Securing the Future of Microbiome Research and Innovation

The Need for Biobanking Infrastructure in the UK

National Biofilms Innovation Centre

CABI

This List of contributors and host institutions (at the time of writing)

Matthew Ryan (lead coordinator)	CABI
Neerupma Bhardwaj	Eagle Genomics
Miguel Camara	University of Nottingham
John Dickie	Royal Botanic Gardens, Kew
Anthony Finbow	Eagle Genomics
David Green	SAMS
Nicola Holden	SRUC
Michael Hoptroff	Unilever
Jethro Johnson	University of Oxford
Samantha Law	NCIMB Ltd
Jacqueline Mackenzie-Dodds	NHM
Tim Mauchline	Rothamsted Research
Alex Mitchell	Eagle Genomics
Andrew Morgan	Independent
Barry Murphy	Unilever
Laura Peachey	University of Bristol
Virginia Pedicord	University of Cambridge
Claire Pearson	University of Oxford
Christian Roghi	Eagle Genomics
Chrysi Sergaki	The Medicines and Healthcare Products Regulatory Agency
Colette Shortt	FoodObservatory
Sven Sewitz	Eagle Genomics
Craig Thomson	HGF
Alan Walker	The Rowett Institute, Aberdeen
Jeremy Webb	University of Southampton
Miriam Gifford (Reviewer)	University of Warwick
Trevor Lawley (Reviewer)	Microbiotica Ltd./ Sanger Inst.
Anne Neville (Reviewer)	Microbiotica Ltd.
Jamie Newbold (Reviewer)	SRUC
Alan Buddie (Reviewer)	CABI

Executive Summary and Key Recommendations

Andrew Morgan

This report has been inspired by the success of the UK Biobank as a global resource for research and innovation that is contributing significantly to the advancement of science and its application to human health. In the Microbiome Strategic Roadmap (KTN, 2021), the microbiome science and innovation community has already signalled the need for microbiome biobanking and now, in this in-depth review of the science and the opportunity this represents, we provide actionable recommendations on how to move this from an ambition to a reality.

The main recommendation of this report is to embark on a highly ambitious enterprise aimed at creating a world-leading microbiome biobank – "UK Microbiome Biobank" – to support academic and industrial scientists in the quest for new medicines and solutions to some of the most significant challenges faced by the world. We now know that microbes play an incredibly important role in sustaining all forms of life on Earth and in supporting the health of humans, animals, plants and the environment as well as (a minority of these) being agents of infection and disease (e.g. see https://microbiologysociety.org/why-microbiology-matters/what-is-microbiology.html).

Microbes have the potential to contribute to solutions for treating cancer, neurodegenerative and metabolic diseases, antimicrobial resistance and pandemics, a sustainable agriculture system that can feed the world's growing population while, at the same time, helping reduce greenhouse gas emissions in support of achieving net zero. To ensure that we have the resources for research and innovation aimed at finding such solutions, we need to conserve raw source materials and microbes (individual strains and communities), sequence their genomes, and make the material and data available for researchers of today and the future on a scale that has, hitherto, not been feasible. The UK Microbiome Biobank will enhance economic competitiveness in the life sciences and the sustainability of our bioeconomy.

To date, such activities have been largely conducted in silos by different research groups, both academic and private enterprises, with very little attempt at ensuring common standards of sampling, DNA extraction, sequencing, data handling and management. The new venture will set the benchmark for such standards, thereby advancing the rigour and applicability of microbiome science more widely and across all main sectors of microbiome research – human, animal, plant and environmental. In effect, this is a **One Health** UK Microbiome Bank that is as much about biobanking as the standards needed to set up and run such an enterprise and for the field at large.

By developing and embedding such standards, the initiative will provide a longitudinal view of microbial diversity, metabolism and function as it potentially changes, enabling us to answer key questions about the impact over time of different human activities, practices, interventions and conservation measures and, not least, of climate change. Moreover, if microbial biodiversity is lost over time, by conserving microbes and communities of microbes from key environments into the future, we shall have the possible means to restore the lost biodiversity for the benefit of societal and planetary health and wellbeing.

In order to establish a UK Microbiome Biobank (UKMB), the report provides the following recommendations:

1. Establish the UKMB as a hub-and-spoke model along the lines of the UK Crop Microbiome Cryobank, (Ryan et al. 2023) ensuring links between sample provenance, the physical sample and its associated metadata using an open access approach to provision of both samples and data to researchers. For finite and irreplaceable specimens, e.g. raw stool/original soil/food lot, it will be necessary to define how access can be requested, reviewed and approved.

   There should be a central coordinating UKMB hub and database with the spokes representing the thematic areas, which will include:

   - Human Health
   - Animal Health
   - Soil, Plants and Environment
   - Food Systems
   - Biofilms – including the National Biofilms Innovation Centre.

2. UKMB will need to:
    - Incorporate existing UK biobanking and culture collection infrastructure (e.g. medical biobanks, UK culture collections, seedbanks, zoological and molecular collections) and link with other UKRI national capabilities, EU and other European infrastructures and international activities, including standards and regulatory organisations.
    - Encourage the exchange of data, technology and research expertise to facilitate improvements across the different thematic areas.
    - Provide links to data infrastructure for storing, querying and backing-up electronic metadata, Encourage use of reporting guidelines: e.g. STORMS checklist (Miryazi et al., 2021); MIGS specification (Field et al., 2008); Minimum Information about any (x) Sequence, (MIxS; https://rdamsc.bath.ac.uk/msc/m108 and https://genomicsstandardsconsortium.github.io/mixs/)
    - Promote the development of standards and encourage common approaches in data handling, management and bioinformatic interpretation.
    - Support ethical, intellectual property (IP) and regulatory compliance, e.g. Nagoya Protocol, The Human Tissues Act and General Data Protection Regulation (GDPR).
    - Support research, education and industry needs and seek to build and protect vital national microbiome assets and ensure that such assets are backed up.
    - Ensure management and timely upgrades of the central database, approved user permissions, management of approved user requests for samples and data access.
    - Identify gaps in provision and recommend/implement solutions.
3. It is proposed that each thematic area will have its own advisory panel to ensure the needs of each of the thematic research user communities are met. A representative of each 'spoke panel' will participate in the UKMB's Hub Steering Group.
4. A Strategic Management Board will need to be appointed to oversee the entire enterprise and will include representatives from UKRI (e.g. BBSRC, MRC, Innovate UK, and Innovate Business Connect) and other potential funding organisations, as well as academia, industry, regulatory bodies and participant organisations.
5. UKMB will need to have a dedicated website, which will promote the UKMB's work as well as provide the portal through which applications for approved user, sample and data requests are processed and provide a link to the UKMB database.

Whilst there is complexity in building on existing infrastructure, it is important to build on the investments already made and existing well-functioning research communities while building an overarching architecture to support a world leading UK Microbiome Biobank enterprise. In essence, the ambition is to create the microbial equivalent of the UK Biobank (Bycroft *et al*. 2018) that can contribute to solving some of the world's greatest challenges. The UKMB will be a beacon for the UK's world-class microbiome science, a catalyst for innovation in the life sciences aligned to the UK's Life Sciences Vision 2021, an accelerator of economic growth, and a resource that will conserve microbial biodiversity for generations to come.

Contents

Acknowledgements

The report authors would like to thank Pedro Carvalho & Gabriela Juarez Martinez, Innovate UK Business Connect, Sarah Hilliar (CABI) for design, NBIC for sponsorship and all the many people involved in supporting the production of the report.

