An assessment of US microbiome research

Elizabeth Stulberg[1]*, Deborah Fravel[2], Lita M. Proctor[3], David M. Murray[4], Jonathan LoTempio[3], Linda Chrisey[5], Jay Garland[6], Kelly Goodwin[7,8], Joseph Graber[9], M. Camille Harris[10], Scott Jackson[11], Michael Mishkind[12], D. Marshall Porterfield[13] and Angela Records[14]

Genome-enabled technologies have supported a dramatic increase in our ability to study microbial communities in environments and hosts. Taking stock of previously funded microbiome research can help to identify common themes, under-represented areas and research priorities to consider moving forward. To assess the status of US microbiome research, a team of government scientists conducted an analysis of federally funded microbiome research. Microbiomes were defined as host-, ecosystem- or habitat-associated communities of microorganisms, and microbiome research was defined as those studies that emphasize community-level analyses using 'omics technologies. Single pathogen, single strain and culture-based studies were not included, except symbiosis studies that served as models for more complex communities. Fourteen governmental organizations participated in the data call. The analysis examined three broad research themes, eight environments and eight microbial categories. Human microbiome research was larger than any other environment studied, and the basic biology research theme accounted for half of the total research activities. Computational biology and bioinformatics, reference databases and biorepositories, standardized protocols and high-throughput tools were commonly identified needs. Longitudinal and functional studies and interdisciplinary research were also identified as needs. This study has implications for the funding of future microbiome research, not only in the United States but beyond.

For decades, the role of microbial communities has been studied in nutrient cycling, food webs, pollution degradation and in plant, animal, human and ecosystem health. Historically, these studies were hampered by the limited ability to grow many microbial members of these communities in the laboratory. The advent of DNA sequencing and other genome-enabled technologies has ushered in a new era for the study of microbial communities in environments and hosts. The use of molecular tools in microbiome research has also been fostered by targeted programmes to advance microbiome understanding.

Medical researchers built on early metagenomics studies of marine microbial communities[1,2] to explore the human ecosystem[3]. These early studies of the human microbiome stimulated large initiatives on human microbiome research around the world. In addition to a focus on human health, a variety of programme efforts have supported improved understanding of the role of microbial communities in the environment. Some past and current programmes in host-associated and environmental microbiome research are summarized in Tables 1 and 2, respectively.

Reflecting the growing recognition of the importance of microbiome research, a National Science and Technology Council Committee of US government scientists was chartered in 2015 to analyse federally supported microbiome research over fiscal years 2012–2014. To promote federal Agency coordination, this assessment involved 14 federal organizations within six departments, four independent agencies and one quasi-governmental entity and covered current (fiscal years 2015–2016) and near-term (5–10 year horizon) gaps, needs and challenges for the microbiome field (see Methods for a description of the assessment approach and agencies involved). A 'microbiome' was defined as a multi-species community of microorganisms in a specific environment (that is, host, habitat or ecosystem) and 'microbiome research' as those studies that emphasize community-level analyses with data derived from genome-enabled technologies. Single species, single pathogen and culture-based studies were not included unless these involved symbiosis studies. The snapshot of federally funded research was apportioned into three broad research themes, eight 'environments' (with multiple 'habitats' identified within each) and eight microbial 'categories', with results delivered in a report to the Life Sciences Subcommittee of the National Science and Technology Council Committee on Science as stipulated by the Committee charter[4]. A general summary of the analysis is reported here.

[1]Office of Science and Technology Policy, Executive Office of the President, Eisenhower Executive Office Building, 1650 Pennsylvania Ave. NW, Washington, DC 20504, USA. [2]United States Department of Agriculture, Plant Health, Crop Production and Protection, Agricultural Research Service, 5601 Sunnyside Ave., Beltsville, Maryland 20705, USA. [3]Department of Health and Human Services, National Human Genome Research Institute, National Institutes of Health, 5635 Fishers Lane, Bethesda, Maryland 20892, USA. [4]Office of Disease Prevention, National Institutes of Health, Department of Health and Human Services, 6100 Executive Boulevard, Rockville, Maryland 20852, USA. [5]Department of Defense, Office of Naval Research, 875 N. Randolph Street, Arlington, Virginia 22205-1995, USA. [6]United States Environmental Protection Agency, Office of Research & Development, 26 West Martin Luther King Drive, Cincinnati, Ohio 45268, USA. [7]Ocean Chemistry and Ecosystems Division, Atlantic Oceanographic & Meteorological Laboratory, National Oceanic and Atmospheric Administration, 4301 Rickenbacker Causeway, Miami, Florida 33149, USA. [8]Stationed at NOAA, Southwest Fisheries Science Center, La Jolla Shores Drive, La Jolla, California 92037, USA. [9]US Department of Energy, Office of Biological & Environmental Research, 1000 Independence Ave. SW, Washington, DC 20585, USA. [10]Department of the Interior, Ecosystems Mission Area, United States Geological Survey, 12201 Sunrise Valley Drive, Reston, Virginia 20192, USA. [11]Biosystems and Biomaterials Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, Maryland 20899, USA. [12]Physiological and Structural Systems Cluster, Integrative Organismal Systems, National Science Foundation, 4201 Wilson Boulevard, Arlington, Virginia 22230, USA. [13]NASA Life and Physical Sciences, Human Exploration and Operations Mission Directorate, National Aeronautics and Space Administration, 300 E Street SW, Washington, DC 20546, USA. [14]Office of Agriculture, Research and Policy, Bureau for Food Security, US Agency for International Development, 1300 Pennsylvania Ave. NW, Washington, DC 20523, USA. *e-mail: estulberg@ostp.eop.gov

Table 1 | Examples of large human microbiome research initiatives.

