Disentangling host–microbiota complexity through hologenomics

Antton Alberdi[1] ✉, *Sandra B. Andersen*[1], *Morten T. Limborg*[1], *Robert R. Dunn*[1,2] *and M. Thomas P. Gilbert*[1,3]

Abstract | Research on animal–microbiota interactions has become a central topic in biological sciences because of its relevance to basic eco-evolutionary processes and applied questions in agriculture and health. However, animal hosts and their associated microbial communities are still seldom studied in a systemic fashion. Hologenomics, the integrated study of the genetic features of a eukaryotic host alongside that of its associated microbes, is becoming a feasible — yet still underexploited — approach that overcomes this limitation. Acknowledging the biological and genetic properties of both hosts and microbes, along with the advantages and disadvantages of implemented techniques, is essential for designing optimal studies that enable some of the major questions in biology to be addressed.

Metagenotype
The specific state of the microbial metagenome characterized at a particular moment and at a given resolution.

Hologenotype
The entire genetic constitution of an individual eukaryotic organism and its associated microorganisms characterized at a given moment and at a given resolution.

[1]*Center for Evolutionary Hologenomics, The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark.*

[2]*Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA.*

[3]*University Museum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.*

✉*e-mail: antton.alberdi@sund.ku.dk*

https://doi.org/10.1038/s41576-021-00421-0

Biological systems composed of a eukaryotic organism and its microbiota are widespread in nature[1–3]. Also known as holobionts[4], host–microbiota systems range in complexity from one-to-one symbiotic associations between a host and a single microorganism (such as the bioluminiscent *Aliivibrio* bacteria in light organs of bobtail squids[5]) to intricate arrangements between a host and a dynamic community of microorganisms (such as mammals and their gut microbiota[6], or plants and their root microbiota[7]) (FIG. 1a). Many such systems play core functional roles in ecosystems or are fundamental to agricultural production (FIG. 1b). Hence, their study and manipulation have become a global strategic priority[8,9].

The study of functional associations of eukaryotic hosts with microbial communities dates to at least the late nineteenth century, when researchers noticed, based on microscopy, that termite guts contained microbial species on which the termites appeared to rely[10] and the intestines of infants hosted bacteria that aided in digestion[11]. Subsequent research on insects[12], farm animals[13] and humans[14] relied on a combination of culture-based approaches, functional assays and microscopy to characterize the microbial taxa associated with these hosts. The more recent development of culture-free molecular technologies has provided new ways to study and understand the complex details of such systems. Today, we are able to generate high-quality eukaryotic genome sequences (at the level of near-complete chromosomes)[15] and recover virtually complete bacterial genomes from samples that contain complex mixtures of DNA (such as faeces or plant roots)[16] relatively quickly and in ways that are replicable within and among host species. Taken together, these advances enable the simultaneous characterization of the so-called hologenome[17], that is, the entirety of the genetic information contained by a eukaryotic host and the microbial community associated with that host.

The terms hologenome and hologenomics used in this Review are primarily instrumental[18,19]. They reflect an approach to the study of microbes and hosts that is holistic rather than atomized and that transcends disciplinary and taxonomic boundaries. The simultaneous analysis of host and microbial genes and genomes naturally allows the study of fundamental evolutionary aspects of host–microorganism associations[20] and, as a result, will contribute to discussions about general evolutionary models for the role of selection on hosts, microbes and their amalgams[21–24]. However, the utility of hologenomics is much broader as it has also proven valuable for the study of diseases[25], the improvement of farming practices[19,26] and in understanding molecular interactions between eukaryotic and prokaryotic cells[27], to cite but a few examples.

Hologenomics leverages the latest advances in genomics and metagenomics to disentangle the complexity of host–microbiota systems by using a systemic approach. This enables not only detailed profiling of the host genotype but also characterization of the microbial metagenotype and, by extension, the definition of the hologenotype. High-resolution characterization of all these units of analysis is only possible through the combination of high-throughput DNA, RNA, protein and metabolite profiling[28], novel mathematical tools to integrate the different omic layers[29,30], and the generation of high-quality reference genomes of both eukaryotic and prokaryotic organisms[15,31]. This hologenomic approach

is unlocking domains of knowledge that have remained inaccessible until recently because of their excessive complexity and the costs of data generation.

In this Review, we first address five fundamental criteria that should ideally be considered in the design of any hologenomic study. In doing so, we explain the relevant methods and model systems available to address a variety of biological questions using a hologenomic approach. Subsequently, we outline how the information extracted using this methodology can be used to address major outstanding questions in life sciences and to design novel interdisciplinary approaches to expand our knowledge horizon.

Designing hologenomic studies

Hologenomics can be used to understand how the combined features of hosts and microorganisms shape biological processes relevant for hosts (such as adaptation), for microorganisms (such as meta-community dynamics) or both (BOX 1). Depending on the aims and features of the study system, hologenomics can be implemented using different study designs, model systems and techniques[18,19,32]. We propose that this landscape of possibilities is shaped around five essential questions that need to be considered when designing and interpreting hologenomic studies (FIG. 2), which relate to five core topics: hologenomic complexity, control of hologenomic variables, hologenomic resolution, spatiotemporal resolution and explanatory versus response variables.

Hologenomic complexity

Hologenomic complexity can be broadly defined as the amount of information relevant to the study that the biological system under analysis contains and it can be decomposed into three major elements: host genomic, microbial metagenomic and environmental complexity. Within each of these elements, two sources of complexity can be defined: the intrinsic complexity of the system under study, including host genome size and number of bacterial genomes, and the complexity introduced by the

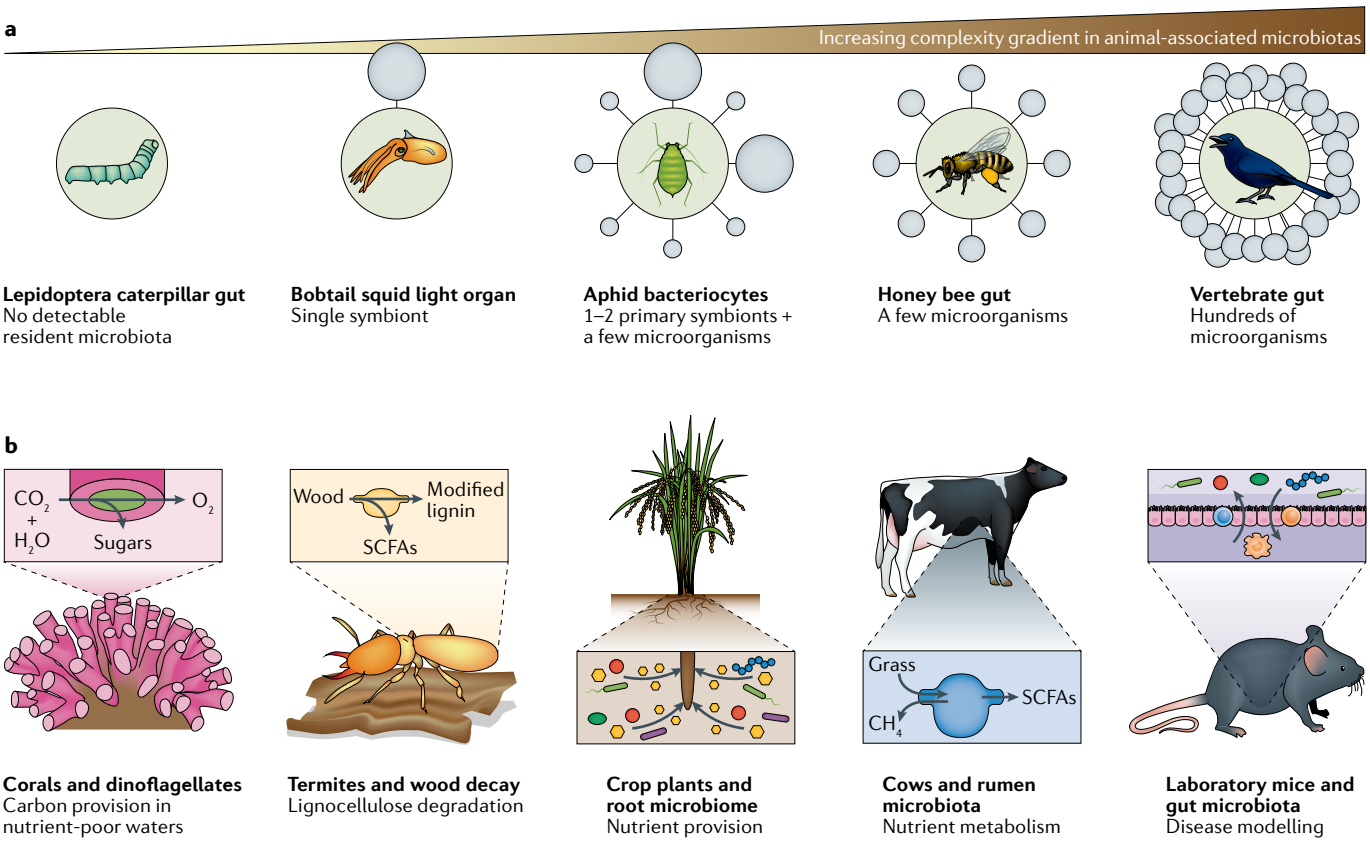


Fig. 1 | **Diversity and ecological relevance of host–microbiota systems.** **a** | Complexity gradient of microbial communities associated with specific body parts of various animals. Many Lepidoptera caterpillars have no detectable resident gut microbiota[163]. Bobtail squids maintain a symbiotic relationship with a single type of bacteria that colonizes their light organs, which provides them with skin luminescence[5]. Pea aphid bacteriocytes harbour one or two dominant bacteria along with a few secondary taxa with considerably lower representation[164]. The gut microbiota of bees is composed of a few types of bacteria[114]. Birds harbour a complex microbiota with hundreds of bacterial species[165]. **b** | Examples of different host–microbiota systems and their relevance for ecosystem or applied processes. Photosynthetic dinoflagellates provide corals with a source of carbon in nutrient-poor tropical waters, forming the basis of coral reefs[166]. Lignocellulose-degrading bacteria in the gut of termites transform complex polysaccharides into short-chain fatty acids (SCFAs) and modify lignin from wood substrates[167]. The root microbiota of plants plays a central role in transforming complex organic compounds into simple nutrients that plant roots can absorb, thus boosting productivity[7]. The rumen microbiota is important for producing SCFAs and other essential compounds through the fermentation of complex plant polysaccharides, which act both as energy sources and signalling elements in animal hosts[168]. The gut microbiota of mice plays a crucial role in laboratory disease models because of their effects on many systemic conditions related to energy metabolism and immune response[169].

Box 1 | **Hologenomics from different perspectives**

While host–microbiota systems cover a range of biological attributes (FIG. 1), the asymmetry between the host genome and the microbial metagenome is a fundamental feature that characterizes them all. The host genome typically belongs to a single multicellular macroorganism, whose generation time is usually much longer than that of its associated microbes. By contrast, the microbial metagenome represents the composite of the genomes of millions of unicellular organisms that may belong to hundreds or thousands of different taxa. In addition to their typically shorter generation times, they often exhibit a capacity for horizontal gene transfer that increases their adaptive potential[170]. As a result, whereas the host genotype is mostly fixed throughout the host's life (although see discussion below), the microbial metagenotype and the resulting hologenotype change continuously[150,171]. Because of this and other differences addressed elsewhere[21], the routes available for the host to affect the microbiota and vice versa differ, as does the information hologenomics provides when applied to describe a system from the microbial versus the host perspective.

The microbial perspective

The host provides the physical basis for the microbiota and defines many properties of their ecosystem[172]. Hence, the genomic features of the host contribute to configuring ecological niches that microorganisms can occupy[173]. For instance, gut morphology shapes oxygen gradients[43], the production of lipids determines the structure of skin microbiota[174] and host macrophages can exert direct selection on the microbial community[175]. In animals, the host genome can also indirectly affect the transmission, acquisition and community dynamics of the microbiota through moulding host behaviour — sociality, nursing behaviour and diet are three of the main factors that shape microbial communities associated with animals[176–178].

Hologenomics is not only useful for studying how host genomic variation changes the conditions that the microbial communities experience but also for understanding how the microbiota alters host processes that shape such a landscape. Microorganisms can modulate the expression of specific genes in the host. In mouse models, for instance, intestinal spore-forming bacteria release metabolites that modify the expression of various host genes involved in the biosynthesis of serotonin[132]. Microorganisms can also trigger more complex systemic changes that induce developmental[179] or behavioural[180,181] modifications in the host, which can indirectly affect the environment in which microorganisms live.