Glossary of terms

aDNA – archive Deoxyribonucleic Acid
AMR – Antimicrobial resistance
ASF – Altered Schaedler Flora
ASV – Amplicon Sequence Variant
ATCC – American Type Culture Collection
BBMRI-ERIC -Biobanking and BioMolecular Resources Research Infrastructure
BBSRC – Biotechnology & Biological Sciences Research Council
BEIS – Department for Business, Energy and Industrial Strategy
BRC – Biological Resource Centre
CABI – CAB International
CAP – College of American Pathologists
CBD – Convention on Biological Diversity
CBS – Centraalbureau voor Schimmelcultures
CCAP – Culture Collection of Algae and Protozoa
CCLG – Children's Cancer and Leukaemia Group Tissue Bank
CECT – Colección Española de Cultivos Tipo
CEN – European Committee for Standardisation
CHAP – Centre for Crop Health and Protection
CIOMS – Council for International Organizations of Medical Sciences
CLIMB – Cloud Infrastructure for Microbial Bioinformatics
CM2BL – Canadian MetaMicroBiome Library
CNCM – Collection Nationale de Cultures de Microorganismes
CRG – Centre for Genomic Regulation
CTRNet – Canadian Tissue Repository Network
DEFRA – Department for Environment, Food and Rural Affairs
DOH – Department of Health
DSMZ – Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
EBI – European Bioinformatics Institute
Eco-tox – Ecotoxicology
eDNA – environmental Deoxyribonucleic acid
EFSA – European Food Safety Authority
EGA- European Genome–phenome Archive
ELIXIR – Distributed infrastructure for life-science information
EMBL – European Molecular Biology Laboratory
ENA – European Nucleotide Archive
FAIR – Findability, Accessibility, Interoperability, and Reusability
FAO – Food and Agriculture Organization of the United Nations
GA4GH – Global Alliance for Genomics and Health
GBE – Global Biobank Engine
GDP – Gross Domestic Product
GDPR – General Data Protection Regulation
GGBN – Global Genome Biodiversity Network
GSC – Genomics Standards Consortium
GWAS – Catalog of human genome-wide association studies
HDR-UK – Health Data Research UK
HTA – Human Tissue Authority
IARC – International Agency for Research on Cancer
IDA – International Depositary Authority
IMG/M – Integrated Microbial Genomes and Microbiomes
IOZ – Institute of Zoology
IP – Intellectual Property
IRCC – Internationally Recognised Certificate of Compliance
ISBER – International Society for Biological and Environmental Repositories
ISO – International Standards Organisation
JGI – Joint Genome Institute
KTN – Knowledge Transfer Network
LBP – Live biotherapeutic products
MAT – Mutually Agreed Terms
MCF – Molecular Collections Facility
MGnify – Microbiome Analysis Resource
MHRA – Medicines and Healthcare products Regulatory Agency
MIMARCS – minimum information about a marker gene sequence
MIRRI – Microbial Resources Research Infrastructure
MRC – Medical Research Council
MTIG – Microbiome Therapeutics Innovation Group
NBIC – National Biofilms Innovation Centre
NCI – National Cancer Institute
NCIMB – National Collection of Industrial and Marine Bacteria
NCPPB – National Collection of Plant Pathogenic Bacteria
NCYC – National Collection of Yeast Cultures
NERC – Natural Environment Research Council
NGS – Next Generation Sequencing
NHM – Natural History Museum
NIB – Northern Ireland Biobank
NIBSC – National Institute of Biological Standards and Control
NIHR – The National Institute for Health and Care Research
OECD – Organisation for Economic Co-operation and Development
OECI – Organisation of European Cancer Institutes
OHHLEP – One health high-level expert panel
OMM – Oligo Mouse Microbiota
OPSS – Office for Product Safety and Standards
OTU – Operational Taxonomic Unit
PCR – Polymerase Chain Reaction
PHE – Public Health England
PHEWAS – Phenome-wide association study
PIC – Prior Informed Consent
PRI – Pharmabiotic Research Institute
QC – Quality control
R&D – Research and Development
RD – Rare disease
RNA – Ribonucleic Acid
SAMS – Scottish Association for Marine Science
SOP – Standard Operating Procedure
SPF – Specific pathogen-free
SR&I – Science Research and Innovation
STEC – Shigatoxigenic Escherichia coli
Syncom – Synthetic Microbial Community
TDCC – Tissue Directory and Coordination Centre
TDWG – Biodiversity Information Standards
TVSP – Thames Valley Science Park
UKBRCN – UK Biological Resource Centre Network
UKCMCB – UK Crop Microbiome CryoBank
UKCRC – UK Clinical Research Collaboration
UKMB – UK Microbiome Biobank
UKRI – UK Research and Innovation
UNEP – United Nations Environment Programme
UNESCO – United Nations Educational, Scientific and Cultural Organization
USP – Unique selling point
WFCC – World Federation of Culture Collections
WHO – World Health Organization

1. Background

The KTN Microbiome Roadmap (KTN, 2021) sets out a mission of "raising visibility of, access to, and investment in UK microbiome science and innovation and fostering an environment that supports the creation of new start-ups, scale-up transitions, industry partnerships and impact on jobs and GDP". To achieve this, a supporting infrastructure is required to underpin and ensure the integrity of the research base, and to further provide the tools, resources and solutions for research and development. Central to this is the need for biobanks, culture collections and data managers to come together in a coordinated approach, but this will require the development of new procedures, standards, scientific approaches and a model to meet the needs of the microbiome community in the UK to ultimately ensure international competitiveness.

2. Current status and provision of biobanking infrastructure in the UK

Anthony Finbow, **Alex Mitchell**, **Christian Roghi**, **Sven Sewitz**, **Neerupma Bhardwaj**, **John Dickie**, **Samantha Law**, **Jacqueline Mackenzie-Dodds** & **Matthew Ryan**

The UK's culture collections, biobanks, and data repositories were established to support the needs of the life science community before the genomics and bioinformatics revolution. This revolution has increased our knowledge of the microbiome by allowing the interrogation of highly complex systems in a systemic manner. This section looks at existing infrastructure and capabilities to support microbiology, medicine, environment, and plant health, and how some of these could be expanded to support microbiome research. It is important to distinguish between biobanks and culture collections. Biobanks incorporate medical tissue collections; museum, fungaria, herbaria or zoological collections of preserved or dried specimens; blood banks; stool banks; and medical gamete / embryo storage facilities. Culture collections, in contrast, tend to be collections of axenically 'pure' cultured living microbes or cell lines.

2.1 Biobanks

2.1.1 UK Biocentre

With accreditation to support clinical sciences, the National Biosample Centre supports the medical research community. It was established by the NIHR in 2014, storing the majority of its biomedical and clinically derived samples in a highly automated repository consisting of -80°C and cryo freezers. According to its website, (ukbiocentre.com), this provides large-scale tank and freezer management services for clients needing to "off-site" collections, providing a backup facility for valuable collections". Although the samples are clinical, in origin, they are likely to be accompanied by mycoplasma, viruses and bacteria.

2.1.2 Natural History Museum – Molecular Collections Facility

The Molecular Collections Facility (MCF) at South Kensington London, is the Natural History Museum's (NHM) biobank, offering purpose-built, secure, centralised, long-term storage to researchers. It holds a wide range of molecular specimens, biological samples/specimens and cell lines from many species across the Tree of Life, along with environmental samples containing microbiomes, in an array of storage formats and a range of temperatures (+4 °C, -20 °C, -80 °C, and at <-180 °C in vapour phase liquid nitrogen; as well as dry storage at ambient temperatures). The focus on ultra-long-term sample storage (i.e. in perpetuity) enables access to unique longitudinal studies from long chronological time series collections from ecosystems and habitats across the globe. The genetic resources currently held include aDNA and eDNA, eco-tox, mixed/filtered samples of soil, water, air, cores (sediments, ice) and whole specimens containing associated organisms and microbiomes. This includes 44,000 UK pollinators (parasites, pollen, gut contents) from the £10M NERC-funded UK IPI programme. As of November 2022, focussed microbiome collections have not yet reached MCF from external research labs. The MCF's current capacity is ~2 million x

0.5 ml samples across all formats, mostly in mechanical freezers at present, but NHM has UK Government funding to expand the biobank by a factor of ~5 times at a new Science and Digitisation Centre at Shinfield, Thames Valley Science Park (TVSP) University of Reading, due to open in 2027. Microbiome collections will be important next generation resources for research, and the new biobank at Reading will have ample capacity for large new microbiome collections. With capacity for >10 million samples, it will provide the necessary space to build up new collections – key to addressing biodiversity challenges in the UK and across the planet. Biodiversity-associated microbiomes are a key part of this. More broadly, the biobank will act as a hub for deposition of biodiversity samples in line with national ambition (UKRI) to increase the capability and capacity of biodiversity research. The biobank is a key piece of infrastructure with NHM's expertise, central to delivering the Museum's strategy to "Transform the study of Natural History" and, in particular, to develop and provide "tools for discovery". It also aligns with development of new genomics facilities, including environmental and ancient DNA laboratories, and living biobank/culture labs at the TVSP facility – key adjacencies linked to the biobank.

2.1.3 Royal Botanic Gardens, Kew – Millennium Seed Bank

Located at Wakehurst Place, the Millennium Seed Bank is a collection of over 2.4 billion seeds from around the world. The seeds are likely to be accompanied by endophytic microbes including fungi and bacteria. The conservation of microbes has not been a key focus as the primary application has been the conservation of the 'seed' and not their associated microbiota. Despite the apparent emphasis on the Millennium Seed Bank collections (which the authors consider primarily as a genetic resource collection rather than a biobank – though it is the latter, too) it is important, to draw attention to DNA and tissue. Indeed, Kew holds a Fungarium of 1.25 million specimens and a herbarium of over 7 million dried specimens. Given the constant improvement in techniques to extract usable DNA from old, dried tissues, the possibility of this DNA being microbially-derived may indicate the presence of phytopathogens and symbiotically-associated microbiota.

2.2 Culture Collections

There are over 700 recognised national culture collections within the World Federation of Culture Collections (WFCC). Culture collections play an important role in the harnessing, description, and long-term preservation of authenticated reference microbial strains for current and future generations of scientists. Public 'open access' collections allow microbes isolated through research to be made available to the worldwide scientific community. The culture collections within the UK (refer to Annex 1) have been in place for decades and therefore have the knowledge to maintain microbes effectively through specific methods of preservation. Currently, UK open access collections collaborate through the UK Biological Resource Centre Network (UKBRCN).

Culture collections can offer a range of deposit and supply services. They may be based within industry (supporting the requirements of the respective company) or they may be smaller collections located in universities / research centres representing both industry and academia. Although resources and accessions may be made available, only collections registered with the World Data Centre for Microorganisms provide a catalogue, allowing users to search their holdings. Public culture collections in the UK include NCIMB, CABI, NCPPB, CCAP, NCYC and PHE. They offer a range of services for the preservation and maintenance of important research and industrial strains. These can include safe deposits, where the culture collection acts as a duplicate collection, holding academic / industrial strains as a back-up. Culture collections can also act as International Depository Authorities (IDAs) under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure (see https://www.wipo.int/treaties/en/registration/budapest). IDAs maintain microbes that are involved in patented processes, ensuring that they remain viable for at least 30 years. Culture collections may also offer a range of ancillary services including microbial identification and contract research.