Programme	Programme period	Programme focus	Countries involved
National Institutes of Health (NIH) 'Human Microbiome Project' (HMP), phase one[23]	2007–2012	Develop community resources to support the field: computational tools for sequence-based analyses; experimental methodologies; clinical protocols; microbial genome, whole metagenome and 16S metagenome reference databases; reference microbial culture collection; as well as ethical, legal, social implications of microbiome research. Included research on the microbiomes of five major body regions (skin, oral, nares, gastrointestinal (GI) tract, urogenital) and some associated diseases/disorders.	USA
European Commission (EC) 'Metagenomics of the Human Intestinal Tract' (MetaHIT) programme[24]	2008–2012	Research on the GI tract microbiome.	China, Denmark, France, Germany, Italy, The Netherlands, Spain, UK
Canadian Microbiome Initiative[25]	2008–2012	Research on a variety of body region microbiomes.	Canada
Irish government 'Metagenomics of the Elderly' (ELDERMET) programme[26]	2008–2013	Research on the elderly GI tract microbiome.	Ireland
Six additional human microbiome research initiatives[27]	2008–2012	Diversity of microbiome studies.	Australia, China, France, Japan, Singapore, South Korea
International Human Microbiome Consortium (IHMC)[28]	2007 to present	Establish standards for the field and coordinate international human microbiome research activities.	Australia, Canada, China, France, Gambia, Germany, Kazakhstan, Ireland, Japan, South Korea, Spain, USA
NIH HMP, phase two[29]	2013 to present	Develop community resources to support the field: computational tools for omics-based functional data analysis; experimental methodologies; clinical protocols; methods to analyse longitudinal functional data from microbiome and host; also bioinformatics approaches to organization of multiple datatype databases. Research on the nares, GI tract and vaginal microbiomes.	USA
EC 'My New Gut' programme[30]	2013–2018	Research on the microbiome's role in energy balance and brain development/function and application to diet-related diseases and behaviour.	Austria, Australia, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, The Netherlands, New Zealand, Serbia, Spain, UK, USA
Irish government Alimentary Pharmacobiotic Center (APC) Microbiome Institute[31]	2013–2019	Industry–academic partnership to support microbiome research and to develop microbiome-based products.	Ireland
French government 'Metagenopolis' programme[32]	2013–2017	Industry–academic 'Demonstration Project' to develop microbiome-based therapeutic products for GI diseases.	France
Canadian government 'Environment, Gene and Chronic Disease' programme[33]	2016 onwards	Diversity of microbiome research.	Canada; open to international collaborations
Genome Canada 'Natural Resources and the Environment: Sector Challenges – Genomic Solutions' programme[34]	2016 onwards	Diversity of microbiome research.	Canada; open to international collaborations
EC 'Joint Action Intestinal Microbiomics' programme[35]	2016 onwards	Multidisciplinary trans-national research on the gut microbiome.	Canada; open to international collaborations

Assessment results

Environments. The sum of funding support provided by all participating organizations for three fiscal years (FY12–14) was approximately US$920 million. Organizational representatives also reported microbiome research investments in a variety of defined 'environments' and 'habitats' within environments (Fig. 1). Although many environments were included in this data call, approximately two-thirds of the research support for this period was for studies of human-associated microbiomes and animal model-based studies of microbiomes (Fig. 1a).

Microbiome research themes. It was admittedly difficult to establish hard and fast boundaries for the three research themes of Tools/Resource Development, Basic Biology and Applied Studies, as one agency's basic biological studies may be another agency's applied studies of the microbiome. However, we extensively discussed and agreed upon broad definitions, developed examples for each of these three research themes, and made good-faith efforts to apply these definitions in order to gauge activities across these themes.

Across the 14 organizations, approximately half of the microbiome research activities involved studies of the basic biology of microbiomes, such as community structure and function, the role of the microbiome in host, habitat or ecosystem health and the relationship between microbiome properties and the properties of the host or surrounding microbial communities (Fig. 2a). Applied studies of the microbiome, examples of which include modulation or

Table 2 | Examples of large environmental microbiome research initiatives.