The host perspective

Genomic features of microorganisms can influence host biology by causing disease[182] or enhancing biological capabilities[125]. Bacteria can significantly shape a myriad of host phenotypic traits due to their capacity to metabolize complex foods[183], produce essential biomolecules[184], modulate host gene expression[132], promote epigenomic changes[185] or trigger hormonal cascades[186]. Hologenomics can be used to understand how host genomic features condition the metabolic functions provided by its microbiota and how the metagenomic features of these microorganisms, along with the genotype of the hosts[128], modulate essential biological processes for the host[18].

degree of difference between the organisms under comparison such as gene expression differences versus distinct genomes (FIG. 3). The combination of host genomic, microbial metagenomic and environmental complexity will determine the relevance of each factor in the study system and the questions that can be answered (FIG. 4). This information could be used either for selecting the most appropriate system to address a specific biological question or, when researchers are bound to a given system, for adjusting the study design to the properties of the system.

Host genomic complexity. The intrinsic complexity of host genomes is primarily determined by features such as genome size, number of chromosomes, ploidy, amount of functional genetic elements and number of repetitive sequences[33]. These features vary across eukaryotes[34,35] and not only determine the degree of complexity under analysis but also how data are generated (for example, by considering complications related to reconstructing high-quality reference genomes). Thus, effective implementation of a hologenomic approach in a host organism whose genome is relatively simple and well characterized (such as *Caenorhabditis elegans* or *Mus musculus*) is more feasible than those with extremely large, duplicate-rich genomes (such as salamanders[36] or polyploid plants[37]) (FIG. 3a).

Host complexity is also determined by the experimental setup itself. Hosts under study may range from individuals with practically identical genomic features (such as monozygotic twins[38] or inbred mice[39]) to intraspecific variants (such as different landraces of plants[40] or reproductively isolated populations of animals[41]) to different species or evolutionary lineages (for example, monocotyledonous versus dicotyledonous plants grown in a common garden experiment[42]) (FIG. 3a). In this gradient of host genomic complexity, the lower extreme considers interindividual genomic variation mainly introduced by SNPs, which normally introduce subtle differences in the morphological (for example, intestinal crypt properties[43]) and biomolecular (for example, antimicrobial peptide and immunoglobulin production[44]) features of the host. The upper extreme, meanwhile, typically addresses major genomic differences that yield important morphological and physicochemical changes that create conspicuous phenotypic differences across the hosts studied[6]. One interesting exception relates to the consideration of how genetically determined sex can introduce large-scale changes to microbial communities within conspecifics[45] through the major structural differences sexual dimorphism can induce on the body[46]. Both types of genomic variation are known to shape the microbial communities associated with the hosts through multiple underlying mechanisms[18].

Microbial metagenomic complexity. One fundamental aspect that determines intrinsic metagenomic complexity is the breadth of taxa included in the definition of the microbial metagenome. Whereas early metabarcoding-based studies of hologenomes often focused on single taxa, such as bacteria (for which both marker genes and databases were available early), shotgun sequencing-based approaches allow the consideration of the entirety of the microbial metagenome. In truth, most microbial metagenomes can also include viruses[47], other prokaryotes (such as archaea[48] and Candidate Phyla Radiation nanobacteria[49]), and eukaryotes such as fungi[50], protozoa[51] and helminths[52]. Even when studied in its entirety, the microbial community of interest can range from a single microbial symbiont to a complex microbial community composed of thousands of taxa (FIG. 1a).

Metagenomic complexity is also shaped by the taxonomic and functional differences among the microbial communities being compared, which can differ in terms of relative abundances of taxa or be compositionally distinct (FIG. 3b). However, gene expression differences caused by the highly dynamic nature of microbial communities do not always lead to significant structural changes and functional redundancy means that microbial turnover is not always translated into an effective functional change in the microbial ecosystem[53].

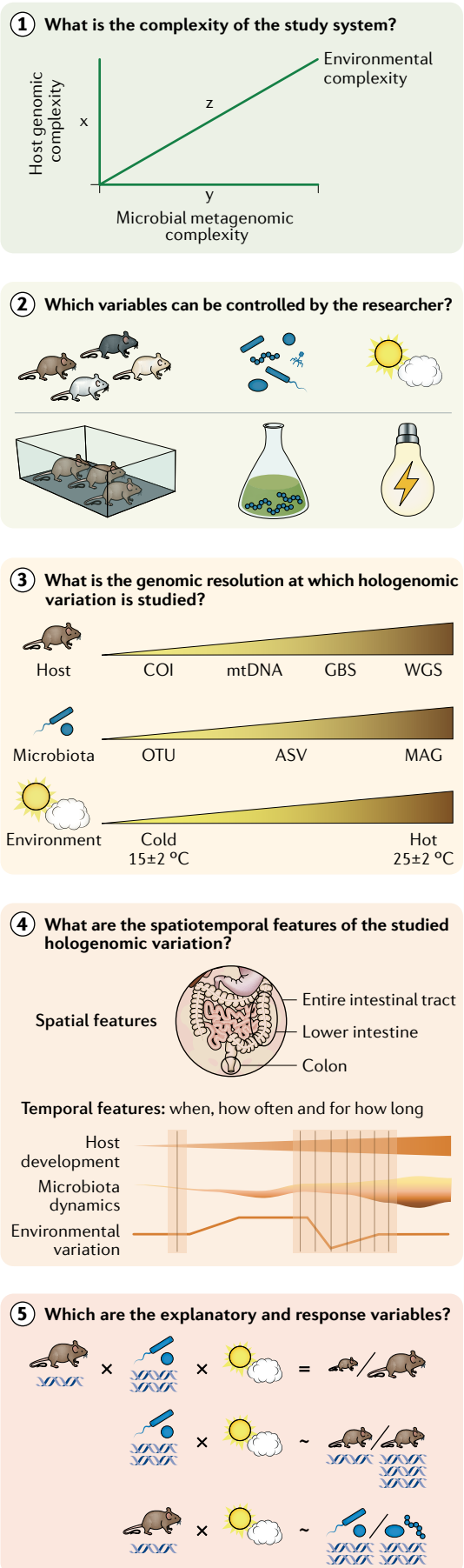


Fig. 2 | **Overview of the five essential criteria for designing and interpreting hologenomic studies.** 1. Complexity. The complexity of a hologenomic system can be decomposed into three elements: host genomic complexity, microbial metagenomic complexity and environmental complexity. 2. Control of variables. What information can be recovered from host–microbiota systems will depend on whether and to what degree host genomic, microbial metagenomic and environmental complexity can be controlled, minimized or removed by the researcher. 3. Genomic resolution. The resolution at which the host genome and the microbial metagenome are characterized and, in consequence, how genotypes, metagenotypes and combined hologenotypes are defined, determine the genomic level of complexity at which to study a host–microbiota system. 4. Spatiotemporal factors. Which questions can be addressed also depends on the intrinsic spatial and temporal features of hosts and microbial communities considered as well as on the spatial and temporal features of the study design. 5. Explanatory and response variables. Due to the multi-layered nature and bi-directional interactions in host–microbiota systems, the definition of explanatory and response variables and, thus, the question researchers want to answer will often be determined by the study design. ASV, amplicon sequence variant; COI, cytochrome oxidase subunit 1 gene; GBS, reduced representation genome sequencing through genotyping-by-sequencing; MAG, metagenome-assembled genome; mtDNA, complete mitochondrial genome; OTU, operational taxonomic unit; WGS, whole-genome sequencing.

Environmental complexity. The complexity of environmental factors, defined as any feature not directly determined by host genomic and microbial metagenomic properties, also plays an essential role when designing and interpreting hologenomic studies. Here, the environment might include abiotic conditions (such as cold versus warm) but also habitat (such as swamp versus forest or zoo versus wild) or diet (such as plant based versus animal based). Gene expression patterns of both hosts and microorganisms as well as other features of hosts, such as behaviour or physiological state, are largely driven by environmental cues[54–56]. A gradient of intrinsic complexity of the environment can be defined, spanning from laboratory experiments, in which all environmental features are kept constant, to natural ecosystems, in which both abiotic and biotic conditions differ from place to place, time to time, and between individual hosts (FIG. 3c). Similar to host and microbial complexity, environmental complexity is also determined by the experimental setup. The degree to which environmental conditions differ across the biological systems being compared will determine the overall environmental complexity of the study system.

Controlling hologenomic variables

Controlling the complexity of hologenomic variables is essential for addressing specific research questions. Broadly speaking, the more detailed and mechanistic the question under study, the greater the required control. For instance, research on specific biomolecular processes using laboratory models will require a higher

