. Select recent findings from diverse biological ecosystems are discussed to inform relevant parallels and potential strategies for research in human health.

Received 26th March 2025

DOI: 10.1039/d5np00021a

rsc.li/npr

1 Introduction

The advent of fast and affordable high-throughput sequencing ignited the development and application of metagenomics and metatranscriptomics methods to decipher host–microbe interactions.[1] The Human Microbiome Project Consortium (HMP) and Metagenomics of the Human Intestinal Tract (MetaHIT) have generated large, complementary datasets of 16S rRNA sequences, whole genome sequences (WGS), shotgun metagenomics sequences, and metagenome-assembled genomes (MAGs) to characterize the human microbiome.[2–4] The HMP was

[a]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. E-mail: neha.garg@chemistry.gatech.edu
[b]Chemistry and Biochemistry, San Diego State University, San Diego, California 92182, USA. E-mail: lmccall@sdu.edu
[c]Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, USA
[d]Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, Georgia 30332, USA
† Equal contribution authors.

started in 2007 looking at the nasal, oral, vaginal, gut, and skin microbiomes of a healthy human cohort, with this effort leading to the discovery of the community compositions at each of the body sites.[5] The second phase of this project (HMP2) expanded the scope to include host–microbe interactions in three disease conditions: preterm birth, onset of IBD and type II diabetes. The focus was also expanded to include transcriptomics, proteomics, and metabolomics data.[6] The studies derived from the HMP and MetaHIT projects initially focused on the bacterial component of the microbiome but have inspired further research exploring the other members of the microbiome, including fungi, archaea and phages.[7] This chronology illustrates the growing awareness in the microbiome research community of the importance of metabolomics in understanding host–microbe interactions. While the focus of this review is targeted towards discussion of human or murine host–bacterial interactions, several examples from diverse environmental hosts are cited to inspire their application in mammalian host-derived microbiome systems.

Metabolomics is the study of all small molecules present in a biological sample, such as those involved in the metabolism of lipids, carbohydrates, nucleotides, and amino acids, the degradation of xenobiotics, and the biosynthesis of specialized metabolites, also referred to as small molecule NPs or specialized metabolites (Fig. 1). In the context of biological life, the central dogma of biology flows from gene to transcript to protein. However, beyond proteins lies the world of metabolites, both primary and specialized metabolites. Specialized metabolites refer to those metabolites that, while not essential for the growth and reproduction of an organism, can confer fitness advantages, which can lead to competitive, predatory, symbiotic or antagonistic interactions between organisms living together in an ecosystem. To investigate chemical interactions between organisms, metabolomics serves as a useful starting point and is applied in an untargeted or targeted fashion. Untargeted metabolomics provides a snapshot of all the detectable metabolites in a sample under the given experimental conditions. The detectable metabolites can vary depending upon the extraction method, the data acquisition strategy, and the instrument used.[8] The development and application of data independent methods to natural product discovery has further improved metabolite coverage and limit of detection (discussed below).[9] In contrast, targeted metabolomics seeks to quantify a predefined subset of metabolites to validate a hypothesis. Both strategies serve complementary roles in hypothesis building, refining, and validation to provide insights into biochemical activities.[10]

There are primarily two analytical techniques applied to interrogate the metabolome: Mass Spectrometry (MS) and Nuclear Magnetic Resonance (NMR).[10] For the purposes of this review, we will focus on MS-based techniques. The higher sensitivity of MS enables detection of specialized metabolites, which are usually found in low concentrations. MS is often coupled to liquid chromatography (LC) or gas chromatography (GC), allowing for separation of analytes in the sample based on their chemical or physical properties such as polarity or volatility. A typical metabolomics workflow will involve sample preparation, data acquisition, and data preprocessing, followed by statistical analysis to rank metabolites, and annotation of metabolites relevant to phenotype being studied (Fig. 1). The metabolites detected using these techniques vary widely in size and structure. Thus, not only the large number of molecules detected, but also the diversity of building blocks and functional groups present in organic molecules presents a significant challenge in compound annotation. Detailed and complementary analyses, such as spectral similarity analysis using networking approaches, substructure discovery, class enrichment, and pathway analysis, aid in compound annotation and understanding the biological significance of the metabolites analyzed.[11–18]

When investigating host–microbe interactions *in vitro*, sample generation can involve co-culturing microbial species and/or mammalian cells of interest in various conditions or against a control or baseline condition (Fig. 1). For complex *in vivo* systems, samples include skin, sweat, sebum, tissues, feces, urine, serum and saliva, directly from a volunteer or from animal models. The samples are homogenized, and extracted using an appropriate combination of organic solvents. These extracts are then dried and resuspended in a solvent compatible with LC or GC.[10] The column used in LC depends on the polarity of the metabolites under investigation. For polar metabolites, HILIC (Hydrophilic interaction chromatography) columns may be used, while for non-polar metabolites a C18 reverse-phase column chromatography is commonly implemented.[19] Newer columns combining the properties of HILIC and reverse phase (polar C18 columns) are now being applied to minimize data acquisition time.[20–22] Since MS operates on the principle of detecting charged molecules, there are several ionization techniques that can be used, with soft ionization techniques such as Electrospray ionization (ESI) and Atmospheric Pressure Chemical Ionization (APCI) often used for specialized metabolites.[23] Fragmentation of precursors generated by the aforementioned ionization methods through Collision-induced Dissociation (CID), higher-energy collisional dissociation (HCD), ultraviolet photodissociation (UVPD) or Electron Capture Dissociation (ECD) generate tandem mass spectra (MS/MS) that aid annotation efforts.[24] An important consideration in the acquisition of MS/MS spectra is the basis on which precursors are chosen for fragmentation. MS/MS experiments can be operated in data-dependent acquisition (DDA) or data-independent acquisition (DIA) mode.[25,26] In DDA, ions are isolated for fragmentation on the basis of abundance, prioritizing fragmentation of ions with higher abundance first. Thus, lower-abundance ions often do not result in acquisition of MS/MS spectra. To overcome this challenge, several vendor-specific strategies, including active exclusion and AcquireX, have been developed.[27] Such data acquisition strategies have been extensively reviewed in Defossez *et al*.[28] In DIA, all ions in a given m/z window are fragmented at a given time, reducing bias, but also making deconvolution of spectra challenging.[9] DIA has been recently applied in NPs research workflows.[29,30] Quality Control (QC) samples, prepared by pooling of all samples, are run every few samples to monitor the performance of the instrument and ensure reproducibility of the results.[31] Data preprocessing involves generation of a 'feature' table consisting of the area under the curves of the extracted ion chromatograms of all detected metabolites, after

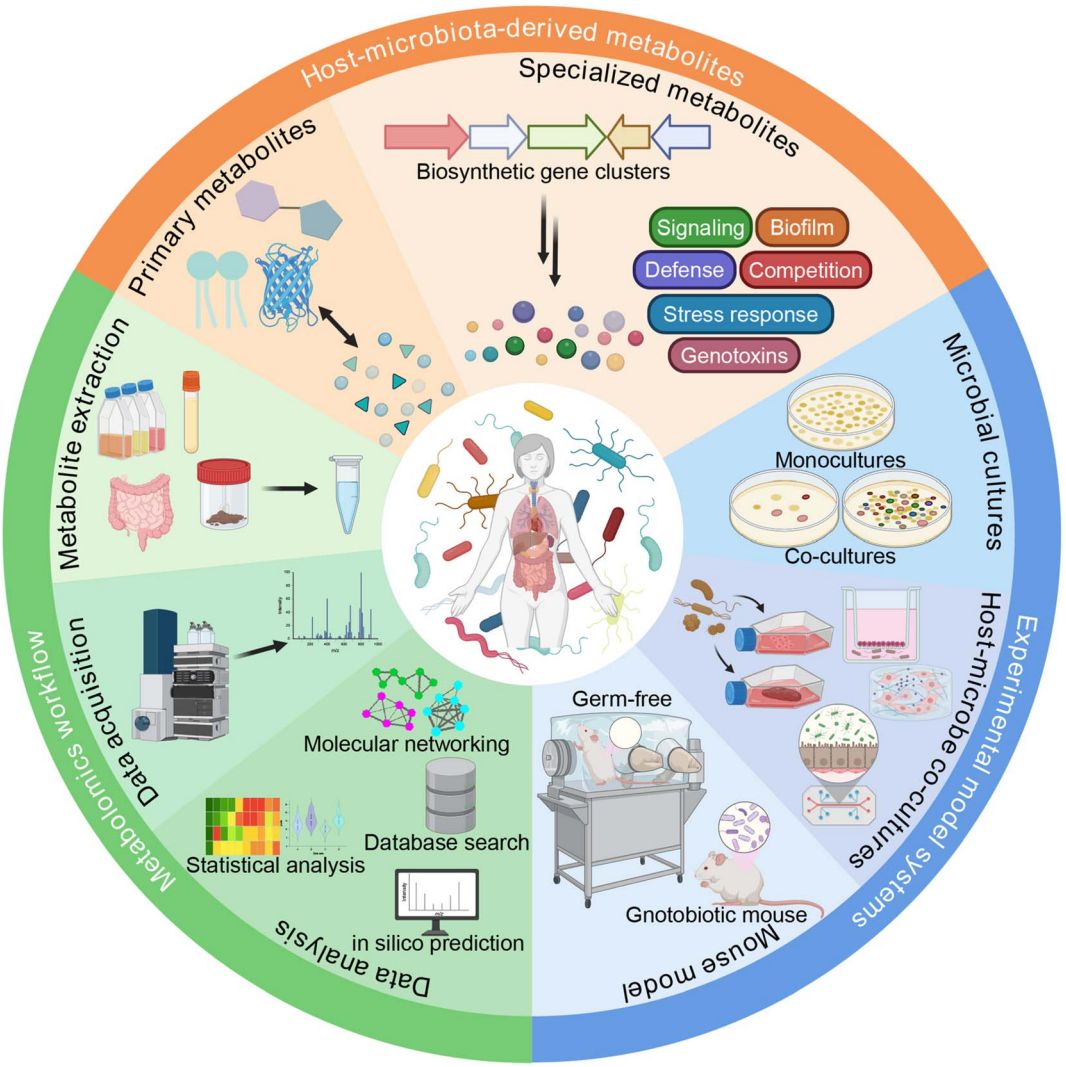

Fig. 1 Overview of metabolite categories and the workflows for investigating their role in host–microbe interactions using LCMS-based metabolomics approaches. Created with https://www.biorender.com.

filtering noise and background, detecting and deconvoluting peaks and aligning for retention time across samples. The effectiveness of this step is crucial in comparing and obtaining accurate statistically significant differences between the samples.[32,33] Integration of metabolomics data with genomes, transcriptomes, and proteomes provides another layer of confidence in data annotation and may facilitate the elucidation of the biochemical functions of discovered metabolites or disease-impacted signals.[34–36]

2 Methods used to decipher host–microbiome interactions and role of metabolomics

The first step in disentangling the host–microbe interactome is to identify the microbial community composition. While bacteria are the most abundant and extensively studied members of the human microbiome, other members like fungi, archaea, and phages are also present and are increasingly recognized for their roles in health and disease.[37–39] Metagenomics has emerged as the ideal technique for studying microbial diversity and composition. A common approach in metagenomics is shotgun sequencing, wherein DNA extracted from samples is fragmented, followed by sequencing. Overlapping reads are assembled into longer composite sequences called contigs, and multiple contigs can be processed to reconstruct a genome and identify the taxa of the microbe associated with it. The creation of these MAGs has enabled the study of previously unknown or unculturable species and has revealed greater phylogenetic diversity in host–microbe environments while also providing access to biosynthetic capability for production of specialized metabolites by these microorganisms (Fig. 2).[40,41] Apart from taxonomic identification, metagenomics can also identify pathway-related genes encoding products involved in host–microbe interactions. For example, in the environmental determinants of diabetes in the young (TEDDY) study for Type 1 diabetes (T1D), metagenomics analysis of stool samples from 783 children susceptible to T1D was

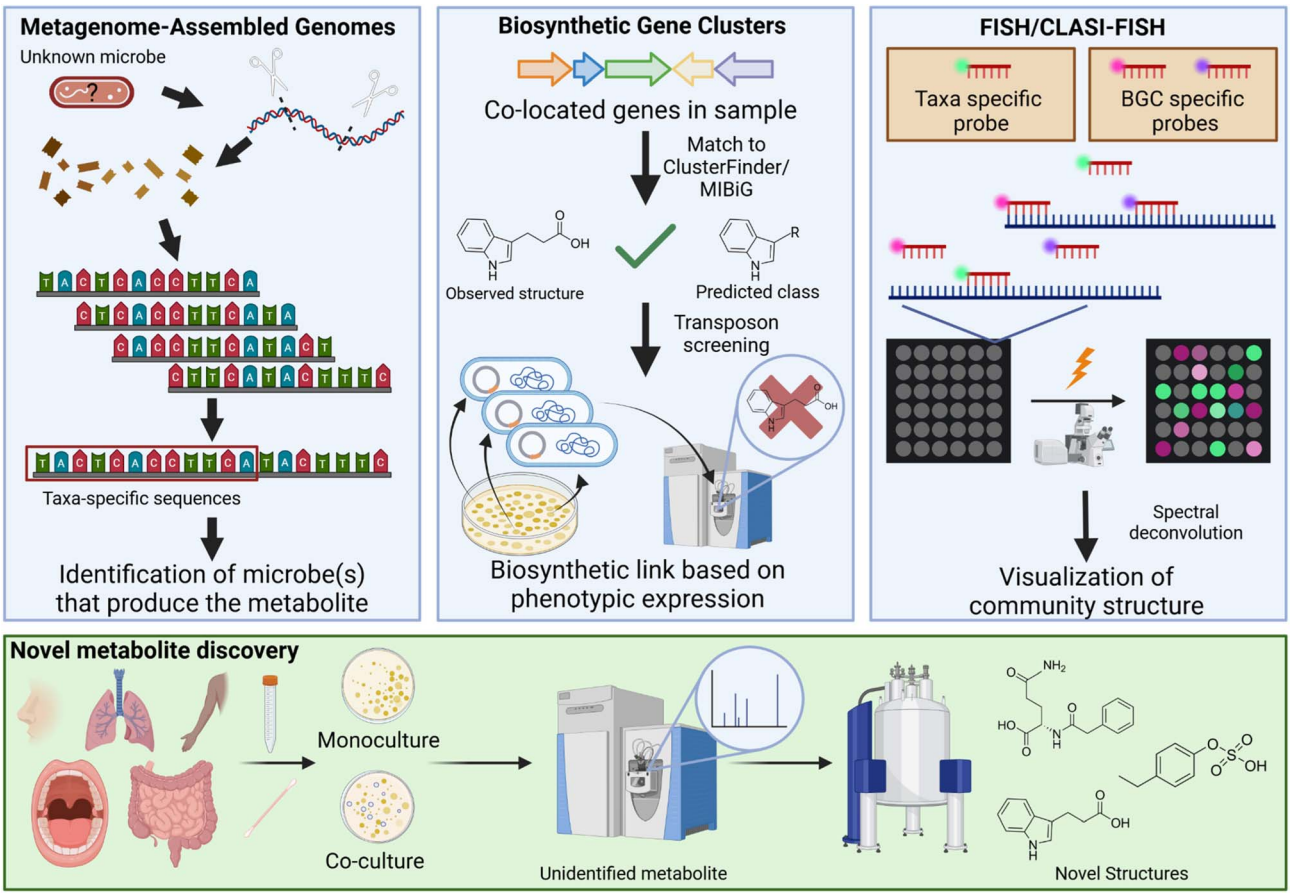

Fig. 2 Multi-omic methods used to identify the relationship between metabolites and microbes. The blue boxes demonstrate different approaches to understanding how specific microbes produce unique metabolites, while the green box represents the methods used to identify metabolites. Created with https://www.biorender.com.

performed.[42] Authors found that the genes mapping to biosynthesis of short-chain fatty acids (SCFAs) and bacterial fermentation were increased in the control group compared to those with who progressed to T1D, furthering support for the protective effect of SCFA. However, underscored in the study was the lack of functional annotations for many gene families that presents a limitation in a purely metagenomics-based approach.[42]

Identification of biosynthetic gene clusters (BGCs) can directly feed into the discovery of novel specialized metabolites. For example, Donia *et al.* used ClusterFinder to predict BGCs of both known and unknown compound classes in human microbiota reference genomes, finding BGCs for a variety of saccharide, polyketide synthase (PKS), ribosomally synthesized and post-translationally modified peptide (RiPP), and non-ribosomal peptide synthase (NRPS) clusters, some of which were widely distributed across species.[43,44] They characterized lactocillin, the product of a RiPP thiopeptide cluster from the vaginal commensal *Lactobacillus gasseri*, by liquid culturing and isolation, using NMR, tandem MS, and isotope labeling (Fig. 3). Lactocillin had antibacterial activity against several Gram-positive bacteria.[44] While the authors could not conclusively prove that lactocillin is made in the human body, they mined publicly available metatranscriptomic data sets to show that the *lcl* BGC encoding its production is expressed in this environment.[44]

An alternative genome-first approach that has been useful to identify the genes involved in producing specialized metabolites impacting host biology is phenotypic screening combined with transposon library screening. For example, colibactin, a genotoxin produced by strains of Enterobacteriaceae, was first discovered in 2006 when certain pathogenic *Escherichia coli* strains isolated from meningitis and urinary tract infections induced megalocytosis in cultured eukaryotic cells (Fig. 4).[45] A genetic screen of transposon mutants pointed towards a 54 kb polyketide synthase genomic island (*pks*). Colibactin induces interstrand crosslinks and double-strand breaks in DNA.[45] While the structure of colibactin remained elusive for many years, its bioactivity and by extension, the activity of the corresponding *pks* island was a topic of great interest, driven by the finding that *pks*+ bacteria were found in a high percentage of patients with IBD and colorectal cancer (CRC).[46–48] Germ-free mice progressed to carcinoma after exposure to a carcinogen if colonized with bacteria that harbored *pks*, compared to controls colonized with *pks*− bacteria, which showed decreased tumor multiplicity and invasion.[48] Moreover, mutational

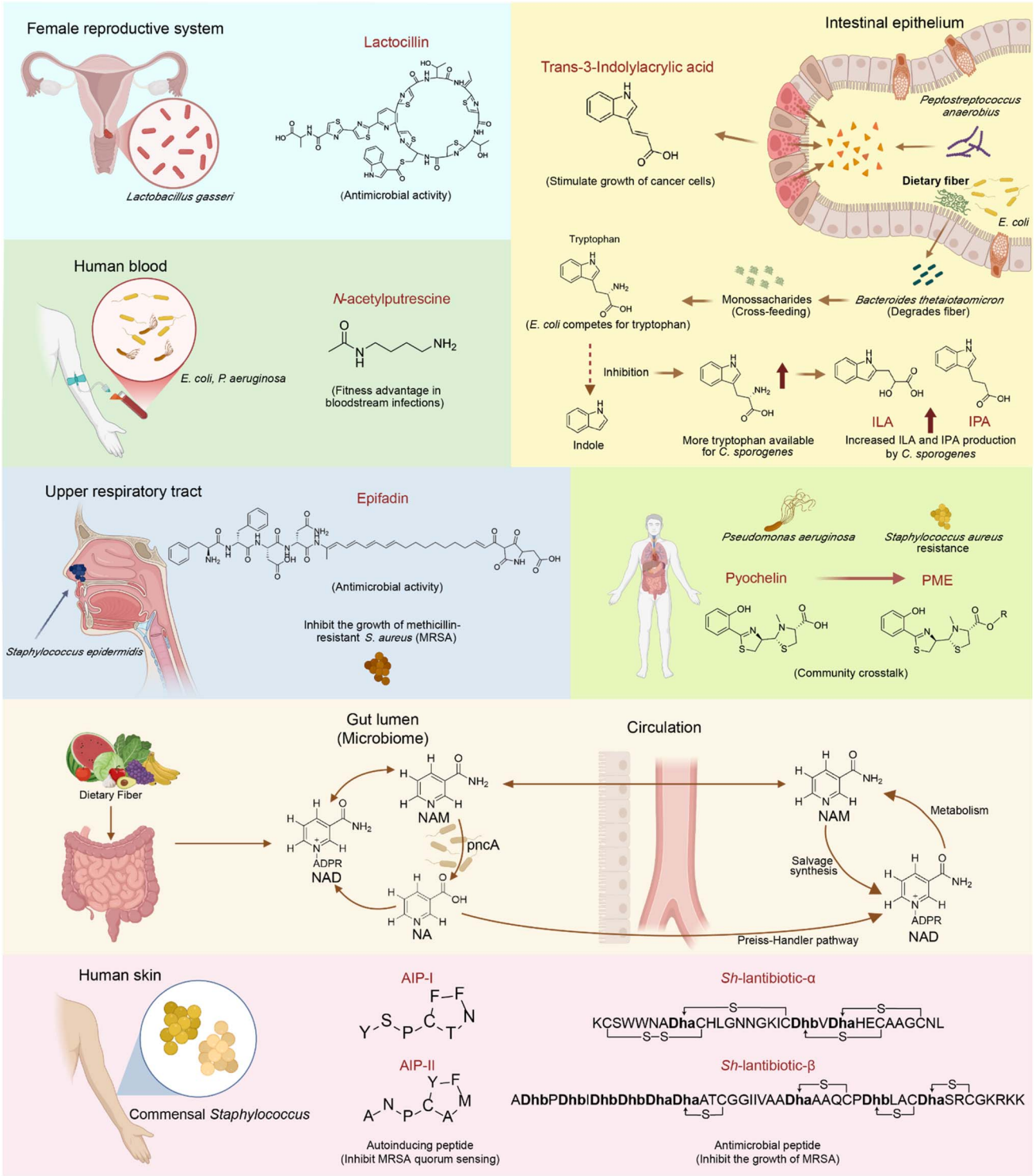

Fig. 3 Overview of select molecules and correlated functional activity related to host–microbe interactions discussed in this review. Created with https://www.biorender.com.