Programme	Programme period	Programme focus	Countries involved
National Science Foundation (NSF) 'Microbial Observatories (MO) and Microbial Interactions and Processes (MIP)' programme[36]	1999–2008	Support activities to discover and characterize novel microorganisms, microbial consortia, communities, activities and other novel properties, and to study their roles in diverse environments. (Joint with USDA)	USA
NSF 'Microbial Genome Sequencing Program'[37]	2001–2009	Support sequencing of environmentally or agronomically important microbes or whole microbial communities. (Joint with USDA)	USA
Alfred P. Sloan & W. M. Keck Foundations, NSF, National Aeronautics and Space Administration (NASA) 'International Census of Marine Microbes'[38]	2004 to present	Promote an agenda and an environment that will accelerate discovery, understanding, and awareness of the global significance of marine microbes.	International, 200 scientists
Gordon and Betty Moore Foundation (GBMF) 'Marine Microbiology Initiative'[39]	2004–2014	Support marine microbial ecology research.	USA
International Soil Metagenome Sequencing Consortium 'TerraGenome'[40]	2008 to present	Coordinate analysis of soil microbiomes.	International, 300 scientists
Department of Energy (DOE) 'Microbial Carbon Cycle' programme[41]	2010 to present	Examine the role of microbiomes in global carbon cycle processes, with particular emphasis on understanding potential impacts of climate change in forest, grassland, and permafrost ecosystems.	USA
John Templeton & W.M. Keck Foundations Earth Microbiome Project[42]	2011 to present	Foundation and industry partnership to construct a microbial map of planet Earth.	International
EC Marine Microbial Biodiversity, Bioinformatics, and Biotechnology (Micro B3) Program[43]	2012 to present	Create international consortium to catalogue the global marine microbiome.	International, 29 countries
GMBF 'Marine Microbiology Initiative' (J. Kaye, Gordon and Betty Moore Foundation, personal communication, 16 July 2015)	2015–2020	Support marine microbial ecology research.	USA

intervention studies to treat or prevent disease, the effect of stressors on the microbiome, and the role of the microbiome in emerging disease, zoonotic infections and antibiotic resistance comprised an additional 28% of these activities (Fig. 2a). A smaller percentage (21%) of studies were in the development of tools, technologies, methods, resources and practices that support microbiome research (Fig. 2a). These proportions remained relatively stable over FY12–14 (Fig. 2a).

Microbial categories. When comparing the distribution of research activities across the eight microbial categories in this analysis, the vast majority of the microbiome research for FY12–14 was in community-level studies of the microbiome, which include metagenomic analyses and other studies that include all members of a microbial community (Fig. 2b). This outcome was also an independent verification that the data call had captured the studies intended. A smaller fraction (23%) focused only on the bacterial members of the microbiome. The remainder focused on studies of Archaea, microeukaryotes, mobile genetic elements, phytoplankton, viruses and other microbial components of the microbiome (Fig. 2b).

Needs and challenges identified by the assessment
The representatives of the participating federal agencies and departments were asked to provide their opinions regarding targeted questions, such as the following. 'What microbiome research area is your program emphasizing currently (in FY15–16)?' 'Where do you see your program going in five years?' 'Over the next 10 years, what crucial type of scientific and technical training will be needed?' These agencies also provided the 'one scientific or technological advance that would enable microbiome research to leap forward quickly' and supplied a 'listing of priority needs in technology,

tools, and infrastructure'. Based on the answers to these questions, we grouped the most common gaps, needs and challenges noted. Most of the needs were in the areas of infrastructure and human resources in computational biology, data science and bioinformatics. Standardization, particularly for sample and data collection protocols, was also cited frequently as a high-priority need, along with high-throughput analytical methods, computer hardware and computational tools for analysing complex data, tools development and vouchered microbial specimens in repositories. These issues are summarized below.

Potential gaps in microbiome research. Some areas that appeared to be less developed than expected and that may benefit from greater emphasis were noted. For example, there were fewer activities in tool and resource development than might be expected, considering that all the participating organizations noted a priority need for new tools, technologies and databases as foundational resources for the field.

We also found that there were fewer studies than might be expected in plant and soil microbiomes in both natural and agricultural environments (8% of total research activities). This was notable considering the importance of these microbiomes in ecosystem services[5,6] and in light of their role in mitigating climate change. Food production-related research, both field and laboratory, and in plants, animals or soil, also appeared less well developed (4% of total research activities).