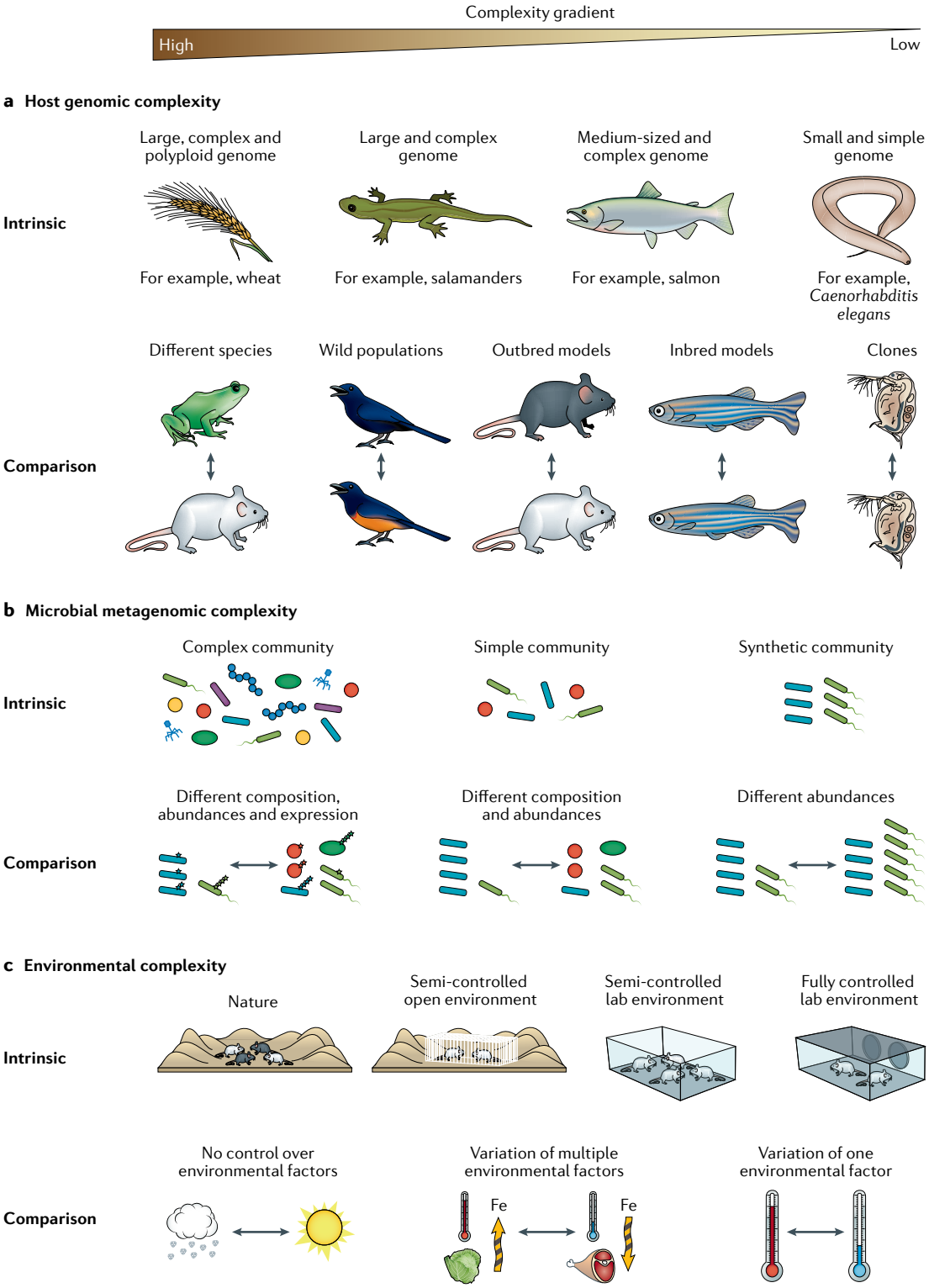
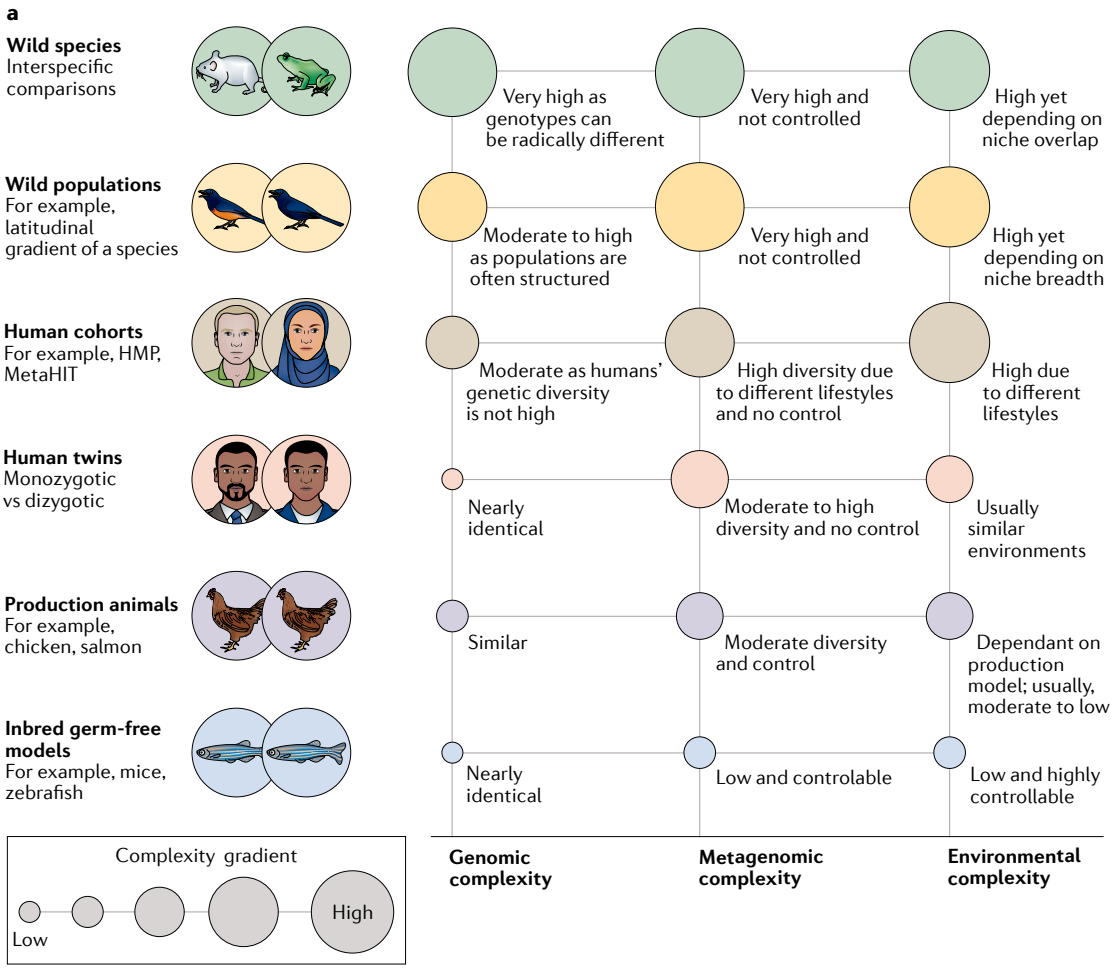
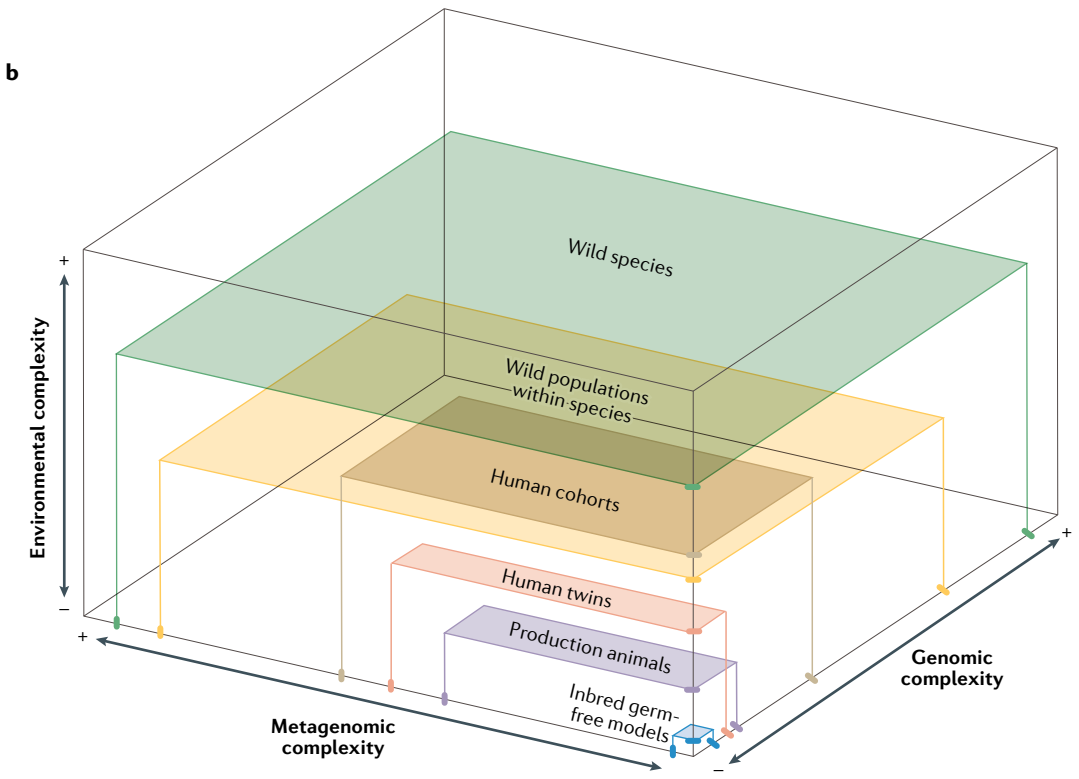


Fig. 3 | **Decomposition of hologenomic complexity.** The design and interpretation of hologenomic studies depend on the host genomic (part **a**), microbial metagenomic (part **b**) and environmental (part **c**) complexity of the system under study. Within each axis of complexity, two types of gradients can be defined based on whether the features are intrinsic to the system or introduced by the researcher through the selection of groups under comparison.

level of control than studying biogeographical patterns of host–microbiota interactions in wild organisms. The control of hologenomic variables can be achieved through a number of strategies.

Controlling host genomes. The control over host genomic complexity largely depends on the model organisms studied and the technical approaches employed. In laboratory organisms that can reproduce asexually, such as


a
Wild species
Interspecific comparisons
Wild populations
For example, latitudinal gradient of a species
Human cohorts
For example, HMP, MetaHIT
Human twins
Monozygotic vs dizygotic
Production animals
For example, chicken, salmon
Inbred germ-free models
For example, mice, zebrafish
Complexity gradient
Low
High
Very high as genotypes can be radically different
Very high and not controlled
High yet depending on niche overlap
Moderate to high as populations are often structured
Very high and not controlled
High yet depending on niche breadth
Moderate as humans' genetic diversity is not high
High diversity due to different lifestyles and no control
High due to different lifestyles
Nearly identical
Moderate to high diversity and no control
Usually similar environments
Similar
Moderate diversity and control
Dependant on production model; usually, moderate to low
Nearly identical
Low and controlable
Low and highly controllable
Genomic complexity
Metagenomic complexity
Environmental complexity


b
Wild species
Wild populations within species
Human cohorts
Human twins
Production animals
Inbred germ-free models
Environmental complexity
Metagenomic complexity
Genomic complexity
+
–

◀ Fig. 4 | **Hologenomic complexity of study systems.** Different study systems hold different levels of complexity, which limit the range of scientific questions that can be addressed. **a** | Six examples of study systems with different levels of genomic, metagenomic and environmental complexity. **b** | Three-dimensional representation of the complexity of the examples. The area of the plain represents the combined host genomic and microbial metagenomic complexity of the system, while the height represents the environmental complexity. The combined three-dimensional volume represents the overall hologenomic complexity of the system. HMP, Human Microbiome Project.

water fleas (*Daphnia*, Crustacea) and Lamiaceae plants, absolute control over host genotypes can be achieved by using clonal organisms[57,58]. When clones cannot be used, inbred laboratory animals can provide a high level of genomic homogeneity. The use of groups of genetically homogeneous hosts allows the effects of contrasting environmental conditions or specific microbial communities to be compared[59,60]. Clonal and inbred models also enable the effects of a specific host genetic factor to be studied in a controlled genomic background through the application of targeted techniques for modulating gene expression (such as RNA-mediated interference) or for genomic engineering (such as CRISPR–Cas9)[60,61]. Working with humans and wild organisms does not enable such a degree of control over the genotypes studied unless in vitro models, such as organ-on-a-chip co-cultures of animal tissues and microbial communities, are generated[62]. When this level of control is not possible, coarse control over host genotypes can be achieved through contrasting animals from different populations or from closely related species[63], while greater control can be achieved through comparing individuals across different degrees of kinship, such as monozygotic versus dizygotic twins[38], and family members to other individuals[64].

Controlling microbial metagenomes. Control over microbial metagenomic complexity is usually achieved through the modulation of microbial communities. Some strategies, such as modification of dietary regimes or the administration of microbiota-targeted additives or prebiotics, aim to modify microbial ecosystems by changing nutrient availability. However, unless compounds that match the unique enzymatic capabilities of specific microorganisms are used[65], it is difficult to accurately modulate the microbiota owing to the complexity of ecological relationships among microorganisms. Alternative approaches to modify microbial communities include the inoculation of target bacteria (such as probiotics)[66] and faecal microbiota transplantation[67]. The efficacy and accuracy of these methods is also variable; there is no guarantee that inoculated bacteria will establish or modulate the microbiota, while transplantation does not enable accurate control over the microbial community introduced or the secondary elements that are transplanted along with bacteria[68]. These issues complicate the interpretation of results; for example, bacteriophages transferred alongside bacteria may severely impact the gut microbiota composition[69]. A higher level of control could potentially be achieved through transplanting synthetic microbial communities[70,71]. While this approach has been successfully implemented in diverse in vitro setups[72,73], the complexity of microbial communities still hinders its efficient use as a routine scientific procedure in live animals[74]. The metagenotypes of these synthetic communities could be further controlled by genetically engineering bacteria with the desired genetic properties[75,76], which could eventually facilitate the assembly and control of interactive microbial consortia[77].

Controlling the environment. In most laboratory studies, environmental complexity is reduced so that no or very few environmental parameters (usually only experimental treatments) vary among groups and subjects. Climate chambers and aquaria provide absolute control of abiotic conditions such as light/dark cycles, humidity and temperature variations. Outdoor common garden experiments[42] do not provide full control over environmental factors but they ensure that the effect on the systems being compared is identical. Some natural systems can also provide special conditions that enable environmental features to be controlled, such as cuckoo nestlings that are bred by other birds[78], or sympatrically occurring but allochronically isolated populations such as salmon populations that breed in the same rivers in alternating years[79]. Research on wild organisms usually incorporates more complex and dynamic environmental conditions; when controlling them is not possible, it is useful to collect relevant environmental metadata, which can be incorporated as covariates in the statistical analyses. A century of ecological research has revealed the advantages of each of these approaches. On the one extreme, laboratory microcosms allow the most reductive control; on the other, studies in the macrocosm of the real world provide perspective on emergent properties of natural ecosystems that cannot be anticipated based solely on microcosms.

Hologenomic resolution

The complexity of a study system is not only determined by its inherent properties and study design but also the techniques and procedures employed to analyse it. Researchers can decide how much a system is simplified by altering the resolution of the hologenomic features under study[80]; in essence, zooming in or zooming out. The biological resolution of hologenomic studies can be increased by not limiting the study of host and microbial community features to the genomic level but rather by incorporating additional molecular data layers associated with both host genotypes and microbial metagenotypes[81] (TABLE 1). As long as appropriate sample preservation and laboratory processing procedures are adopted, high-throughput DNA sequencing platforms and mass spectrometry instruments can be used to generate multiple layers of biological information that enhance the resolution when trying to ascertain biological interactions between hosts and microbial communities. Multi-omic approaches that incorporate data from hosts and associated microorganisms, defined as holo-omics[18], generate hypercomplex datasets that may require the dimensionality of the data to be reduced to gain power. However, it is crucial to acknowledge that reducing resolution comes with an increased risk of overlooking essential parts of the total variability that

Table 1 | **Examples of techniques used in hologenomics and the level of information they generate**