signatures were found in an organoid study resulting from past exposure to bacteria containing the *pks* island, furthering the link with colorectal cancer.[49] Structural characterization of biosynthetic intermediates and DNA-crosslinked adducts provided important clues into the structure of colibactin.[50–57] Eventually, the structure of colibactin was elucidated using a combination of genetic deletions in the colibactin BGC, stable isotope labeling of precursors, tandem mass spectrometry, and chemical synthesis.[55,56,58–62] Patients with cystic fibrosis (CF), a disease known to have dysfunctional dense mucus, who were

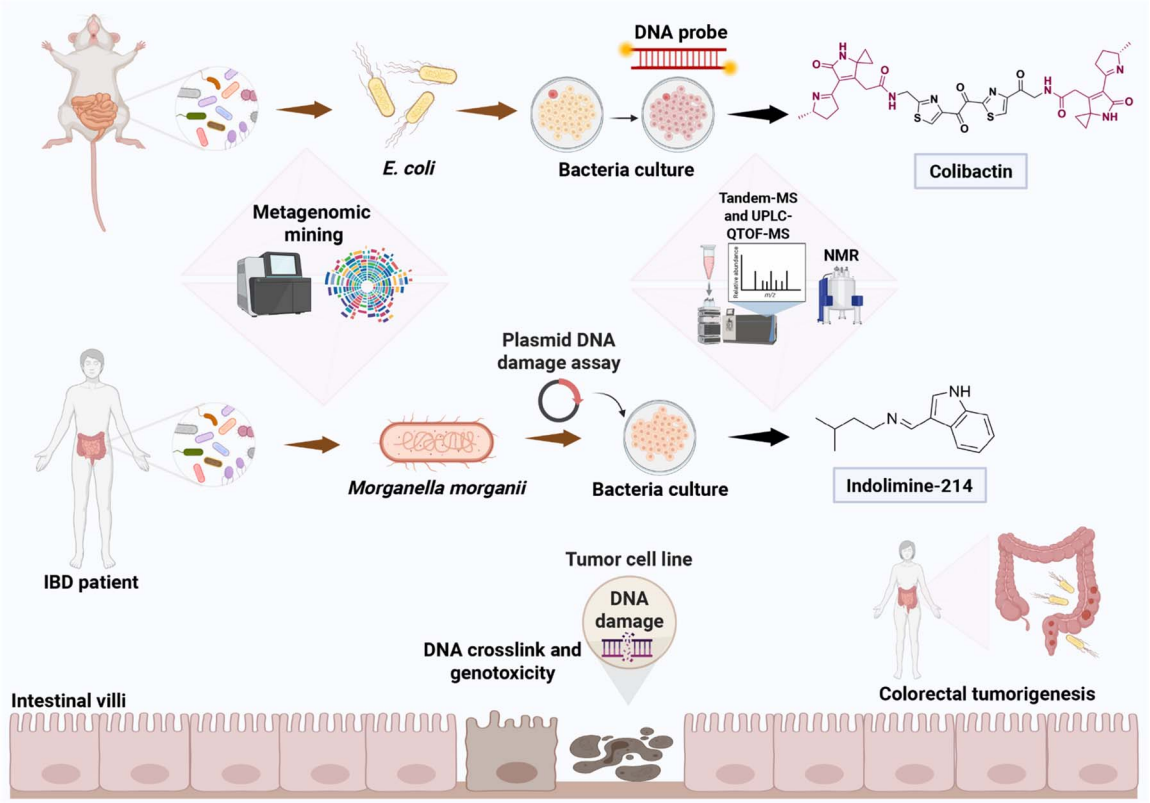

Fig. 4 Workflow of colibactin and indolimine-214 discovery from *E. coli* and *Morganella morganii* cultures and their implication in colorectal cancer tumorigenesis. This figure presents the workflow used to discover colibactin and indolimine-214, genotoxins produced by certain *E. coli* and *M. morganii* strains linked to colorectal cancer. Colibactin and Indolimine-214 induce DNA interstrand crosslinks and double-strand breaks, thereby contributing to tumorigenesis through genotoxic mechanisms. Created with https://www.biorender.com.

not taking CF transmembrane conductance regulator (CFTR) modulators, were found to have a higher frequency of colibactin-induced DNA adduct formation compared to other CF and non-CF patients. Notably, a patient with both CF and colon cancer harbored *E. coli* with the BGC for colibactin, further extending the implications of the prominent discovery of colibactin.[63]

Advances in metabolomics tools, such as Global Natural Product Social Molecular Networking (GNPS, discussed below and extensively reviewed elsewhere),[64,65] has made it possible to harness the power of database annotations to aid in discovery of novel microbial biotransformations.[11,66] Pyochelin, a siderophore produced by Pseudomonas and certain Burkholderia species, is a well-established virulence factor.[67,68] Molina-Santiago *et al.* found the presence of the methyl ester form of pyochelin (PME) in the chemical interactions of plant pathogens *Pseudomonas chlororaphis* PCL1606 and *Bacillus amyloliquefaciens* FZB42 (Fig. 3).[69] The annotation was deposited to GNPS for PME and was encountered a few years later by two separate groups studying *Pseudomonas aeruginosa* and *Staphylococcus aureus* from the perspective of co-infection and co-existence, respectively.[70,71] Jenul *et al.* went on to use transposon library screening of 576 mutants *via* LC-MS to identify the gene encoding a methyltransferase in *S. aureus* that catalyzed the formation of PME from pyochelin, thereby abrogating its ability to chelate iron and conferring protection to *S. aureus* against *P. aeruginosa*.[70] Methods using transposon libraries have been further advanced to identify production of cryptic NPs.[72] Currently, screening of transposon libraries by LC-MS is limited due to the associated cost and time required to screen thousands of clones in the absence of a visual or measurable phenotypic change. This limitation is alleviated by applications of methods such as Desorption Electrospray Ionization-Mass Spectrometry Imaging (DESI-MSI) capable of acquisition speeds of under a minute per sample, which was recently applied to profile free-fatty acids produced by engineered *E. coli*.[73] A comprehensive discussion of other high-throughput approaches for screening mutant libraries has been published by Shepherd *et al.*[74]

Once microbial communities and candidate genes producing specialized metabolites of interest are identified, techniques incorporating spatial visualization of these communities in their environment can aid in correlating the site of BGC expression to its cognate microbial producer. Fluorescence *in situ* hybridization (FISH) is a technique that consists of fluorescent DNA probes that can bind specific DNA, including ribosomal DNA that enables species identification.[75] The colocalization of these probes with microbes can be visualized by microscopy. FISH was employed to confirm the cyanobacterial origin of a polychlorinated NP, barbamide, in the

Table 1 List of selected databases supporting analyses of microbial products

Database	Common acronym (if applicable)	Type of data	Recent example
Dictionary of natural products[211]	DNP	Database containing data for over 340 000 natural products, with the ability to search using structures, melting/boiling point data, chemical names, molecular formulas, and CAS numbers	Observed the metabolite variation and bioactivity of a co-culture of two marine Aspergillus species[212]
Global natural product social molecular networking[11]	GNPS	User-submitted MS/MS spectra of tissue/microbial samples and their associated metadata	GNPS-guided discovery of madurastatin siderophores from the termite-associated *Actinomadura* sp.[213]
AntiBase[214]	N/A	Compendium of natural products and their compound properties, physicochemical data, and NMR/MS spectra	Analyzed aromatic polyketides produced by *Streptomyces* sp. PU-MM59 (ref. 215)
SciFinder[216]	N/A	Extensive database of compound properties, references, and structures for a large range of compounds, including those of microbial origin	Identified the activity and novelty of microbially-derived metabolites from *Crocodylus porosus* gut flora[217]
Human metabolome database[218,219]	HMDB	Comprehensive catalogue of metabolites known to be present in humans, including those derived from microbes native to the human microbiome	Linked metabolites present in the breast muscles of chickens to their genotype-dependent cecal microbes[220]
MicrobeMASST[90]	N/A	Query language for GNPS that identifies microbially-derived metabolite MS/MS spectra from microbial monoculture extracts. Linked to taxonomic data of the original sample	Performed functional metabolomics of the human scalp[91]
Human microbial metabolome database[221]	MiMeDB	Multi-omics database combining the taxonomic, phenotypic, and genotypic data related to a microbe with the metabolites produced by each microbe as well as any suspected or proven health effects on humans. Limited to microbes known to reside in humans	Associated gut microbiota features and circulating metabolites with systemic inflammation in children[222]
Streptome DB[82,223]	N/A	Database of natural products derived from Streptomyces, phylogenetic trees, predicted BGCs, MS spectra, and natural product–organism relationships	Identified potential inhibitors of a key protein of *H. pylori*[97]
Natural product atlas[224,225]	NPAtlas	Collection of known microbial products connected with their known source organisms, structure, and associated publications. Integration with MIBiG, NPClassifier, and ClassyFire	Identified potential inhibitors of Bcl-2 protein[98]
Natural product activity and species source[226,227]	NPASS	Database focusing on the known/estimated activity against targets of interest, native concentrations, and drug-likeness of natural products	Determined new inhibitors of MurG[228]
Pseudomonas aeruginosa metabolome database[229]	PAMDB	Compiled resource containing compound descriptions, MS spectra, molecular pathways, associated genes/proteins, and physicochemical data for metabolites derived from *P. aeruginosa*	Analyzed the different metabotypes of *P. aeruginosa* associated with different types of cystic fibrosis virulence[88]
COlleCtion of Open NatUral prodUcTs[230]	COCONUT	Web-based compendium of natural product data from open-source databases, like NPAtlas and NPASS	Determined the effect of media conditions on metabolites produced by *Colletotrichum* spp.[89]

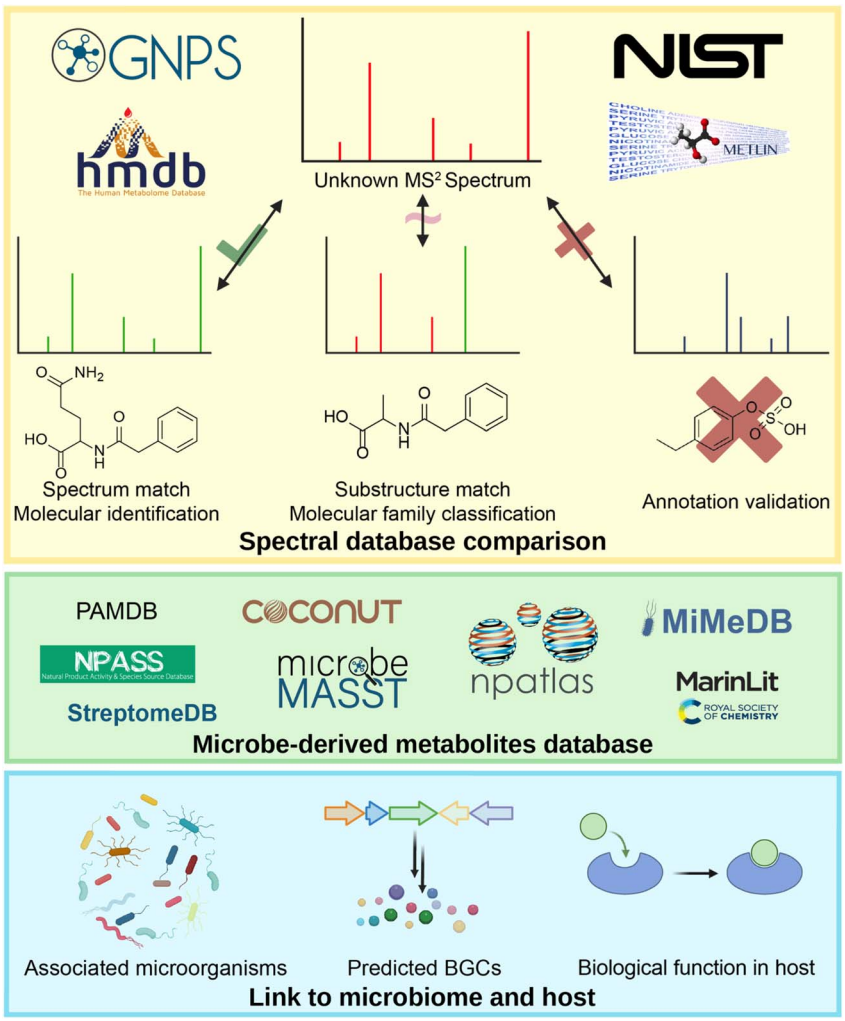

Fig. 5 Identification of metabolites and their links to the microbiome and host *via* online metabolomic databases. Created with https://www.biorender.com.

holobiont of the marine sponge *Dysidea herbacea*.[76] When combined with spatial metabolomics (discussed in the 'Spatial metabolomics in host–microbiome studies' section), FISH can help link the metabolome of a microbe to its cognate producer as exemplified for the investigation of chemical interactions between a deep-sea mussel and its intracellular symbiont community.[77,78] These examples from the marine microbiome provide parallels and serve as an inspiration for their application in human microbiome research, particularly due to the ease in acquiring metagenomic data and availability of robust BGC predictive tools.[44,79] Traditionally, FISH was limited to the analysis of two to three microbes or genes, due to lack of dyes with non-overlapping colors. However, advances in spectral imaging and algorithms for spectral unmixing of fluorophores with overlapping spectra have made it possible to incorporate two or more fluorophores per microbe, achieving combinatorial labeling and greatly increasing the number of species that can be distinguished. This method, also called Combinatorial Labeling and Spectral Imaging (CLASI), combined with FISH, has been used to analyze the spatial organization of microbe communities in complex microbiomes such as shown for human oral plaques.[80] These methods demonstrate the presence of an organism, its biosynthetic potential to produce a molecule and the neighboring microbial community structure that may play a role in the expression of the BGC, while mass spectrometry approaches such DESI-MSI can prove the direct production of a molecule in the context of interest. Lastly, cell sorting and single cell genomics can also enable linking of a molecule to its BGC.[81]

3 Database-aided identification of microbiome-derived metabolites

After extraction of metabolites from a sample and MS-based data acquisition, dereplication is a crucial step in metabolomics analysis. Dereplication is the classification and structural annotation of unknown metabolites using compound characteristics such as m/z, retention time, UV/Vis absorption spectrum, isotope patterns, collision cross section, and spectral fragmentation patterns. One way of performing dereplication of

known compounds is through the use of MS/MS databases (Fig. 5). In recent years, online MS/MS databases have rapidly expanded in terms of library size and associated tools, due to openly available, user-submitted MS data, and advances in computational power and machine learning-based analytical techniques in the analysis of mass spectrometry data.[82–87] Commonly used databases include Metlin, HMDB, DNP, GNPS, NIST, AntiBase, and MarinLit, to name a few. Each database has a variety of information, including spectra and supporting metadata, making them a valuable resource for both structural and functional annotation as well as source tracking (identify biological producer or biotic source of a given metabolite) when provided (Table 1).

As recognition of the role of microbial small molecules in human health and the bioeconomy increased, the demand for databases specific to microbially-derived metabolites led to the development of MiMeDB, COCONUT, StreptomeDB, NPAtlas, NPASS, and PAMDB (Table 1). These databases offer detailed information on microbial NPs and their producers, such as predicted BGCs, phylogenetic information, and known activity in humans (Fig. 5). PAMDB has been used to identify molecular pathways significantly altered in *P. aeruginosa* isolated from the sputum of cystic fibrosis patients.[88] These *P. aeruginosa* isolates were categorized into "metabotypes" that correlated with virulence based on which molecular pathways were altered. Using COCONUT and NPAtlas, Reveglia *et al.* identified microbially-derived metabolites affected by different media conditions in three strains of *Colletotrichum* spp. known to cause legume disease.[89] Furthermore, development of the search query tool microbeMASST, which searches MS/MS data against datasets for thousands of microbial monocultures present in the GNPS ecosystem, has provided researchers with the ability to query a MS/MS spectrum of interest and receive an output with interactive taxonomic trees without the need for prior molecular annotation.[90] The application of this search query tool led to the discovery of microbially-derived molecules in human scalp samples that correlate with scalp oiliness and in skin samples from patients with atopic dermatitis (AD).[91,92] Online repositories of data can also be used to identify undiscovered modifications of known microbially-derived molecules. After creating a library of MS/MS spectra of bile acids and their potential modifications from publicly available data in GNPS, Mohanty *et al.* utilized microbeMASST to identify monocultures of bacteria that produce previously unidentified modifications of microbially-derived bile amidates.[93] Indeed, 92 potential modifications of bile amidates were matched with a quarter of the monocultures present in microbeMASST, with two-thirds of these matches belonging to previously uncharacterized bile amidates.

Molecular networks, such as those generated by GNPS, while enabling dereplication of known NPs, can also provide a systems-level overview of the metabolites from metabolomics experiments.[11,94] Molecular networks rely on the similarities of MS/MS fragmentation patterns between molecules belonging to the same molecular family, such as peptides, glycolipids, and quinolones.[95] As a result, molecules can be grouped into molecular families based on their MS/MS patterns, and mass differences between annotated features and unannotated features within a molecular family can be used to annotate previously unidentified molecules.[11,95] For example, using AntiBase and SciFinder to dereplicate MS data and GNPS to create molecular networks, Rak Lee *et al.* showed that a termite-associated *Actinomadura* sp. RB99 produces families of polybrominated and polychlorinated isoflavonoids with antimicrobial activity.[96]

Beyond simple annotation and structure identification, databases have also proven useful in identifying potential roles and applications of microbe-derived products. Noori Goodarzi *et al.* used specialized metabolites present in StreptomeDB to perform docking simulations on *Helicobacter pylori*, an increasingly antibiotic-resistant bacterial species with a role in gastrointestinal disease.[97] 36 potential inhibitors of a critical protein present in *H. pylori* but not in humans were identified using this docking approach, two of which showed favorable properties for *in vivo* testing. Similarly, Almansour *et al.* screened compounds from NPAtlas to identify potential inhibitors of Bcl-2 protein, a critical anti-apoptotic protein involved in carcinomas.[98] This approach led to the discovery of the *Streptomyces* spp. NP saquayamycin F as a potential candidate for Bcl-2 inhibition.[99]

Newer methods such as reverse metabolomics utilize the large repositories of publicly available untargeted metabolomics data in databases such as GNPS, to allow association of a given phenotype with compounds *via* matching the mass spectral signature of a known class of compounds. For example, Gentry *et al.* used combinatorial chemistry to synthesize classes of metabolites such as *N*-acyl amide and bile acid conjugates, and acquired MS/MS spectra on the synthetic library.[22] Bioinformatic analysis using GNPS revealed the abundance of these molecules in existing public metabolomics datasets, and found associations of bile acid profiles with IBD subtypes. For compound classes where large numbers of compounds can be synthesized using combinatorial synthesis, reverse metabolomics will lead to increased annotation rates and enable discovery of new microbiome-related metabolites relevant to human health and disease.[22]

4 Spatial metabolomics in host–microbiome studies

The LC-MS-based approaches described above are inherently agnostic to spatial distribution of metabolites, which can provide another layer of confidence in metabolite annotations and their role in host–microbe, and microbe–microbe interactions. By integrating the analytical power of metabolomics with spatial imaging methods, one can uncover intricate chemical landscapes and gain critical insights into the role of metabolites in cellular crosstalk. Owing to these advantages and advancements in instrumentation, mass spectrometry imaging techniques using ionization methods such as Matrix-Assisted Laser Desorption/Ionization (MALDI), DESI, Secondary-Ion Mass Spectrometry (SIMS), and Laser Ablation Electrospray Ionization (LAESI) have grown in popularity. These methods have

been extensively reviewed elsewhere and provide spatial resolutions as high as 0.1 μm for SIMS, <100 μm for LAESI, 10–200 μm for MALDI and DESI (with atmospheric pressure MALDI and nanoDESI achieving <10 μm).[100–103]

Direct application of MALDI-MSI to cultures grown on solid media has become a widely used method.[104–107] Aiosa *et al.* recently used this method to study the interactions between a candidate airway probiotic from the *Bacillus* genus and *B. thailandensis*, a surrogate for the causative agent of melioidosis *B. pseudomallei*.[108] This technique, coupled with untargeted metabolomics, mapped molecules made by both the pathogen and potential probiotic in and around a zone of inhibition observed in an agar diffusion assay. Amongst the molecules found were *Bacillus*-produced fengycins, surfactants, and bacilomycins, and pathogen-produced capistruins. While many compounds were produced in both mono and co-culture, capistruins, known to have antibacterial activity, were induced in co-culture, and their diffusion captured using MALDI-MSI. These results demonstrate the utility of visualizing microbial interactions at the compound level. Geier *et al.* took MSI one step further by incorporating FISH to label members of a bacterial community in a deep sea mussel.[77,78] The authors optimized their MALDI method to minimize sample ablation so they could use the same tissue section for FISH labeling. Their high spatial resolution facilitated differentiation of host and symbionts. It enabled them to find differences in the localization of hopanoid molecules in the gills of fish at a micrometer scale, and to discover a new molecular family co-localizing with the symbiont communities. While the authors were unable to elucidate the structures of the molecules, molecular networking provided insight into the transformations that linked each molecule in the network. Single-cell metabolomics has been used to study the effects of *Trypansoma cruzi* intracellular parasite infection on parasite-containing and neighbouring host cells.[109] A similar approach could be implemented in the context of host–microbiome interactions, including application of the new organ-on-a-chip model system, as described later in the 'Host–microbe co-culture metabolomics and tissue culture models' section.