Another area of limited research support was in symbiosis studies, whether in plants, animals or insects. Symbioses serve as important model systems for more complex host-associated microbiomes and, although this research area received some support from several agencies, this area would also benefit from more attention.
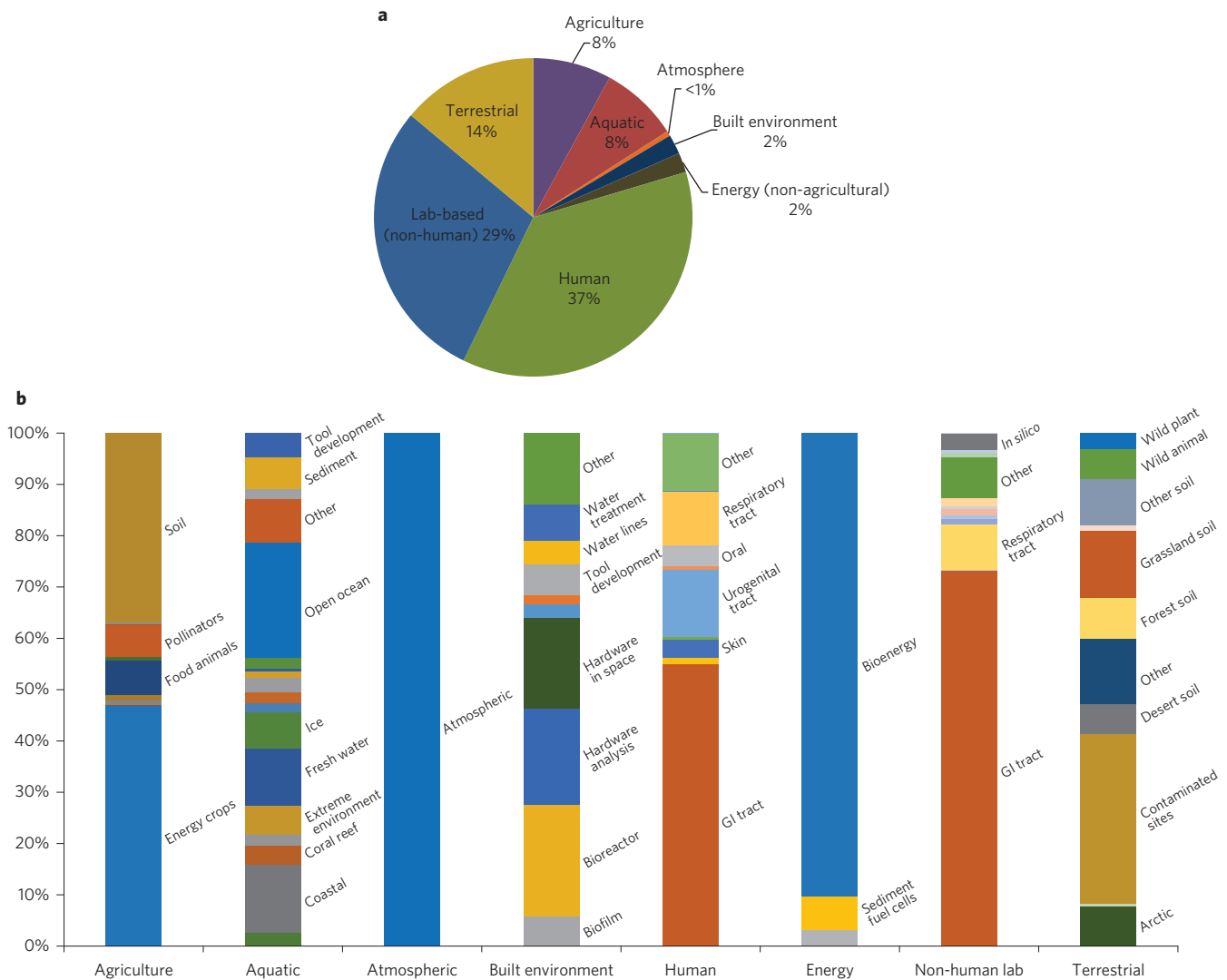
Figure 1 | Microbiome research by ecosystem. a,b, Distribution of microbiome research activities apportioned to categories of environment (**a**) or sub-environment, habitat or sub-habitat for each environment (**b**). Only sub-environments, habitats or sub-habitats representing more than 3% of each environment total are labelled (see Methods).

Finally, another major gap identified was that there were relatively few studies reported for the non-bacterial components of the microbiome, particularly viruses. Researchers are beginning to demonstrate the pivotal role of the virome in human health[7], the oceans[8], plants[9], wildlife[10], soils[11] and in the origins of emerging infectious diseases[10]. Given such findings, attention on host- and ecosystem-associated viromes would seem appropriate. Yet, studies of viral communities made up only 3% of the microbiome research activities, although some of the whole community studies may or may not have included viruses.

Need for methods, references and standards. The participating agencies cited a priority need for higher-throughput, more accurate and less expensive methods for data acquisition and data analysis pipelines, particularly for metagenomic and other 'omic data. Representatives from most agencies specifically mentioned the need for less expensive, long-read sequencing and automated software for the assembly of genetic, genomic and metagenomic sequence. A Smithsonian respondent nicely summarized that access to such sequencing capabilities 'would eliminate the taxonomic biases associated with PCR and other capture methods' and would enable a more accurate assessment of microbial

community composition. A National Oceanic and Atmospheric Administration (NOAA) scientist underscored this need by explaining that 'resolving differences in method biases for sequencing would permit honest comparisons between studies and for longitudinal studies'.