Culture collections were established for the maintenance of microbes in pure "axenic" format and have long supported research in microbiology and associated fields. While some collections store samples in association with their hosts (e.g., mycoparasites) or those with symbiotic relationships (e.g., algal examples), none of the UK's current open access culture collections (refer to Annex 1) curate environmental samples or synthetic communities of more than two distinguishable microbial components. Indeed, no collection will undertake storage of complex microbiome material outside project-funded work*.

Most of the services offered by culture collections depend on the ability to culture and propagate microbes. While purified microbial isolates from an environmental or human-derived biological material can be propagated and expanded for use in research or development, original source microbiome samples cannot be expanded without introducing bias. Consequently, appropriate storage and management of these finite and irreplaceable source microbiome materials is an important goal, and poses some challenges. Furthermore, while some collections do provide for *ex situ* conservation (which is expensive), many of the UK's collections have business models that rely on income generation to support the services they provide. Figure 1 is an overview of the current infrastructure of biobanks and sequencing capability within the UK. Even though the culture collections within the UK have considerable experience in the preservation of microbes, currently they have little capacity to support microbiome research and innovation. This is however mirrored in culture collections throughout the world. Through research and innovation, the UK could become the world leader in the preservation, storage and supply of original source microbiome materials in open access, safe deposit and patent deposit collections, as well as a world class bioinformatic infrastructure.

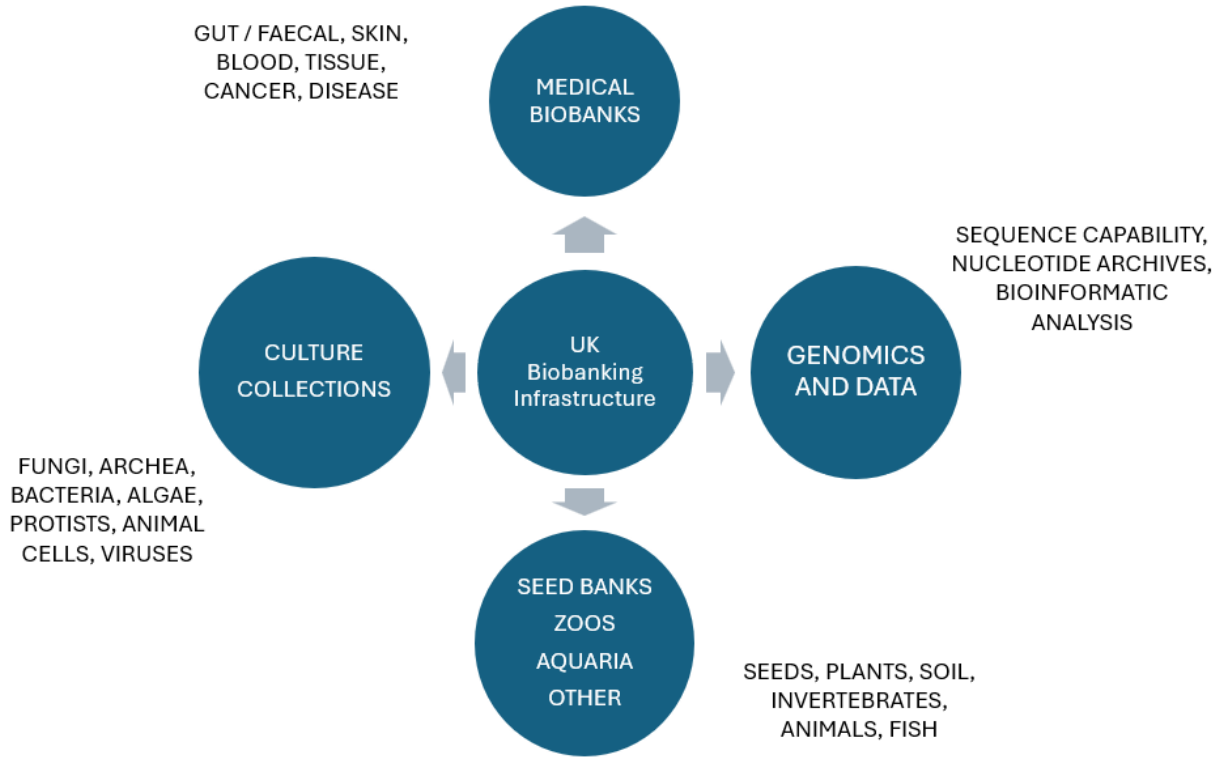


Figure 1 – The status of current infrastructure in the UK, with Microbiomes not specifically catered for.

*An exception is the UK Crop Microbiome Cryobank: a collaboration between CABI, Rothamsted Research, The John Innes Centre, Scotland's Rural College and the James Hutton Institute. (See Section 2.4)

2.3 Data and Bioinformatics

Currently, microbiome biobanking and One Health approaches are not very well developed in the UK. To date, only the UK Crop Microbiome CryoBank (UK-CMCB), a project led by CABI, has been established in this area and is being actively developed. This initiative provides the comprehensive banking of cultures, resources and information to enable scientific research in optimising plant yields using a sustainable agricultural approach. Through this project, a validated sequence resources database, 'AgMicrobiome Base' was developed and linked to the European Bioinformatics Institute (EBI) with tools and data made available to the Agritech sector. However, more data and tools are needed to build the supporting foundation and to provide data harmonisation, interoperability and internationalisation.

Unfortunately, in the UK, vast amounts of microbiome research data remain in silos, held and managed by individual researchers/biobanks, operating under different standards and governance structures; a framework that impedes sharing and effective reuse of data. To drive translational microbiome research, microbiome biobank data infrastructure needs to be developed to integrate with other data sources (for example, clinical, epidemiological, and agricultural), supporting novel data-intensive research and applications in the microbiome sector. However, despite the overall lack of microbiome biobank initiatives, the UK has taken the initiative to boost the coordination and collaboration between UK biobanks and organisations both at national and international level.

The UK Health Data Research Alliance developed and coordinated the adoption of tools, techniques, conventions, technologies, and designs that enable the use of health data in a trustworthy and ethical way for research and innovation. They formulated best practices and standards in areas such as privacy, transparency, public engagement, inclusivity, and governance to ensure that health data are shared and used responsibly by researchers and innovators.

The UK is also among the founding Members of BBMRI-ERIC, the largest infrastructure supporting health-related research in Europe. BBMRI-ERIC has built a reputation for developing open-source tooling that can support the integration of biobanks into a wider discovery framework. BBMRI.UK, the UK node of BBMRI, along with BBMRI-ERIC has launched an interoperability forum to discuss and develop open standards. BBMRI.UK, also known as the UKCRC Tissue Directory and Coordination Centre, has a track record of working with some of the world's leading suppliers of biobank software to ensure the software they provide complies with data standards that come from infrastructures such as BBMRI-ERIC. The tools for data generation and microbiome data analysis need to be demonstrably robust and fit-for-purpose, and should be accompanied by standard operating procedures. The analysis parameters specified, and the software/database versions used must be documented alongside the results when these tools are used to generate data and produce analysis outputs. Such practices will improve traceability and reproducibility of results. Moreover, approaches to improve interoperability and harmonisation of microbiome analysis tools with microbiome biobanks' infrastructures should be supported by the research community and funders . BBMRI-ERIC is a full member of RD-Connect too. RD-Connect was a six-year global infrastructure project initiated in November 2012 that further links genomic data with patient registries, biobanks, and clinical bioinformatics tools to create a central research resource for rare diseases (RDs). It was a multidisciplinary project that united partners from the EU and beyond to create an integrated global infrastructure for RD research.

2.4 Microbiome specific resources

The UK currently host two microbiome specific resources. Firstly, the BBSRC UKRI funded UK-CMCB project is building a comprehensive bank of cultures, resources and information to help facilitate research into optimising plant yields using a sustainable agricultural approach. The project is establishing a cryopreserved and characterised crop microbiome resource to underpin UK and international crop research, building on the UK agritech capability provided through the Centre for Crop Health and Protection (CHAP; now part of the UK AgriTech Centre). The resource will provide a facility for researchers to source data and samples for their work, including living microbial material from the rhizoplane (as axenic cultures, microbiome samples and in 96-well plates) and genomic sequences from different microbiome environments. This will enable soil scientists and plant researchers to assess and compare their work against validated datasets generated by the project. The focus will be on the microbiomes of six major UK crops (barley, oats, oilseed rape, potato, sugar beet and wheat) from three different soil types obtained from across the UK. Through CABI and the project's partners, research will be presented to the public as well as ways in which they can engage with this unique resource. Secondly, The Science Research and Innovation (SR&I) part of MHRA (formerly NIBSC; National Institute of Biological Standards and Control), with decades of experience on biological standardization had developed and established the first WHO International Reference Reagents for the Microbiome field (WHO 2022). The MHRA Microbiome Group has developed DNA Reference Reagents to standardise the sequencing and bioinformatics analysis of microbiome samples (Amos *et al*. 2020) and Whole Cell Reference Reagents to standardise the DNA extraction part of the process (Sergaki *et al*. 2022). These reagents are complementary to each other and allow users to assess the current limits of accuracy in their methodologies. The team has evaluated the reagents through international collaborative studies, revealing the variability amongst methodologies across the work, highlighting the urgent need for appropriate standardisation.