Complementing MALDI at spatial scales of the centimeter and above, spatial metabolomics techniques involve a three-dimensional reconstruction of organs with an overlay of metabolites detected using LC-MS. Spatial resolution is achieved by analyzing tissue sections. Spatial mapping played an important role in understanding the relationship between the location of the microbiome and metabolome in SPF mice across all organ systems, which set up the discovery of novel bile acid conjugates by Quinn *et al.*, as described below.[110] The lung of a late-stage CF patient was used to create a 3D model to visualize microbial localization and antibiotic penetration, and to map metabolites made by the host, microbes, and those derived from metabolism of antibiotics.[111] The lung samples, dominated by *Pseudomonas*, showed a diversity of quinolone molecules, including the 4-hydroxy-2-alkylquinolines (AHQ), but did not show the specific *Pseudomonas* quinolone signal (PQS) molecule produced downstream of the AHQs, pointing towards differences between results observed in culture models compared to *in situ* models. Moreover, the study showed a negative spatial correlation of *Achromobacter* species with the antibiotic meropenem, showing how pathogens in a diverse environment can find their ecological niche and survive. A follow-up study looked at six explanted lungs and found patient-specific differences in microbial and metabolome composition, but surprisingly found more intrapatient than inter-patient differences in distribution of microbial NPs. The localization and abundance of certain microbial quinolones correlated with the presence of *Staphylococcus*.[112]

Spatial mapping also provided insights into the metabolomic signatures associated with *Trypanosoma cruzi* eukaryotic parasites in the heart and gastrointestinal tract. *T. cruzi*, a protozoan and the causative agent of Chagas disease, was preferentially localized at the base of the heart. Metabolomics analysis of uninfected mice revealed metabolite gradients that either correlated with the parasite presence (eicosanoids) or anti-correlated with the parasite (adenosine monophosphate (AMP)).[113] Acylcarnitines were also perturbed by infection in a spatially-defined manner in both the heart and GI tract.[114,115] The hypothesis that perturbations in acylcarnitine levels play a key role in the disease inspired the authors to evaluate carnitine supplementation as a treatment strategy. Supplementation of carnitine in drinking water prevented mouse mortality in the acute stage of infection.[114] This study highlights the value of spatial methods in studying host–microbe metabolic interactions and informing possible therapeutic interventions.

5 Monoculture, microbe–microbe co-culture, and synthetic community metabolomics

To identify the source of microbial products, microbes can either be cultured in isolation (axenic cultures) or in the presence of one or more co-occurring microbes (co-cultures and synthetic communities).[116–118] The metabolite output of a microbe in isolation differs significantly from when it is associated with a host or is a part of a microbial community, which is why many BGCs are silent under laboratory conditions.[119–121] Cell-to-cell signaling mechanisms such as quorum sensing play an important role in regulation of gene expression in a cell density-dependent manner, and control metabolite production.[122] In the absence of the appropriate chemical signals, production of many metabolites remains below the limit of detection.[123] In mixed cultures, NPs control community structure and BGC expression, as shown for the synthetic 'The Hitchhikers Of the Rhizosphere' (THOR) *in vitro* community designed from the soil microbiome.[118] Notable *in vivo* synthetic communities that have been developed to model the mammalian microbiome include Oligo Mouse Microbiota (OMM) and hCom2.[124,125] Simulator of Human Intestinal Microbial Ecosystem (SHIME®) is an example of an *ex vivo* model of the gut synthetic microbial community that uses multi-compartment reactors to mimic the entire gastrointestinal tract.[126]

Biotransformations and conjugations between metabolites of two bacteria can further produce additional metabolites with unique biological functions, emphasizing the importance of mixed cultures. For instance, when *Pseudomonas aeruginosa* and *Burkholderia thailandensis* are cultured in isolation, the NP malleonitrone is not detected. The production of this compound is observed only in co-culture.[119] When the monoculture of *B. thailandensis* is supplemented with IQS (2-(2-hydroxylphenyl)-thiazole-4-carbaldehyde), a quorum sensing small molecule signal produced by *Pseudomonas aeruginosa*, production of malleonitrone is observed. Structural characterization revealed that malleonitrone is derived by conjugation between IQS produced by *P. aeruginosa* and malleobactin produced by *B. thailandensis*. Similarly, incubating soil *Streptomyces* with compounds containing catechol moieties induces the production of catechol-containing siderophores by *Streptomyces*.[127] Similar to the phenomenon of substrate-dependent regulation of specific metabolite pathways, indole-producing *E. coli* and indolelactic acid (ILA)- and indolepropionic acid (IPA)-producing *Clostridium sporogenes* competed for tryptophan within a three-species community including *Bacteroides thetaiotaomicron* (Fig. 3). Fiber-degrading *B. thetaiotaomicron* influenced the competition by cross-feeding monosaccharides to *E. coli*. Thus, indole production through catabolite repression was inhibited, which in turn made more tryptophan available to *C. sporogenes*. This observation was confirmed with increased ILA and IPA production by *C. sporogenes*.[128] Thus, cross-feeding between microbes and host plays an important role in synthesis of NPs, which further structure the composition of the microbial community present. Additionally, genomics and metabolomics were applied to reveal the genes involved in the production of IPA (Fig. 3). The authors compared metabolomes of a genetically modified strain with a disruption in a phenyllactate dehydratase subunit *fldC*, hypothesized to be involved in IPA production from tryptophan, with the wild-type strain using LC-MS/MS.[129] It was shown that in addition to being involved in IPA production, FldC was also involved in reductive metabolism of tyrosine and phenylalanine, indicating a shared pathway for reductive aromatic amino acid metabolism in *C. sporogenes*. Thus, tryptophan metabolism by the microbiome plays an important role in host–microbiome interactions. Co-culture-based interactions can also be dependent on the environmental context. Recently, Jin *et al.* showed that expression of the *ham* BGC in *Burkholderia cenocepacia*, in the presence of clinically prescribed antibiotic trimethoprim, affected its chemical interactions with *Aspergillus fumigatus*.[130] The induction of *B. cenocepacia* NPs was linked to increased production of the fungal siderophore triacetylfusarinin C, biotransformation of fungal membrane ergosterol, and inactivation of the *ham* BGC NP fragin *via* methylation by the fungus. In studies of the coral holobiont, Moree and coworkers demonstrated the impact of light exposure on the metabolic output of the commensal coral microbiota and underlying mechanisms to prevent growth of fungal pathogens.[131] A *Pseudoalteromonas* sp. isolated from a healthy octocoral was cocultured with a marine fungus, *Penicillium citrinicum*, isolated from the necrotic tissue of a diseased gorgonian octocoral. MALDI-IMS revealed that Pseudoalteromonas produce antifungal polyketide alteramides at higher levels in the dark than in light. The alteramides undergo a photoinduced intramolecular cyclization, forming an inactive congener in the presence of light. As corals are nocturnal feeders, it is hypothesized that they may be more vulnerable to fungal infections at night and thus may benefit from specific production of an antifungal molecule in dark conditions.

6 Host–microbe co-culture metabolomics and tissue culture models

Beyond microbial co-culture systems, microbe–host co-cultures can further elucidate the specific metabolite signals involved in health and disease. Such co-culture systems have been useful to identify signals differentially produced in the presence of hosts, and to interrogate immunometabolic crosstalk. For example, the brain and muscles of the ant *Camponotus pennsylvanicus* were dissected and kept alive in an insect cell culture medium. The vegetative mycelia of two fungal species, *Beauvaria* and *Metarhizium*, were introduced to this ant *ex vivo* tissue culture to simulate an ant infection. A different repertoire of specialized metabolites were produced in co-culture with insect tissue rather than in regular fungal media, resulting in the description of several new analogs of previously known specialized metabolites responsible for ant killing activity, namely destruxins and beauverolides.[132]

Another study employed co-culturing of *Burkholderia thailandensis*, a model bacterium for the disease melioidosis, intracellularly within murine epithelial and macrophage lines.[133] Monocultures, including naïve host cells and *B. thailandensis*, were used as controls. Metabolites structurally similar to romidepsin, namely burkholdacs and spiruchostatin C, a potent inhibitor of histone deacetylases (HDACs), were detected in both epithelial cells and macrophage cells.[134–136] HDACs regulate gene expression by removing acetyl groups from histone proteins, which makes the DNA more compact and less accessible for transcription. HDACs are known to inhibit phagocytosis of pathogens by macrophages.[137] The BGC for these compounds are lowly expressed under laboratory conditions and require manipulations such as addition of elicitors in culture or over expression of transcriptional regulators.[138,139] This study suggests the use of mammalian cell co-culture as a relevant and useful approach to discover specialized metabolites from pathogens with specific phenotype such as inhibition of HDACs, and that *in vitro* mammalian cell culture provides the right environment for the inductions of virulence factors.

Additional examples of such systems emerge from the studies involving the gut microbiome. The communities of microorganisms in the small intestine play an important role in host physiology. *Bifidobacterium* is one of the most dominant groups of bacterial genera residing in human gut microbiota and has been used industrially as probiotics. A two-chamber co-culture device containing *Bifidobacterium breve* MCC 1274 and

intestinal epithelial cells derived from induced pluripotent stem (iPS) cells to mimic the environment of small intestinal tract was used to investigate immunometabolic crosstalk. Based on metabolite analysis, a significant increase in immunomodulatory metabolites, including microbially-derived indole-3-lactic acid and phenyllactic acid, were observed, representing a synergistic interaction between probiotics and intestinal epithelial cells.[140]

Other successful iterations of this method are organ-on-a-chip platforms aiming to recapitulate the 3D topography of organs and their physicochemical cues.[141] Lucchetti *et al.* combined *E. coli* bacteria with a human microbial-crosstalk (HuMIX) gut-on-chip and a liver-on-chip, representing a multi-organ-on-a-chip platform, to assess the metabolism of the colorectal cancer drug irinotecan. Based on LC-MS/MS analysis, the presence of *E. coli* enabled transformation of the inactive irinotecan metabolite SN-38 G into the toxic metabolite SN-38.[142] Wang *et al.* likewise performed gut-on-a-chip analysis, co-culturing aerobic human intestinal epithelial cells and anaerobic gut microbiota from individuals with depression.[143] Analysis of metabolites in the co-culture revealed a significant increase in tyrosine, and a significant decrease in L-proline, L-threonine, and the secondary bile acid hyodeoxycholic acid. Further, reduced gut barrier function, chronic low-grade inflammatory responses and decreased neurotransmitter 5-hydroxytryptamine levels were also discovered, signatures that are indicative of depression.[143] Continued development of new *in vitro* systems, including optimized organs-on-a-chip and organoid systems mirroring the complex multicellular organization of tissues will facilitate the elucidation of host–microbe cross-talk. In this context, human immune organoids may be particularly promising.[144]

7 Linking phenotype to microbial chemistry

Host or microbial phenotypes are observable characteristics that can be a representation of the biological response to a stimulus. Thus, investigation of disease and beneficial phenotypes by mass spectrometry metabolomics can provide useful hypotheses about the role of metabolites in association with host–microbe interactions. Correlative analysis between a specific disease phenotype and a bioactive compound can be performed by screening a library of compounds, or a library of niche and disease relevant microbial isolates, or by using samples from disease-specific animal models or human samples. Select examples from the last decade that combined these approaches with metabolomics are discussed in this section. For example, the study of ferroptosis, an iron-dependent cell death, by Cui *et al.* brings forth the usefulness of screening endogenous metabolite libraries.[145] The authors applied a ferroptosis assay to screen a library containing hundreds of endogenous human host and gut microbe metabolites using 786-O human kidney adenocarcinoma cells. This screen identified a specific tryptophan metabolite derived from *Peptostreptococcus anaerobius*, *trans*-3-indoleacrylic acid (IDA), which mediated ferroptosis resistance in cancer cells (Fig. 3). 16S rRNA sequencing of human faecal samples found bacteria enriched in CRC, which, combined with LC-MS based metabolomics on candidate bacterial isolates, was used to identify the microbial source of IDA. The observations from these analyses were validated using xenograft models and in $Apc^{Min/+}$ mice wherein supplementation of IDA in either model or gavage of *P. anaerobius* into the mice promoted progression of tumors.

Another approach screened a library of bacterial cultures to identify small-molecule metabolites associated with DNA damage from a diverse set of gut microbes. Using a plasmid-based DNA damage assay, Cao *et al.* discovered a new class of genotoxins, namely indolimines, produced by the gut bacterium *Morganella morganii* in patients with inflammatory bowel disease (IBD).[146] DNA damage by gut microbes can increase risk of colorectal cancer, exacerbated by the presence of IBD. After detecting genotoxicity from a *Morganella morganii* isolate that was not due to known genotoxins such as colibactin, the authors used untargeted metabolomics and bioactivity guided-fractionation of cultures to isolate and characterize indolimines (Fig. 4). MS-based quantification showed that *M. morganii* produced a high level of indolimines, which led to increased expression of the double strand break marker y-H2AX.[146]

In addition to screening bacterial and compound libraries, direct investigation of disease phenotype-relevant samples can advance the understanding of the link between disease and host microbiome. For example, type 2 Diabetes mellitus (T2DM) is associated with a higher risk of cardiovascular disease. However, control of blood glucose is not effective in reducing adverse cardiac events.[147,148] Nemet *et al.* analyzed a cohort of human subjects undergoing elective cardiac evaluation over a three-year period, using untargeted plasma metabolomics.[149] Metabolites were prioritized based on their increased abundance in those with T2DM and an association with major adverse cardiac events, while being poorly correlated with indices of glycemic control such as fasting glucose levels, HbA1C levels or glucose/insulin ratio. A metabolite with *m*/*z* 265.1188 was identified and annotated as phenylacetylglutamine (PAGln) through comparison of MS/MS fragmentation pattern and retention time with a standard. The role of the gut microbiome in generating phenylacetic acid, the precursor of PAGln, was confirmed by metabolomics in human subjects who presented substantially reduced levels of PAGln following an antibiotic course. Three weeks after the discontinuation of antibiotics, restoration of PAGln levels was observed. This initial discovery using untargeted metabolomics was further verified using stable isotope dilution LC-MS/MS with deuterated PAGIn as an internal standard.[149]

Recently, a similar approach was used to identify the fungal specialized metabolite AT-C1, which was linked to polycystic ovary syndrome (PCOS). The fungus *Aspergillus tubingensis* was found at higher levels in patients with PCOS and induced a PCOS-like phenotype in murine models. *A. tubingensis* targets interleukin-22 (IL-22) secretion by inhibiting the aryl hydrocarbon receptor (AhR). Untargeted metabolomics data from 23 *A. tubingensis* strains was correlated with AhR inhibition, to identify a candidate *m*/*z* responsible for this inhibition. This

analysis was followed by isolation, MS/MS analysis and NMR for structure elucidation. The identified compound was named AT-C1 and was found at higher levels in the feces of PCOS patients, with correlation to immunological and clinical markers of PCOS.[150]

The study of the gut microbiota–brain axis has also brought forward several examples of the role of gut metabolites in modulating brain health.[151] For example, during progression of the neurodegenerative disorder amyotrophic lateral sclerosis (ALS), changes in the composition of the gut microbiome are observed.[152] Using a superoxide dismutase mutant *Sod1*-Tg mice model, the levels of *Akkermansia muciniphila* (AM) were found to be decreased and AM supplementation mitigated motor degeneration in these mice. Untargeted serum metabolomics of AM-supplemented and control *Sod1*-Tg mice found nicotinamide to be significantly increased in the AM-supplemented group. Based on metagenomic data, nicotinamide had the highest probability to be synthesized by a wild-type microbiome but not the *Sod1*-Tg microbiome, suggesting that AM was responsible for increased nicotinamide levels. Targeted metabolomics showed significantly higher levels of nicotinamide in AM cultures compared to other commensals and further supported the link between AM supplementation and increased levels of nicotinamide in the cerebrospinal fluid of *Sod1*-Tg mice. An analysis of a cohort of human ALS patients using untargeted serum metabolomics further revealed significant changes in the tryptophan-nicotinamide pathway, including indoleacetate, kynurenine, serotonin and nicotinamide itself.[153]

Gut microbes may also play an important role in maintaining host nicotinamide adenine dinucleotide (NAD) homeostasis. To investigate *de novo* gut microbial NAD synthesis, Chellapa *et al.* fed mice a diet including stable isotope-labeled fiber and protein, and found that dietary fiber accounted for 37% of total carbons in NAD in the cecum, while protein only accounted for 3%.[154] This indicated that most of the dietary precursors of NAD are absorbed prior to reaching the end of the small intestine, pointing to an alternate precursor. Indeed, labeled NAM injected intravenously into mice was detected in the lumen along with labeled NAD. Interestingly, NAM was converted to nicotinic acid (NA) by luminal microbes *via* a deamidase activity that mammals lack. NA that in turn could feed into the host Preiss–Handler pathway to synthesize NAD, suggesting a synergistic host–microbe relationship (Fig. 3).[154]

Many pathogens can enter the bloodstream through invasive procedures such as surgery, patient care equipment (IV lines, ventilators), and even through skin and mucous membrane breaches. Mayers *et al.* developed an iterative metabolomics pipeline to prioritize metabolite features relevant to bloodstream infections by Gram-negative bacteria.[155] Their pipeline involved comparative metabolomics on human plasma samples of healthy controls *versus* infected individuals, to obtain a list of significant metabolite features. These features were filtered based on disease severity using a clinically-relevant score (Acute Physiology and Chronic Health Evaluation (APACHE II)). The authors then moved from this observational cohort to an experimental study using a mouse model of sterile injury *versus* live infection, to identify metabolite features that showed differential abundance in the same direction as the human plasma samples. As a final filtration step, to focus on features that would have a microbial origin, culture supernatants of *E. coli* were examined by metabolomics. Initially, this pipeline was used on a targeted set of features, where the authors found *N*-acetylputrescine as a metabolite significantly abundant in the infection condition that was likely microbially-derived (Fig. 3). Using the same pipeline but with an untargeted approach, they additionally found diacetylspermidine, another acetylated polyamine, which led them to investigate the possible role of polyamine acetylation. This strategy allowed the authors to link the enzyme, spermidine *N*-acetyltransferase (SpeG), encoded in the genome of *E. coli*, with production of acetylated putrescine. Notably, inhibition of these enzymes re-sensitized antimicrobial-resistant Gram-negative bacteria to antibiotics by increasing membrane permeability.[155] These findings demonstrate how an initial discovery made using metabolomics allowed hypothesis building, with subsequent validation extending the implication of this discovery in a clinical application *via* the development of alternate treatment strategies.