Importantly, DNA-based data from microbiomes were not the only data deemed necessary. The respondents specifically identified the need for widely available tools that measure multiple functional properties of a microbiome, including the ability to measure the transcriptomes, proteomes, lipidomes and metabolomes of microbiomes in a spatial and temporal context. Department of Energy (DOE) programme managers stressed the need for developing 'high-resolution analytical technologies that would allow quantitative measurements of microbial community activities at the nano-, micro-, and mesoscale'. When paired with multi-omics data sets, these tools would provide new approaches for dynamic analysis of microbiome functional attributes and the scaling of this information to the level of the host, ecosystem or environment. National Aeronautic and Space Administration (NASA) respondents described a desire for such a tool as a 'robust, mobile, high-throughput diagnostic instrument', which could be used to identify the basis of 'dysbioses, diseases, and environmental fouling'.
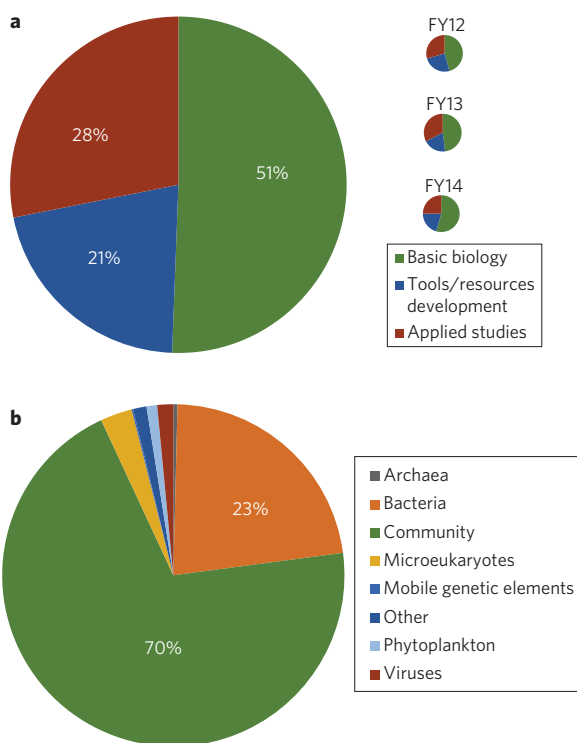
Figure 2 | Microbiome research by theme and microbe. a,b, Distribution of research activities by research theme (**a**) or microbial category (**b**) (see Methods). Only wedges representing 1% or more of the total are shown.

The participating agencies emphasized the need for sample and data collection protocol standards. A US Agency for International Development (USAID) representative listed 'standardized protocols' as its most highly prioritized technical need. A National Institute of Standards and Technology (NIST) scientist reported that such standard protocols are needed because 'the interlab comparability of measurements on microbiomes is generally poor. Biases exist along every step of the measurement process, from sample collection, extraction techniques, measurement technology employed (e.g. NGS, mass spec, NMR), and, finally, to data analysis and interpretation. There is a need for the adoption of reference materials, reference data, and reference protocols in order to identify and eliminate measurement bias.' A National Institutes of Health (NIH) programme manager also stated 'tissue and sample storage' capability as another need that would benefit from protocol standardization.

Furthermore, many organizations reported a need for standard reference materials, such as reference microbial genomes and 'mock community' metagenome sequences, to facilitate comparisons between different laboratories and for longitudinal studies. A Department of the Interior (DOI) respondent highlighted the need for 'robust reference genomes' and 'vouchered specimens' for advancing the field. Echoing this need, representatives from the Food and Drug Administration (FDA), the NIH and NIST also noted a need for 'improved methods for cultivating currently uncultivable bacteria and *in vitro* platforms to cultivate simple or complex communities'. A National Science Foundation (NSF) representative supported this need and suggested a 'National Microbial DNA and Culture Collection'.

Need for databases and repositories. Because much of microbiome research is based on analyses using genome-enabled technologies, this research generates enormous volumes of complex data of many types. Therefore, the ability to organize, store, access, share, combine and integrate these data with other data sets is of crucial

importance. Most organizations cited a pressing need for the hardware, software and data science for managing microbiome data. This was well expressed by an FDA scientist, who recognized the need for 'high performance computation resources that are flexible to accommodate a rapidly evolving analytical pipeline' and 'cloud computing solutions that overcome confidentiality and privacy concerns'. US Environmental Protection Agency (EPA) and NIH representatives particularly expressed a need for 'open-source, supervised, quality-controlled software for data analysis' and a NSF representative for 'well-curated databases that are interoperable and easy to use'. Many of the agencies identified a need for comprehensive, publically available data sets housed either within a single repository or database or in a federated database.

Importantly, programme staff from many of the participating organizations connected the need for a database with the potential for data integration of microbial community structure and function data across diverse ecosystems. Such software would link meta-genomics sequences with microbiome phenotype and function, enabling high-level comparisons and analyses across ecosystems. For example, a US Department of Agriculture (USDA) scientist called for a 'national database for the soil biome'. It was generally noted that collaborative and multidisciplinary microbiome research centres could be useful for supporting this kind of research.

Need for human resources. Despite the wide-ranging diversity of scientific disciplines and expertise surveyed in the data call, the participating organizations unanimously identified education and training in data science, computational biology and bioinformatics as crucial for advancing the field of microbiome research. As one NIH representative reported, 'There is a desperate need for individuals who can develop creative approaches to bioinformatics problems' and a USDA scientist wrote that 'students should be trained to use "big data" and computer modelling to describe networks within microbiomes, such as chemical signaling between microbes and their hosts (plant or animal) or the flow of genetic material'.

Training biologists in computational biology, data science and bioinformatics was identified as a key need. In addition, representatives from the participating agencies repeatedly described a need for generalized interdisciplinary education and training. A DOE respondent explained that 'multidisciplinary training opportunities will be critical, particularly in developing a larger pool of researchers that more effectively bridge experimental and computational biology'. A NIH programme manager took this a step further and suggested that a diversity of education and training should be mirrored in a diversity of scientists and encouraged 'more diversity in the field as it relates to access for underserved populations in research training'. In fact, another NIH respondent proposed the development of an intriguing new discipline, 'human microbiome counselors'. These specialists would function somewhat at the intersection of genetic counselors and nutritionists and would be trained to translate the complex results of microbiome studies into practical advice for the public on drug responses, or improved microbiome-targeted nutrition to support health.