3. The International Landscape

Internationally, biobanking and collection infrastructures to support microbiome research are somewhat limited, despite being key to a sustainable international bioeconomy. Biobanks that preserve stool (faeces) samples for medical intervention are becoming well established in the developed world (e.g. The Netherlands Donor Faeces Bank (The Netherlands), OpenBiome (USA) and Metagenopolis (France)) and have developed operating protocols focussed on handling of human-derived material. As with the situation in the UK, few international culture collections have the capacity to handle complex microbial samples. A notable exception is DSMZ, the German National Culture Collection, which houses putative collections of single-cultured samples, including strains isolated from mouse, human intestinal and *Arabidopsis* microbiomes. There are many seed-banks and soil archives, but the maintenance of the microbial material associated with samples has not been a prime objective. Indeed, the process or preparation may involve sterilisation or storage methods that would render microbial cells inactive.

At a European level, there is no research infrastructure covering microbiome research. Elements are covered by BBMRI, MIRRI and ELIXIR but the microbiome tends to fall in the gap between them. This particular aspect was highlighted by the EU Microbiome Support project, which concluded that "Microbiome science signals a paradigm shift in the scientific approach from preservation of axenic samples in culture collections towards preservation of complex communities, which requires the supporting infrastructure to be developed" (Ryan *et al*., 2021). To drive innovation and translation of microbiome research, especially for drug development, coalitions of interested stakeholders have formed in Europe and in the United States. In Europe, the Pharmabiotic Research Institute (PRI, https://www.pharmabiotic.org/), provides regulatory support for the development and registration of "therapeutic and diagnostic products emerging from microbiome science ". In the United States, the Microbiome Therapeutics Innovation Group, (MTIG, https://microbiometig.org/about/) champions "regulatory policies and legislation that support innovation in the microbiome arena".

For culture collections, the needs of the microbiome community are still to be addressed. Interestingly, the topic was addressed by the European Confederation of Culture Collections in 2021 and again in 2024, and is on the agenda in North America but coordinated action is slow. In addition to the physical infrastructure to preserve and store samples, data management is a critical concern. There are few standardised approaches to metadata management for data originating from microbiome samples, partly because something that is important metadata for one sample type may be completely irrelevant for a different sample type. Furthermore, sample-relevant metadata differs from biologically-relevant metadata and both must be recorded appropriately and managed. Ten Hoopen *et al*. (2017) considered the overarching standards and best practices for the metagenomic lifecycle. Several initiatives are seeking to address these problems by providing publicly-accessible collections of metagenomic libraries. These include the U.S. National Microbiome Data Collaborative, and The Canadian MetaMicroBiome Library (CM2BL; http://cm2bl.org), while the mission of the Integrated Microbial Genomes and Microbiomes (IMG/M) system is to "support the annotation, analysis and distribution of microbial genome and microbiome datasets" sequenced at the US DOE's Joint Genome Institute (JGI) (Chen *et al*. 2019). Similarly, a standard has been developed by the Genomic Standards Consortium (GSC) for reporting marker gene sequences – "the minimum information about a marker gene sequence", MIMARKS (Yilmaz, *et al*. 2011). The lack of a coordinated, common approach to standards provides an opportunity for the UK to lead globally by providing infrastructure for the physical samples, their metadata and reference standards to assess the quality of both the organisms themselves and the methods used to evaluate them.

4. The Underpinning Science

Anthony Finbow, **Alex Mitchell**, **Christian Roghi**, **Sven Sewitz**, **Neerupma Bhardwaj**, and **Matthew Ryan**

In considering the need to support microbiome research, there are some underlying microbiome concepts that are essential to understand.

4.1 Synthetic Communities (SynComs)

These are microbial communities engineered to enhance or alter the wider microbiome system in order to achieve a beneficial impact. For example, a SynCom added to a soil may enhance plant nutrient uptake, mitigate against climate change or protect against disease (e.g. see Shayanthan *et al*. 2022). However, the impact of addition of SynComs to a given microbiome is largely untested and there is a need to establish not only the impact on the indigenous microbiome, but also how SynComs change the relevant microbiome in the longer term. While some microbiomes may consist of cultured/formulated microbes, there is a shift to next generation approaches using complex and continually evolved SynComs. There is a need to characterise and conserve SynComs for regulation and business continuity.

In the microbiome context, a **Keystone Species** is a microorganism that contributes to the existence of an entire ecosystem, often disproportionately in relation to others (e.g. see Tudela *et al*. 2021). There is however a need to establish what the key species in a system are, along with their individual and collective role within it. Discovering which parts of a microbial community can be taken away without altering its overall functioning would inform us about which features are important. In practice, this requires a combination of metagenomic analysis combined with laboratory-based investigations of function and behaviour. Understanding how the microbiome changes from the baseline is essential to understand the impact of removing a species. In all cases, preserving the physical components of the ecosystem, recording their provenance and linking with their metadata, are each essential.

Systems-Based Approaches e.g. the Phytobiome, the Holobiont and One Health are crucial since microbes do not occur in isolation: they are part of broader communities made up of many millions of organisms. However, although some microbes can be cultured and studied in isolation, they should be considered in the context of the ecosystem in which they naturally occur. For example, the phytobiome consists of plants, their environment, and their associated micro- and macroorganisms (American Phytopathological Society 2016). These organisms, which may be inside (including endophytes), on the surface, or adjacent to plants, include a wide diversity of microbes (viruses, bacteria, fungi, oomycetes, and algae), animals (arthropods, worms, nematodes, and rodents), and other plants. The environment includes the physical and chemical environment influencing plants and their associated organisms, and therefore, the soil, air, water, and climate. Interactions within phytobiomes are dynamic and have profound effects on soil, plant, and agroecosystem health. The sphere of relevance of phytobiomes is quite broad, spanning from crops (commodity crops, fruits, vegetables, forest, and specialty and bioenergy crops), rangelands, grasslands, and natural ecosystems to consumer products, including the quality, nutritional value, and safety of our food's source: (American Phytopathological Society 2016). In the One Health context, systems such as the phytobiome or animal microbiome may interlink with each other. For example, in the food system, the microbiome in the soil may directly influence the nutritional quality of a food or crop. When this is consumed by a ruminant, it may then directly influence its health and associated gut microbiome while the associated microbes may allow for the transfer of mobile genetic elements between them. For example, the transfer of antimicrobial resistance genes from the soil to animal via the plant. (Zhang *et al*. 2019)

4.2 Live biotherapeutic products (LBPs)

Live biotherapeutic products are considered by the European Pharmacopoeia as 'medicinal products containing live micro-organisms (bacteria or yeasts) for human use' and by the US Food and Drug Administration as a 'biological product that (1) contains live organisms; (2) is applicable to the prevention, treatment or cure of disease or condition of human beings; and (3) is not a vaccine' (European Pharmacopoeia 2019; FDA 2016). LBPs may consist of a single strain of a therapeutic microbe, of small

or large consortia of purified microorganisms (typically bacteria), and as whole-community or whole-community derived products that attempt to leverage the high taxonomic and functional diversity therein to restore an entire ecosystem. LBPs are in development for the treatment of several diseases and conditions, including inflammatory bowel disease, elimination of multi-drug resistant organisms and for immune-oncology applications, amongst others. Recently, three LBPs (BIOMICTRA, REBYOTA, VOWST have been approved for treatment of recurrent *C. difficile* infections (see Yu *et al*. 2023).

4.3 Cryotechnology

Key to the successful preservation and storage of environmental samples or synthetic communities is the application of cryobiological techniques that retain stability and community structure. Cryopreservation expertise and infrastructure within the UK is primarily centred around biomedical applications (e.g. fertility clinics), though some expertise also exists in the university and research sectors. While biobanks and culture collections have infrastructure for cryopreservation, their research expertise is limited. There is therefore a need to develop cryopreservation methodologies and protocols to meet the needs of the microbiome research community. This requires the application of cryobiological principles to ensure that samples are not compromised or destroyed by the damaging effects of freezing. Key cryo-infrastructure includes ultra-low temperature cryo-refrigerators, controlled rate cooling, nitrogen-free approaches, monitoring mechanisms and associated technologies to assess the impact of freezing on cells which itself includes, but which is not limited to, differential scanning calorimetry and cryomicroscopy. Methods to assess viability and functionality, include genomic and proteomic technologies combined with methodologies for assessing viability e.g. spectroscopy.

For cultures of microorganisms, the aim of microbial preservation is to conserve an organism without change to its physiological and genomic integrity so that it can be revived for later study or application (see Smith *et al*. 2023). Microbes are frequently archived using traditional freezing methods that were first developed in the 1950's and 1960's (see Smith *et al*., 2023)[1]. Such approaches are often effective and are used commonly by culture collections in the UK and globally. Understanding the properties of the microbial strain or species chosen for cryopreservation informs the selection of an appropriate method. Complex microbial communities, such as those found in original source microbiome materials and fractions thereof present a challenge for optimal cryopreservation of all constituent taxa. Finding appropriate solutions will require an informed approach based on what is already known and by incorporating new methods that are likely to achieve a high level of post-preservation recovery while maintaining the functional potential of the mixed-community, whether it is a synthetic community, environmental, or human-derived microbiome sample. The preservation approach chosen will likely vary between samples, and will need to be optimised for the precise type of microbiome being conserved. Importantly, the UK Microbiome Biobank could facilitate technology transfer between different research disciplines. While medical approaches are usually ahead of other sectors, in terms of scientific development, in the Agrifood sector the UK Crop Microbiome Cryobank is pioneering approaches for the conservation of the soil / rhizoplane microbiome utilising Stirling cycle cooling approaches. Successful cryopreservation of samples will also provide a mechanism to standardise samples 'at a snapshot in time' and preservation should potentially minimise microbiome drift prior to DNA extraction and sequencing.