Data type	Hologenomic domain	Molecular details	Information revealed
(Meta)genomic layer			
Targeted sequencing of markers	Host	Typical markers in animals include the mitochondrial cytochrome oxidase 1 gene (a standardized animal DNA barcode) and the whole mitochondrial genome sequence; typical markers in plants include chloroplast rRNA genes and the entire chloroplast genome sequence	Species identification based on conserved organelle markers and potential inference of phylogenetic lineages within species
	Microbiota	Typical markers include the 16S rRNA ribosomal gene (a standardized bacterial DNA barcode) and the internal transcribed spacer gene (a standardized fungal DNA barcode)	Basic association of population structure with broad microbial community taxonomic patterns; preliminary assignment of inferred function to the microbiota
Partial genome sequencing	Host	Approaches include genotyping-by-sequencing (in which a reduced representation of the genome, often selected via the use of restriction enzymes, is sequenced) and SNPchip (which uses a pre-designed chip with a given number of known SNPs and genotypes are discerned using allele-specific primers)	The reduced representation of the genome-wide variation of the eukaryotic host is sufficient to allow low-resolution GWAS to associate host genotype with variation in microbiota profiles
	Microbiota	Shallow shotgun metagenomics	As for targeted sequencing of markers but with improved functional resolution of microbial genes present in a sample
Whole-genome sequencing	Host	All DNA fragments in a sample are sequenced followed by mapping against a reference genome from the species or group being studied	Screening of all segregating SNPs in the genome of the study host organisms enables the identification of causative genes and regions associated with microbiota profiles
	Microbiota	Deep shotgun metagenomics	Provides greatly improved resolution compared with partial genome sequencing, including recovery of more complete and less redundant MAGs to potentially allow strain-level analyses, identification of causal genetic variants in MWAS and discovery of potential function in rare microbial community members
(Meta)epigenomic layer			
DNA methylation profiling	Host	Whole-genome bisulfite sequencing is used to characterize methylation patterns of single cytosine nucleotides across the genome[158]	DNA methylation patterns can be used to infer gene expression patterns
	Microbiota	Single-molecule real-time sequencing of metagenomic DNA reads is used to generate long DNA reads	DNA methylation patterns can be used to infer gene expression patterns and to improve metagenomic binning[159]
Histone modification profiling	Host	MS is used to characterize the histone components of the epigenome and their combinatorial PTMs; SERS enables the sensitive detection of histone demethylase activity by observation of formaldehyde by-products[160]	Characterization of histone proteins and their PTMs enables the inference of gene expression patterns and of the host genome–transcriptome–microbial metagenome axis
	Microbiota	NA	NA
Genome conformation profiling	Host	Hi-C can be used to reconstruct the three-dimensional folding of chromosomes and to improve scaffolding of contigs into genomes[91]	Enables the association of host genomic variation with microbiota profiles at levels beyond SNP and basic copy number variation, for example, the 3D association of distantly spaced loci
	Microbiota	Hi-C can be used to accurately resolve MAGs[161]	Enables MWAS accounting for interactions between spatially distant host genomic loci; allows analysis of horizontal gene transfer dynamics in different hosts[162]
(Meta)transcriptomic layer			
Targeted RNA sequencing	Host	Real-time PCR can be used to detect and measure gene expression levels of a subset of predefined genes	Generates single gene expression profiles
	Microbiota		Generates single gene expression profiles in which the gene may be present in multiple microbial species with identical gene functions
Non-targeted RNA sequencing	Host	Shotgun RNA sequencing of genome-wide host transcripts	Enables profiling of all transcribed genes in the host genome from a specific tissue sample such as liver or gut epithelium; enables the detection of differentially expressed genes among study groups and associations with, for example, epigenetic profiles and metagenotypes

Table 1 (cont.) | **Examples of techniques used in hologenomics and the level of information they generate**

Data type	Hologenomic domain	Molecular details	Information revealed
(Meta)transcriptomic layer (cont.)			
Non-targeted RNA sequencing (cont.)	Microbiota	Shotgun RNA sequencing of metagenome-wide host transcripts	Enables profiling of all expressed genes from the microbial domain irrespective of metagenomic origin; reveals information about the functional properties of host-associated microbiomes and is directly related to the metabolomics landscape
(Meta)proteomic and metabolomic layers			
Targeted profiling	Host	LC-MS, GC-MS, IC-MS and NMR can be used to identify different substances, such as metabolites or protein molecules, within a test sample; NMR is less sensitive than MS-based methods	Provides absolute quantification of a small number of predefined proteins and metabolites to help discern functional relationships between host and microbiota; proteins can be assigned to coding genes originating from either the host genome or the metagenome; metabolites reveal active pathways that are controlled by host and/or microbiota gene functions in a given environment such as the gut or inside host tissues
	Microbiota		
Non-targeted profiling	Host		Provides a semi-quantitative global overview of proteins and metabolites to help discern functional relationships between host and microbiota as described above for the targeted approaches
	Microbiota		

GC-MS, gas chromatography–mass spectrometry; GWAS, genome-wide association study; IC-MS, ion chromatography–mass spectrometry; LC-MS, liquid chromatography–mass spectrometry; MAGs, metagenome-assembled genomes; MS, mass spectrometry; MWAS, metagenome-wide association study; NA, not applicable; PTMs, post-translational modifications; rRNA, ribosomal RNA; SERS, surface-enhanced Raman scattering.

lie under a certain grouping threshold (for example, strain-level variation of microbiota or interindividual variation in hosts). The use of automatic dimensionality reduction techniques, such as principal components analysis or partial least squares regression, can also complicate the biological interpretation of the results[82].

Host genotypes. In host–microbiota studies, host genotypes can be defined at different levels, including species, breeds, populations[83,84], strains[85], sex or individuals[86]. Genotypes can be defined as categorical variables, without analysing the differences between them, or can be studied in more detail by considering their actual genetic content and establishing correlations among them[86]. When using an evolutionary perspective, phylogenetic relationships between genotypes are established based on phylogenomic markers[87], which usually vary above population and species level but not among individuals. This implies that genomic variability among the individuals included within each genotype is overlooked. Studying the effect of interindividual genomic variability on host–microbiota systems, such as identifying candidate host genomic variants associated with microbial features, requires a higher level of resolution. This is achieved by defining genotypes at the individual level and using techniques based on whole-genome resequencing that enable the complexity of host genomes to be screened at a much finer level such that differences between the individuals contrasted are not only defined based on their kinship but also on the functional properties of their genomic variants. Currently, this approach requires high-quality reference genomes from which high-density SNP profiles of individuals can be generated, for example, through SNPchip or resequencing studies[15]. The genomic resolution could be further refined by incorporating structural variants[88], methylation patterns[89,90] or, even, we hypothesize, chromosome 3D folding structure as revealed through techniques such as Hi-C[91]. In doing so, researchers can identify associations between SNPs or gene variants and specific microbiota traits, such as the relative abundance of certain taxa or the enrichment of a given function, and thus identify mechanisms by which a host exerts control over the composition and function of its associated microbiota[92].

Microbial metagenotypes. The structure and resolution at which microbial metagenotypes are defined also affects the complexity of the metagenome under analysis. Metagenotypes can be defined as arrays of microbial taxa, microbial genes or a combination of both. The most common approach to define them is to rely on short marker sequences targeted for metabarcoding purposes such as the 16S ribosomal RNA (rRNA) or the internal transcribed spacer[93]. Based on sequence similarity, the microbial sequences detected can be clustered into operational taxonomic units that approximate the species-level identity of microorganisms (usually using a 97% similarity threshold) or can be analysed at a finer level as amplicon sequence variants with the aim of approaching strain-level resolution[93]. However, these procedures often do not enable reliable taxonomic assignment at genus or species level[94], do not capture sub-amplicon sequence variant strain-level community dynamics (as shown by a recent non-peer-reviewed preprint[95]) and are prone to generating biased functional inferences as bacteria with identical marker genes (particularly those associated with wild taxa) might carry very different catalogues of genes[96]. Thus, while useful for estimating microbial diversity and obtaining preliminary insights into functionality[97–99], targeted sequencing approaches do not provide conclusive evidence about the metabolic

Metagenome-assembled genomes
(MAGs). Partial or semi-complete draft bacterial genomes reconstructed through metagenomic assembly and binning from samples containing mixtures of microbial taxa.

capabilities of the microbiota, particularly when working with non-human systems[100].

By contrast, if appropriate strategies and adequate sequencing depths are employed, shotgun metagenomics enables bacterial genome sequences to be recovered from which genes can be predicted and annotated to create a gene catalogue that can define a metagenotype[101,102]. However, these genes are not randomly distributed but enclosed within genomes of specific bacteria or other microorganisms, with a particular combination of genes that shape their expression and the specific biological features (such as oxygen affinity, reproduction time, metabolic capacity) that determine their ecology. Hence, a more refined characterization of microbial metagenotypes can be achieved through binning algorithms that enable bacterial genome reconstruction from metagenomic mixtures[16], yielding metagenome-assembled genomes (MAGs). Nevertheless, unless short-read sequencing is combined with long-read approaches, it is challenging to capture multi-copy genes such as the 16S rRNA marker gene[103], which is often employed in metabarcoding studies and therefore represents a useful link to a large number of existing studies. However, machine learning-based solutions to link 16S rRNA marker gene sequences with MAGs are being developed[104]. Finally, regardless of the approach used to define the microbial metagenotype, the complexity of microbial communities will often require a dimensionality reduction to increase statistical power[105,106]. This can be achieved by defining co-abundance clusters, ecological guilds or more complex strategies that also consider temporal features of microbiota variation such as compositional tensor factorization[107].

Envirotypes. The characterization of environmental factors that affect the host–microbiota system under study enables the definition of envirotypes, a term drawn from crop sciences[108] that is useful in accounting for environmental factors in the hologenomic context. Any different physical place or a place sampled at different time points will be exposed to a different environment as conditions will seldom be identical between two spatial and temporal points. Hence, the resolution at which the composite of environmental factors is considered will define whether these two environments will be thought of different envirotypes or not. For example, if only considering water temperature, killer whales sampled in the Arctic and the Antarctic seas experience the same envirotype. However, if the biotic composition is also considered in the definition of the environment, the Arctic and the Antarctic will need to be split into two distinct envirotypes as some killer whales will have access to penguins as a food source while others will not[109]. The same principle applies to laboratory setups or mesocosm experiments: a temperature shift of 2–3 °C might not be considered relevant under some experimental setups, while it can define different envirotypes under other study designs. Finally, the failure to recognize environmental factors that affect host–microbiota interactions and thus define relevant envirotypes can lead to increased noise and a decreased capacity to achieve statistical significance.

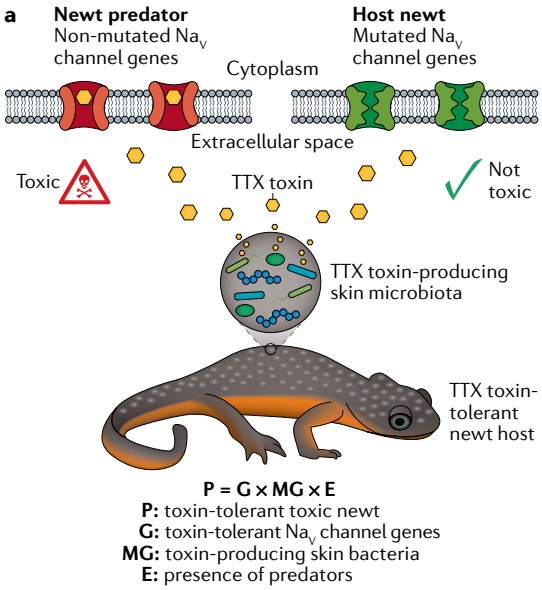
Spatiotemporal resolution

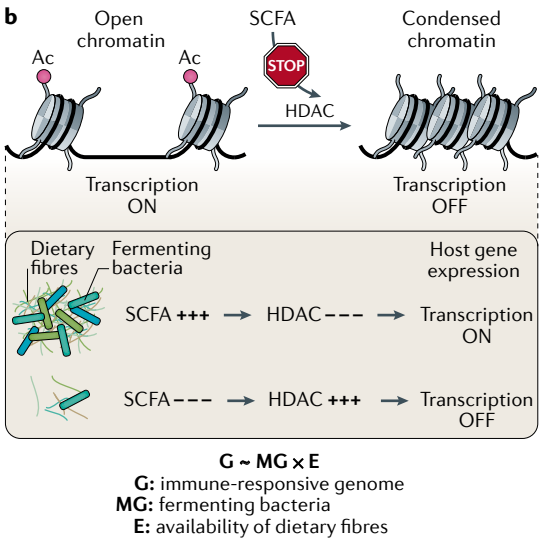
Spatial resolution. Microbial communities associated with animal and plant hosts vary not only across coarse body parts[110] but also at the microscale[111,112] such as between the lumen and the intestinal crypts[113]. Thus, the resolution at which a body site is defined will also determine how a hologenomic system is characterized. For example, the animal gastrointestinal tract can be considered a single sampling unit[114], 4–5 units[65,115] or hundreds of micro-units[111,112] depending on the sampling and data processing strategies employed. Naturally, each level of resolution will allow different questions to be addressed and will require the use of different technologies and analytical approaches.

Temporal resolution. Temporal features to be considered include when, how often and for how long host–microbiota systems are to be analysed. Researchers must consider when a host is first exposed to microbes with regard to temporal benchmarks (number of days or years), as well as the order in which it is exposed to them. Priority effects relate to how the order of species arrivals in an ecosystem shape the potential for subsequently arriving taxa to establish themselves[116]. Although originally discussed at the macroorganismal level in the context of plant communities[117], the phenomenon is also relevant for building host-associated microorganism communities[118], for example, as documented in the human gut[119,120]. In addition, microbial communities are known to vary daily[121], seasonally[122] and relative to life-stage patterns[123]. Hence, the extent and frequency of sampling determine which of these dynamics will be observed or, conversely, missed. Finally, it is important to consider that the consequences of changes at one time period or life stage may appear only later in time and thus detection of such effects obviously requires that the subsequent period is also studied. For example, interventional animal experiments show that, when the immune system develops early in life, there is a window of opportunity where the gut microbiota composition shapes the risk of developing diseases in the future[123,124].