Finally, the participating agencies reported the need for more hypothesis-driven research that will move the field to the development of intervention strategies to improve human and ecosystem health. A NSF programme director reported that 'although many are expert at generating sequence data', the field needs more 'investigators asking functional and biological questions about microbiomes. Training needs to move the focus away from data and method towards biological inquiry' and also noted a particular need for training in the 'intersection of microbiology, systems biology and information science'. Many of the agencies echoed the pressing need for interdisciplinary research, education and training, with suggestions that comparative microbiome research would benefit all fields.

Concluding remarks

This report provides a snapshot in time of the broad scope of intramural and extramural federally supported microbiome research activities during three fiscal years, while also documenting the opinions of government scientists in wide-ranging disciplines about the most pressing needs in microbiome research. Each surveyed agency or department in the Government reported strong interest in microbiome research as a means to solving problems, particularly those related to the production of food, the improvement of human health and ecosystem health, the production of clean, renewable energy and the manufacture of microbiome-based therapeutics and products. Emerging human health issues, such as the interplay of environmental microbes and human pathogens in the spread of antibiotic resistance[13], the surprising role of the microbiome in unexpected conditions, such as graft-versus-host disease in transplant patients[14, 15], and even new understanding of the role of the microbiome in traditionally well-studied diseases like HIV/AIDS[16], will benefit from microbiome research. Critical issues such as the need for improved food production for a growing global population will depend on new discoveries in the microbiomes of food crops such as wheat, corn, rice, barley and cassava[17]. To support this need, an international community of scientists is advocating for global databases on soil biodiversity data[18]. Even larger issues, such as efforts to predict and adapt to global climate change impacts or develop more sustainable renewable energy resources, will need to incorporate the role of ecosystem-associated microbial communities[19–21]. Remarkably, the gaps, needs and challenges identified by the broad group of scientists participating in this data call were quite similar, calling for standardization of protocols and data; hardware and user-friendly software; training and human resources for inter-disciplinary research and for computational biology and data science; baseline reference data; open access databases for microbiome data and repositories for samples and strains. One respondent to the data call noted that '"Big Science" requires proportionally big investments to generate translational outcomes'. Investment in common infrastructure for microbiome research would support and enable many types of microbiome research and allow this field to move from a descriptive phase to a predictive phase to the development of new interventions and applications grounded in microbiome science[22]. Indeed, there seems to be truly no difference between skin and soil when it comes to what is needed for the future of microbiome research.

Methods

The Fast Track Action Committee on Mapping the Microbiome (Committee) was a 14-member, inter-agency Committee chartered in February 2015 under the Life Sciences Subcommittee of the National Science and Technology Council Committee on Science for the US Government to analyse the current portfolio of federally funded microbiome research activities and to evaluate the cross-cutting gaps, needs and challenges that must be addressed to enable a predictive understanding of microbiomes. The portfolio analysis could be used to identify priority areas for coordination and to support the development of an integrated federal plan for microbiome research. Through extensive discussion, the inter-agency Committee established a set of definitions used to structure a spreadsheet and issue a data call designed to capture the diverse research portfolios of the participating organizations. The Committee distributed a set of explanatory slides that provided explanations, examples and contact information for questions. Multiple tutorials were held with the participating organizations over the six-week data-gathering period. An interactive version of the data call spreadsheet (worksheet 1) is included in the Supplementary Information, together with the elements of the drop-down menu (worksheet 2) and the logic flow of the drop-down menu (worksheet 3).

For the purposes here, 'microbiome' was defined as a multi-species community of microbes in a specific environment (that is, host, habitat or ecosystem), and 'microbiome research' was defined as the study of these communities with regard to phylogenetic and genetic composition, structure and function, and interactions with their hosts or in ecosystems, with emphasis on data derived from community sequencing and other genome-enabled technologies. The framework was community-level, with single species, single pathogen and culture-based studies excluded. All forms of microbial life and communities were included in the analysis. In addition, host–microbe interactions were considered in the context of a larger

microbial community or if the microbe served as a model for the microbiome, such as with symbiosis studies.

Included in the spreadsheet were fields for research investment apportioned by the following three Committee-defined broad research themes: (1) tools/technology/resource development; (2) basic biology; or (3) applied studies of the microbiome. Categorization by these themes required extensive discussion as each agency had different perspectives, as represented by the Committee members, particularly with regard to how best to differentiate along the continuum of basic versus applied research. Although these three themes do not have hard and fast boundaries, the Committee used a combination of definitions and examples to guide respondent data input.

The spreadsheet was further structured so that research investment was apportioned among Committee-defined 'environments'. After analysis, some of the original environments were combined to produce the following eight environments: (1) Agriculture, (2) Aquatic, (3) Atmosphere, (4) Built environment, (5) Human, (6) Non-agricultural plants for energy production, (7) Non-human laboratory studies and (8) Terrestrial. Sub-environments, habitats and sub-habitats or locations within habitats were additionally provided as choices within each of the environments.