Important note – Storage at ultra-low temperatures is the preeminent approach to long term storage. While storage in mechanical freezers at -80°C may fulfil a purpose, it was discarded as a method of choice for culture collections as far back as the 1960s, primarily because it is sub-optimal (ice recrystalises at temperatures warmer than -136 °C), and it relies on electrical supply which, if interrupted, can be disastrous. Often researchers simply place samples in a -20°C or -80°C freezer, failing to control cooling rate which causes ice damage to samples, resulting in poor viability, compromising genomic integrity.(ref) This was reinforced in a study on human faecal samples where freezing was highly detrimental, leading to the loss of some samples.(Bahl *et al*. 2012)

1. Existing UK cryobiology infrastructure is dispersed throughout existing UK Biobanks and Culture Collections (Annex 1). Although physical capacity for storage exists, the number of centres proactively undertaking cryopreservation research is limited and has reduced substantially in recent years. Institutes actively undertaking cryo-research include the Millennium SeedBank, CABI, SAMS, IOZ and a few academic groups. However, this is fragmented and uncoordinated and, largely without exception, does not focus on complex microbial material.

As a consequence of the above, back-up power supplies and a duplicate microbiome collection at a separate location should also be considered for disaster recovery. This issue transcends the microbiome, in that the UK does not have a back-up facility like the National Germplasm reserve in Fort Collins, US, should there be an ecological or environmental disaster.

Lyophilisation, a procedure also known as "freeze-drying", involves the preservation of the biospecimen by dehydration. Many commercial culture collections, including NCTC and DSMZ, store and provide microbial strains to customers in lyophilised format. The strains can be stored for years in this format and can be revived subsequently by rehydration in culture media. Despite offering long-term stability, in general, lyophilisation subjects the microorganism being preserved to harsh conditions which may negatively impact its viability (see Smith *et al*. 2023). As such, lyophilisation procedures often need to be optimised for each different type of specimen and strain, and are not suitable for algae and more delicate fungi (Smith *et al.* 2023).

4.4 Genomics and Bioinformatics

Historically, the clinical, environmental and microbiology fields have relied on classical culture collections and their services for the study of the so-called "cultivable" microbes. The advent of next-generation DNA sequencing and its falling costs has yielded an affordable and accessible mechanism to study whole microbiotas longitudinally and across sample sets in a culture-independent manner. Several studies such as the Human Microbiome Project, the TARA Oceans Project, the TerraGenome soil project and the Earth Microbiome Project, (Gevers *et al.* 2012; Sunagawa *et al.* 2020; Vogel *et al.* 2009; Thompson *et al.* 2017) have begun to give us that broader view, providing vast amounts of data which can be used for processing into information and knowledge. However, besides the growth of our knowledge and understanding of microbiomes, research in this field is still in its relative infancy, with scientists only now beginning to fully understand the diversity and functions of these complex ecosystems. With these emerging data sets, we need new computational tools for analysis and microbiome standards will be essential to compare data sets.

A common issue in the field of microbiome/phytobiome research is the lack of standard operating procedures not only for sample collection and processing, but also for data analysis. Developing best practices and standard methods for sampling, analysis and data collection is essential for robust microbiome research, given the large degree of variability and the need to compare and combine datasets. Without standardisation of processes, there are limits to reproducibility and comparison between studies. Although the solution should not be to require everyone to use the same methods, we should consider method/technology-specific effects that can be monitored and accounted for in cross-study comparisons; the use of standard strains is particularly useful in this regard. This has been identified by scientists as a barrier to advancing research, particularly when trying to validate potential links between disease and changes in the human microbiome. Standard operating practices also have an important role to play in the development of regulatory frameworks to enable commercialisation, for example to reliably assess novel microbiome therapies, tools or products. Robust, agreed standards** will therefore both support basic research by giving greater authority to results, and facilitate more rapid translation into commercially available products and therapies. There is a need, however, to balance the requirement for standards against the need to enable new discoveries and account for rapid advances in technology, which may mean that methods become outdated quickly. In addition, determining what are the best standard protocols presents a challenge due to the inherent complexity and variability of microbiomes. It is important that these issues are considered early on in research and biobanking through close collaboration between researchers, funders and regulators. The international microbiome research community, funders, publishers and regulators should work together to agree on standards. There would be an added and longer-term benefit to this, in that it could help lead to standards being determined for the impact of microbes on other organisms, such as for assessing the impact of beneficial endophytic microbes on crop and soil, or the impact of microbial treatments in disease.

**The need for standards is a matter for debate, and there needs to be a distinction between standards used as part of QC / ISO accreditation and experimental method standardisation. There is no perfect methodology. Every method has biases and limitations, meaning that, if everyone used the same approach, we would all be equally blind to its limitations and biases. There is also the problem of deciding which standard approach is "best" and should therefore be forced on everyone else. The problem with this is that the best method depends fundamentally on the

underlying structure of the microbiome to be studied which varies hugely between environments, and even between individual humans. Some consider true standardisation is therefore an impossible and undesirable pipe dream. While, optimisation, and extensive metadata collection, is better, at least when it comes to sample collection and processing, data generation and analysis. Biological standardisation is and should be a fundamental principle underpinning all science (e.g. see Coxon *et al*. 2019)

5. Rationale – why Microbiome Biobanking is Required

The requirement for biobanks was identified in the Microbiome Roadmap (KTN, 2021). However, as illustrated in Section 1, supporting microbiome infrastructure in the UK and indeed internationally is very fragmented, with neither existing culture collections nor biobanks providing the facility for the deposit and supply of resources. The need is pressing, and the particular requirements of different sectors are mapped out in section 6. There are however some common drivers that are prerequisites if the needs of the microbiome research community are to be met through the establishment of a UK wide microbiome cryobank:

- To ensure the development of i) biological reference standards through the provision of strains and DNA and ii) frameworks and methods to support microbiome research;
- To ensure compliance of deposits with legislation such as the Nagoya Protocol of the Convention on Biological Diversity (CBD) or for deposit of samples used for medical application or environmental release etc.;
- To provide a source of new potential products for industry, medical and environmental applications;
- To protect IP and outputs from academia and industry e.g. stable storage of Syncoms, LBPs;
- To facilitate biodiversity conservation;
- To provide resources to underpin research, furthering our scientific knowledge but also to ensure the reproducibility and stringency of research. To ensure the link between provenance, sample and sample-origin (e.g. patient, environment) metadata;
- To protect the public investment in research.

5.1 Protecting public investment – deposit

According to the Innovate UK Microbiome Landscape Map (Innovate UK KTN, 2021), UKRI invested a total of £154.12M in microbiome R&D projects between 2006 and 2021, but there is no measure of whether biological samples have been deposited in collections and biobanks, that would allow other researchers to validate, repeat and scrutinise research output. Maximising the reproducibility and stringency of research is paramount, and measures should be taken to ensure that key strains are not lost, but also ensuring the link between sample and metadata is retained. At a recent microbiome conference, the audience of 80 assembled experts were asked if they had deposited sequence information into ENA or Genbank. Around 75% had done so, but just one individual had then gone on to deposit either the DNA or associated organism into a culture collection (Ryan, pers. comm.) This poses a significant question: how can the work of a researcher be repeated? The simple answer is it cannot always easily be done if source material or DNA is not available. Although some scientists will share material, others are deterred from doing so, due to the complexities of implementing MTAs. This is contradictory to the prime spirit and focus of publicly-funded research which for UKRI's mission is to "convene, catalyse and invest in close collaboration with others to build a thriving, inclusive research and innovation system that connects discovery to prosperity and public good". Without relevant resources, it is hard to ensure that discovery can deliver prosperity and public good because the very connection with the research output may be lost. Therefore, at the conclusion of a study,

research materials (with appropriate consent) should be deposited for future use which not only serves to protect the investment by UKRI and UK government departments but also provides a baseline 'reference point' and a springboard for future use and research and development.

5.2 Resource provision

While culture collections propagate axenic cultures for study, the scope for supply of microbiome material or DNA derived from it may be finite. However, microbes[2] may still be isolated from such material, if it has been optimally preserved or maintained. Importantly, this would include microbes that cannot be cultured at present', but which, as methods develop, may be culturable in the future. Resources, such as microbial cultures, DNA, extracts, Syncoms or environmental material may be provided to industry for the development of new products and solutions, including but not limited to, new LBPs, fertilisers or plant protection products. The resources provided may additionally include sequence information, provenance information, knowledge and know-how, SOPs, protocols and standards The development of 'next generation biobanking', as identified in the KTN Microbiome Roadmap, through culture-dependent, high throughput approaches, involves the storage of huge numbers of culturable material from environmental or human samples. As reported previously, there needs to be a rapid development of standards to improve the quality, efficacy and stringency of microbiome research. While this has started in the human-medical area, through the activities of the WHO and MHRA there remains a need to support and provide standards in other sectors such as agriculture, food and the environment.