Explanatory and response variables

Host genomic and microbial metagenomic data generated under hologenomic setups can take on different roles when generating statistical models. While the environment is most often considered as an explanatory variable (though one can also study how the hologenome affects the environment), the host genome and the microbial metagenome are sometimes viewed as explanatory and sometimes as response variables, depending on the aim of the research. In many cases, directionality is set by the researcher rather than the biological system itself, as host–microbiota systems contain many bi-directional interactions and circular processes, which complicate the establishment of causal relationships. Here, we define three basic models in which the three main variables (genome, metagenome and environment) are assigned different roles to address different types of fundamental questions (FIG. 5).





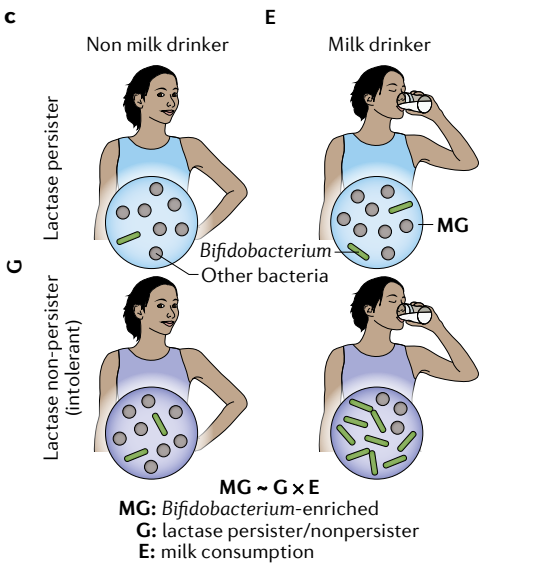
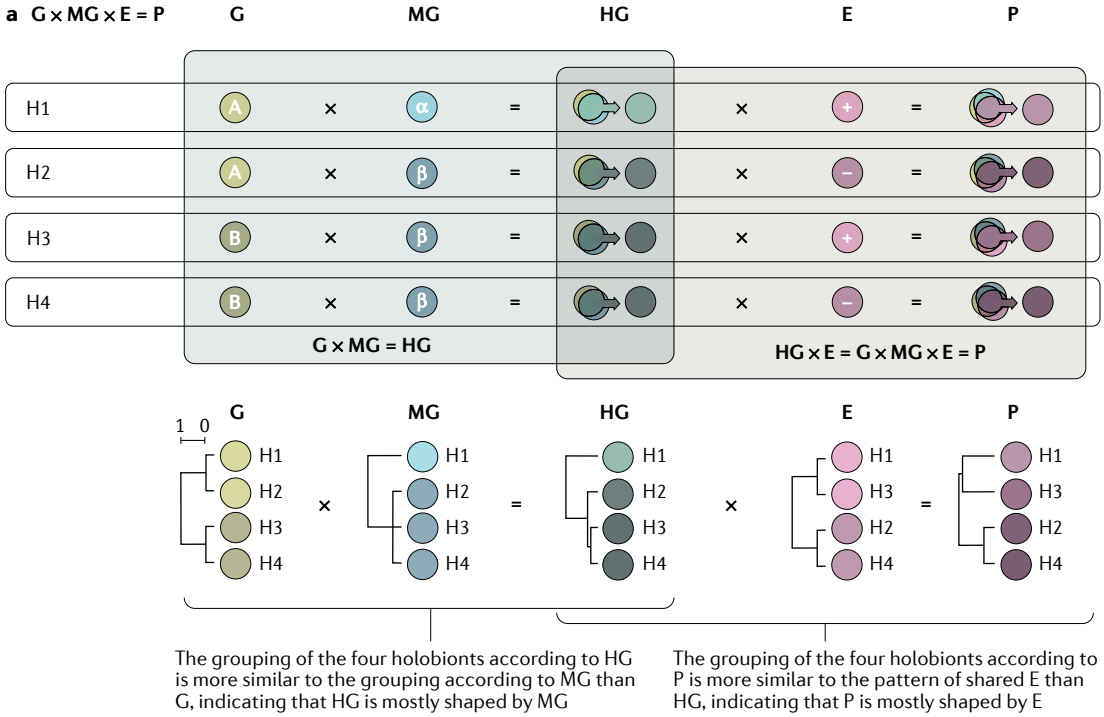
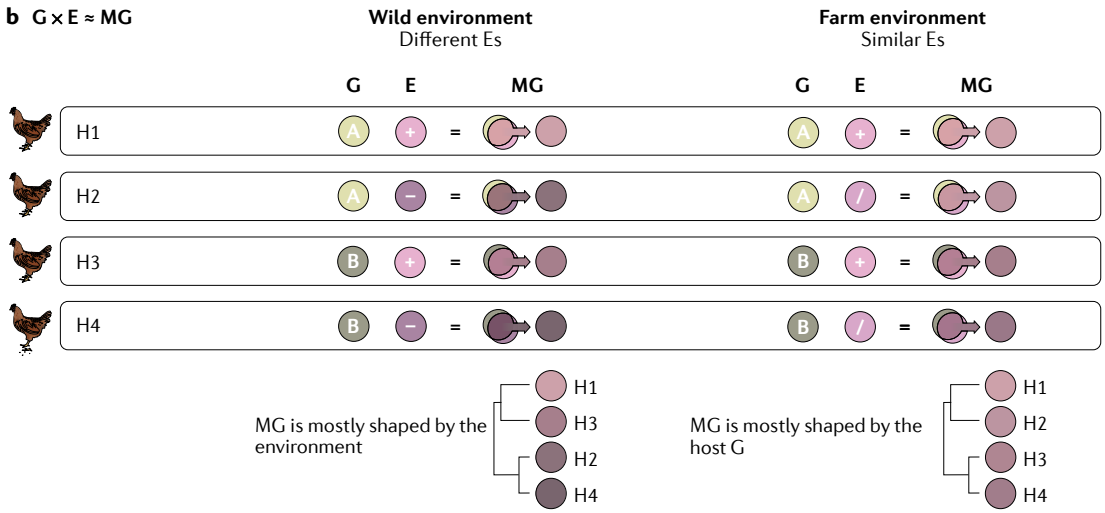


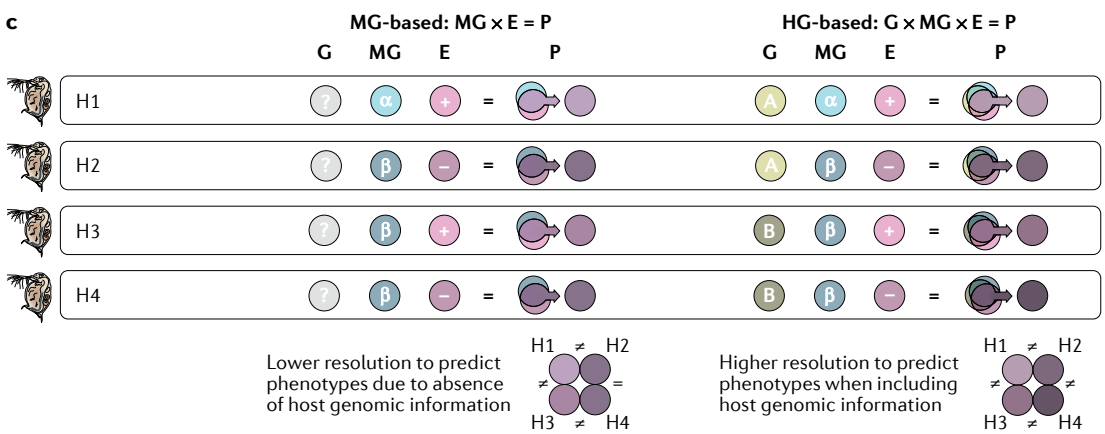
Fig. 5 | **Examples of biological processes addressed by the different models of host–microbiota interactions.** **a** | How does the hologenome shape animal phenotypes? Only the combination of specific host genotypes (G) and microbial metagenotypes (MG), probably resulting from a selective force exerted by the presence of predators (that is, a specific envirotype (E)), enables rough-skinned newts to have skin toxicity[128], an ecologically relevant phenotypic trait. **b** | How do the microbial metagenome and environment shape host genomic features? Short-chain fatty acid (SCFA)-producing bacteria along with a fibre-rich diet enhance chromatin accessibility and thus activate immune gene expression[135]. **c** | How do the host genome and the environment shape microbial genomic features? Only the combination of a lactase non-persister genotype combined with the milk-drinking envirotype generates a microbial metagenotype characterized by the enrichment of *Bifidobacterium*. HDAC, histone deacetylase; Na_V, voltage-gated sodium; P, phenotype.

Phenotype as a product of genotype, metagenotype and envirotype. This is the main model used when hologenomics is conducted to ascertain how genome–metagenome–environment interactions affect the biological properties of a host such as disease susceptibility, performance or fitness. It is an especially common and relevant model for health, agricultural, ecological and evolutionary research[19,125–127]. One clear example of a phenotype shaped by host genomic, microbial metagenomic and environmental factors was recently reported for rough-skinned newts[128]. The study showed that bacteria on the skin of the newts produce a deadly neurotoxin from which the newt is protected by mutations in five host genes that encode the voltage-gated sodium (Na_V) channels normally targeted by the toxin. Thus, this 'toxic newt' phenotype is the result of both host and microbial genes, which likely evolved under the pressure exerted by an environmental factor, namely the presence of predators (FIG. 5a).

Genotype expression influenced by metagenotype and envirotype. When studying how core host genomic features, which contribute to shaping phenotypes, are affected by the microbiota, host genomic features become the response variable. Unlike the microbial metagenome, the genome sequence of the host organism is not variable but microorganisms can induce chromatin remodelling[129] and DNA methylation of the host genome (reviewed in REF.[130]) (FIG. 5b) and thus modulate the bioactivity of molecular receptors[131] and host gene expression[66,132,133]. A well-studied pathway that links the microbiota with host gene expression involves modulation of the activity of host histone deacetylases (HDACs) by short-chain fatty acids (SCFAs) produced by intestinal microorganisms. HDACs remove histone lysine acetyl groups, which leads to chromatin condensation and transcriptional silencing of genes[134]. Increased SCFA concentrations inhibit HDACs, thereby enhancing chromatin accessibility and activating gene expression. Thus, a metagenotype with a higher capacity to produce SCFAs combined with an envirotype required to produce SCFAs (that is, a fibre-rich diet), contributes to boosting the host immune


a G × MG × E = P
G
MG
HG
E
P
H1
H2
H3
H4
G × MG = HG
HG × E = G × MG × E = P
1 0
The grouping of the four holobionts according to HG is more similar to the grouping according to MG than G, indicating that HG is mostly shaped by MG
The grouping of the four holobionts according to P is more similar to the pattern of shared E than HG, indicating that P is mostly shaped by E


b G × E ≈ MG
Wild environment
Different Es
Farm environment
Similar Es
G
E
MG
H1
H2
H3
H4
MG is mostly shaped by the environment
MG is mostly shaped by the host G


c
MG-based: MG × E = P
HG-based: G × MG × E = P
G
MG
E
P
H1
H2
H3
H4
Lower resolution to predict phenotypes due to absence of host genomic information
Higher resolution to predict phenotypes when including host genomic information

◀ Fig. 6 | **Decomposition of hologenomic complexity and its impact on results.** **a** | Variation of different properties that determine hologenomic systems among four example holobionts (H1–H4). The similarity of different properties is reflected by colour tonality; the closer the colour tone, the more similar the considered property. **b** | Representation of the relative impact of host genomic and environmental factors in determining microbial metagenotypes under different scenarios. The relative effect of the environment with respect to that of the host genome in shaping microbial metagenotypes is expected to be larger in the wild, where the environment is highly variable, than in a farm setup, in which environmental conditions are controlled. **c** | Representation of the importance of incorporating host genomic information to increase the resolution to predict phenotypes. The inclusion of host genomic information will provide additional detail that can help explain the phenotypic variability observed. E, envirotype; G, genotype; HG, hologenotype; MG, metagenotype; P, phenotype.

response through the activation of host immune gene expression[135].