The spreadsheet called for research investment to be apportioned among the following seven 'habitats': (1) Wild animal, (2) Domestic animal, (3) Environmental sample, (4) Human, (5) Wild plant, (6) Domestic plant and (7) Other. Categories of sub-environments, habitats or sub-habitats that represented less than 3% of the total research activities in each environment were not labelled in figures. These included the following: (a) Agriculture: Aquaculture, Food crops, Food safety, Other, Smallholder farms, Tool development; (b) Aquatic: Estuary, Invasive species, Lake community, Lake sediment, Marine sediment, Marine wild animals, Oil spill, River, Service animals; (c) Built environment: Home plumbing, Hospital; (d) Human: Blood, Decomposition, Immune System, Nares, Tool development, Wound; (e) Energy: Fuel cells; (f) Non-human lab environment: Basic biofilm Research, Bioinformatics, Immune system, *In vitro* gastrointestinal, Lab plants and soil, Skin, Synthetic microbiome, Tool development, Toxicology, Urogenital tract; and (g) Terrestrial: Cave community, Invasive species, Mountain soil, Nutrient cycling.

Finally, the spreadsheet asked respondents to identify research investment according to the following eight 'microbial categories': (1) Bacteria, (2) Archaea, (3) Microeukaryotes (fungi, protists, diatoms), (4) Phytoplankton, (5) Mobile genetic elements, (6) Viruses (eukaryotic viruses and bacteriophage), (7) Community-level studies (for example, total community analyses) and (8) Other microbial components of the microbiome.

Participating agencies were also asked for brief assessments of their current microbiome research activities, their current and future technology and training needs, and their outlook on a five- to ten-year horizon. The survey question responses were binned to cluster the most common gaps needs or challenges identified in the survey.

Included in the spreadsheet were fields for levels of support per fiscal year for FY12, FY13 or FY14. Data were collected for both intramural and extramural research, with extramural research defined here as including research by non-governmental organizations, such as universities and DOE National Laboratories. Specific funding amounts were collected for internal use; a total funded amount across all participating organizations for this period is included here.

This report provides the proportion of microbiome research investment categorized into the various themes, environments and microbial categories. For each analysis, percentages were derived from the sum of all activities reported within a classification scheme (theme, environment or microbial category). In many cases, a microbiome research activity included in this assessment represents a portion of a larger project. Although each participating organization recognized the essential role microbiome research will play in advancing its mission, many did not support stand-alone programmes or initiatives for microbiome research. To be comprehensive, the data call accepted entries for that portion of any study that included microbiome research. As such, a research activity identified in any one bin should not be interpreted as equivalent to a grant, project, programme or initiative. The entire data set was compiled, reviewed for incomplete or unclear entries and errors, respondents were contacted, and errors were corrected.

Data obtained from the six departments represented 16 separate agencies. These included the National Institute of Standards and Technology (NIST) and the National Oceanic and Atmospheric Administration (NOAA) in the Department of Commerce (DOC); the Army, Air Force, Navy/Marine Corps, the Defense Advanced Research Projects Agency (DARPA) and joint programmes within the Department of Defense (DoD); the Department of Energy (DOE); the US Geological Survey (USGS), the National Park Service (NPS) and the Bureau of Ocean Energy Management (BOEM) in the Department of the Interior (DOI); the Centers for Disease Control and Prevention (CDC), the Food and Drug Administration (FDA) and the National Institutes of Health (NIH) in the Department of Health and Human Services (HHS) and the Agricultural Research Service (ARS), the US Forest Service (FS) and the National Institute of Food and Agriculture (NIFA) in the Department of Agriculture (USDA). We also obtained data from four independent agencies. These included the US Environmental Protection Agency (EPA); the National Aeronautics and Space Administration (NASA); the National Science Foundation (NSF); and the US Agency for International Development (USAID). A quasi-governmental entity, the Smithsonian Institution (SI), was also included. Some data were reported separately per an agency; other data were pooled across

agencies or programmes. Ultimately, the data call included 14 participating organizations, representing six departments, four independent agencies and one quasi-governmental entity (the Smithsonian Institution).