Another common phenotypic method involves bioactivity-guided assays to screen a potential specialized metabolite producer against other members of its community from a host-relevant environment. This serves two purposes. First, growing bacteria in an environment closer to their natural setting increases the chances of eliciting a silent BGC. Second, it can provide an accurate assessment of the biochemical interactions that occur in that environment. For example, *S. epidermidis* isolates from the human nose were identified as useful sources of antimicrobial compounds with activity against other members of the nose microbiome.[156] A particularly active strain, IVK83, inhibited a broad range of strains from various phyla. The authors generated and screened a transposon mutant library to find the BGC responsible for activity. This led to the discovery of a BGC that encoded for an unstable broad-spectrum antimicrobial peptide-polyene that could inhibit methicillin-resistant *Staphylococcus aureus* (MRSA).[157] While the culture extracts from IVK83 displayed activity, this activity was lost within a few hours upon exposure to light, pH and temperature conditions mimicking *S. epidermidis in vivo* habitats, suggesting the bioactive antimicrobial was unstable. Using reverse-phase HPLC-MS-UV, the authors correlated the fraction with the antimicrobial activity to the UV absorption wavelength, found the *m*/*z*, and purified the compound, which they called epifadin (Fig. 3). While its instability precludes its use as a drug, the authors proposed the commensal producers of epifadin as potential probiotics. *S. lugdunensis* IVK28 had similarly been found by the same group to have potent activity to inhibit *S. aureus* under iron-limiting conditions on solid agar, leading to the discovery of the peptide antibiotic lugdunin.[158] Broadly, coagulase-negative Staphylococcus (CoNS) commensals are an important part of the skin microbiome. Healthy patients have *S. epidermidis* and *S. hominis* clones on the skin with strain-specific activity against *S. aureus* at a higher frequency than those with AD.[159] These commensal species produced lantibiotics that could synergize with host antimicrobial peptides.

Additionally, autoinducing peptides from *S. hominis*, *S. simulans*, and *S. warneri* inhibit the *agr* quorum sensing system in MRSA (Fig. 3).[160–162] Notably, a phase 1 randomized clinical trial of AD patients consisting of topical therapy with *S. hominis* A9 was shown to be safe and beneficial for decreasing *S. aureus* abundance.[163]

Thus, with the appreciation that human microbiomes are a useful resource for medically important compounds, large bioinformatics studies have been carried out, which seek to mine available genomic datasets for possible NP BGCs. The type of NP being made can be predicted based on homology to known BGC types (*e.g.* RiPP, NRPS *etc.*). These BGCs can be expressed en masse in a heterologous host such as *E. coli*, and then their products can be purified, isolated, and screened for bioactivity. For instance, 70 unique lanthipeptide and lassopeptide gene clusters from the HMP were chosen by King *et al.*, their genes cloned and expressed heterologously in *E. coli*, leading to the purification of 21 lanthipeptides and 4 lasso peptides.[164] Of the 112 possible RiPPs in all the BGCs that the authors attempted to produce, modified peptides could only be detected in 27 of them. This highlights the challenges associated with optimizing conditions needed to produce a NP even after identifying the BGC. Furthermore, structural characterization of the modifications using mass shift data from LC-MS/MS was successful for 18 of the 27. Bioactivity against MRSA or vancomycin-resistant Enterococci was observed only in a subset of 6 RiPPs. However, the study does provide a framework that can be used for rapid discovery of NPs with undiscovered biological functions in the future.

8 Germ-free approaches to identify microbiome-derived metabolites

Germ-free (GF) mice are useful animal models to study the effects of microbiome absence and of microbiota on health, as they can be inoculated with host-derived and defined microbiomes. When colonized with defined communities (gnotobiotic mice), this system offers a unique way to control and tailor the microbiome to the specific requirements of a study. In particular, combining gnotobiotic animals and cultured microbial isolates allows for re-colonization experiments that help to unequivocally assign function to specific microbes and the produced metabolites (Fig. 6). On the other hand, the stringent conditions under which they need to be housed isolated from the environment, coupled with prohibitive costs, present a limitation in their use. Furthermore, GF mice have an under-developed immune system that may not fully replicate the host–microbe interactions observed in normal colonized individuals, even if subsequently colonized.[165] Nevertheless, GF mice have been extensively used to enhance our understanding of microbiota-produced and microbiota-modulated metabolites.[166] Lai *et al.* used comparative metabolomics on germ-free *versus* conventional specific pathogen free (SPF) mice to study the effect of microbiota on the metabolites involved in the gut–brain axis, using samples from feces, serum, and the brain.[167] The gut microbiota enriched numerous pathways involved in aromatic amino acid and neurotransmitter synthesis, and altered the transsulfuration, redox homeostasis, and neuroinflammation pathways in the brain. These interactions were correlated with the production of microbial-derived metabolites such as indoxyl sulfate and trimethylamine-*N*-oxide (TMAO).[167,168] TMAO, made from gut microbial metabolism of trimethyl quaternary ammonium compounds such as choline and L-carnitine to trimethylamine, followed by oxidation *via* host metabolism, was associated with an increased risk of atherosclerosis.[169,170] δ-Valerobetaine, another precursor of TMAO, was at a higher abundance in liver and liver mitochondria samples of conventionalized mice derived from GF mice housed with conventional mice, compared to mice that remained GF. Elevated levels of δ-valerobetaine inhibited mitochondrial fatty acid oxidation, leading to increased lipid accumulation in the adipose tissue of mice, suggesting a link to obesity. Notably, this effect was dependent on the type of diet fed to the mice, since supplementation of δ-valerobetaine in a western diet led to more weight gain than a control diet, indicating a complex regulatory network involving host, microbiota, and diet.[171]

Quinn *et al.* investigated the impact of the microbiome on the whole animal by analyzing various organs, skin, blood and stool of GF *versus* SPF mice using untargeted metabolomics, finding unique chemical profiles and significant differences in the metabolome at most of the sampled locations.[110] They discovered a previously uncharacterized conjugation of phenylalanine, tyrosine or leucine with cholic acid, in the duodenum, jejunum and ileum of SPF mice only. Interestingly, Mass Spectrometry Search Tool (MASST) searches showed that these transformations were also present in humans, with a high frequency in patients with IBD and CF. *In vitro* culturing of the Clostridium species correlated to conjugated bile acid levels in mice demonstrated the microbial origin of the metabolites, while exogenous administration of labeled amino acid to mice on a high-fat diet confirmed that diet-derived amino acids could be incorporated into bile acids. This study has inspired the discovery of hundreds of conjugated bile acids through *in silico* methods combined with searching in spectral databases and synthesis of standards.[22,66,93,172–179] The discovery of bile salt hydrolases with acyltransferase activity has led to the identification of additional bile acid modifications.[180,181]

Dohnalová *et al.* profiled the genomes, microbiomes and metabolomes of genetically diverse mice and compared them with regards to performance while running on treadmills and wheels.[182] They found that the gut microbial community by itself could predict performance. The authors investigated the molecular basis for these differences and found that there was activation of the striatum and an increase in striatal dopamine levels in the diversity outbred (DO) mice, coupled with decrease in monoamine oxidase (MAO) levels, which is important in motivating physical activity. GF mice had lower dopamine levels after exercise, which could be restored by supplementing with microbiome transplants. Upon discovering that neuronal calcium signaling in the dorsal root ganglion (DRG) was correlated with exercise performance, the authors carried out untargeted metabolomics on mice ceca and screened candidate

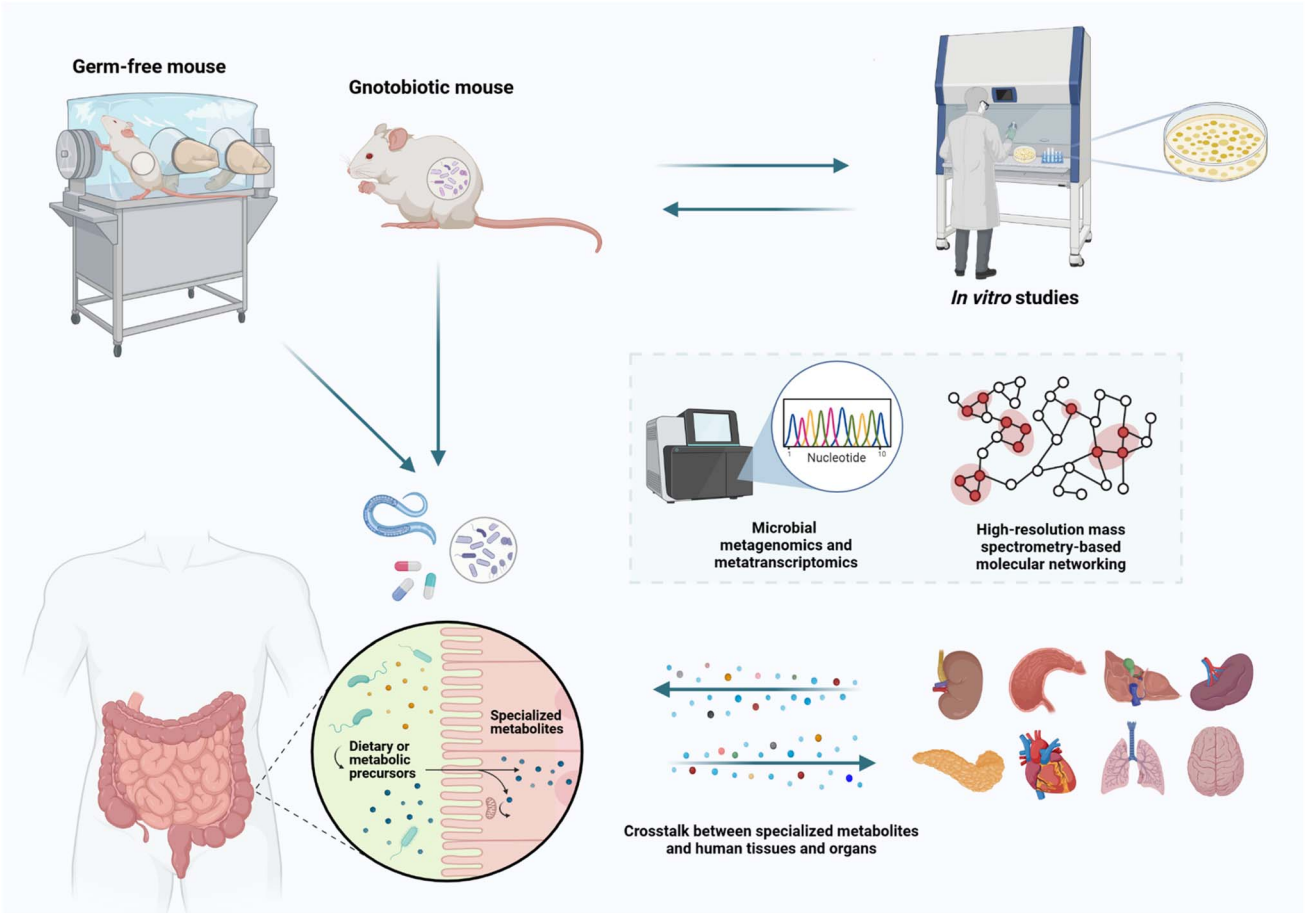

Fig. 6 Overview of germ-free (GF) and gnotobiotic mouse models as tools for studying microbiota-derived metabolites and their effects on host physiology. This figure illustrates the use of germ-free (GF) and gnotobiotic mice in both *in vitro* and *in vivo* studies, enabling the discovery and characterization of specialized metabolites produced by microorganisms that influence the function and physiology of various organs and tissues. GF mice, devoid of any microbial community, provide a valuable model to investigate the role of the microbiota in health and disease. In contrast, gnotobiotic mice, derived from GF mice colonized with a controlled microbial community, serve as robust systems for examining both host–microbe and microbe–microbe interactions. These models can be further analyzed using microbial metagenomics and metatranscriptomics to determine community composition and active genes, respectively, and mass spectrometry-based molecular networking to predict the structures and roles of metabolites within biological systems. Created with https://www.biorender.com.

molecules that could activate calcium signaling in DRG neurons. They found several fatty acid amides (FAA), including *N*-oleoylethanolamide (OEA), correlated with exercise performance. FAA dietary supplementation restored the exercise performance of mice, even in the absence of a microbiome. Moreover, colonizing GF mice with a strain of *E. coli* expressing FAA biosynthesis genes enhanced exercise performance compared to the empty vector control.[183] Further examination showed that fatty acid amides bind to the endocannabinoid CB1 receptor on vanilloid receptor (TRPV1+)-expressing DRG sensory neurons, to transmit signals to the brain and lead to downregulation of MAO levels in the striatum.

It is intriguing that microbial OEA plays a role in promoting physical activity, given that it is also produced endogenously and is known to regulate feeding.[184] Indeed, studies looking at the effects of a specialized diet on malnourished children found increased levels of *Faecalibacterium prausnitzii* in the gut microbiome of children with improved outcomes, and showed that the bacterium encodes a fatty acid amide hydrolase (FAAH) that can reduce levels of OEA.[185,186] MAGs assembled from stool samples of the study population were used to build two consortia of bacteria differing in the presence or absence of a strain of *F. prausnitzii* Bg7063 and were used to colonize gnotobiotic mice. Using untargeted metabolomics, *F. prausnitzii*-colonized gnotobiotic mice were found to have lower levels of *N*-acylethanolamides such as OEA and palmitoylethanolamide (PEA). Supplementation of stable isotope-labeled PEA in the exponential phase monocultures of *F. prausnitzii* led to degradation of PEA. Moreover, *F. prausnitzii* monocultures supplemented with PEA accumulated a product annotated as *N*-palmitoylarginine using tandem MS. Fractionation of active protein extracts and bioinformatic analysis led to the discovery of an FAAH enzyme with bidirectional activity that could both hydrolyze *N*-acyl amides and synthesize *N*-acylamino acids.[186]

N-Acyl amides are subdivided in many categories based on their amine headgroups and acyl tails (Fig. 7). They are

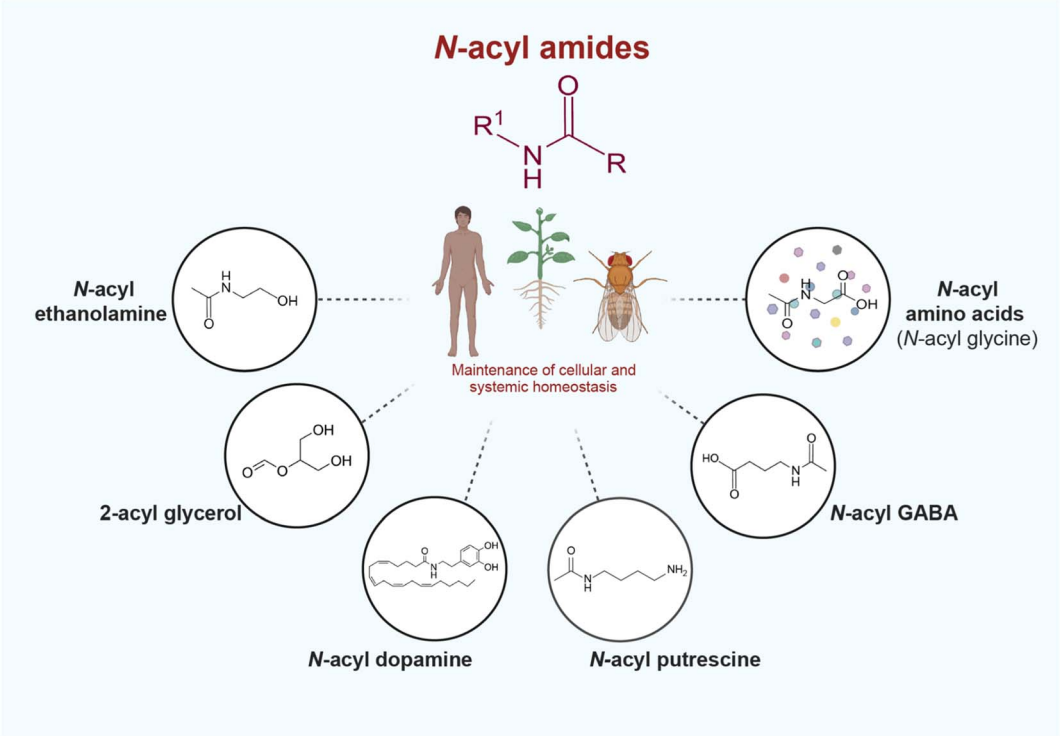

Fig. 7 Subdivision of classes of *N*-acyl amides that can be found in vertebrates, invertebrates and plants. *N*-Acyl amides comprise a general class of endogenous fatty acid conjugates with a fatty acyl group linked to a primary amine *via* an amide bond. *N*-Acyl amides participate in several biochemical pathways related to diverse physiological and pathological processes. Created with https://www.biorender.com.

produced by the host and function as endogenous signaling molecules.[187] However, they are also produced by microbes. For instance, Cohen *et al.* conducted a large metagenomics-guided, high-throughput fluorescence microscopy screen for nuclear factor-κB activators.[188] Active metagenomic clones were further subjected to transposon mutagenesis to find genes responsible for activity. One of the genes encoded an acetyltransferase-5 family enzyme, which produced the small molecule 3-hydroxypalmitoylglycine (commendamide). Commendamide activates GPCR G2A/GPR132, thereby affecting native GPCR signaling and playing a role in modulation of autoimmunity and atherosclerosis.[188]

Burkholderia cenocepacia, a deadly CF pathogen, constitutively produces ornithine lipids (OL) containing an ornithine linked to a 3-hydroxy fatty acyl group further esterified by a second fatty acyl group which is often enzymatically converted to the 2-hydroxy form.[189,190] Moreover, the amide-linked fatty acyl moiety can itself undergo enzymatic hydroxylation.[189] While the role of OLs in *B. cenocepacia* are yet to be fully explained, mutants lacking OL production showed reduced virulence in a *Galleria mellonella* infection model.[191] OLs may also aid Gram-negative bacteria to increase inflammation in mammalian hosts and evade the immune system through prevention of bacterial lipopolysaccharide recognition by Toll-like receptor 4.[192] Increasing interest in microbiome-derived *N*-acyl amides has led to the development of a spectral library resource using MASSQL to query MS/MS spectra against untargeted GNPS datasets.[193]

Gnotobiotic mice, derived from GF mice administered a controlled microbial inoculum, serve as reliable model systems to study both host–microbe and microbe–microbe interactions (Fig. 6). For example, *Limosilactobacillus reuteri*, a gut symbiont species used as a probiotic, exhibits intraspecies competition. A rodent isolate *L. reuteri* R2lc induced the production of IL-22 by activating the aryl hydrocarbon receptor (AhR). Using genomic approaches, the authors found a PKS cluster *pks* encoding for antimicrobial activity that was also responsible for AhR activation. *In vitro* competition assays showed that R2lc inhibited most of the other *L. reuteri* strains, while mutants with the *pks* cluster deleted resulted in loss of antimicrobial activity. Likewise, gnotobiotic mice with 1 : 1 mixtures of R2lc and a competitor strain showed that R2lc outcompeted the other strains throughout the GI tract. UPLC-MS and UV analysis were used to infer the molecular formula of the compound and that it was a polyene, though the exact structure remains unknown. Interestingly, certain *L. reuteri* strains resistant to the antimicrobial effect of R2lc were found to contain an acyltransferase gene that was responsible for acetylation of the C6 hydroxyl of *N*-acylmuramic acid in peptidoglycan that makes up the cell wall. Future studies would be needed to elucidate the mechanism by which cell wall acetylation confers protection against R2lc antimicrobial activity.[194]

A study using the maternal immune activation (MIA) model for atypical neurodevelopment in autism spectrum disorder (ASD) observed gut dysbiosis in the MIA mice. Treatment of mice with *Bacteroides fragilis* ameliorated GI defects and

corrected autism spectrum disorder-related behaviors. Comparison of the serum metabolome of treated *versus* untreated MIA mice with untargeted metabolomics using gas chromatography coupled with MS revealed that the metabolite 4-ethylphenyl sulfate (4EPS) was driving the significant differences between the two groups. Notably, serum levels of 4EPS were almost undetectable in GF mice compared to conventional SPF mice, supporting a microbial contribution. Systemic administration of 4EPS was sufficient to induce anxiety-like behaviour in MIA mice.[195] 4EPS was similarly found at elevated levels in the plasma of individuals with ASD.[196] 4EPS is biosynthesized from tyrosine, a common source of mammalian neurotransmitters, through a combination of gut microbial enzymes to yield 4-ethylphenol (4 EP) and then to 4EPS by host metabolism. GF mice were colonized with candidate microbial strains engineered to produce 4 EP, leading to detection of 4EPS in the serum and urine. Additionally, when these mice were treated with probenecid to inhibit small molecule efflux from the blood–brain barrier, 4EPS accumulated in the brain and disrupted the myelination of axons and maturation of oligodendrocytes, highlighting the importance of microbial metabolism in the gut–brain axis.[195–197]