Metagenotype as a product of genotype and envirotype. This model assumes the inverse causal directionality between the host genome and microbial metagenome to that described above. Candidate host genes related to microbiota features can be identified through genome-wide association studies (GWAS) in which the metagenotype (or derived metrics, such as diversity or abundance of specific microbial taxa, genes or metabolic functions[92]) are treated as a phenotypic trait. For instance, the increased abundance of the lactose degrader Bifidobacteria in humans has been shown to be associated with the lactase non-persister genotype and the consumption of milk (envirotype)[136]. Once candidate genes are known, targeted analyses in which natural or human-controlled genomic variability (such as the number of copies of the amylase-encoding gene in humans) can be contrasted under controlled environmental conditions to ascertain the effect on metagenotypes (such as the abundance of Ruminococcaceae bacteria in the gut microbiota)[137] (FIG. 5c).

Biology through the hologenomics lens

The concepts, techniques and designs introduced so far should provide the necessary background to identify the study designs and associated methodologies most useful for addressing specific scientific questions as well as for devising interdisciplinary approaches to open up new research avenues.

Outstanding questions

Below, building on the work of others[18,19,32], we outline how hologenomics is already being used to address some of the major questions concerning animal–microbiota interactions. We also provide our view on how hologenomics could be used in both the near and far future.

Does environment outweigh host genetics in determining the microbiota? A recurrent question in host–microbiota research is the relative contribution of host genetics versus the environment in shaping the microbiota. Researchers have attempted to answer this question based on humans[136,138], laboratory models[139] and wild organisms[140], with no overall consensus. We argue that the lack of consensus is because of the often poorly described range of genomic and environmental complexity considered in such studies. As a result, comparisons among studies or even within studies are rarely like-to-like. Take, for example, the interindividual genomic variation of red junglefowls, the wild ancestors and closest living relatives of the domestic chicken *Gallus gallus*[141]. The environmental complexity in the wild habitat in which these animals live may, under normal circumstances, almost completely mask the effect of host genomic variation[142] because the effect size of extrinsic factors is much larger than any intrinsic features of the host or associated microorganisms (wild populations example in FIG. 4). However, the same level of genetic variation may have a very different impact in the case of conventionally shed-raised farm chickens that live under highly controlled and uniform environments[18,19] (production animal example in FIG. 4). Hence, considering the three axes of complexity (host genomic, microbial metagenomic and environmental) defined in this article is essential to understand the relative nature of host–microbiota interactions as identical metrics in one axis of complexity (for example, the host genome) can have different consequences or relevance depending on the complexity of the other axes (for example, the environment) (FIG. 6).

Do host–microbiota interactions shape host fitness? It is known that, in at least some species and in particular contexts, microbial metagenotypes can shape host fitness[125,143]. How the interactions between host genomic and microbial metagenomic features modulate fitness is less clear as are the conditions under which microbial communities are expected to have the biggest effects on host fitness. Hologenomics could be a valuable asset in clarifying these points and to test theory as it develops. For instance, recent studies have shown that the fitness of *Drosophila melanogaster* is influenced by the specific composition of their gut microbial communities[144], that laboratory mice reconstituted with natural microbiota exhibit reduced inflammation and increased survival following infection by influenza virus[145], and that the enrichment or depletion of specific microorganisms can differentially affect the fitness of water fleas as can environmental factors such as temperature[146]. However, none of these studies analysed host genomic variability. Incorporating information on individual host genotypes could not only be helpful to better explain fitness variation (FIG. 6c) but could also cast light on the mechanistic processes through which microorganisms affect host fitness by identifying the host genetic variants associated with microbial metagenomic features. Here, the hologenomic perspective provides insight not just into whether microbes influence phenotype but also under what context they are most likely to influence it.

Do microorganisms shape host evolution? The fact that microorganisms can shape host fitness renders the microbiota a potential modulator of host evolution. An essential step in determining whether (and, perhaps more usefully, when) this is the case is to unveil the degree of interdependence between microbial metagenotypes and host genotypes when shaping core phenotypic features that impact fitness. For example,

major host dietary shifts are often believed to have been facilitated by the functional characteristics of associated microbiota[147]. Host phylogenetic approaches have revealed a general pattern across animal gut microbiotas whereby the acquisition of ancient and large microbial lineages coincides with changes in host diet, whereas more recently diverged bacterial lineages correlate with host phylogeny[87]. More detailed approaches in which functional traits of hosts and microbiota are analysed jointly can provide (and indeed have provided) insights into how microorganisms shape host evolution. For instance, a hologenomic approach revealed that the genomes of vampire bats as well as the metagenome of their gut microbiota encode key traits that together enable them to overcome the nutritional and metabolic challenges of sanguivory[20]. The full exploitation of such functional approaches requires high-quality reference genomes, which are now being generated for a wide range of eukaryotic organisms by a range of international consortia[15]. Hologenomics based on the resequencing of fully annotated host genomes and MAGs would reveal whether hosts have developed adaptations to microorganism-mediated signals (such as differential activity of SCFA receptors[148]) or whether functional traits have been transferred from the host genome to the microbial metagenome and vice versa. The adaptive implications of human genomic and microbial metagenomic variation have been recently reviewed[149].

Does the microbiota provide an adaptive buffer to their hosts? Joint analysis of host genomic and microbial metagenomic features could reveal whether microorganisms enable initial responses to environmental changes, thus buffering host genomic adaptations that may require multiple generations[150]. Such challenges include rapid climate-linked environmental change but also the often radical change of conditions faced when wild species enter the human environment; for example, during domestication or feralization an organism may have to adapt to a very different physical environment and change its diet and behaviour. Hologenomics provides the ideal means to test this hypothesis through observational approaches that compare the hologenotypes of wild and domesticated animals with intermediate-state counterparts that might have undergone metagenomic adaptation before developing any genomic adaptation or through the experimental manipulation of laboratory populations subjected to environmental changes. For instance, a recent study showed that multi-generational sub-toxic exposure of *Nasonia vitripennis* wasps (and their microbiome) to a pesticide resulted in metagenomic adaptation that increased the rate of host genome selection[151].

Do hosts shape microbial evolution? Given their tight interactions, it is possible that the evolution of microorganisms with high heritability (that is, the proportion of the variability in the microbiota across host individuals that is attributable to host genetic effects) is shaped by their hosts[152,153]. In a recent study[154], a hologenomic approach was used to show that polymorphisms in innate and adaptive immune genes affect microbiota composition and that the microbial taxa most affected by mice genomic background (that is, those with the highest heritability) are those that get bound by host IgA such as *Mucispirillum*[155]. These results suggest that some microbes may have adapted to host immune mechanisms to thrive in host environments. While widespread and long-lasting co-evolution between a specific host and most microorganisms seems unlikely[156,157], hologenomic approaches could explore, for example, whether co-diversification and co-phylogeny patterns of IgA-bound and unbound bacteria differ and thus contribute to our understanding on how animals can shape the evolutionary pathways of microorganisms.

Future perspectives

Hologenomic researchers will continue to leverage the novel laboratory technologies developed in diverse fields of science (such as spatial metagenomics or CRISPR–Cas9 approaches) and the ever-increasing availability of annotated genes and genomes. While doing so will improve the resolution at which the biomolecular properties of both host and bacterial communities are characterized, the maximum potential of hologenomics will only be achieved if statistical tools are further developed that enable host genomic and microbial metagenomic features to be fully integrated. Meanwhile, the efficiency of hologenomic research will increase when technology enables an optimal combination of in vivo and in vitro approaches. Although in vitro modelling of the complex intestinal environments of diverse hosts is challenging, the development of organ-on-a-chip organoid models is starting to enable more sophisticated in vitro studies of interactions between microbial communities and eukaryotic tissues.

Besides technological improvement, a major challenge in the use of hologenomics will be thinking clearly about those aspects of the story of life that can be best studied using hologenomics. What are the biggest questions that can now be resolved that could not before? Additionally, what are the predictions for the outcomes of particular studies? We suspect that cross-disciplinary collaborations will be critical to identifying the most important and relevant applied and theoretical research questions to which hologenomics can contribute substantial advances in understanding.

While the field of hologenomics continues to advance, the complexity of most host–microbiota systems and the substantial costs of the methodologies used to characterize them will always require careful assessment of the biological features of the system under study and a critical evaluation of the capacities and limitations of the available techniques. Ultimately, hologenomics will be most useful if it is used in a way that leverages its strengths and makes clear its limitations. Only a balanced combination of innovative and critical thinking will enable the design of fully powered studies to address some of the most challenging and insightful questions about the interconnectedness of life forms on Earth.