5.3 Facilitating R&D – Deposit

The need for deposit of microbiome material to protect public investment (Section 5.1) and to ensure the integrity and reproducibility of research is covered elsewhere, as is the need to underpin industry and processes related to IP (Section 5.3). However, there are several scientific prerequisites that are worthy of consideration, including, but not limited to:

i. Providing a 'snapshot in time' – allowing future analysis of evolutionary trends, understanding what is in a sample and utilising this knowledge to mitigate against climate change – or understand the impacts of agricultural practice.

ii. Deposit of medical examples – for example to assess the prevalence of antimicrobial resistance genes in a human or an environmental sample.

Section 5.4 considers IP matters in the context of both the activities of the proposed UK Microbiome Biobank itself, as well as for those seeking to commercialise the deposits of the Biobank.

5.4 IP Considerations for the UK Microbiome Biobank

In seeking to create standards in preservation and analysis of the stored material, it will be important that the UK Microbiome Biobank has freedom to operate those preservation and analytical techniques that are chosen to define the standards. Consequently, apart from those aspects of the preservation and analytical techniques that have been known for more than 20 years (and so can no longer be the subject of a pending patent right), analysis of the chosen techniques will be required to confirm that no third-party UK patent rights exist that could be used to restrict their use in the UK. Of course, in arriving at the standard procedure, the Biobank themselves may develop inventive proprietary techniques that possess significant commercial advantages (e.g. enhanced post-storage viability, accurate retention of species diversity for mixed-species deposits, enhanced accuracy or precision of bioinformatics to define the deposits). It would then be possible for the UK Microbiome Biobank to protect such techniques by patent filings (if kept secret) or as know-how. Recognising and protecting such techniques as IP may help the Biobank to maintain a

2. It may be considered that a microbe may be "culturable" in the lab or "unculturable". However, if it can grow in nature, it can be potentially grown in a lab, but technology and expertise is not developed sufficiently to achieve this.

differentiation between themselves and competitor depository institutions based on the ability uniquely to provide the commercial advantage associated with the proprietary technique. Developing a brand strategy should also be employed to ensure such differentiation is clear in the marketplace, underpinned by a brand protection strategy; likely with the filing of Trade Mark applications to protect company name, service/ product names and key logos. If the UK Microbiome Biobank aim to accept microorganism deposits for the purpose of patent procedures, the Biobank must also achieve the status of an "International Depository Authority" by following the requirements of Article 7 of the Budapest Treaty. Whilst these requirements do not appear to represent a significant hurdle, it should be noted that, by becoming an International Depository Authority, the Biobank would have to accept the requirements of Article 6(2)(iv) of the Budapest Treaty: "be available for the purpose of deposit to any depositor under the same conditions". It should be noted that elements related to IDA's could be handled on behalf of any new 'cryobank' by UKBRCN associated collections.

IP Considerations for Users of the Biobank: For users of the Biobank to build successful businesses from work derived from the deposits, these users need freedom to use and exploit their developments based on samples or data provided by the Biobank. They will also likely need an ability to easily demonstrate this freedom to use in order to capture the investment required for development work. Consequently, it would be advisable that any patent rights held by the depositor or Biobank with respect to the deposited material, or contractual obligations on use of the deposit or matters derived therefrom, are clearly logged for analysis by third-parties who wish to develop commercial products based on those deposits. For example, if a depositor were to impose a contractual obligation that their deposit may only be used for non-commercial research purposes, this obligation should be clear to all viewing a deposit. It is important that a party seeking to develop a commercial venture based on work derived from a deposit can easily avoid wasted research and funding on those deposits that they can never relay on for the intended purpose. Those developing products obtained from the Biobank will also need to be able to protect their developments in order to raise finance and establish an exclusivity in the market for their business. It will be important, therefore, for a developer to be able to obtain their own IP for their work (e.g. use of the deposited material for treating disease X, when the developers work leads to the understanding of utility with respect to X). Further, it may be worth considering compelling those developing products from a deposit that they extract from the Biobank to log any patent rights they obtain based on their work. In this way, any company or institution wishing to develop a commercial product based on material from the Biobank should clearly be able to establish what they had freedom to do commercially without impinging on others' patent rights or breaching any contractual obligations. A fee may be charged by the Biobank, which could be linked to the point when the deposit is obtained, or to a commercially relevant milestone (e.g. patent filing or sale of product). Part of the fee charged by the Biobank for a deposit could include payment for an agreed license for any IP to which the user of the deposit may need to have access and that is owned by the Biobank and/or depositor. It may even be possible to arrange a simplified patent pool model around some deposits (a patent pool being a collection of patent rights relating to a specific technical area and that are owned by a number of parties, but pooled together and controlled by a single licensing entity). Finally, as we see IP due diligence questions increasingly including questions with regard to the Nagoya Protocol, it will be important that the Biobank is able to provide full details of the origin of each deposit to those wishing to exploit the deposits. For human-derived original source microbiome materials, e.g. stool samples, skin swabs, it will be necessary to verify that appropriate informed consent was provided and the boundaries of the consent given should be clearly documented. For example, it is possible that consent in place at the time of signature may not cover all potential future-use scenarios.

Researchers are required to undertake due diligence to ensure that obligations under the Nagoya Protocol are met. Through the deposit of strains, Biobanks can help researchers to ensure that appropriate agreements are in place related to the genetic resource being handled. This would include ensuring Prior Informed Consent (PIC) and Mutually Agreed Terms (MAT) are met through the provision of appropriate material transfer agreements. Without such documentation, Biobanks would not be able to store and provide microbial material. The need for researchers to make appropriate due diligence declaration to the the Office for Product Safety and Standards (OPSS) is a requirement and researchers should be aware of the regulations and seek advice as necessary. (Defra 20xx). Similarly, depositing material to be released into the environment is of upmost importance as it provides a permanent comparative baseline and may help to fulfil obligations related to environmental regulations in different jurisdictions.

Microbiome biobanks could also support companies to ensure business continuity. Millions of pounds could be spent to develop a product such as LBP or Syncom. If samples are not backed up and stored optimally, or if samples are compromised as a result of problems in manufacturing or production, or even through disaster, the cost of investment could be lost if activity or efficacy cannot be regained.

5.5 Access to data and knowledge

Open access to published data originating from microbiome research is essential. However, the usability of such data is dependent on access to, and the quality and completeness of, its associated metadata that should include details of its heritage and provenance. Microbiome biobanks will be well placed to capture and provide these details. The evolution of biobanks is rapidly enabling massive stores of human biological material and its associated information along with non-human materials: plants, animals, microbes, and more, to be collected for use in agriculture, ecology, medical research, and drug development. There is a hope that the establishment and evolution of microbiome biobanks and data sharing capabilities will revolutionise research leading to personalised solutions in healthcare and more. It is important that researchers and organisations can easily access well-maintained, integrated microbiome datasets. There are some positive examples upon which to build, where platforms for microbial genomics have been improved, for example the European ELIXIR project to develop resources to handle data, and the Medical Research Council (MRC) Cloud Infrastructure for Microbial Bioinformatics (CLIMB), which provides free microbial informatics storage and analysis tools for microbiologists in the UK.

To handle the wealth of microbiome data being generated, there is the need for:

- Development of biobanks for original source microbiome material, and improved access to existing biobanks for microbiome samples.
- Improved data quality, completeness, access, integration and infrastructure.
- More robust, more complete and higher quality databases, with improved interoperability between datasets and easier sharing of data.
- Increased discoverability and access to data and associated standards and tools for data sharing whilst protecting privacy, and harmonising data and IT across biobanks.
- Fully developed access to procedures and services for researchers, common services to all ethical, legal and societal issues related to access to personal data, and involving all relevant sectors.
- Enhanced opportunity for translation of basic research through clinical and real-world evidence, particularly meeting the needs of industry.

6. Sector specific requirements

Miguel Camara, **David Green**, **Jethro Johnson**, **Tim Mauchline**, **Claire Pearson**, **Virginia Pedicord**, **Mike Hoptoft**, **Barry Murphy**, **Laura Peachey**, **Colette Shortt**, **Alan Walker**, **Jeremy Webb**, and **Nicola Holden**

6.1 Plants and Soil – including crop health

Soil microbes, including those selected by plants from the bulk soil reservoir to form the root microbiome, offer potential to contribute to sustainable intensification of agriculture and to reduce the use of environmentally-damaging agrochemicals. Soil microbes have been shown to benefit plants by increasing nutrient availability (e.g. macro- and micro-nutrients), providing abiotic stress tolerance (e.g. salt and drought) as well as disease suppression (e.g. fungal root diseases). However, to date, microbes, with the exception of arbuscular mycorrhizal fungi and rhizobial bacteria, have been underemployed in agricultural systems as alternatives to agrochemicals such as synthetic fertilisers, fungicides and pesticides.