8 Metabolomics in host–microbe–exogenous metabolite interactions

Traditionally, metabolism studies within the preclinical drug development pipeline have focused on host drug metabolism by liver-associated cytochrome P450 enzymes, for example using liver microsomes.[198] However, recent findings have revealed that metabolism by the microbiome also plays a critical but under-appreciated role in drug inactivation, detoxification, and in the production toxic by-products (Fig. 8).[199] For example, a study by Guthrie *et al.* examined the prodrug irinotecan, used for treatment of colorectal cancer.[200] While irinotecan is activated in the liver and subsequently inactivated *via* glucuronidation, reactivation can occur due to β-glucuronidases (BG) produced by the microbiome. Presence of the reactivated drug in the gut can lead to adverse effects, including diarrhea. Using an integrative metagenomics and metabolomics approach, an individual's unique gut microbial community composition could be related to the capacity of the microbiota to reactivate the inactive form of irinotecan. After categorizing subjects' microbiome as either high-metabolizing or low-metabolizing using targeted LC-MS/MS for the inactive metabolite, the authors moved to look for significant differential abundance of BG genes between the two groups. This approach led to the discovery of BGs, which could be linked to a high-metabolizing metabotype for inactive irinotecan.[200]

Drugs such as antibiotics can themselves affect the function of host enzymes and the composition of the gut microbiome, thereby also altering both host metabolism and drug metabolism, a factor that may be important to consider when administering multiple drugs at a time. For instance, vancomycin, an antibiotic of last resort, increases levels of Proteobacteria, Firmicutes, and Fusobacteriota. The increased level of these bacteria was correlated to increased production of intestinal BG enzymes based on increased levels of D-glucuronic acid, a by-product of glucuronidation, found by GC-MS.[201] However, vancomycin also caused dysbiosis in the gut microbiota of healthy human subjects, with increased levels of *Escherichia*, *Enterobacter*, and *Lactobacillus*. This was associated with increased activity of the host enzyme dihydropyrimidine dehydrogenase (DPD) that can metabolize pyrimidines such as uracil, and their analogs used in chemotherapy. The decreased levels of uracil and increased levels of its fecal metabolites, as measured using GC-TOF-MS, correlated with the vancomycin-induced dysbiosis of the gut microbiome.[202] Untargeted urine metabolomics on human subjects with uncontrolled hypertension found that the beta-blocker metoprolol could likewise affect microbial metabolism. Higher levels of gut microbiome-derived acylglycine metabolites (hippuric acid and hydroxyhippuric acid) were detected in the treatment group.[203] Hippurate, a product of phenylalanine metabolism, was found to be a metabolic marker for gut microbiome diversity, and was associated with reduced odds for metabolic syndrome.[204]

To highlight the application of untargeted metabolomics and networking approaches in understanding the role of the gut microbiome in drug metabolism, Jarmusch *et al.* investigated the effect of cefprozil on the microbiome, pharmacokinetics and metabolome, in a prospective study consisting of 14 human subjects.[205] Cefprozil, a β-lactam antibiotic, decreased CYP450 enzyme activity, and increased the microbiome beta diversity while reducing alpha diversity. To study the effect cefprozil was having on drug metabolism, a drug cocktail consisting of caffeine, omeprazole, dextromethorphan and midazolam was administered on day 1 and 8, and 500 mg cefprozil administered orally twice every day from day 2 to day 7. Fecal, urine and plasma samples were collected for untargeted LC-MS/MS metabolomics. Omeprazole, 5-OH omeprazole, and 5-OH omeprazole sulfide, were associated with decreased alpha diversity. The hydroxylation of a reported bacterial metabolite, omeprazole sulfide, was annotated using network propagation, suggesting that both the host and microbiome play a role in drug metabolism. Furthermore, the disruption of the microbiome caused by cefprozil led to the detection of elevated levels of drug and drug metabolites in fecal samples.[205] This study highlights the need to consider emergence of previously unexpected drug metabolites as a result of drug–microbiome interactions and their implications on the structure of gut microbiome.

Untargeted metabolomics also played a crucial role in the search for the enzymes from the gut microbiota responsible for degrading 5-aminosalicylic acid (5-ASA), an oral drug used for treatment of inflammatory bowel disease. When homology-based methods using gut metagenomes to identify potential *N*-acetyltransferases failed to yield candidates, a combined metatranscriptomics and metabolomics approach was followed, correlating the levels of fecal *N*-acetylated-5-ASA with the presence or absence of microbial transcripts. This enabled the discovery of the candidate degrading enzymes.[206]

Zimmermann *et al.* built a model evaluating the impact of oral bioavailability, enzyme activity (host and microbiome), metabolite absorption, and intestinal transit, to quantitatively

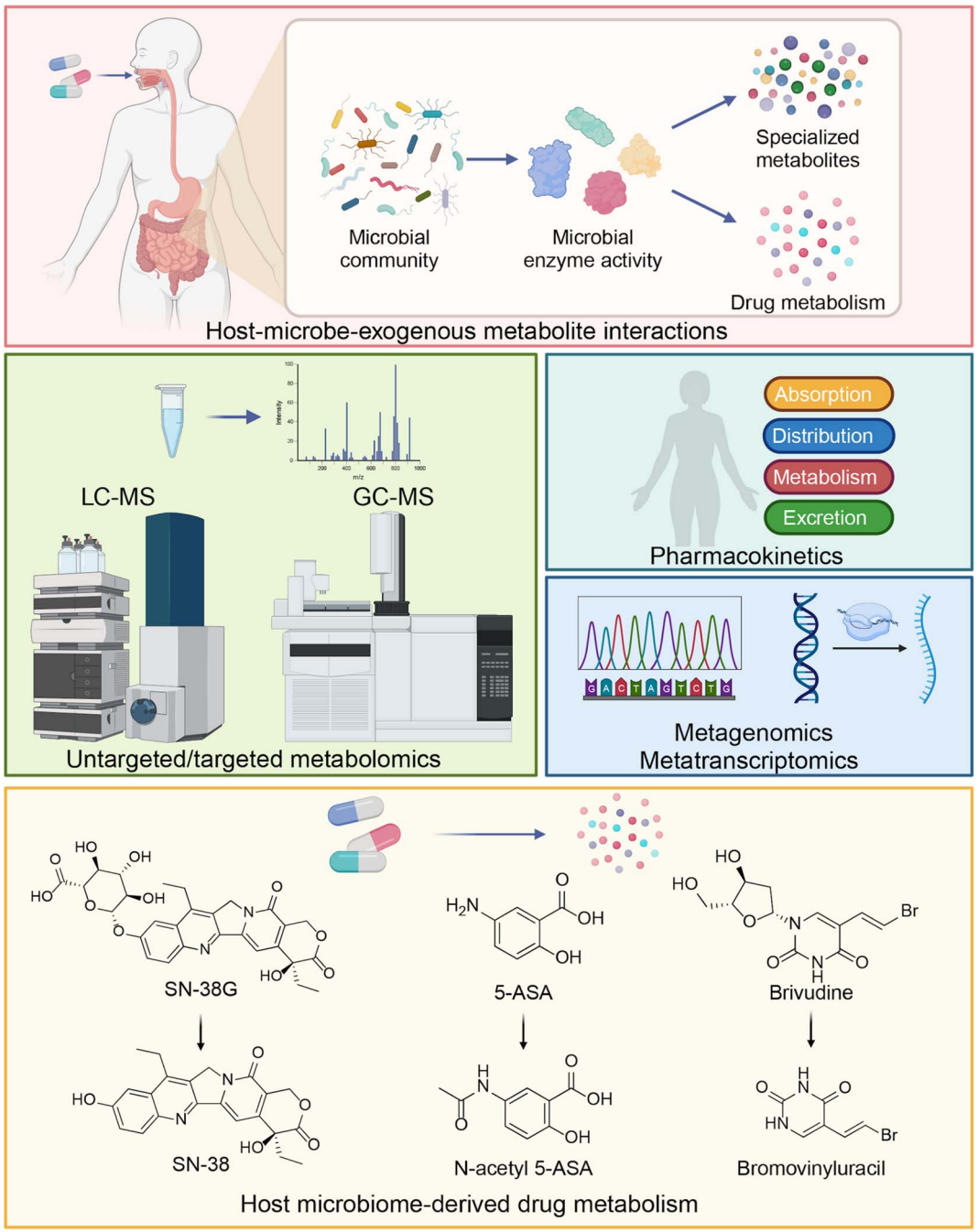

Fig. 8 Overview of host–microbe–exogenous metabolite interactions. Drugs administered to the host can change the microbiome community composition and affect microbial enzyme activity, thereby influencing specialized metabolite biosynthesis in the microbes and leading to changes in drug metabolism by the host microbiome. Metabolomics facilitates the discovery of drug metabolites; pharmacokinetics provides understanding of the body's interaction with the drug; metagenomics and metatranscriptomics aid in identification of microbial species and enzymes responsible for specific drug metabolism. Created with https://www.biorender.com.

predict the contribution of the microbiome on the metabolism of the antiviral compound brivudine.[207] Brivudine is metabolized by both the host and microbiota to bromovinyluracil. Comparisons between conventional and germ-free mice showed increased levels of bromovinyluracil in serum and liver for conventional mice. Drug conversion assays using MS were used to search candidate microbial enzymes in transposon libraries of high-metabolizing bacterial species. After identifying *Bacteroides thetaiotaomicron* and *Bacteroides ovatus* as responsible for the metabolism of brivudine, gnotobiotic mice containing the wild-type strain or a mutant lacking a predicted purine nucleoside phosphorylase (*bt4554*) were established. MS-based metabolomics was then used to quantify the abundances of drug and metabolite in the serum and in various compartments

along the intestinal tract, to make a pharmacokinetic model that showed a microbial contribution of 71% to serum bromovinyluracil levels. The modeling approach was then successfully extended to sorivudine and clonazepam, opening the possibility for broader use in other drugs and exogenous metabolites.[207]

10 Challenges and future opportunities

The study of host–microbe interactions is rapidly evolving due to the convergence of powerful technologies such as NGS, metaproteomics, and metabolomics and the integration of knowledge from diverse scientific disciplines including biochemistry, microbiology, immunometabolism, high-resolution imaging, and chemoinformatics. An interdisciplinary approach is essential to unravel the complexities of these interactions and for translating the knowledge gained into advancements in human health and disease prevention. In this review, through examples (Table 2), we have highlighted the steps involved and the role that mass spectrometry-based metabolomics plays in elucidating function of specialized metabolites in host–microbe interactions. A large portion of the literature covered here focuses on bacterial interactions with the host, often from the perspective of the gut microbiome. However, there is a clear need to expand these studies to fungi, protozoa, and yeasts, which also play diverse roles in the microbiome. Better integration of host–inspired model systems such as mammalian-microbial co-culture systems, *ex vivo* tissue models, and models mimicking multi-organ interfaces such as the gut–liver–brain axis are needed. Advances in organs-on-a-chip and organoid systems, particularly through developments in bioengineering and cell biology, present new opportunities for biochemists and analytical chemists in microbial specialized metabolite discovery.

We also highlight the development of computational approaches to compare metabolomics data acquired from cultured microbes, animal models, and host-derived *in vivo* and *ex vivo* model samples. Methods such as microbeMASST enable extraction of microbial metabolites from a complex milieu of host-derived, environment-derived and diet-derived metabolites. Continued expansion of mass spectral databases and advanced methods for *in silico* structure prediction and determination from small amount of materials such as microED present exciting opportunities to unravel chemical underpinnings of host–microbiome interactions.[208,209] The development of computational methods for integration with other "-omics" data such as metagenomics and metatranscriptomics data is

Table 2 List of selected metabolite discoveries relating to host–microbe interactions discussed in this review

Compound	Discovery	Biochemical role	Microbial source	Ref.
4-Ethylphenylsulfate	Comparative metabolomics	Autism spectrum disorder	Combined host–microbe metabolism	195 and 197
Alteramides	Microbe–microbe co-culture metabolomics	Antifungal activity	Pseudoalteromonas	131
AT-C1	Phenotype-guided	PCOS	*A. tubingensis*	150
Colibactin	Transposon mutagenesis screening	CRC	*E. coli*	45
Commendamide	Functional metagenomics	Autoimmunity, atherosclerosis	Bacteroides	188
Epifadin	Transposon mutagenesis screening	Antibacterial activity	*S. epidermidis*	157
Indolepropionic acid	GF mice metabolomics, genomics	Antioxidant, neuroprotective	*C. sporogenes*	129
Indolimines	Monoculture metabolomics and bioactivity-guided fractionation	CRC	*M. morganii*	146
Lactocillin	*In silico* BGC prediction followed by culturing and isolation	Antibacterial activity	*L. gasseri*	44
Lugdunin	Transposon mutagenesis screening	Antibacterial activity	*S. lugdunensis*	158
Malleonitrone	Microbial co-culture metabolomics	Antibacterial activity	*B. thailandensis*	119
N-Acetylputrescine	Iterative and comparative metabolomics	Fitness advantage to bacteria in bloodstream infections	*E. coli*, *P. aeruginosa*	155
OEA	GF mice metabolomics	Regulating motivation for exercise and meditation satiety	*Eubacterium*, *Coprococcus*	182
TMAO	Comparative metabolomics	Cardiovascular disease	Combined host–microbe metabolism	169
trans-3-Indoleacrylic acid	Bioassay screening	CRC	*P. anaerobius*	145
δ-Valerobetaine	GF mice metabolomics	Obesity	Multiple gut microbial species	171

essential to identify microbial pathways and link them to the metabolites produced. Thus, further development of accessible and user-friendly computational tools for metabolomics data analysis is crucial to facilitate wider adoption of this field.[197,210]

11 Data availability

There is no additional data associated with this article.

12 Conflicts of interest

There are no conflicts to declare.

13 Acknowledgements

Research on microbial natural products in the Garg lab is supported by NIH R35GM150870 and in the McCall lab is supported by R01GM145649.

14 References

1 S. V. Lynch and O. Pedersen, *N. Engl. J. Med.*, 2016, **375**, 2369–2379.
2 Human Microbiome Project Consortium, *Nature*, 2012, **486**, 215–221.
3 J. Qin, R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J.-M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y. Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, M. H. I. T. Consortium, P. Bork, S. D. Ehrlich and J. Wang, *Nature*, 2010, **464**, 59–65.
4 G. W. Tyson, J. Chapman, P. Hugenholtz, E. E. Allen, R. J. Ram, P. M. Richardson, V. V. Solovyev, E. M. Rubin, D. S. Rokhsar and J. F. Banfield, *Nature*, 2004, **428**, 37–43.
5 P. J. Turnbaugh, R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight and J. I. Gordon, *Nature*, 2007, **449**, 804–810.
6 Integrative HMP (iHMP) Research Network Consortium, *Nature*, 2019, **569**, 641–648.
7 NIH Human Microbiome Portfolio Analysis Team, *Microbiome*, 2019, **7**, 31.
8 T. N. Clark, J. Houriet, W. S. Vidar, J. J. Kellogg, D. A. Todd, N. B. Cech and R. G. Linington, *J. Nat. Prod.*, 2021, **84**, 824–835.
9 R. Wang, Y. Yin and Z.-J. Zhu, *Anal. Bioanal. Chem.*, 2019, **411**, 4349–4357.
10 J. Roach, R. Mital, J. J. Haffner, N. Colwell, R. Coats, H. M. Palacios, Z. Liu, J. L. P. Godinho, M. Ness, T. Peramuna and L. I. McCall, *Methods*, 2024, **222**, 81–99.
11 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. B. P, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. O. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein and N. Bandeira, *Nat. Biotechnol.*, 2016, **34**, 828–837.
12 Z. Pang, Y. Lu, G. Zhou, F. Hui, L. Xu, C. Viau, A. F. Spigelman, P. E. MacDonald, D. S. Wishart, S. Li and J. Xia, *Nucleic Acids Res.*, 2024, **52**, W398–W406.
13 M. Kanehisa, M. Furumichi, Y. Sato, M. Kawashima and M. Ishiguro-Watanabe, *Nucleic Acids Res.*, 2023, **51**, D587–D592.
14 F. Huber, L. Ridder, S. Verhoeven, J. H. Spaaks, F. Diblen, S. Rogers and J. J. J. van der Hooft, *PLoS Comput. Biol.*, 2021, **17**, e1008724.
15 F. Huber, S. van der Burg, J. J. J. van der Hooft and L. Ridder, *J. Cheminform.*, 2021, **13**, 84.
16 A. Aisporna, H. P. Benton, A. Chen, R. J. E. Derks, J. M. Galano, M. Giera and G. Siuzdak, *J. Am. Soc. Mass Spectrom.*, 2022, **33**, 530–534.
17 J. Wandy, Y. Zhu, J. J. J. van der Hooft, R. Daly, M. P. Barrett and S. Rogers, *Bioinformatics*, 2017, **34**, 317.
18 D. K. Barupal and O. Fiehn, *Sci. Rep.*, 2017, **7**, 1–11.
19 S. T. Ovbude, S. Sharmeen, I. Kyei, H. Olupathage, J. Jones, R. J. Bell, R. Powers and D. S. Hage, *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.*, 2024, **1239**, 124124.
20 H. B. Skogvold, E. M. Sandås, A. Østeby, C. Løkken, H. Rootwelt, P. O. Rønning, S. R. Wilson and K. B. P. Elgstøen, *J. Proteome Res.*, 2021, **20**, 4010–4021.
21 G. P. O'Sullivan, N. M. Scully and J. D. Glennon, *Anal. Lett.*, 2010, **43**, 1609–1629.
22 E. C. Gentry, S. L. Collins, M. Panitchpakdi, P. Belda-Ferre, A. K. Stewart, M. Carrillo Terrazas, H.-H. Lu, S. Zuffa, T. Yan, J. Avila-Pacheco, D. R. Plichta, A. T. Aron,

M. Wang, A. K. Jarmusch, F. Hao, M. Syrkin-Nikolau, H. Vlamakis, A. N. Ananthakrishnan, B. S. Boland, A. Hemperly, N. Vande Casteele, F. J. Gonzalez, C. B. Clish, R. J. Xavier, H. Chu, E. S. Baker, A. D. Patterson, R. Knight, D. Siegel and P. C. Dorrestein, *Nature*, 2023, **626**, 419–426.

23 H. Awad, M. M. Khamis and A. El-Aneed, *Appl. Spectrosc. Rev.*, 2014, **50**, 158–175.

24 D. P. Demarque, A. E. M. Crotti, R. Vessecchi, J. L. C. Lopes and N. P. Lopes, *Nat. Prod. Rep.*, 2016, **33**, 432–455.

25 J. Guo and T. Huan, *Anal. Chem.*, 2020, **92**, 8072–8080.

26 F. Fenaille, P. Barbier Saint-Hilaire, K. Rousseau and C. Junot, *J. Chromatogr. A*, 2017, **1526**, 1–12.

27 L. E. Razafindrabenja, L. Suc, F. Véran, A. Verbaere, A. L. de Sousa Dias, E. Meudec, N. Galy, D. Tixador, C. Loisel, L. Mouls and N. Sommerer, *Food Chem.*, **483**, 144207.

28 E. Defossez, J. Bourquin, S. von Reuss, S. Rasmann and G. Glauser, *Mass Spectrom. Rev.*, 2023, **42**, 131–143.

29 S. K. Hight, T. N. Clark, K. L. Kurita, E. A. McMillan, W. Bray, A. F. Shaikh, A. Khadilkar, F. P. J. Haeckl, F. Carnevale-Neto, S. La, A. Lohith, R. M. Vaden, J. Lee, S. Wei, R. S. Lokey, M. A. White, R. G. Linington and J. B. MacMillan, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2208458119.

30 F. C. Neto, T. N. Clark, N. P. Lopes and R. G. Linington, *J. Nat. Prod.*, 2022, **85**, 519–529.

31 D. Broadhurst, R. Goodacre, S. N. Reinke, J. Kuligowski, I. D. Wilson, M. R. Lewis and W. B. Dunn, *Metabolomics*, 2018, **14**, 1–17.