Published online: 21 October 2021

1. McFall-Ngai, M. et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl Acad. Sci. USA* **110**, 3229–3236 (2013).
2. Bosch, T. C. G. & Miller, D. J. *The Holobiont Imperative: Perspective from Early Emerging Animals* (Springer-Verlag, 2016).
3. Müller, D. B., Vogel, C., Bai, Y. & Vorholt, J. A. The plant microbiota: systems-level insights and perspectives. *Annu. Rev. Genet.* **50**, 211–234 (2016).
4. Baedke, J., Fábregas-Tejeda, A. & Nieves Delgado, A. The holobiont concept before Margulis. *J. Exp. Zool. B Mol. Dev. Evol.* **334**, 149–155 (2020).
5. Nyholm, S. V. & McFall-Ngai, M. The winnowing: establishing the squid–vibrio symbiosis. *Nat. Rev. Microbiol.* **2**, 632–642 (2004).
6. Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. I. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* **6**, 776–788 (2008).
7. Sasse, J., Martinoia, E. & Northen, T. Feed your friends: do plant exudates shape the root microbiome? *Trends Plant. Sci.* **23**, 25–41 (2018).
8. Małyska, A., Markakis, M. N., Pereira, C. F. & Cornelissen, M. The microbiome: a life science opportunity for our society and our planet. *Trends Biotechnol.* **37**, 1269–1272 (2019).
9. Sessitsch, A. & Mitter, B. 21st century agriculture: integration of plant microbiomes for improved crop production and food security. *Microb. Biotechnol.* **8**, 32 (2015).
10. Leidy, J. *Parasites of the Termites* (Collins, Printer, 1881).
11. Escherich, T. The intestinal bacteria of the neonate and breast-fed infant. *Rev. Infect. Dis.* **10**, 1220–1225 (1988).
12. Beckwith, T. D. & Rose, E. J. Cellulose digestion by organisms from the termite gut. *Proc. Soc. Exp. Biol. Med.* **27**, 4–6 (1929).
13. Margolin, S. Methods for the cultivation of cattle ciliates. *Biol. Bull.* **59**, 301–305 (1930).
14. Bergeim, O., Hanszen, A. & Arnold, L. The influence of fruit ingestion before meals upon the bacterial flora of stomach and large intestine and on food allergins. *Am. J. Dig. Dis. Nutr.* **3**, 45–52 (1936).
15. Lewin, H. A. et al. Earth BioGenome Project: sequencing life for the future of life. *Proc. Natl Acad. Sci. USA* **115**, 4325–4333 (2018).
16. Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* **35**, 833–844 (2017).
17. Jefferson, R. *The Hologenome. Agriculture, Environment and the Developing World: A Future of PCR* (Cold Spring Harbor Press, 1994).
18. Nyholm, L. et al. Holo-omics: integrated host-microbiota multi-omics for basic and applied biological research. *iScience* **23**, 101414 (2020).
This article provides an overview of possible applications of hologenomics across basic and applied biological sciences.
19. Limborg, M. T. et al. Applied hologenomics: feasibility and potential in aquaculture. *Trends Biotechnol.* **36**, 252–264 (2018).
20. Zepeda Mendoza, M. L. et al. Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat. Ecol. Evol.* **2**, 659–668 (2018).
21. Foster, K. R., Schluter, J., Coyte, K. Z. & Rakoff-Nahoum, S. The evolution of the host microbiome as an ecosystem on a leash. *Nature* **548**, 43–51 (2017).
22. Theis, K. R. et al. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**, e00028-16 (2016).
23. Moran, N. A. & Sloan, D. B. The hologenome concept: helpful or hollow? *PLoS Biol.* **13**, e1002311 (2015).
24. Lloyd, E. A. & Wade, M. J. Criteria for holobionts from community genetics. *Biol. Theory* **14**, 151–170 (2019).
25. Srinivas, G. et al. Genome-wide mapping of gene–microbiota interactions in susceptibility to autoimmune skin blistering. *Nat. Commun.* **4**, 2462 (2013).
26. Bashiardes, S., Godneva, A., Elinav, E. & Segal, E. Towards utilization of the human genome and microbiome for personalized nutrition. *Curr. Opin. Biotechnol.* **51**, 57–63 (2018).
27. Bermudez-Brito, M., Plaza-Díaz, J., Fontana, L., Muñoz-Quezada, S. & Gil, A. In vitro cell and tissue models for studying host–microbe interactions: a review. *Br. J. Nutr.* **109**, S27–S34 (2013).
28. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome Biol.* **18**, 83 (2017).
29. Argelaguet, R. et al. Multi-Omics Factor Analysis — a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **14**, e8124 (2018).
30. Singh, A. et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).
31. Almeida, A. et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).
32. Xu, L. et al. Holo-omics for deciphering plant-microbiome interactions. *Microbiome* **9**, 69 (2021).
33. Lynch, M. & Walsh, B. *The Origins of Genome Architecture*. Vol. 98 (Sinauer Associates, 2007).
34. Donoghue, P. C. J. & Purnell, M. A. Genome duplication, extinction and vertebrate evolution. *Trends Ecol. Evol.* **20**, 312–319 (2005).
35. De Bodt, S., Maere, S. & Van de Peer, Y. Genome duplication and the origin of angiosperms. *Trends Ecol. Evol.* **20**, 591–597 (2005).
36. Liedtke, H. C., Gower, D. J., Wilkinson, M. & Gomez-Mestre, I. Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. *Nat. Ecol. Evol.* **2**, 1792–1799 (2018).
37. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **6**, 836–846 (2005).
38. Goodrich, J. K. et al. Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).
39. Turnbaugh, P. J., Bäckhed, F., Fulton, L. & Gordon, J. I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **3**, 213–223 (2008).
40. Peiffer, J. A. et al. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl Acad. Sci. USA* **110**, 6548–6553 (2013).
41. Fietz, K. et al. Mind the gut: genomic insights to population divergence and gut microbial composition of two marine keystone species. *Microbiome* **6**, 82 (2018).
42. Ofek-Lalzar, M. et al. Niche and host-associated functional signatures of the root surface microbiome. *Nat. Commun.* **5**, 4950 (2014).
43. Sumigray, K. D., Terwilliger, M. & Lechler, T. Morphogenesis and compartmentalization of the intestinal crypt. *Dev. Cell* **45**, 183–197.e5 (2018).
44. Pabst, O. & Slack, E. IgA and the intestinal microbiota: the importance of being specific. *Mucosal Immunol.* **13**, 12–21 (2020).
45. Bolnick, D. I. et al. Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat. Commun.* **5**, 4500 (2014).
46. Karp, N. A. et al. Prevalence of sexual dimorphism in mammalian phenotypic traits. *Nat. Commun.* **8**, 15475 (2017).
47. Dahlman, S., Avellaneda-Franco, L. & Barr, J. J. Phages to shape the gut microbiota? *Curr. Opin. Biotechnol.* **68**, 89–95 (2020).
48. Edwards, J. A. et al. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biol.* **16**, e2003862 (2018).
49. Danczak, R. E. et al. Members of the candidate phyla radiation are functionally differentiated by carbon- and nitrogen-cycling capabilities. *Microbiome* **5**, 112 (2017).
50. Siriyappagouder, P. et al. Exposure to yeast shapes the intestinal bacterial community assembly in zebrafish larvae. *Front. Microbiol.* **9**, 1868 (2018).
51. Chabé, M., Lokmer, A. & Ségurel, L. Gut protozoa: friends or foes of the human gut microbiota? *Trends Parasitol.* **33**, 925–934 (2017).
52. Leung, J. M., Graham, A. L. & Knowles, S. C. L. Parasite-microbiota interactions with the vertebrate gut: synthesis through an ecological lens. *Front. Microbiol.* **9**, 843 (2018).
53. Moya, A. & Ferrer, M. Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends Microbiol.* **24**, 402–413 (2016).
54. Stanton-Geddes, J. et al. Thermal reactionomes reveal divergent responses to thermal extremes in warm and cool-climate ant species. *BMC Genomics* **17**, 171 (2016).
55. Megha, S., Basu, U. & Kav, N. N. V. Regulation of low temperature stress in plants by microRNAs. *Plant. Cell Env.* **41**, 1–15 (2018).
56. Yu, T. & Chen, Y. Effects of elevated carbon dioxide on environmental microbes and its mechanisms: a review. *Sci. Total. Environ.* **655**, 865–879 (2019).
57. Mushegian, A. A., Arbore, R., Walser, J.-C. & Ebert, D. Environmental sources of bacteria and genetic variation in behavior influence host-associated microbiota. *Appl. Environ. Microbiol.* **85**, e01547–18 (2019).
58. Vannier, N. et al. A microorganisms' journey between plant generations. *Microbiome* **6**, 79 (2018).
59. Sommer, F. et al. The gut microbiota modulates energy metabolism in the hibernating brown bear ursus arctos. *Cell Rep.* **14**, 1655–1661 (2016).
60. Fung, T. C. et al. Intestinal serotonin and fluoxetine exposure modulate bacterial colonization in the gut. *Nat. Microbiol.* **4**, 2064–2073 (2019).
61. Li, Y. et al. Gut microbiota dependent anti-tumor immunity restricts melanoma growth in Rnf5$^{-/-}$ mice. *Nat. Commun.* **10**, 1492 (2019).
62. Bein, A. et al. Microfluidic organ-on-a-chip models of human intestine. *Cell Mol. Gastroenterol. Hepatol.* **5**, 659–668 (2018).
63. Knowles, S. C. L., Eccles, R. M. & Baltrūnaitė, L. Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol. Lett.* **22**, 826–837 (2019).
64. Brito, I. L. et al. Transmission of human-associated microbiota along family and social networks. *Nat. Microbiol.* **4**, 964–971 (2019).
65. Michalak, L. et al. Microbiota-directed fibre activates both targeted and secondary metabolic shifts in the distal gut. *Nat. Commun.* **11**, 5773 (2020).
66. Muller, P. A. et al. Microbiota modulate sympathetic neurons via a gut–brain circuit. *Nature* **583**, 441–446 (2020).
This paper shows that the microbiota modulates the expression of the neuronal transcription factor cFos through SCFAs in the gut sympathetic ganglia.
67. Khoruts, A. & Sadowsky, M. J. Understanding the mechanisms of faecal microbiota transplantation. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 508–516 (2016).
68. Bojanova, D. P. & Bordenstein, S. R. Fecal transplants: what is being transferred? *PLoS Biol.* **14**, e1002503 (2016).
69. Rasmussen, T. S. et al. Bacteriophage-mediated manipulation of the gut microbiome - promises and presents limitations. *FEMS Microbiol. Rev.* **44**, 507–521 (2020).
70. Vázquez-Castellanos, J. F., Biclot, A., Vrancken, G., Huys, G. R. B. & Raes, J. Design of synthetic microbial consortia for gut microbiota modulation. *Curr. Opin. Pharmacol.* **49**, 52–59 (2019).
71. Kehe, J. et al. Massively parallel screening of synthetic microbial communities. *Proc. Natl Acad. Sci. USA* **116**, 12804–12809 (2019).
72. Kim, H. J., Li, H., Collins, J. J. & Ingber, D. E. Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip. *Proc. Natl Acad. Sci. USA* **113**, E7–E15 (2016).
73. D'hoe, K. et al. Integrated culturing, modeling and transcriptomics uncovers complex interactions and emergent behavior in a three-species synthetic gut community. *Elife* **8**, e37090 (2018).
74. Lawson, C. E. et al. Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* **17**, 725–741 (2019).
75. Mimee, M. et al. An ingestible bacterial-electronic system to monitor gastrointestinal health. *Science* **360**, 915–918 (2018).
76. Riglar, D. T. & Silver, P. A. Engineering bacteria for diagnostic and therapeutic applications. *Nat. Rev. Microbiol.* **16**, 214–225 (2018).
77. McCarty, N. S. & Ledesma-Amaro, R. Synthetic biology tools to engineer microbial communities for biotechnology. *Trends Biotechnol.* **37**, 181–197 (2019).
78. Ruiz-Rodríguez, M., Martín-Vivaldi, M., Martínez-Bueno, M. & Soler, J. J. Gut microbiota of great spotted cuckoo nestlings is a mixture of those of their foster magpie siblings and of cuckoo adults. *Genes* **9**, 381 (2018).
79. Taylor, R. S. & Friesen, V. L. The role of allochrony in speciation. *Mol. Ecol.* **26**, 3330–3342 (2017).
80. Wang, Q. et al. Host and microbiome multi-omics integration: applications and methodologies. *Biophys. Rev.* **11**, 55–65 (2019).
81. Nielsen, R. L. et al. Data integration for prediction of weight loss in randomized controlled dietary trials. *Sci. Rep.* **10**, 20103 (2020).
82. Sankaran, K. & Holmes, S. P. Multitable methods for microbiome data integration. *Front. Genet.* **10**, 627 (2019).
83. Frankel-Bricker, J., Song, M. J., Benner, M. J. & Schaack, S. Variation in the microbiota associated with Daphnia magna across genotypes, populations, and temperature. *Microb. Ecol.* **79**, 731–742 (2020).
84. Dirksen, P. et al. The native microbiome of the nematode Caenorhabditis elegans: gateway to a new host-microbiome model. *BMC Biol.* **14**, 38 (2016).