The development of cryobanks is one strategy that can be adopted to capture microbial diversity as well as to understand and exploit the microbiome resource. This is important as the prospect of losing microbial diversity through climate change (and other anthropogenic impacts) is a clear possibility. More generically, the development of biobanks results in a coordinated and unfragmented resource which safeguards future research and can be exploited both by academia and the private sector alike. Although the concept of a microbial cryobank is not a new one, to date, perhaps surprisingly, there has been no publicly available crop microbiome resource. As mentioned previously in the report, the UK Crop Microbiome Cryobank, is an ongoing BBSRC BBR-funded project (2019-2024) that aims to fill this gap. It is assembling a UK arable agriculture-focussed root and soil microbiome resource derived from key crops cultured in UK agriculture. The collection is derived from a network of nine well characterised UK soils, broadly classified into three contrasting soil types with known agronomic history, and is generating an indexed microbial collection with associated metadata in the form of microbial phenotypes of benefit to plants (e.g. nutrient solubilisation and disease suppression), metagenomics, 16S rRNA gene and ITS region amplicon community profiles, microbial genomic as well as soil chemical and textural data. In addition, the resource houses 'state of the art' controlled rate coolers for cryopreserving soil samples, which means that whole, intact soil samples can be used to replenish microbes in further research projects. All of this information will be held on the publicly available and searchable AgMicrobiome Base – see Ryan *et al*. 2023). It is hoped that the resource will be exploited going forward and that it will contribute to the underpinning of microbiome derived solutions to crop production. Furthermore, the model described for the UKCMB can be used as a blueprint for similar projects either from different substrates (e.g. forests, grasslands) or for arable agriculture-based collections with other soil types and crops across the globe of crop microbiome cryobanks.

Crop microbiome cryobanks with strong metadata can be exploited in several ways. Firstly, they can be used to gain a descriptive understanding of a healthy microbial community structure and the relative importance of factors, such as cropping rotations, in shaping this community. This knowledge is important as it is crucial for agronomists to gain an understanding of the likelihood of microbes establishing in a given environment. This is especially important when considering the development of associative microbial inoculants, which often have disappointing success rates, presumably due to incompatibility issues with the host and environment that they are being applied to (e.g. soil type, crop species, niche). If the use of microbiome cryobanks can help overcome this challenge, then the use of microbes including those derived from biobanks as inoculants could contribute to the microbiome assisted agriculture, which is less dependent on agrochemical inputs. It is possible that microbes can be deployed either as single inoculants or perhaps as mixtures through the development of Syncoms, with desirable phenotypic traits, which may provide a robust inoculation solution, and one that has clear commercialisation potential. Furthermore, the use of biobanks to screen for microbes with disease suppression phenotypes, or to screen for microbes that improve nutrient acquisition can lead to the discovery of the next generation of fungicides, pesticides and fertilisers as well as the development of drugs for the suppression of medically important pathogens. The latter point demonstrates that the potential use of soil microbe cryobanks goes beyond agriculture.

6.2 Human/Medicine

6.2.1 Biobanks for Human Medical Research

Biobanks are a valuable resource to collect, curate and share samples in a standardised way with the broadest group of researchers. Biobanks already play an important role in human research, with the UK Biobank (https://www.ukbiobank.ac.uk/) showcasing what can be achieved when human samples are collected in a standard fashion with sufficient donor epidemiological information, enabling a multitude of different studies within the UK and around the world. Currently there is no UK-wide biobank for human microbiome samples. Such biobanks already exist in Europe (The Netherlands Donor Faeces Bank and Metagenopolis, France) and the USA (OpenBiome) for both research purposes and as repositories for faecal microbiome transplant (FMT). Although it is possible for UK researchers to apply for samples from these existing biobanks, many of the samples currently available only represent healthy populations with "healthy" stool microbiomes.

6.2.2 Need for a Centralised UK Microbiome Biobank for Human Medicine

A UK-wide microbiome biobank would go beyond what is currently available, allowing study scientists to deposit microbiome samples for cryopreservation and research purposes from different patient cohorts representing both health and disease. Such samples could include not only stools, but also sputum, mucous and other human-derived samples containing microorganisms.

While a principal purpose of a microbiome biobank would be to preserve and access microbes ex vivo, such human-derived raw materials will inevitably contain some human cells and hence may require compliance with Human Tissue Authority (HTA) regulation. A huge advantage of a UK microbiome biobank would be that the biobank itself could apply for research ethics approval and a HTA license to permit more than short-term storage of human tissues. In line with existing human tissue biobanks, access to samples for research purposes could require an application to be made that explains the desired use and methodology, which would then be assessed by a review panel. A UK Microbiome Biobank would centralise storage of human tissues, allow oversight of sample use and user requirements and would minimise the administration needed if human tissues were to be held at decentralised facilities.

A human microbiome biobank would serve as a repository for collections of samples to be made accessible to the research community at the end of a study. This would require collection ethics and participant consent that allows future use beyond the initial study. Collecting samples that might otherwise be destroyed would also generate a highly valuable resource. In particular, researchers could address novel research questions through accessing samples and existing clinical metadata across multiple studies. Methods could be optimised to allow raw materials to be appropriately preserved and stored to facilitate both nucleic acid extraction and microbial culturing downstream whenever possible.

6.2.3 A resource for the microbiome as licensed medicine

While clinical use of human microbiome samples is still nascent, examples such as the University of Birmingham's Microbiome Treatment Centre demonstrate the potential for Medicines and Healthcare products Regulatory Agency (MHRA) licensed facilities to provide samples for clinical trials and, in the case of treatment for *C. difficile* infections, therapeutic use. Such initiatives have different requirements compared to the collection of samples for basic and translational research. They typically recruit healthy participants with detailed information about their lifestyle and medical conditions, and these microbiome samples require careful screening for pathogens. Although this would be beyond the present scope and mission of a UK Microbiome Biobank, the proposed biobank could facilitate such endeavours, in future, by providing population-level data, linking existing resources and helping to streamline sample collection, storage and access for clinical trials.

6.3 Healthcare (personal microbiome) and well-being

There is a well-established relationship connecting human health/wellbeing and the microbiomes of the skin and of the oral cavity (ref). In dermatology, the microbiome is critical to the aetiology of medically important conditions such as atopic eczema, rosacea and psoriasis, which affect over 125 million people globally (Adamson 2017; Weil *et al*. 2022).

In dentistry, the microbiome is the primary driver of periodontitis, a condition which affects over 10% people globally and incurs a global economic burden of over $300Bn (Botelho *et al*. 2022)] and caries, which is the most common dental disease in the world, impacting over 25% of children in the UK (Public Health England 2020). The microbiome is also the critical biological driver of conditions as diverse as body odour, dandruff, acne and gingivitis which, taken together, represent over £400Bn of global business and are significant contributors to the UK economy (British Beauty Council, 2020).

6.3.1 Value of Biobanks to Health & Wellbeing Research

Currently, innovation in the rapidly expanding area of microbiome technologies for skin and oral health is stymied by a lack of access to relevant microbiome isolates. This limits access to clinically relevant isolates with appropriate metadata and results in a situation that has direct and negative consequences for the wellbeing of both patients and consumers and leads to unrealised innovation and growth for the wider UK economy.

6.3.2 Existing Biobank Resources

Currently no microbiome biobank exists to provide access for skin or oral microbiome samples/isolates. Furthermore, no structured repositories of individual pure-culture isolates characterised by relevant demographics such as age, ethnicity, diet, or clinical pathology exist, thereby compounding the restriction of innovation by both academia and industry. To exemplify the current situation, a University researcher presently seeking to investigate the skin microbiome of infants, the elderly, or people exhibiting eczematic pathologies must commission and conduct a dedicated clinical study, likely costing many tens or hundreds of thousands of pounds, thus creating an enormous barrier to academic investigation. The costs and endeavour to gain access to raw materials to launch new programs of investigation could therefore be reduced substantially by having access to a large, centralised bank of relevant starting materials and associated metadata.

6.4 Environment

6.4.1 Value of Biobanks to Environmental Research

A great deal of environmental microbiome research is ecologically-oriented and non-commercial. It typically focuses on characterising biodiversity and investigating the environmental drivers of community structure and function. An active and growing branch of environmental microbiome research is based around bioprospecting for exploitable novelty, such as extremophiles, secondary metabolites and bioreporters. There are no centralised biobanks for the preservation of natural-environmental samples from terrestrial (excluding agriculture-related) and aquatic sources. Cultivated strains derived as a product of environmental research, such as bacteria, archaea, fungi and algae, may be deposited in culture collections. In some cases, peer-review publication stipulates that isolates (bacteria typically) are made available for academic research. The value in biobanking of environmental samples relates largely to preservation of biodiversity, reproducibility of analysis and retrospective (bio)prospecting for novel organisms (viable) or metagenomes. The preservation of environmental samples may be especially valuable for hard-to-get samples derived from expensive and/or unique research missions, such as exploration of extreme environments (e.g. deep sea, sub-surface soil, hadal (deepest) depths, polar regions). It could be argued that public funding bodies might choose to demand that natural materials derived from extreme, ecologically fragile or expensive research-missions must be preserved for safekeeping, as the work is publicly funded and is analogous to ensuring open-access to data and publications.

6.4.2 Existing Biobank Resources

Existing biobanks that will take cultivated isolates/strains derived from environmental research, include: microbial culture collections such as CABI (UK), the DSMZ (Germany), NCIMB (UK) and many others around the globe and algal collections such as CCAP (UK). None of these culture collections however directly accept natural environmental microbiome samples for deposit to the collections. Inadvertent microbiome preservation does occur when a cultivated isolate is xenic (typical when the material contains eukaryotic organisms such as algae or fungi) and deposited to collections such as CABI and CCAP. However, the stability of the incumbent microbiome of axenic hosts is relatively understudied. Evidence from CCAP is that algal cultures have a reduced microbiome (cf. the natural environment from which they were isolated), but that this microbiome is stable over decades of continuous cultivation Nevertheless, the effects of continuous cultivation or cryostorage on the microbiome composition of xenic hosts would need to be investigated to validate its use.