32 K. J. Abram and D. McCloskey, *Metabolites*, 2022, **12**, 202.

33 R. Schmid, S. Heuckeroth, A. Korf, A. Smirnov, O. Myers, T. S. Dyrlund, R. Bushuiev, K. J. Murray, N. Hoffmann, M. Lu, A. Sarvepalli, Z. Zhang, M. Fleischauer, K. Dührkop, M. Wesner, S. J. Hoogstra, E. Rudt, O. Mokshyna, C. Brungs, K. Ponomarov, L. Mutabdžija, T. Damiani, C. J. Pudney, M. Earll, P. O. Helmer, T. R. Fallon, T. Schulze, A. Rivas-Ubach, A. Bilbao, H. Richter, L.-F. Nothias, M. Wang, M. Orešič, J.-K. Weng, S. Böcker, A. Jeibmann, H. Hayen, U. Karst, P. C. Dorrestein, D. Petras, X. Du and T. Pluskal, *Nat. Biotechnol.*, 2023, **41**, 447–449.

34 Z. Yan, B. Chen, Y. Yang, X. Yi, M. Wei, G. Ecklu-Mensah, M. M. Buschmann, H. Liu, J. Gao, W. Liang, X. Liu, J. Yang, W. Ma, Z. Liang, F. Wang, D. Chen, L. Wang, W. Shi, M. R. Stampfli, P. Li, S. Gong, X. Chen, W. Shu, E. M. El-Omar, J. A. Gilbert, M. J. Blaser, H. Zhou, R. Chen and Z. Wang, *Nat. Microbiol.*, 2022, **7**, 1361–1375.

35 W. Zhou, M. R. Sailani, K. Contrepois, Y. Zhou, S. Ahadi, S. R. Leopold, M. J. Zhang, V. Rao, M. Avina, T. Mishra, J. Johnson, B. Lee-McMullen, S. Chen, A. A. Metwally, T. D. B. Tran, H. Nguyen, X. Zhou, B. Albright, B.-Y. Hong, L. Petersen, E. Bautista, B. Hanson, L. Chen, D. Spakowicz, A. Bahmani, D. Salins, B. Leopold, M. Ashland, O. Dagan-Rosenfeld, S. Rego, P. Limcaoco, E. Colbert, C. Allister, D. Perelman, C. Craig, E. Wei, H. Chaib, D. Hornburg, J. Dunn, L. Liang, S. M. S.-F. Rose, K. Kukurba, B. Piening, H. Rost, D. Tse, T. McLaughlin, E. Sodergren, G. M. Weinstock and M. Snyder, *Nature*, 2019, **569**, 663–671.

36 R. H. Mills, P. S. Dulai, Y. Vázquez-Baeza, C. Sauceda, N. Daniel, R. R. Gerner, L. E. Batachari, M. Malfavon, Q. Zhu, K. Weldon, G. Humphrey, M. Carrillo-Terrazas, L. D. Goldasich, M. Bryant, M. Raffatellu, R. A. Quinn, A. T. Gewirtz, B. Chassaing, H. Chu, W. J. Sandborn, P. C. Dorrestein, R. Knight and D. J. Gonzalez, *Nat. Microbiol.*, 2022, **7**, 262–276.

37 N. Begum, A. Harzandi, S. Lee, M. Uhlen, D. L. Moyes and S. Shoaie, *Gut Microbes*, 2022, **14**, 2121576.

38 T. Kuehnast, C. Kumpitsch, R. Mohammadzadeh, T. Weichhart, C. Moissl-Eichinger and H. Heine, *FEBS J.*, 2024, **292**, 1316–1329.

39 I. Rybicka and Z. Kaźmierczak, *Appl. Environ. Microbiol.*, 2025, e0178824.

40 S. Zeng, D. Patangia, A. Almeida, Z. Zhou, D. Mu, R. Paul Ross, C. Stanton and S. Wang, *Nat. Commun.*, 2022, **13**, 5139.

41 E. Pasolli, F. Asnicar, S. Manara, M. Zolfo, N. Karcher, F. Armanini, F. Beghini, P. Manghi, A. Tett, P. Ghensi, M. C. Collado, B. L. Rice, C. DuLong, X. C. Morgan, C. D. Golden, C. Quince, C. Huttenhower and N. Segata, *Cell*, 2019, **176**, 649–662.

42 T. Vatanen, E. A. Franzosa, R. Schwager, S. Tripathi, T. D. Arthur, K. Vehik, Å. Lernmark, W. A. Hagopian, M. J. Rewers, J.-X. She, J. Toppari, A.-G. Ziegler, B. Akolkar, J. P. Krischer, C. J. Stewart, N. J. Ajami, J. F. Petrosino, D. Gevers, H. Lähdesmäki, H. Vlamakis, C. Huttenhower and R. J. Xavier, *Nature*, 2018, **562**, 589–594.

43 P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy, B. W. Birren, E. Takano, A. Sali, R. G. Linington and M. A. Fischbach, *Cell*, 2014, **158**, 412–421.

44 M. S. Donia, P. Cimermancic, C. J. Schulze, L. C. Wieland Brown, J. Martin, M. Mitreva, J. Clardy, R. G. Linington and M. A. Fischbach, *Cell*, 2014, **158**, 1402–1414.

45 J.-P. Nougayrède, S. Homburg, F. Taieb, M. Boury, E. Brzuszkiewicz, G. Gottschalk, C. Buchrieser, J. Hacker, U. Dobrindt and E. Oswald, *Science*, 2006, **313**, 848–851.

46 E. P. Balskus, *Nat. Prod. Rep.*, 2015, **32**, 1534–1540.

47 H. B. Bode, *Angew Chem. Int. Ed. Engl.*, 2015, **54**, 10408–10411.

48 J. C. Arthur, E. Perez-Chanona, M. Mühlbauer, S. Tomkovich, J. M. Uronis, T.-J. Fan, B. J. Campbell, T. Abujamel, B. Dogan, A. B. Rogers, J. M. Rhodes, A. Stintzi, K. W. Simpson, J. J. Hansen, T. O. Keku, A. A. Fodor and C. Jobin, *Science*, 2012, **338**, 120–123.

49 C. Pleguezuelos-Manzano, J. Puschhof, A. Rosendahl Huber, A. van Hoeck, H. M. Wood, J. Nomburg, C. Gurjao, F. Manders, G. Dalmasso, P. B. Stege, F. L. Paganelli, M. H. Geurts, J. Beumer, T. Mizutani, Y. Miao, R. van der Linden, S. van der Elst, K. C. Garcia, J. Top, R. J. L. Willems, M. Giannakis, R. Bonnet, P. Quirke,

M. Meyerson, E. Cuppen, R. van Boxtel and H. Clevers, *Nature*, 2020, **580**, 269–273.
50 X. Bian, J. Fu, A. Plaza, J. Herrmann, D. Pistorius, A. F. Stewart, Y. Zhang and R. Müller, *ChemBioChem*, 2013, **14**, 1194–1197.
51 C. A. Brotherton and E. P. Balskus, *J. Am. Chem. Soc.*, 2013, **135**, 3359–3362.
52 X. Bian, A. Plaza, Y. Zhang and R. Müller, *Chem. Sci.*, 2015, **6**, 3154–3160.
53 C. A. Brotherton, M. Wilson, G. Byrd and E. P. Balskus, *Org. Lett.*, 2015, **17**, 1545–1548.
54 A. R. Healy, M. I. Vizcaino, J. M. Crawford and S. B. Herzon, *J. Am. Chem. Soc.*, 2016, **138**, 5426–5432.
55 Z.-R. Li, Y. Li, J. Y. H. Lai, J. Tang, B. Wang, L. Lu, G. Zhu, X. Wu, Y. Xu and P.-Y. Qian, *ChemBioChem*, 2015, **16**, 1715–1719.
56 M. I. Vizcaino and J. M. Crawford, *Nat. Chem.*, 2015, **7**, 411–417.
57 M. I. Vizcaino, P. Engel, E. Trautman and J. M. Crawford, *J. Am. Chem. Soc.*, 2014, **136**, 9244–9247.
58 L. Zha, M. R. Wilson, C. A. Brotherton and E. P. Balskus, *ACS Chem. Biol.*, 2016, **11**, 1287–1295.
59 Z.-R. Li, J. Li, J.-P. Gu, J. Y. H. Lai, B. M. Duggan, W.-P. Zhang, Z.-L. Li, Y.-X. Li, R.-B. Tong, Y. Xu, D.-H. Lin, B. S. Moore and P.-Y. Qian, *Nat. Chem. Biol.*, 2016, **12**, 773–775.
60 M. Xue, E. Shine, W. Wang, J. M. Crawford and S. B. Herzon, *Biochemistry*, 2018, **57**, 6391–6394.
61 M. R. Wilson, Y. Jiang, P. W. Villalta, A. Stornetta, P. D. Boudreau, A. Carrá, C. A. Brennan, E. Chun, L. Ngo, L. D. Samson, B. P. Engelward, W. S. Garrett, S. Balbo and E. P. Balskus, *Science*, 2019, **363**, eaar7785.
62 M. Xue, C. S. Kim, A. R. Healy, K. M. Wernke, Z. Wang, M. C. Frischling, E. E. Shine, W. Wang, S. B. Herzon and J. M. Crawford, *Science*, 2019, **365**, eaax2685.
63 A. Mandarino Alves, C. Lecchi, S. Lopez, A. Stornetta, P. P. Mathai, P. W. Villalta, S. Ishii, E. P. Balskus, S. Balbo and A. Khoruts, *Gut Microbes*, 2024, **16**, 2387877.
64 S. A. Jarmusch, J. J. J. van der Hooft, P. C. Dorrestein and A. K. Jarmusch, *Nat. Prod. Rep.*, 2021, **38**, 2066–2082.
65 A. Bauermeister, H. Mannochio-Russo, L. V. Costa-Lotufo, A. K. Jarmusch and P. C. Dorrestein, *Nat. Rev. Microbiol.*, 2021, **20**, 143–160.
66 M. Wang, A. K. Jarmusch, F. Vargas, A. A. Aksenov, J. M. Gauglitz, K. Weldon, D. Petras, R. da Silva, R. Quinn, A. V. Melnik, J. J. J. van der Hooft, A. M. Caraballo-Rodríguez, L. F. Nothias, C. M. Aceves, M. Panitchpakdi, E. Brown, F. Di Ottavio, N. Sikora, E. O. Elijah, L. Labarta-Bajo, E. C. Gentry, S. Shalapour, K. E. Kyle, S. P. Puckett, J. D. Watrous, C. S. Carpenter, A. Bouslimani, M. Ernst, A. D. Swafford, E. I. Zúñiga, M. J. Balunas, J. L. Klassen, R. Loomba, R. Knight, N. Bandeira and P. C. Dorrestein, *Nat. Biotechnol.*, 2020, **38**, 23–26.
67 G. J. Jeong, F. Khan, N. Tabassum, D. M. Jo, W. K. Jung and Y. M. Kim, *Res. Microbiol.*, 2024, **175**, 104211.
68 A. T. Butt and M. S. Thomas, *Front. Cell. Infect. Microbiol.*, 2017, **7**, 304015.
69 C. Molina-Santiago, D. Vela-Corcía, D. Petras, L. Díaz-Martínez, A. I. Pérez-Lorente, S. Sopeña-Torres, J. Pearson, A. M. Caraballo-Rodríguez, P. Dorrestein, A. de Vicente and D. Romero, *Cell Rep.*, 2021, **36**, 109449.
70 C. Jenul, K. C. Keim, J. N. Jens, M. J. Zeiler, K. Schilcher, M. J. Schurr, C. Melander, V. V. Phelan and A. R. Horswill, *Cell Rep.*, 2023, **42**, 112540.
71 S. Uzi-Gavrilov, Z. Tik, O. Sabti and M. M. Meijler, *Angew Chem. Int. Ed. Engl.*, 2023, **62**, e202300585.
72 B. C. Covington and M. R. Seyedsayamdost, *Methods Enzymol.*, 2021, **665**, 305–323.
73 H. S. Shepard, J. C. May, B. E. Zuniga, J. P. Abraham, B. F. Pfleger, J. D. Young and J. A. McLean, *J. Am. Soc. Mass Spectrom.*, 2025, **36**, 400–408.
74 R. A. Shepherd, C. A. Fihn, A. J. Tabag, S. M. K. McKinnie and L. M. Sanchez, *Nat. Prod. Rep.*, 2025, DOI: **10.1039/D4NP00053F**.
75 T. Luke Simmons, R. Cameron Coates, B. R. Clark, N. Engene, D. Gonzalez, E. Esquenazi, P. C. Dorrestein and W. H. Gerwick, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 4587–4594.
76 P. M. Flatt, J. T. Gautschi, R. W. Thacker, M. Musafija-Girt, P. Crews and W. H. Gerwick, *Mar. Biol.*, 2005, **147**, 761–774.
77 B. Geier, E. M. Sogin, D. Michellod, M. Janda, M. Kompauer, B. Spengler, N. Dubilier and M. Liebeke, *Nat. Microbiol.*, 2020, **5**, 498–510.
78 P. Bourceau, B. Geier, V. Suerdieck, T. Bien, J. Soltwisch, K. Dreisewerd and M. Liebeke, *Nat. Protoc.*, 2023, **18**, 3050–3079.
79 Y. Sugimoto, F. R. Camacho, S. Wang, P. Chankhamjon, A. Odabas, A. Biswas, P. D. Jeffrey and M. S. Donia, *Science*, 2019, **366**, eaax9176.
80 J. L. M. Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst and G. G. Borisy, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E791–E800.
81 M. C. Wilson, T. Mori, C. Rückert, A. R. Uria, M. J. Helf, K. Takada, C. Gernert, U. A. E. Steffens, N. Heycke, S. Schmitt, C. Rinke, E. J. N. Helfrich, A. O. Brachmann, C. Gurgui, T. Wakimoto, M. Kracht, M. Crüsemann, U. Hentschel, I. Abe, S. Matsunaga, J. Kalinowski, H. Takeyama and J. Piel, *Nature*, 2014, **506**, 58–62.
82 Y. Feng, A. Qaseem, A. F. A. Moumbock, S. Pan, P. A. Kirchner, C. V. Simoben, Y. I. Malange, S. B. Babiaka, M. Gao and S. Günther, *Nucleic Acids Res.*, 2025, **53**, D724–D729.
83 H. W. Kim, M. Wang, C. A. Leber, L.-F. Nothias, R. Reher, K. B. Kang, J. J. J. van der Hooft, P. C. Dorrestein, W. H. Gerwick and G. W. Cottrell, *J. Nat. Prod.*, 2021, **84**, 2795–2807.
84 A. A. Aksenov, I. Laponogov, Z. Zhang, S. L. F. Doran, I. Belluomo, D. Veselkov, W. Bittremieux, L. F. Nothias, M. Nothias-Esposito, K. N. Maloney, B. B. Misra, A. V. Melnik, A. Smirnov, X. Du, K. L. Jones 2nd, K. Dorrestein, M. Panitchpakdi, M. Ernst, J. J. J. van der Hooft, M. Gonzalez, C. Carazzone, A. Amézquita, C. Callewaert, J. T. Morton, R. A. Quinn, A. Bouslimani, A. A. Orio, D. Petras, A. M. Smania, S. P. Couvillion,

M. C. Burnet, C. D. Nicora, E. Zink, T. O. Metz, V. Artaev, E. Humston-Fulmer, R. Gregor, M. M. Meijler, I. Mizrahi, S. Eyal, B. Anderson, R. Dutton, R. Lugan, P. L. Boulch, Y. Guitton, S. Prevost, A. Poirier, G. Dervilly, B. Le Bizec, A. Fait, N. S. Persi, C. Song, K. Gashu, R. Coras, M. Guma, J. Manasson, J. U. Scher, D. K. Barupal, S. Alseekh, A. R. Fernie, R. Mirnezami, V. Vasiliou, R. Schmid, R. S. Borisov, L. N. Kulikova, R. Knight, M. Wang, G. B. Hanna, P. C. Dorrestein and K. Veselkov, *Nat. Biotechnol.*, 2021, **39**, 169–173.

85 Z. Tian, F. Liu, D. Li, A. R. Fernie and W. Chen, *Comput. Struct. Biotechnol. J.*, 2022, **20**, 5085–5097.

86 M. W. Mullowney, K. R. Duncan, S. S. Elsayed, N. Garg, J. J. J. van der Hooft, N. I. Martin, D. Meijer, B. R. Terlouw, F. Biermann, K. Blin, J. Durairaj, G. M. Gorostiola, E. J. N. Helfrich, F. Huber, S. Leopold-Messer, K. Rajan, T. de Rond, J. A. van Santen, M. Sorokina, M. J. Balunas, M. A. Beniddir, D. A. van Bergeijk, L. M. Carroll, C. M. Clark, D. A. Clevert, C. A. Dejong, C. Du, S. Ferrinho, F. Grisoni, A. Hofstetter, W. Jespers, O. V. Kalinina, S. A. Kautsar, H. Kim, T. F. Leao, J. Masschelein, E. R. Rees, R. Reher, D. Reker, P. Schwaller, M. Segler, M. A. Skinnider, A. S. Walker, E. L. Willighagen, B. Zdrazil, N. Ziemert, R. J. M. Goss, P. Guyomard, A. Volkamer, W. H. Gerwick, H. U. Kim, R. Müller, G. P. van Wezel, G. J. P. van Westen, A. K. H. Hirsch, R. G. Linington, S. L. Robinson and M. H. Medema, *Nat. Rev. Drug Discovery*, 2023, **22**, 895–916.

87 J. J. R. Louwen and J. J. J. van der Hooft, *mSystems*, 2021, **6**, e0072621.

88 O. Moyne, F. Castelli, D. J. Bicout, J. Boccard, B. Camara, B. Cournoyer, E. Faudry, S. Terrier, D. Hannani, S. Huot-Marchand, C. Léger, M. Maurin, T.-D. Ngo, C. Plazy, R. A. Quinn, I. Attree, F. Fenaille, B. Toussaint and A. Le Gouëllec, *Metabolites*, 2021, **11**, 63.

89 P. Reveglia, F. J. Agudo-Jurado, E. Barilli, M. Masi, A. Evidente and D. Rubiales, *J. Fungi*, 2023, **9**, 610.

90 S. Zuffa, R. Schmid, A. Bauermeister, P. W. P Gomes, A. M. Caraballo-Rodriguez, Y. El Abiead, A. T. Aron, E. C. Gentry, J. Zemlin, M. J. Meehan, N. E. Avalon, R. H. Cichewicz, E. Buzun, M. C. Terrazas, C.-Y. Hsu, R. Oles, A. V. Ayala, J. Zhao, H. Chu, M. C. M. Kuijpers, S. L. Jackrel, F. Tugizimana, L. P. Nephali, I. A. Dubery, N. E. Madala, E. A. Moreira, L. V. Costa-Lotufo, N. P. Lopes, P. Rezende-Teixeira, P. C. Jimenez, B. Rimal, A. D. Patterson, M. F. Traxler, R. de C. Pessotti, D. Alvarado-Villalobos, G. Tamayo-Castillo, P. Chaverri, E. Escudero-Leyva, L.-M. Quiros-Guerrero, A. J. Bory, J. Joubert, A. Rutz, J.-L. Wolfender, P.-M. Allard, A. Sichert, S. Pontrelli, B. S. Pullman, N. Bandeira, W. H. Gerwick, K. Gindro, J. Massana-Codina, B. C. Wagner, K. Forchhammer, D. Petras, N. Aiosa, N. Garg, M. Liebeke, P. Bourceau, K. B. Kang, H. Gadhavi, L. P. S. de Carvalho, M. Silva Dos Santos, A. I. Pérez-Lorente, C. Molina-Santiago, D. Romero, R. Franke, M. Brönstrup, A. Vera Ponce de León, P. B. Pope, S. L. La Rosa, G. La Barbera, H. M. Roager, M. F. Laursen, F. Hammerle, B. Siewert, U. Peintner, C. Licona-Cassani, L. Rodriguez-Orduña, E. Rampler, F. Hildebrand, G. Koellensperger, H. Schoeny, K. Hohenwallner, L. Panzenboeck, R. Gregor, E. C. O'Neill, E. T. Roxborough, J. Odoi, N. J. Bale, S. Ding, J. S. Sinninghe Damsté, X. L. Guan, J. J. Cui, K.-S. Ju, D. B. Silva, F. M. R. Silva, G. F. da Silva, H. H. F. Koolen, C. Grundmann, J. A. Clement, H. Mohimani, K. Broders, K. L. McPhail, S. E. Ober-Singleton, C. M. Rath, D. McDonald, R. Knight, M. Wang and P. C. Dorrestein, *Nat. Microbiol.*, 2024, **9**, 336–345.

91 L.-F. Nothias, R. Schmid, A. Garlet, H. Cameron, S. Leoty-Okombi, V. André-Frei, R. Fuchs, P. C. Dorrestein and P. Ternes, *mSystems*, 2024, **9**, e0035623.