85. Org, E. et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res.* **25**, 1558–1569 (2015).
86. Suzuki, T. A. et al. Host genetic determinants of the gut microbiota of wild mice. *Mol. Ecol.* **28**, 3197–3207 (2019).
87. Groussin, M. et al. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nat. Commun.* **8**, 14319 (2017).
88. Chiang, C. et al. The impact of structural variation on human gene expression. *Nat. Genet.* **49**, 692–699 (2017).
89. Ansari, I. et al. The microbiota programs DNA methylation to control intestinal homeostasis and inflammation. *Nat. Microbiol.* **5**, 610–619 (2020). **This article demonstrates that intestinal microorganisms induce epigenetic changes in regulatory features of host genes.**
90. Woo, V. & Alenghat, T. Host-microbiota interactions: epigenomic regulation. *Curr. Opin. Immunol.* **44**, 52–60 (2017).
91. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
92. Awany, D. et al. Host and microbiome genome-wide association studies: current state and challenges. *Front. Genet.* **9**, 637 (2019).
93. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
94. Johnson, J. S. et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat. Commun.* **10**, 5029 (2019).
95. Goyal, A., Bittleston, L. S., Leventhal, G. E., Lu, L. & Cordero, O. X. Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *bioRxiv* https://doi.org/10.1101/2021.01.04.425224 (2021).
96. Antony-Babu, S. et al. Multiple Streptomyces species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci. Rep.* **7**, 11089 (2017).
97. Aβhauer, K. P., Wemheuer, B., Daniel, R. & Meinicke, P. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **31**, 2882–2884 (2015).
98. Iwai, S. et al. Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. *PLoS One* **11**, e0166104 (2016).
99. Langille, M. G. I. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
100. Sun, S., Jones, R. B. & Fodor, A. A. Inference-based accuracy of metagenome prediction tools varies across sample types and functional categories. *Microbiome* **8**, 46 (2020).
101. Lesker, T. R. et al. An integrated metagenome catalog reveals new insights into the murine gut microbiome. *Cell Rep.* **30**, 2909–2922.e6 (2020).
102. Nayfach, S. et al. A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* **39**, 499–509 (2021).
103. Overholt, W. A. et al. Inclusion of Oxford Nanopore long reads improves all microbial and viral metagenome-assembled genomes from a complex aquifer system. *Environ. Microbiol.* **22**, 4000–4013 (2020).
104. Jiang, Y., Balaban, M., Zhu, Q. & Mirarab, S. DEPP: deep learning enables extending species trees using single genes. *bioRxiv* https://doi.org/10.1101/2021.01.22.427808 (2021).
105. Zhao, L. et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science* **359**, 1151–1156 (2018).
106. Minot, S. S. & Willis, A. D. Clustering co-abundant genes identifies components of the gut microbiome that are reproducibly associated with colorectal cancer and inflammatory bowel disease. *Microbiome* **7**, 110 (2019).
107. Martino, C. et al. Context-aware dimensionality reduction deconvolutes gut microbial community dynamics. *Nat. Biotechnol.* **39**, 165–168 (2021).
108. Xu, Y. Envirotyping for deciphering environmental impacts on crop plants. *Theor. Appl. Genet.* **129**, 653–673 (2016).
109. Foote, A. D. et al. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nat. Commun.* **7**, 11693 (2016).
110. Rojas, C. A., Holekamp, K. E., Winters, A. D. & Theis, K. R. Body site-specific microbiota reflect sex and age-class among wild spotted hyenas. *FEMS Microbiol. Ecol.* **96**, fiaa007 (2020).
111. Riva, A. et al. A fiber-deprived diet disturbs the fine-scale spatial architecture of the murine colon microbiome. *Nat. Commun.* **10**, 4366 (2019).
112. Sheth, R. U. et al. Spatial metagenomic characterization of microbial biogeography in the gut. *Nat. Biotechnol.* **37**, 877–883 (2019).
113. Zaborin, A. et al. Spatial compartmentalization of the microbiome between the lumen and crypts is lost in the murine cecum following the process of surgery, including overnight fasting and exposure to antibiotics. *mSystems* **5**, e00377 (2020).
114. Ellegaard, K. M. & Engel, P. Genomic diversity landscape of the honey bee gut microbiota. *Nat. Commun.* **10**, 446 (2019).
115. Shah, T. M., Patel, J. G., Gohil, T. P., Blake, D. P. & Joshi, C. G. Host transcriptome and microbiome interaction modulates physiology of full-sibs broilers with divergent feed conversion ratio. *NPJ Biofilms Microb.* **5**, 24 (2019).
116. Fukami, T. Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annu. Rev. Ecol. Evol. Syst.* **46**, 1–23 (2015).
117. Gleason, H. A. Further views on the succession-concept. *Ecology* **8**, 299–326 (1927).
118. Debray, R. et al. Priority effects in microbiome assembly. *Nat. Rev. Microbiol.* https://doi.org/10.1038/s41579-021-00604-w (2021).
119. Sprockett, D., Fukami, T. & Relman, D. A. Role of priority effects in the early-life assembly of the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 197–205 (2018).
120. Martínez, I. et al. Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly. *eLife* **7**, e36521 (2018).
121. Johnson, A. J. et al. Daily sampling reveals personalized diet-microbiome associations in humans. *Cell Host Microbe* **25**, 789–802.e5 (2019).
122. Smits, S. A. et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* **357**, 802–806 (2017).
123. Knutie, S. A., Wilkinson, C. L., Kohl, K. D. & Rohr, J. R. Early-life disruption of amphibian microbiota decreases later-life resistance to parasites. *Nat. Commun.* **8**, 86 (2017).
124. Arnold, I. C. et al. Helicobacter pylori infection prevents allergic asthma in mouse models through the induction of regulatory T cells. *J. Clin. Invest.* **121**, 3088–3093 (2011).
125. Lynch, J. B. & Hsiao, E. Y. Microbiomes as sources of emergent host phenotypes. *Science* **365**, 1405–1409 (2019).
126. Mayoral-Peña, Z., Álvarez-Martínez, R., Fornoni, J. & Garrido, E. In *Evolutionary Ecology of Plant-Herbivore Interaction* (eds Núñez-Farfán, J. & Valverde, P. L.) 135–146 (Springer International Publishing, 2020).
127. Lindsay, E. C., Metcalfe, N. B. & Llewellyn, M. S. The potential role of the gut microbiota in shaping host energetics and metabolic rate. *J. Anim. Ecol.* **89**, 2415–2426 (2020).
128. Vaelli, P. M. et al. The skin microbiome facilitates adaptive tetrodotoxin production in poisonous newts. *eLife* **9**, e53898 (2020). **This work demonstrates that the toxic newt phenotype is shaped through the combination of toxin-producing skin microorganisms and toxin-tolerant host genotypes.**
129. Thaiss, C. A. et al. Microbiota diurnal rhythmicity programs host transcriptome oscillations. *Cell* **167**, 1495–1510.e12 (2016).
130. Nichols, R. G. & Davenport, E. R. The relationship between the gut microbiome and host gene expression: a review. *Hum. Genet.* **140**, 747–760 (2021).
131. Chen, H. et al. A forward chemical genetic screen reveals gut microbiota metabolites that modulate host physiology. *Cell* **177**, 1217–1231.e18 (2019).
132. Yano, J. M. et al. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264–276 (2015).
133. Meisel, J. S. et al. Commensal microbiota modulate gene expression in the skin. *Microbiome* **6**, 20 (2018).
134. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016).
135. Sanchez, H. N. et al. B cell-intrinsic epigenetic modulation of antibody responses by dietary fiber-derived short-chain fatty acids. *Nat. Commun.* **11**, 60 (2020). **This study shows that microbiota-produced SCFAs impair intestinal and systemic antibody responses upregulating micro RNAs that target immune genes through inhibition of histone deacetylation.**
136. Bonder, M. J. et al. The effect of host genetics on the gut microbiome. *Nat. Genet.* **48**, 1407–1412 (2016).
137. Poole, A. C. et al. Human Salivary amylase gene copy number impacts oral and gut microbiomes. *Cell Host Microbe* **25**, 553–564.e7 (2019).
138. Rothschild, D. et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).
139. Carmody, R. N. et al. Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* **17**, 72–84 (2015).
140. Wagner, M. R. et al. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat. Commun.* **7**, 12151 (2016).
141. Qanbari, S. et al. Genetics of adaptation in modern chicken. *PLoS Genet.* **15**, e1007989 (2019).
142. Weissbrod, O., Rothschild, D., Barkan, E. & Segal, E. Host genetics and microbiome associations through the lens of genome wide association studies. *Curr. Opin. Microbiol.* **44**, 9–19 (2018).
143. Suzuki, T. A. Links between natural variation in the microbiome and host fitness in wild mammals. *Integr. Comp. Biol.* **57**, 756–769 (2017).
144. Gould, A. L. et al. Microbiome interactions shape host fitness. *Proc. Natl Acad. Sci. U. S. A.* **115**, E11951–E11960 (2018).
145. Rosshart, S. P. et al. Wild mouse gut microbiota promotes host fitness and improves disease resistance. *Cell* **171**, 1015–1021.e13 (2017).
146. Cooper, R. O., Vavra, J. M. & Cressler, C. E. Targeted manipulation of abundant and rare taxa in the daphnia magna microbiota with antibiotics impacts host fitness differentially. *mSystems* **6**, e00916-20 (2021).
147. Moeller, A. H. & Sanders, J. G. Roles of the gut microbiota in the adaptive evolution of mammalian species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190597 (2020).
148. Fontaine, S. S. & Kohl, K. D. Optimal integration between host physiology and functions of the gut microbiome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190594 (2020).
149. Suzuki, T. A. & Ley, R. E. The role of the microbiota in human genetic adaptation. *Science* **370**, aaz6827 (2020).
150. Alberdi, A., Aizpurua, O., Bohmann, K., Zepeda-Mendoza, M. L. & Gilbert, M. T. P. Do vertebrate gut metagenomes confer rapid ecological adaptation? *Trends Ecol. Evol.* **31**, 689–699 (2016).
151. Wang, G.-H. et al. Changes in microbiome confer multigenerational host resistance after sub-toxic pesticide exposure. *Cell Host Microbe* **27**, 213–224.e7 (2020). **This paper shows that the gut microbiota confers *Nasonia* wasps with resistance towards pesticides and that this acquired trait shapes their genomic evolutionary features across multiple generations.**
152. van Opstal, E. J. & Bordenstein, S. R. Rethinking heritability of the microbiome. *Science* **349**, 1172–1173 (2015).
153. Douglas, G. M., Bielawski, J. P. & Langille, M. G. I. Re-evaluating the relationship between missing heritability and the microbiome. *Microbiome* **8**, 87 (2020).
154. Khan, A. A. et al. Polymorphic immune mechanisms regulate commensal repertoire. *Cell Rep.* **29**, 541–550.e4 (2019). **This study shows that genetic variants of innate and adaptive immunity genes shape distinct microbial communities through the production of distinct immunoglobulins some microorganisms use for colonizing the mucosal barrier.**
155. Davenport, E. R. Genetic variation shapes murine gut microbiota via immunity. *Trends Immunol.* **41**, 1–3 (2020).
156. Van Vliet, S. & Doebeli, M. The role of multilevel selection in host microbiome evolution. *Proc. Natl Acad. Sci. USA* **116**, 20591–20597 (2019).
157. Groussin, M., Mazel, F. & Alm, E. J. Co-evolution and co-speciation of host-gut bacteria systems. *Cell Host Microbe* **28**, 12–22 (2020).
158. Meissner, A. et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
159. Beaulaurier, J. et al. Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat. Biotechnol.* **36**, 61–69 (2018).

160. Ma, F., Jiang, S. & Zhang, C.-Y. Recent advances in histone modification and histone modifying enzyme assays. *Expert Rev. Mol. Diagn.* **19**, 27–36 (2019).
161. DeMaere, M. Z. & Darling, A. E. bin3C: exploiting Hi-C sequencing data to accurately resolve metagenome-assembled genomes. *Genome Biol.* **20**, 46 (2019).
162. Kent, A. G., Vill, A. C., Shi, Q., Satlin, M. J. & Brito, I. L. Widespread transfer of mobile antibiotic resistance genes within individual gut microbiomes revealed through bacterial Hi-C. *Nat. Commun.* **11**, 4379 (2020).
163. Hammer, T. J., Janzen, D. H., Hallwachs, W., Jaffe, S. P. & Fierer, N. Caterpillars lack a resident gut microbiome. *Proc. Natl Acad. Sci. USA* **114**, 9641–9646 (2017).
164. Oliver, K. M., Moran, N. A. & Hunter, M. S. Costs and benefits of a superinfection of facultative symbionts in aphids. *Proc. Biol. Sci.* **273**, 1273–1280 (2006).
165. Song, S. J. et al. Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *MBio* **11**, e02901-19 (2020).
166. Krediet, C. J., Ritchie, K. B., Paul, V. J. & Teplitski, M. Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases. *Proc. Biol. Sci.* **280**, 20122328 (2013).
167. Brune, A. Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* **12**, 168–180 (2014).
168. McCann, J. C., Wickersham, T. A. & Loor, J. J. High-throughput methods redefine the rumen microbiome and its relationship with nutrition and metabolism. *Bioinform. Biol. Insights* **8**, 109–125 (2014).
169. Rosshart, S. P. et al. Laboratory mice born to wild mice have natural microbiota and model human immune responses. *Science* **365**, aaw4361 (2019).
170. Niehus, R., Mitri, S., Fletcher, A. G. & Foster, K. R. Migration and horizontal gene transfer divide microbial genomes into multiple niches. *Nat. Commun.* **6**, 8924 (2015).
171. Carrier, T. J. & Reitzel, A. M. The hologenome across environments and the implications of a host-associated microbial repertoire. *Front. Microbiol.* **8**, 802 (2017).
172. Mushegian, A. A. & Ebert, D. Rethinking 'mutualism' in diverse host-symbiont communities. *Bioessays* **38**, 100–108 (2016).
173. Pereira, F. C. & Berry, D. Microbial nutrient niches in the gut. *Environ. Microbiol.* **19**, 1366–1378 (2017).
174. Erin Chen, Y., Fischbach, M. A. & Belkaid, Y. Skin microbiota–host interactions. *Nature* **553**, 427–436 (2018).
175. Earley, A. M., Graves, C. L. & Shiau, C. E. Critical role for a subset of intestinal macrophages in shaping gut microbiota in adult Zebrafish. *Cell Rep.* **25**, 424–436 (2018).
176. Kartzinel, T. R., Hsing, J. C. & Musili, P. M. Covariation of diet and gut microbiome in African megafauna. *Proc. Natl Acad. Sci. USA* **116**, 23588–23593 (2019).
177. Perofsky, A. C., Lewis, R. J., Abondano, L. A., Di Fiore, A. & Meyers, L. A. Hierarchical social networks shape gut microbial composition in wild Verreaux's sifaka. *Proc. Biol. Sci.* **284**, 20172274 (2017).
178. Asnicar, F. et al. Studying vertical microbiome transmission from mothers to infants by strain-level metagenomic profiling. *mSystems* **2**, e00164-16 (2017).
179. Storelli, G. et al. Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. *Cell Metab.* **14**, 403–414 (2011).
180. Johnson, K. V.-A. & Foster, K. R. Why does the microbiome affect behaviour? *Nat. Rev. Microbiol.* **16**, 647–655 (2018).
181. Davidson, G. L., Raulo, A. & Knowles, S. C. L. Identifying microbiome-mediated behaviour in wild vertebrates. *Trends Ecol. Evol.* **35**, 972–980 (2020).
182. Zhang, Y.-J. et al. Impacts of gut bacteria on human health and diseases. *Int. J. Mol. Sci.* **16**, 7493–7519 (2015).
183. Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**, 289–306 (2012).
184. Nicholson, J. K. et al. Host-gut microbiota metabolic interactions. *Science* **336**, 1262–1267 (2012).
185. Fellows, R. et al. Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases. *Nat. Commun.* **9**, 105 (2018).
186. Neuman, H., Debelius, J. W., Knight, R. & Koren, O. Microbial endocrinology: the interplay between the microbiota and the endocrine system. *FEMS Microbiol. Rev.* **39**, 509–521 (2015).

Acknowledgements
The authors thank the Danish National Research Foundation award DNRF143 'A Center for Evolutionary Hologenomics' for funding their research. A.A. acknowledges Lundbeckfonden grant R250-2017-1351. A.A., M.T.L. and M.T.P.G. were supported by the European Union (H2020-SFS-2018-1 project HoloFood-817729). M.T.L. and M.T.P.G. acknowledge the FHF (Norwegian Seafood Research Fund; "HoloFish", grant No. 901436). S.B.A. acknowledges Lundbeckfonden Fellowship R335-2019-1513 'Understanding the Health Effects of Microbial Interactions' and Independent Research Council Denmark Sapere Aude grant 9064-00029B 'Understanding the Effects of Microbial Interactions on Host Health'.

Author contributions
A.A. researched data for the article. All authors made substantial contributions to discussions of the content and writing the article and reviewed and/or edited the manuscript before submission.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Genetics thanks S. Borderstein, T. Bosch and M. van Oppen for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.