References

1. Breitbart, M. *et al.* Genomic analysis of uncultured marine viral communities. *Proc. Natl Acad. Sci. USA* **99,** 14250–14255 (2002).
2. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304,** 66–74 (2004).
3. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312,** 1355–1359 (2006).
4. *Fast-Track Action Committee on Mapping the Microbiome* (National Science and Technology Council, 2015); https://www.whitehouse.gov/sites/default/files/microsites/ostp/NSTC/ftac-mm_report_final_112015_0.pdf
5. Schenk, P. M., Carvalhais, L. C. & Kazan, K. Unraveling plant–microbe interactions: can multi-species transcriptomics help? *Trends Biotechnol.* **30,** 177–184 (2012).
6. Bakker, M. G., Schlatter, D. C., Otto-Hansen, L. & Kinkel, L. L. Diffuse symbioses: roles of plant–plant, plant–microbe and microbe–microbe interactions in structuring the soil microbiome. *Mol. Ecol.* **23,** 1571–1583 (2014).
7. Virgin, H. W. The virome in mammalian physiology and disease. *Cell* **157,** 142–150 (2014).
8. Brum, J. R. & Sullivan, M. B. Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nature Rev. Microbiol.* **13,** 147–159 (2015).
9. Roosinck, M. J. Move over bacteria! Viruses make their mark as mutualistic microbial symbionts. *J. Virol.* http://dx.doi.org/10.1128/JVI.02974–14 (2015).
10. Hewson, I. *et al.* Densovirus associated with sea-star wasting disease and mass mortality. *Proc. Natl Acad. Sci. USA* **111,** 17278–17283 (2014).
11. Srinivasiah, S. *et al.* Dynamics of autochthonous soil viral communities parallels dynamics of host communities under nutrient stimulation. *FEMS Micro. Ecol.* http://dx.doi.org/10.1093/femsec/fiv063 (2015).
12. Wu, Z. *et al.* Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME J.* http://dx.doi.org/10.1038/ismej.2015.138 (2015).
13. Forsberg, K. J. *et al.* The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337,** 1107–1111 (2012).
14. Jeng, R. R. *et al.* Regulation of intestinal inflammation by microbiota following allogeneic bone marrow transplantation. *J. Exp. Med.* **209,** 903–911 (2012).
15. Vossen, J. M. *et al.* Complete suppression of the gut microbiome prevents acute graft-versus-host disease following allogenic bone marrow transplantation. *PLoS ONE* **9,** e105706 (2014).
16. Salas, J. T. & Chang, T. L. Microbiome in human immunodeficiency virus infection. *Clin. Lab. Med.* **34,** 733–745 (2014).
17. Bulgarelli, G. *et al.* Structure and functions of the bacterial microbiota of plants. *Annu. Rev. Plant Biol.* **64,** 807–838 (2013).
18. Ramirez, K. M. *et al.* Towards a global platform for linking soil biodiversity data. *Front. Ecol. Evol.* **3,** 91 (2015).
19. Batterman, S. A. *et al.* Key role of symbiotic N_2 fixation in tropical forest secondary succession. *Nature* **502,** 224–227 (2013).
20. Jansson, J. & Tas, N. The microbial ecology of permafrost. *Nature Rev. Microbiol.* **12,** 414–425 (2014).
21. Karhu, K. *et al.* Temperature sensitivity of soil respiration rates enhanced by microbial community response. *Nature* **513,** 81–84 (2014).
22. Alivisatos, A. P. *et al.* A unified initiative to harness Earth's microbiomes. *Science* **350,** 507–508 (2015).
23. The NIH Human Microbiome Project, phase one; www.commonfund.nih.gov/hmp/
24. The EC MetaHIT programme; http://www.metahit.eu/
25. The Canadian Human Microbiome Initiative; http://www.cihr-irsc.gc.ca/e/38534.html
26. The Irish ELDERMET programme; http://eldermet.ucc.ie/
27. Mullard, A. Microbiology: the inside story. *Nature* **453,** 578–579 (2008).
28. The International Human Microbiome Consortium; www.human-microbiome.org
29. The NIH Human Microbiome Project, phase two; http://ihmpdcc.org/
30. The EC 'My New Gut' programme; http://cordis.europa.eu/project/rcn/111044_en.html
31. The Irish APC Microbiome Institute; http://www.sfi.ie/assets/media/files/downloads/Investments/APC.pdf
32. The EC 'Metagenopolis' programme; http://mgps.eu/index.php?id=accueil
33. The Canadian 'Environment, Gene and Chronic Diseases' programme; https://www.researchnet-recherchenet.ca/rnr16/vwOpprtntyDtls.do?prog=2176&resultCount=25&sponsor=CIHR-10&type=EXACT&view=browseActive&language=E
34. The Canadian 'Natural Resources and the Environment: Sector Challenges – Genomic Solutions' programme; http://www.genomecanada.ca/en/portfolio/research/2015-competition.aspx
35. The EC 'Joint Action Intestinal Microbiomics' programme; https://www.healthydietforhealthylife.eu/index.php/joint-actions/microbiomics
36. The NSF Microbial Observatories and Microbial Interactions and Processes programme; https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=6166
37. The NSF Microbial Genome Sequencing Program; http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5688
38. The International Census on Marine Microbes; http://icomm.mbl.edu/
39. The Gordon and Betty Moore Foundation Marine Microbiology Initiative; https://www.moore.org/programs/science/marine-microbiology-initiative
40. TerraGenome, the International Soil Metagenome Sequencing Consortium; http://www.terragenome.org/
41. The DOE Microbial Carbon Cycle programme; http://genomicscience.energy.gov/carboncycle/index.shtml
42. Gilbert, J. A., Jansson, J. K. & Knight, R. The Earth Microbiome Project: successes and aspirations. *BMC Biol.* **12,** 69 (2014).
43. Kopf, A. *et al.* The ocean sampling day consortium. *GigaScience* **4,** 66 (2015).

Author contributions

E.S., D.F., L.P. and D.M. conceived the study. All authors jointly planned the data collection effort. All authors except E.S. and J.L. managed the data collection at their agency or department and provided their data to the central database. E.S. and J.L. conducted the analysis, and L.P. and E.S. prepared the initial draft of the manuscript. All authors revised the manuscript and agreed on the final version.

Additional information

Supplementary information is available online. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.S.

Competing interests

All authors are federal employees and the preparation of this manuscript was done as part of their official duties.