92 P. W. P Gomes, H. Mannochio-Russo, J. Mao, H. N. Zhao, J. Ancira, C. D. Tipton, P. C. Dorrestein and M. Li, *mSystems*, 2024, **9**, e0111923.

93 I. Mohanty, H. Mannochio-Russo, J. V. Schweer, Y. El Abiead, W. Bittremieux, S. Xing, R. Schmid, S. Zuffa, F. Vasquez, V. B. Muti, J. Zemlin, O. E. Tovar-Herrera, S. Moraïs, D. Desai, S. Amin, I. Koo, C. W. Turck, I. Mizrahi, P. M. Kris-Etherton, K. S. Petersen, J. A. Fleming, T. Huan, A. D. Patterson, D. Siegel, L. R. Hagey, M. Wang, A. T. Aron and P. C. Dorrestein, *Cell*, 2024, **187**, 1801–1818.

94 A. T. Aron, E. C. Gentry, K. L. McPhail, L.-F. Nothias, M. Nothias-Esposito, A. Bouslimani, D. Petras, J. M. Gauglitz, N. Sikora, F. Vargas, J. J. J. van der Hooft, M. Ernst, K. B. Kang, C. M. Aceves, A. M. Caraballo-Rodríguez, I. Koester, K. C. Weldon, S. Bertrand, C. Roullier, K. Sun, R. M. Tehan, C. A. Boya, M. H. Christian, M. Gutiérrez, A. M. Ulloa, J. A. Tejeda Mora, R. Mojica-Flores, J. Lakey-Beitia, V. Vásquez-Chaves, Y. Zhang, A. I. Calderón, N. Tayler, R. A. Keyzers, F. Tugizimana, N. Ndlovu, A. A. Aksenov, A. K. Jarmusch, R. Schmid, A. W. Truman, N. Bandeira, M. Wang and P. C. Dorrestein, *Nat. Protoc.*, 2020, **15**, 1954–1991.

95 J. Watrous, P. Roach, T. Alexandrov, B. S. Heath, J. Y. Yang, R. D. Kersten, M. van der Voort, K. Pogliano, H. Gross, J. M. Raaijmakers, B. S. Moore, J. Laskin, N. Bandeira and P. C. Dorrestein, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, E1743–E1752.

96 S. Rak Lee, F. Schalk, J. W. Schwitalla, R. Benndorf, J. Vollmers, A.-K. Kaster, Z. W. de Beer, M. Park, M.-J. Ahn, W. H. Jung, C. Beemelmanns and K. H. Kim, *J. Nat. Prod.*, 2020, **83**, 3102–3110.

97 N. Noori Goodarzi, M. Khazani Asforooshani, B. Shahbazi, N. Rezaie Rahimi and F. Badmasti, *Front. Bioinform.*, 2024, **4**, 1482338.

98 N. M. Almansour, K. S. Allemailem, A. A. Abd El Aty, E. I. F. Ismail and M. A. A. Ibrahim, *Molecules*, 2023, **28**, 783.

99 K. A. Shaaban, T. A. Ahmed, M. Leggas and J. Rohr, *J. Nat. Prod.*, 2012, **75**, 1383–1392.

100 D. Petras, A. K. Jarmusch and P. C. Dorrestein, *Curr. Opin. Chem. Biol.*, 2017, **36**, 24–31.

101 X. Ma and F. M. Fernández, *Mass Spectrom. Rev.*, 2024, **43**, 235–268.

102 X. Min, Y. Zhao, M. Yu, W. Zhang, X. Jiang, K. Guo, X. Wang, J. Huang, T. Li, L. Sun and J. He, *Clin. Transl. Med.*, 2024, **14**, e70031.
103 K. Wheeler, C. Gosmanov, M. J. Sandoval, Z. Yang and L. I. McCall, *TrAC Trends Anal. Chem.*, 2024, **175**, 117713.
104 S. J. B. Dunham, J. F. Ellis, B. Li and J. V. Sweedler, *Acc. Chem. Res.*, 2017, **50**, 96–104.
105 J. D. Watrous and P. C. Dorrestein, *Nat. Rev. Microbiol.*, 2011, **9**, 683–694.
106 M. Feucherolles and G. Frache, *Cells*, 2022, **11**, 3900.
107 N. Garg, Y. Zeng, A. Edlund, A. V. Melnik, L. M. Sanchez, H. Mohimani, A. Gurevich, V. Miao, S. Schiffler, Y. W. Lim, T. Luzzatto-Knaan, S. Cai, F. Rohwer, P. A. Pevzner, R. H. Cichewicz, T. Alexandrov and P. C. Dorrestein, *mSystems*, 2016, **1**, e00139.
108 N. Aiosa, A. Sinha, H. Albataineh, A. M. Phillips, C. M. Mageeney, D. S. Wilde, K. P. Williams, N. M. Collette, S. S. Branda and N. Garg, *ChemBioChem*, 2023, **24**, e202200802.
109 T. D. Nguyen, Y. Lan, S. S. Kane, J. J. Haffner, R. Liu, L. I. McCall and Z. Yang, *Anal. Chem.*, 2022, **94**, 10567–10572.
110 R. A. Quinn, A. V. Melnik, A. Vrbanac, T. Fu, K. A. Patras, M. P. Christy, Z. Bodai, P. Belda-Ferre, A. Tripathi, L. K. Chung, M. Downes, R. D. Welch, M. Quinn, G. Humphrey, M. Panitchpakdi, K. C. Weldon, A. Aksenov, R. da Silva, J. Avila-Pacheco, C. Clish, S. Bae, H. Mallick, E. A. Franzosa, J. Lloyd-Price, R. Bussell, T. Thron, A. T. Nelson, M. Wang, E. Leszczynski, F. Vargas, J. M. Gauglitz, M. J. Meehan, E. Gentry, T. D. Arthur, A. C. Komor, O. Poulsen, B. S. Boland, J. T. Chang, W. J. Sandborn, M. Lim, N. Garg, J. C. Lumeng, R. J. Xavier, B. I. Kazmierczak, R. Jain, M. Egan, K. E. Rhee, D. Ferguson, M. Raffatellu, H. Vlamakis, G. G. Haddad, D. Siegel, C. Huttenhower, S. K. Mazmanian, R. M. Evans, V. Nizet, R. Knight and P. C. Dorrestein, *Nature*, 2020, **579**, 123–129.
111 N. Garg, M. Wang, E. Hyde, R. R. da Silva, A. V. Melnik, I. Protsyuk, A. Bouslimani, Y. W. Lim, R. Wong, G. Humphrey, G. Ackermann, T. Spivey, S. S. Brouha, N. Bandeira, G. Y. Lin, F. Rohwer, D. J. Conrad, T. Alexandrov, R. Knight and P. C. Dorrestein, *Cell Host Microbe*, 2017, **22**, 705–716.
112 A. V. Melnik, Y. Vázquez-Baeza, A. A. Aksenov, E. Hyde, A. C. McAvoy, M. Wang, R. R. da Silva, I. Protsyuk, J. V. Wu, A. Bouslimani, Y. W. Lim, T. Luzzatto-Knaan, W. Comstock, R. A. Quinn, R. Wong, G. Humphrey, G. Ackermann, T. Spivey, S. S. Brouha, N. Bandeira, G. Y. Lin, F. Rohwer, D. J. Conrad, T. Alexandrov, R. Knight, P. C. Dorrestein and N. Garg, *mSystems*, 2019, **4**, e00375.
113 L.-I. McCall, J. T. Morton, J. A. Bernatchez, J. L. de Siqueira-Neto, R. Knight, P. C. Dorrestein and J. H. McKerrow, *Anal. Chem.*, 2017, **89**, 10414–10421.
114 E. Hossain, S. Khanam, D. A. Dean, C. Wu, S. Lostracco-Johnson, D. Thomas, S. S. Kane, A. R. Parab, K. Flores, M. Katemauswa, C. Gosmanov, S. E. Hayes, Y. Zhang, D. Li, C. Woelfel-Monsivais, K. Sankaranarayanan and L.-I. McCall, *Sci. Adv.*, 2020, **6**, eaaz2015.
115 D. A. Dean, G. Gautham, J. L. Siqueira-Neto, J. H. McKerrow, P. C. Dorrestein and L.-I. McCall, *PLoS Negl. Trop. Dis.*, 2021, **15**, e0009819.
116 J. Jing, P. Garbeva, J. M. Raaijmakers and M. H. Medema, *ISME J.*, 2024, **18**, wrae049.
117 A. Hurley, M. G. Chevrette, N. Rosario-Meléndez and J. Handelsman, *mBio*, 2022, **13**, e0248621.
118 M. G. Chevrette, C. S. Thomas, A. Hurley, N. Rosario-Meléndez, K. Sankaran, Y. Tu, A. Hall, S. Magesh and J. Handelsman, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2212930119.
119 F. Trottmann, J. Franke, K. Ishida, M. García-Altares and C. Hertweck, *Angew Chem. Int. Ed. Engl.*, 2019, **58**, 200–204.
120 Y. Sun, W.-C. Liu, X. Shi, H.-Z. Zheng, Z.-H. Zheng, X.-H. Lu, Y. Xing, K. Ji, M. Liu and Y.-S. Dong, *Microb. Cell Fact.*, 2021, **20**, 42.
121 A. G. Albarracín Orio, D. Petras, R. A. Tobares, A. A. Aksenov, M. Wang, F. Juncosa, P. Sayago, A. J. Moyano, P. C. Dorrestein and A. M. Smania, *Commun. Biol.*, 2020, **3**, 670.
122 M. B. Miller and B. L. Bassler, *Annu. Rev. Microbiol.*, 2001, **55**, 165–199.
123 J. Meinwald, *J. Nat. Prod.*, 2011, **74**, 305–309.
124 S. Brugiroux, M. Beutler, C. Pfann, D. Garzetti, H.-J. Ruscheweyh, D. Ring, M. Diehl, S. Herp, Y. Lötscher, S. Hussain, B. Bunk, R. Pukall, D. H. Huson, P. C. Münch, A. C. McHardy, K. D. McCoy, A. J. Macpherson, A. Loy, T. Clavel, D. Berry and B. Stecher, *Nat. Microbiol.*, 2016, **2**, 1–12.
125 A. G. Cheng, P. Y. Ho, A. Aranda-Díaz, S. Jain, F. B. Yu, X. Meng, M. Wang, M. Iakiviak, K. Nagashima, A. Zhao, P. Murugkar, A. Patil, K. Atabakhsh, A. Weakley, J. Yan, A. R. Brumbaugh, S. Higginbottom, A. Dimas, A. L. Shiver, A. Deutschbauer, N. Neff, J. L. Sonnenburg, K. C. Huang and M. A. Fischbach, *Cell*, 2022, **185**, 3617–3636.
126 T. Van de Wiele, P. Van den Abbeele, W. Ossieur, S. Possemiers and M. Marzorati, in *The Impact of Food Bioactives on Health: in vitro and ex vivo models*, ed. K. Verhoeckx, P. Cotter, I. López-Expósito, C. Kleiveland, T. Lea, A. Mackie, T. Requena, D. Swiatecka and H. Wichers, Springer, Cham, 2015, ch. 27, pp. 305–317.
127 D. A. van Bergeijk, S. S. Elsayed, C. Du, I. N. Santiago, A. M. Roseboom, L. Zhang, V. J. Carrión, H. P. Spaink and G. P. van Wezel, *Commun. Chem.*, 2022, **5**, 1–12.
128 A. K. Sinha, M. F. Laursen, J. E. Brinck, M. L. Rybtke, A. P. Hjørne, N. Procházková, M. Pedersen, H. M. Roager and T. R. Licht, *Nat. Microbiol.*, 2024, **9**, 1964–1978.
129 D. Dodd, M. H. Spitzer, W. Van Treuren, B. D. Merrill, A. J. Hryckowian, S. K. Higginbottom, A. Le, T. M. Cowan, G. P. Nolan, M. A. Fischbach and J. L. Sonnenburg, *Nature*, 2017, **551**, 648–652.
130 J. Jin, A. S. Kulkarni, A. C. McAvoy and N. Garg, *ACS Chem. Biol.*, 2025, DOI: **10.1021/acschembio.4c00562**.

131 W. J. Moree, O. J. McConnell, D. D. Nguyen, L. M. Sanchez, Y.-L. Yang, X. Zhao, W.-T. Liu, P. D. Boudreau, J. Srinivasan, L. Atencio, J. Ballesteros, R. G. Gavilán, D. Torres-Mendoza, H. M. Guzmán, W. H. Gerwick, M. Gutiérrez and P. C. Dorrestein, *ACS Chem. Biol.*, 2014, **9**, 2300–2308.

132 C. de Bekker, P. B. Smith, A. D. Patterson and D. P. Hughes, *PLoS One*, 2013, **8**, e70609.

133 N. Aiosa, A. Sinha, O. A. Jaiyesimi, R. R. da Silva, S. S. Branda and N. Garg, *ACS Infect. Dis.*, 2022, **8**, 1646–1662.

134 D. Mao, L. B. Bushin, K. Moon, Y. Wu and M. R. Seyedsayamdost, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E2920–E2928.

135 T. Roger, J. Lugrin, D. Le Roy, G. Goy, M. Mombelli, T. Koessler, X. C. Ding, A.-L. Chanson, M. K. Reymond, I. Miconnet, J. Schrenzel, P. François and T. Calandra, *Blood*, 2011, **117**, 1205–1217.

136 S. N. Micheva-Viteva, M. Shakya, S. H. Adikari, C. D. Gleasner, N. Velappan, J. R. Mourant, P. S. G. Chain and E. Hong-Geller, *mSystems*, 2020, **5**, e00609–e00619.

137 M. Mombelli, J. Lugrin, I. Rubino, A.-L. Chanson, M. Giddey, T. Calandra and T. Roger, *J. Infect. Dis.*, 2011, **204**, 1367–1374.

138 M. R. Seyedsayamdost, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 7266–7271.

139 J. B. Biggins, C. D. Gleber and S. F. Brady, *Org. Lett.*, 2011, **13**, 1536–1539.

140 A. Sen, T. Nishimura, S. Yoshimoto, K. Yoshida, A. Gotoh, T. Katoh, Y. Yoneda, T. Hashimoto, J.-Z. Xiao, T. Katayama and T. Odamaki, *Front. Microbiol.*, 2023, **14**, 1155438.

141 A. Valiei, J. Aminian-Dehkordi and M. R. K. Mofrad, *APL Bioeng.*, 2023, **7**, 011502.

142 M. Lucchetti, K. O. Aina, L. Grandmougin, C. Jäger, P. P. Escriva, E. Letellier, A. S. Mosig and P. Wilmes, *Adv. Healthc. Mater.*, 2024, **13**, 2303943.

143 W. Wang, Y. Liu, Z. Yao, D. Chen, Y. Tang, J. Cui, J. Zhang, H. Liu and Z. Hao, *Lab Chip*, 2024, **24**, 2537–2550.

144 Z. Zhong, M. Quiñones-Pérez, Z. Dai, V. M. Juarez, E. Bhatia, C. R. Carlson, S. B. Shah, A. Patel, Z. Fang, T. Hu, M. Allam, S. L. Hicks, M. Gupta, S. L. Gupta, E. Weeks, S. D. Vagelos, A. Molina, A. Mulero-Russe, A. Mora-Boza, D. J. Joshi, R. P. Sekaly, T. Sulchek, S. L. Goudy, J. Wrammert, K. Roy, J. M. Boss, A. F. Coskun, C. D. Scharer, A. J. García, J. L. Koff and A. Singh, *Nat. Mater.*, 2025, **24**, 297–311.

145 W. Cui, M. Guo, D. Liu, P. Xiao, C. Yang, H. Huang, C. Liang, Y. Yang, X. Fu, Y. Zhang, J. Liu, S. Shi, J. Cong, Z. Han, Y. Xu, L. Du, C. Yin, Y. Zhang, J. Sun, W. Gu, R. Chai, S. Zhu and B. Chu, *Nat. Cell Biol.*, 2024, **26**, 124–137.

146 Y. Cao, J. Oh, M. Xue, W. J. Huh, J. Wang, J. A. Gonzalez-Hernandez, T. A. Rice, A. L. Martin, D. Song, J. M. Crawford, S. B. Herzon and N. W. Palm, *Science*, 2022, **378**, eabm3233.

147 The Action to Control Cardiovascular Risk in Diabetes Study Group, *N. Engl. J. Med.*, 2008, **358**, 2545–2559.

148 W. Duckworth, C. Abraira, T. Moritz, D. Reda, N. Emanuele, P. D. Reaven, F. J. Zieve, J. Marks, S. N. Davis, R. Hayward, S. R. Warren, S. Goldman, M. McCarren, M. E. Vitek, W. G. Henderson and G. D. Huang, *N. Engl. J. Med.*, 2009, **360**, 129–139.

149 I. Nemet, P. P. Saha, N. Gupta, W. Zhu, K. A. Romano, S. M. Skye, T. Cajka, M. L. Mohan, L. Li, Y. Wu, M. Funabashi, A. E. Ramer-Tait, S. V. Naga Prasad, O. Fiehn, F. E. Rey, W. H. W. Tang, M. A. Fischbach, J. A. DiDonato and S. L. Hazen, *Cell*, 2020, **180**, 862–877.

150 J. Wu, K. Wang, X. Qi, S. Zhou, S. Zhao, M. Lu, Q. Nie, M. Li, M. Han, X. Luo, C. Yun, P. Wang, R. Li, C. Zhong, X. Yu, W.-B. Yin, C. Jiang, J. Qiao and Y. Pang, *Cell Host Microbe*, 2025, **33**, 119–136.

151 K. J. O'Riordan, G. M. Moloney, L. Keane, G. Clarke and J. F. Cryan, *Cell Rep. Med*, 2025, **6**, 101982.

152 S. L. Boddy, I. Giovannelli, M. Sassani, J. Cooper-Knock, M. P. Snyder, E. Segal, E. Elinav, L. A. Barker, P. J. Shaw and C. J. McDermott, *BMC Med.*, 2021, **19**, 1–14.

153 E. Blacher, S. Bashiardes, H. Shapiro, D. Rothschild, U. Mor, M. Dori-Bachash, C. Kleimeyer, C. Moresi, Y. Harnik, M. Zur, M. Zabari, R. B.-Z. Brik, D. Kviatcovsky, N. Zmora, Y. Cohen, N. Bar, I. Levi, N. Amar, T. Mehlman, A. Brandis, I. Biton, Y. Kuperman, M. Tsoory, L. Alfahel, A. Harmelin, M. Schwartz, A. Israelson, L. Arike, M. E. V. Johansson, G. C. Hansson, M. Gotkine, E. Segal and E. Elinav, *Nature*, 2019, **572**, 474–480.

154 K. Chellappa, M. R. McReynolds, W. Lu, X. Zeng, M. Makarov, F. Hayat, S. Mukherjee, Y. R. Bhat, S. R. Lingala, R. T. Shima, H. C. Descamps, T. Cox, L. Ji, C. Jankowski, Q. Chu, S. M. Davidson, C. A. Thaiss, M. E. Migaud, J. D. Rabinowitz and J. A. Baur, *Cell Metab.*, 2022, **34**, 1947–1959.

155 J. R. Mayers, J. Varon, R. R. Zhou, M. Daniel-Ivad, C. Beaulieu, A. Bhosle, N. R. Glasser, F. M. Lichtenauer, J. Ng, M. P. Vera, C. Huttenhower, M. A. Perrella, C. B. Clish, S. D. Zhao, R. M. Baron and E. P. Balskus, *Cell*, 2024, **187**, 4095–4112.

156 D. Janek, A. Zipperer, A. Kulik, B. Krismer and A. Peschel, *PLoS Pathog.*, 2016, **12**, e1005812.

157 B. O. Torres Salazar, T. Dema, N. A. Schilling, D. Janek, J. Bornikoel, A. Berscheid, A. M. A. Elsherbini, S. Krauss, S. J. Jaag, M. Lämmerhofer, M. Li, N. Alqahtani, M. J. Horsburgh, T. Weber, J. M. Beltrán-Beleña, H. Brötz-Oesterhelt, S. Grond, B. Krismer and A. Peschel, *Nat. Microbiol.*, 2024, **9**, 200–213.

158 A. Zipperer, M. C. Konnerth, C. Laux, A. Berscheid, D. Janek, C. Weidenmaier, M. Burian, N. A. Schilling, C. Slavetinsky, M. Marschal, M. Willmann, H. Kalbacher, B. Schittek, H. Brötz-Oesterhelt, S. Grond, A. Peschel and B. Krismer, *Nature*, 2016, **535**, 511–516.

159 T. Nakatsuji, T. H. Chen, S. Narala, K. A. Chun, A. M. Two, T. Yun, F. Shafiq, P. F. Kotol, A. Bouslimani, A. V. Melnik, H. Latif, J. N. Kim, A. Lockhart, K. Artis, G. David, P. Taylor, J. Streib, P. C. Dorrestein, A. Grier, S. R. Gill,

K. Zengler, T. R. Hata, D. Y. Leung and R. L. Gallo, *Sci. Transl. Med.*, 2017, **9**, eaah4680.

160 M. M. Severn, M. R. Williams, A. Shahbandi, Z. L. Bunch, L. M. Lyon, A. Nguyen, L. S. Zaramela, D. A. Todd, K. Zengler, N. B. Cech, R. L. Gallo and A. R. Horswill, *mBio*, 2022, **13**, e00930.

161 M. M. Brown, J. M. Kwiecinski, L. M. Cruz, A. Shahbandi, D. A. Todd, N. B. Cech and A. R. Horswill, *Antimicrob. Agents Chemother.*, 2020, **64**, e00172.

162 M. M. Severn, Y. K. Cho, H. S. Manzer, Z. L. Bunch, A. Shahbandi, D. A. Todd, N. B. Cech and A. R. Horswill, *J. Invest. Dermatol.*, 2022, **142**, 3349–3352.

163 T. Nakatsuji, T. R. Hata, Y. Tong, J. Y. Cheng, F. Shafiq, A. M. Butcher, S. S. Salem, S. L. Brinton, A. K. Rudman Spergel, K. Johnson, B. Jepson, A. Calatroni, G. David, M. Ramirez-Gama, P. Taylor, D. Y. M. Leung and R. L. Gallo, *Nat. Med.*, 2021, **27**, 700–709.

164 A. M. King, Z. Zhang, E. Glassey, P. Siuti, J. Clardy and C. A. Voigt, *Nat. Microbiol.*, 2023, **8**, 2420–2434.

165 T. Gensollen, S. S. Iyer, D. L. Kasper and R. S. Blumberg, *Science*, 2016, **352**, 539–544.

166 W. R. Wikoff, A. T. Anfora, J. Liu, P. G. Schultz, S. A. Lesley, E. C. Peters and G. Siuzdak, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 3698–3703.

167 Y. Lai, C.-W. Liu, Y. Yang, Y.-C. Hsiao, H. Ru and K. Lu, *Nat. Commun.*, 2021, **12**, 6000.

168 J. Ni, G. D. Wu, L. Albenberg and V. T. Tomov, *Nat. Rev. Gastroenterol. Hepatol.*, 2017, **14**, 573–584.

169 Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y.-M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. W. Tang, J. A. DiDonato, A. J. Lusis and S. L. Hazen, *Nature*, 2011, **472**, 57–63.

170 R. A. Koeth, Z. Wang, B. S. Levison, J. A. Buffa, E. Org, B. T. Sheehy, E. B. Britt, X. Fu, Y. Wu, L. Li, J. D. Smith, J. A. DiDonato, J. Chen, H. Li, G. D. Wu, J. D. Lewis, M. Warrier, J. M. Brown, R. M. Krauss, W. H. W. Tang, F. D. Bushman, A. J. Lusis and S. L. Hazen, *Nat. Med.*, 2013, **19**, 576–585.

171 K. H. Liu, J. A. Owens, B. Saeedi, C. E. Cohen, M. P. Bellissimo, C. Naudin, T. Darby, S. Druzak, K. Maner-Smith, M. Orr, X. Hu, J. Fernandes, M. C. Camacho, S. Hunter-Chang, D. VanInsberghe, C. Ma, T. Ganesh, S. M. Yeligar, K. Uppal, Y.-M. Go, J. A. Alvarez, M. B. Vos, T. R. Ziegler, M. H. Woodworth, C. S. Kraft, R. M. Jones, E. Ortlund, A. S. Neish and D. P. Jones, *Nat. Metab.*, 2021, **3**, 1694–1705.

172 M. A. Hoffmann, L.-F. Nothias, M. Ludwig, M. Fleischauer, E. C. Gentry, M. Witting, P. C. Dorrestein, K. Dührkop and S. Böcker, *Nat. Biotechnol.*, 2021, **40**, 411–421.

173 I. Mohanty, C. Allaband, H. Mannochio-Russo, Y. El Abiead, L. R. Hagey, R. Knight and P. C. Dorrestein, *Nat. Rev. Gastroenterol. Hepatol.*, 2024, **21**, 493–516.

174 D. Petras, A. M. Caraballo-Rodríguez, A. K. Jarmusch, C. Molina-Santiago, J. M. Gauglitz, E. C. Gentry, P. Belda-Ferre, D. Romero, S. M. Tsunoda, P. C. Dorrestein and M. Wang, *Anal. Chem.*, 2021, **93**, 12833–12839.

175 Y.-Z. Wang, P.-C. Mei, P.-R. Bai, N. An, J.-G. He, J. Wang, Q.-F. Zhu and Y.-Q. Feng, *Anal. Chim. Acta*, 2023, **1239**, 340691.

176 Q.-F. Zhu, Y.-Z. Wang, N. An, J.-D. Hao, P.-C. Mei, Y.-L. Bai, Y.-N. Hu, P.-R. Bai and Y.-Q. Feng, *Anal. Chem.*, 2022, **94**, 2655–2664.

177 Y. Zhang, Y. Huang, J. Fan, M. Zhang, A. Hasan, Y. Yi, R. Yu, X. Zhou, M. Ye and X. Qiao, *Anal. Chem.*, 2022, **94**, 16596–16603.

178 L. N. Lucas, K. Barrett, R. L. Kerby, Q. Zhang, L. E. Cattaneo, D. Stevenson, F. E. Rey and D. Amador-Noguez, *mSystems*, 2021, **6**, e0080521.

179 C. J. Garcia, V. Kosek, D. Beltrán, F. A. Tomás-Barberán and J. Hajslova, *Biomolecules*, 2022, **12**, 687.

180 D. V. Guzior, M. Okros, M. Shivel, B. Armwald, C. Bridges, Y. Fu, C. Martin, A. L. Schilmiller, W. M. Miller, K. M. Ziegler, M. D. Sims, M. E. Maddens, S. F. Graham, R. P. Hausinger and R. A. Quinn, *Nature*, 2024, **626**, 852–858.

181 B. Rimal, S. L. Collins, C. E. Tanes, E. R. Rocha, M. A. Granda, S. Solanki, N. J. Hoque, E. C. Gentry, I. Koo, E. R. Reilly, F. Hao, D. Paudel, V. Singh, T. Yan, M. S. Kim, K. Bittinger, J. P. Zackular, K. W. Krausz, D. Desai, S. Amin, J. P. Coleman, Y. M. Shah, J. E. Bisanz, F. J. Gonzalez, J. P. Vanden Heuvel, G. D. Wu, B. S. Zemel, P. C. Dorrestein, E. E. Weinert and A. D. Patterson, *Nature*, 2024, **626**, 859–863.

182 L. Dohnalová, P. Lundgren, J. R. E. Carty, N. Goldstein, S. L. Wenski, P. Nanudorn, S. Thiengmag, K.-P. Huang, L. Litichevskiy, H. C. Descamps, K. Chellappa, A. Glassman, S. Kessler, J. Kim, T. O. Cox, O. Dmitrieva-Posocco, A. C. Wong, E. L. Allman, S. Ghosh, N. Sharma, K. Sengupta, B. Cornes, N. Dean, G. A. Churchill, T. S. Khurana, M. A. Sellmyer, G. A. FitzGerald, A. D. Patterson, J. A. Baur, A. L. Alhadeff, E. J. N. Helfrich, M. Levy, J. N. Betley and C. A. Thaiss, *Nature*, 2022, **612**, 739–747.

183 F.-Y. Chang, P. Siuti, S. Laurent, T. Williams, E. Glassey, A. W. Sailer, D. B. Gordon, H. Hemmerle and C. A. Voigt, *Nat. Microbiol.*, 2021, **6**, 792–805.

184 S. Gaetani, F. Oveisi and D. Piomelli, *Neuropsychopharmacology*, 2003, **28**, 1311–1316.

185 R. Y. Chen, I. Mostafa, M. C. Hibberd, S. Das, M. Mahfuz, N. N. Naila, M. M. Islam, S. Huq, M. A. Alam, M. U. Zaman, A. S. Raman, D. Webber, C. Zhou, V. Sundaresan, K. Ahsan, M. F. Meier, M. J. Barratt, T. Ahmed and J. I. Gordon, *N. Engl. J. Med.*, 2021, **384**, 1517–1528.

186 J. Cheng, S. Venkatesh, K. Ke, M. J. Barratt and J. I. Gordon, *Science*, 2024, **386**, eado6828.

187 N. Battista, M. Bari and T. Bisogno, *Biomolecules*, 2019, **9**, 822.

188 L. J. Cohen, H.-S. Kang, J. Chu, Y.-H. Huang, E. A. Gordon, B. V. B. Reddy, M. A. Ternei, J. W. Craig and S. F. Brady, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E4825–E4834.

189 N. González-Silva, I. M. López-Lara, R. Reyes-Lamothe, A. M. Taylor, D. Sumpton, J. Thomas-Oates and O. Geiger, *Biochemistry*, 2011, **50**, 6396–6408.
190 V. C. Scoffone, L. R. Chiarelli, G. Trespidi, M. Mentasti, G. Riccardi and S. Buroni, *Front. Microbiol.*, 2017, **8**, 1592.
191 L. A. Córdoba-Castro, R. Salgado-Morales, M. Torres, L. Martínez-Aguilar, L. Lozano, M. Á. Vences-Guzmán, Z. Guan, E. Dantán-González, M. Serrano and C. Sohlenkamp, *Front. Mol. Biosci.*, 2021, **7**, 610932.
192 M. Pizzuto, L. Hurtado-Navarro, C. Molina-Lopez, J. Soubhye, M. Gelbcke, S. Rodriguez-Lopez, J.-M. Ruysschaert, K. Schroder and P. Pelegrin, *Cell. Rep.*, 2024, **43**, 114788.
193 H. Mannochio-Russo, V. Charron-Lamoureux, M. van Faassen, S. Lamichhane, W. D. Gonçalves Nunes, V. Deleray, A. V. Ayala, Y. Tanaka, A. Patan, K. Vittali, P. Rajkumar, Y. El Abiead, H. N. Zhao, P. W. P. Gomes, I. Mohanty, C. Lee, A. Sund, M. Sharma, Y. Liu, D. Pattynama, G. T. Walker, G. J. Norton, L. Khatib, M. S. Andalibi, C. X. Wang, R. J. Ellis, D. J. Moore, J. E. Iudicello, D. Franklin, S. Letendre, L. Chin, C. Walker, S. Renwick, J. Zemlin, M. J. Meehan, X. Song, D. Kasper, Z. Burcham, J. J. Kim, S. Kadakia, M. Raffatellu, L. Bode, H. Chu, K. Zengler, M. Wang, D. Siegel, R. Knight and P. C. Dorrestein, *Cell*, 2025, DOI: **10.1016/j.cell.2025.05.015**.
194 M. Özçam, J. H. Oh, R. Tocmo, D. Acharya, S. Zhang, T. J. Astmann, M. Heggen, S. Ruiz-Ramírez, F. Li, C. C. Cheng, E. Vivas, F. E. Rey, J. Claesen, T. S. Bugni, J. Walter and J. P. van Pijkeren, *Cell Host Microbe*, 2022, **30**, 824–835.
195 E. Y. Hsiao, S. W. McBride, S. Hsien, G. Sharon, E. R. Hyde, T. McCue, J. A. Codelli, J. Chow, S. E. Reisman, J. F. Petrosino, P. H. Patterson and S. K. Mazmanian, *Cell*, 2013, **155**, 1451–1463.
196 B. D. Needham, M. D. Adame, G. Serena, D. R. Rose, G. M. Preston, M. C. Conrad, A. Stewart Campbell, D. H. Donabedian, A. Fasano, P. Ashwood and S. K. Mazmanian, *Biol. Psychiatry*, 2020, **89**, 451–462.
197 B. D. Needham, M. Funabashi, M. D. Adame, Z. Wang, J. C. Boktor, J. Haney, W.-L. Wu, C. Rabut, M. S. Ladinsky, S.-J. Hwang, Y. Guo, Q. Zhu, J. A. Griffiths, R. Knight, P. J. Bjorkman, M. G. Shapiro, D. H. Geschwind, D. P. Holschneider, M. A. Fischbach and S. K. Mazmanian, *Nature*, 2022, **602**, 647–653.
198 K. M. Knights, D. M. Stresser, J. O. Miners and C. L. Crespi, *Curr. Protoc. Pharmacol.*, 2016, **74**, 7.
199 L. Guthrie and L. Kelly, *EBioMedicine*, 2019, **44**, 708–715.
200 L. Guthrie, S. Gupta, J. Daily and L. Kelly, *NPJ Biofilms Microbi.*, 2017, **3**, 27.
201 J. Sunwoo, S. C. Ji, A. H. Kim, K.-S. Yu, J.-Y. Cho, I.-J. Jang and S. Lee, *Clin. Transl. Sci.*, 2020, **13**, 752–760.
202 A. H. Kim, Y. Lee, E. Kim, S. C. Ji, J.-Y. Chung and J.-Y. Cho, *Front. Cell. Infect. Microbiol.*, 2021, **11**, 629438.
203 C. N. Brocker, T. Velenosi, H. K. Flaten, G. McWilliams, K. McDaniel, S. K. Shelton, J. Saben, K. W. Krausz, F. J. Gonzalez and A. A. Monte, *Hum. Genomics*, 2020, **14**, 10.
204 T. Pallister, M. A. Jackson, T. C. Martin, J. Zierer, A. Jennings, R. P. Mohney, A. MacGregor, C. J. Steves, A. Cassidy, T. D. Spector and C. Menni, *Sci. Rep.*, 2017, **7**, 1–9.
205 A. K. Jarmusch, A. Vrbanac, J. D. Momper, J. D. Ma, M. Alhaja, M. Liyanage, R. Knight, P. C. Dorrestein and S. M. Tsunoda, *Clin. Transl. Sci.*, 2020, **13**, 972–984.
206 R. S. Mehta, J. R. Mayers, Y. Zhang, A. Bhosle, N. R. Glasser, L. H. Nguyen, W. Ma, S. Bae, T. Branck, K. Song, L. Sebastian, J. A. Pacheco, H.-S. Seo, C. Clish, S. Dhe-Paganon, A. N. Ananthakrishnan, E. A. Franzosa, E. P. Balskus, A. T. Chan and C. Huttenhower, *Nat. Med.*, 2023, **29**, 700–709.
207 M. Zimmermann, M. Zimmermann-Kogadeeva, R. Wegmann and A. L. Goodman, *Science*, 2019, **363**, eaat9931.
208 C. G. Jones, M. W. Martynowycz, J. Hattne, T. J. Fulton, B. M. Stoltz, J. A. Rodriguez, H. M. Nelson and T. Gonen, *ACS Cent. Sci.*, 2018, **4**, 1587–1592.
209 E. Danelius, S. Halaby, W. A. van der Donk and T. Gonen, *Nat. Prod. Rep.*, 2020, **38**, 423–431.
210 K. M. Pruss, H. Chen, Y. Liu, W. Van Treuren, S. K. Higginbottom, J. B. Jarman, C. R. Fischer, J. Mak, B. Wong, T. M. Cowan, M. A. Fischbach, J. L. Sonnenburg and D. Dodd, *Nat. Commun.*, 2023, **14**, 512.
211 J. Buckingham, *Dictionary of Natural Products*, CRC Press, Boca Raton, 1993.
212 Y. Wang, E. Glukhov, Y. He, Y. Liu, L. Zhou, X. Ma, X. Hu, P. Hong, W. H. Gerwick and Y. Zhang, *Antibiotics*, 2022, **11**, 513.
213 S. R. Lee, F. Schalk, J. W. Schwitalla, H. Guo, J. S. Yu, M. Song, W. H. Jung, Z. W. de Beer, C. Beemelmanns and K. H. Kim, *Chemistry*, 2022, **28**, e202200612.
214 H. Laatsch, *AntiBase: The Natural Compound Identifier*, Wiley-VCH, Weinheim, 2017.
215 Y. Zhang, M. T. Cheema, L. V. Ponomareva, Q. Ye, T. Liu, I. Sajid, J. Rohr, Q.-B. She, S. R. Voss, J. S. Thorson and K. A. Shaaban, *J. Nat. Prod.*, 2021, **84**, 1930–1940.
216 S. W. Gabrielson, *J. Med. Libr. Assoc.*, 2018, **106**, 588–590.
217 N. A. Khan, M. Soopramanien, S. K. Maciver, T. S. Anuar, K. Sagathevan and R. Siddiqui, *Molecules*, 2021, **26**, 4999.
218 D. S. Wishart, D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.-A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. Macinnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel and L. Querengesser, *Nucleic Acids Res.*, 2007, **35**, D521–D526.
219 D. S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B. L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V. W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao,

Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H. B. Schiöth, R. Greiner and V. Gautam, *Nucleic Acids Res.*, 2022, **50**, D622–D631.

220 Y. Feng, D. Liu, Y. Liu, X. Yang, M. Zhang, F. Wei, D. Li, Y. Hu and Y. Guo, *iScience*, 2022, **25**, 104469.

221 D. S. Wishart, E. Oler, H. Peters, A. Guo, S. Girod, S. Han, S. Saha, V. W. Lui, M. LeVatte, V. Gautam, R. Kaddurah-Daouk and N. Karu, *Nucleic Acids Res.*, 2023, **51**, D611–D620.

222 B. Bohn, C. Tilves, Y. Chen, M. Doyon, L. Bouchard, P. Perron, R. Guérin, É. Massé, M.-F. Hivert and N. T. Mueller, *BMJ Open Gastroenterol.*, 2024, **11**, e001470.

223 X. Lucas, C. Senger, A. Erxleben, B. A. Grüning, K. Döring, J. Mosch, S. Flemming and S. Günther, *Nucleic Acids Res.*, 2013, **41**, D1130–D1136.

224 J. A. van Santen, G. Jacob, A. L. Singh, V. Aniebok, M. J. Balunas, D. Bunsko, F. C. Neto, L. Castaño-Espriu, C. Chang, T. N. Clark, J. L. Cleary Little, D. A. Delgadillo, P. C. Dorrestein, K. R. Duncan, J. M. Egan, M. M. Galey, F. P. J. Haeckl, A. Hua, A. H. Hughes, D. Iskakova, A. Khadilkar, J.-H. Lee, S. Lee, N. LeGrow, D. Y. Liu, J. M. Macho, C. S. McCaughey, M. H. Medema, R. P. Neupane, T. J. O'Donnell, J. S. Paula, L. M. Sanchez, A. F. Shaikh, S. Soldatou, B. R. Terlouw, T. A. Tran, M. Valentine, J. J. J. van der Hooft, D. A. Vo, M. Wang, D. Wilson, K. E. Zink and R. G. Linington, *ACS Cent. Sci.*, 2019, **5**, 1824–1833.

225 J. A. van Santen, E. F. Poynton, D. Iskakova, E. McMann, T. A. Alsup, T. N. Clark, C. H. Fergusson, D. P. Fewer, A. H. Hughes, C. A. McCadden, J. Parra, S. Soldatou, J. D. Rudolf, E. M.-L. Janssen, K. R. Duncan and R. G. Linington, *Nucleic Acids Res.*, 2022, **50**, D1317–D1323.

226 X. Zeng, P. Zhang, W. He, C. Qin, S. Chen, L. Tao, Y. Wang, Y. Tan, D. Gao, B. Wang, Z. Chen, W. Chen, Y. Y. Jiang and Y. Z. Chen, *Nucleic Acids Res.*, 2018, **46**, D1217–D1222.

227 H. Zhao, Y. Yang, S. Wang, X. Yang, K. Zhou, C. Xu, X. Zhang, J. Fan, D. Hou, X. Li, H. Lin, Y. Tan, S. Wang, X.-Y. Chu, D. Zhuoma, F. Zhang, D. Ju, X. Zeng and Y. Z. Chen, *Nucleic Acids Res.*, 2023, **51**, D621–D628.

228 A. Shtaiwi, S. U. Khan, M. Khedraoui, M. Alaraj, A. Samadi and S. Chtita, *Sci. Rep.*, 2024, **14**, 7098.

229 W. Huang, L. K. Brewer, J. W. Jones, A. T. Nguyen, A. Marcu, D. S. Wishart, A. G. Oglesby-Sherrouse, M. A. Kane and A. Wilks, *Nucleic Acids Res.*, 2018, **46**, D575–D580.

230 M. Sorokina, P. Merseburger, K. Rajan, M. A. Yirik and C. Steinbeck, *J. Cheminform.*, 2021, **13**, 2.