6.4.3 Drivers for a Centralised UK Microbiome Biobank for Environmental Research

There is no officially mandated requirement for a UK microbiome biobank to accommodate environmental microbiome research. As stated above, there may be a justifiable requirement to ensure preservation of (i) biodiversity, notably if collected from hard-to-access regions, expensive research campaigns, or ethically, ecological sensitive or fragile environments; (ii) as part of a long-term monitoring programme; (iii) to ensure reproducibility and open-access to publicly funded research; and (iv) retrospective bioprospecting for exploitable novelty.

6.4.4 Centralised oversight and management of sample storage and use

A clear question emerges if there were to be a centralised biobank for environmental samples is, would the cost of storage, maintenance and oversight be proportionate to the future potential and value of preserving environmental samples? This value is clearer if the samples have an immediate intrinsic value (e.g. rarity, extreme origin, or as part of a bioprospecting campaign). For the bulk of samples collected, the costs and convenience to the researcher and biobank could initially appear disproportionately large, given that many stored environmental samples may never be accessed thereafter. An easier cost justification for environmental biobanking is to utilise existing UK biobanks such as CABI, CCAP and NCIMB, which have expertise and infrastructure for cryopreservation or continuous cultivation of xenic hosts. This requires a common (and centralised) format and data trail to ensure sample visibility. **Ultimately, what price is biodiversity? Beyond 'commercial' value, preservation of biodiversity is a 'public good' for the commons and humanity, and public funding possibilities should be explored**.

6.5 Veterinary including companion animals and livestock

The benefits of a UK Biobank resource for veterinary medicine mirror many of those listed above for human health. Progress in animal microbiome research (particularly in companion-animal species) lags significantly behind that of human research, because a lack of large-scale funding has led to a paucity of well-powered studies. The inconsistent methodology across a large number of smaller studies has led to variable outcomes and a lack of consensus. Consequently, biobanked samples with standardised associated metadata could provide an invaluable resource for many different types of study, including cultivation of constituent microbes. In comparison to the human microbiome, the microbiome of animals, including key livestock and companion animals, is far less well-represented in culture collections. This lack of cultured isolates is a major barrier to carrying out important research to identify animal microbiome species with the greatest potential benefits. Similarly, stored samples could be used to help answer key societal problems, such as microbiome-mediated improvements in food productivity, greenhouse gas emissions, and in enhancing the resilience of animals against infectious disease. The latter holds great promise from an animal welfare perspective, but also from a human (and one-) health perspective, as reducing carriage of zoonotic pathogens in livestock and companion animals could reduce onward transmission to humans, and hence significantly benefit human health. As discussed in the human health section above, the stored animal microbiome samples also have similar promise for therapeutic interventions such as faecal microbiota transplants in animals.

6.5.1 Accounting for the Microbiome in Animal Models

The use of animal models to understand disease mechanisms and phenotypes is widespread and includes a variety of model organisms, such as zebrafish, rodents and non-human primates. These animals are maintained in facilities throughout the UK, which often exclude specific microbes to avoid infections (specific pathogen-free, SPF). However, the normal resident microbes that form the microbiome are rarely characterised and can vary substantially between animal research facilities. As host-microbiome interactions shape many aspects of host physiology, inter-facility variation can impact model phenotypes in ways that reduce the reproducibility and translatability of pre-clinical research.

6.5.2 Need for a Centralised UK Microbiome Biobank for Model Systems

Current efforts to archive microbiomes from animal models are primarily based around collections of individual bacterial strains, which are often held by individual researchers or not-for-profit institutions such as the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ) or the University of Gothenburg Culture Collection in Sweden. While valuable, such resources often lack a single point of access and are not typically supported by a data infrastructure that links them to detailed information about the models and experimental conditions from which the samples were derived. A UK Microbiome Biobank would broaden access to detailed sample information and preserve both individual strains and intact microbial communities that could be utilised for research involving microbiome profiling, bacterial characterisation and phenotyping and faecal microbiome transfer.

6.5.2.1 Biobanking microbiomes for efficient research

In addition to banking samples from directed microbiome studies, in models where the microbiome is a potential determinant of host phenotype, a UK microbiome biobank would allow researchers to archive samples from expensive and/or difficult experiments. This would enable retrospective access to these microbiomes to study their impact on key experimental outcomes (e.g. model kinetics, disease severity) without the need to repeat the experiments solely for the collection of microbiome samples. The recently-funded MRC National Mouse Genetics Network includes plans to bank tissues from mouse experiments in which the microbiome has been tightly controlled, or excluded (see Section 6.5.2.2). Biobanking of microbiome samples alongside host tissues would likely increase synergistically the value of both resources.

6.5.2.2 Developing use of gnotobiology to dissect and refine animal model phenotypes

To define the role of the microbiome in health and disease, research has increasingly turned to the use of germ-free and gnotobiotic animal models in which microorganisms are either excluded or controlled, respectively. Growing interest in this field is reflected by the recent establishment of a germ-free facility at the Mary Lyon Centre, Harwell, which aims to provide germ-free mouse rederivation and related services to researchers across the UK.

Whilst germ-free models are useful to assess the contribution of the microbiome to model phenotypes, development of many physiological systems in germ-free animals is altered compared to that of the normally-colonised equivalent, which can impact the physiological relevance of any observations made in these animals. These alterations can be ameliorated by the controlled addition of specific bacterial species to produce gnotobiotic models via mono- and dual-colonisations, use of complex Syncoms (e.g., the Altered Schaedler Flora (ASF) and the Oligo-Mouse Microbiota (OMM)) or the introduction of microbes from a different, normally colonised host (e.g., human-to-mouse or mouse-to-mouse faecal microbiome transfer). A UK biobank would provide a valuable resource for the continued development of gnotobiotic models, enabling researchers to assess the aetiological impact of different microbial strains, develop novel synthetic microbial communities and/or transfer intact microbiomes to reproduce previously observed phenotypes and demonstrate causality of the microbiome in specific contexts.

6.6 Food Production and Culture Collections

Many food industries have long-standing established collections of microbial cultures. These in-house collections are used for a variety of reasons: to preserve cultures *ex situ*; to carry out research on species and strains; to monitor pathogen and spoilage organisms; for use in challenge testing, product design and process validation; to make use of promising microbial strains to produce foods, food supplements and additives; and to comply with regulatory and patent requirements. Information collected on microorganisms includes taxonomy, source of isolation, evolution, metabolic and functional traits, and ecological relationships. Biological resource centres (BRC), often non-public or with restricted access, are also used routinely for deposition, characterisation, authentication, and preservation of food-related microbial cultures to international standards, such as the World Federation for Culture Collections (WFCC, 2022). Examples of BRC commonly used by food companies are shown in **Table 1**. The Food Industry also interacts with other organisations who have food-relevant collections, such as: Campden BRI, the Quadram Institute, and other Research Organisations. When food regulatory submissions are made, authorities require species identification and sufficient characterisation (e.g. genetic typing) at strain level using internationally-accepted molecular DNA-based methods. Although there is no direct regulatory requirement on deposition of a particular strain in an internationally recognised culture collection, the FAO/WHO and EFSA recommend that strains should be deposited in an internationally recognised culture collection that has acquired the status of International Depositary Authority under the Budapest Treaty (FAO/WHO, 2006; EFSA, 2018). This is to ensure that tracking and access to the strain and related information is possible, should it be necessary. The importance of characterisation and deposition of strains was recently highlighted in the context of Health Claim authorisation (Regulation (EC) No. 1924/2006). Both have proven to be areas where applicants have failed to comply, and this has frequently led to the rejection of regulatory submissions related to microbial cultures (EFSA, 2009). This highlights the need for greater promotion of the role of BRC and awareness of the importance of BRC within the food industry. It is evident that BRC plays a critical role in food security and in driving sustainable innovation in the Food Industry and thus it is imperative that UK resources are provided to preserve and advance culture collections.

Table 1. Examples of Biological Resource Centres/Microbial Culture Collections used by Food Companies

Name	Holdings	Website	Acronym	*WDCM No.
National Collection of Industrial and Marine Bacteria (UK)	Includes a collection of food related bacteria, used for production of pigments and more general use.	https://www.ncimb.com	NCIMB	653
Collection Nationale de Cultures de Microorganismes	Broad collection of medical related bacteria	https://www.pasteur.fr/en/cncm	CNCM	174
German Collection of Microorganisms and Cell Cultures	A diverse collection of bacteria, widely known as 'DSMZ' contains many bacteria of relevance to food.	https://www.dsmz.de	DSMZ	274
Centraalbureau voor Schimmelcultures	Fungi, includes food strains and yeasts.	https://wi.knaw.nl	CBS	133
Colección Española de Cultivos Tipo. Spain	Collection of Fungi and bactera	https://www.uv.es/uvweb/coleccion-espanola-cultivos-tipo	CECT	412
American Type Culture Collection	Commercial focussed collection of fungi and bacteria, contains many test strains and microbial standards	https://www.atcc.org	ATCC	1

*** World Data Centre for Microorganisms Culture Collection Information**

6.7 Biofilms

Microbial biofilms present numerous translational opportunities across sectors. They exert a major impact on human and animal health, pose food safety challenges, disrupt production from oil and gas wells and contaminate drinking water supplies, but they can be beneficial in other areas such as waste-water treatment processes or increasing the bioavailability of nutrients in the soil and remediating oil spillages. The wide impact of biofilms across different sectors represents an area of opportunity for translational research as they have an economic impact in excess of $5tn a year. (Cámara *et al*. 2022).

Biobanks and collections of biofilm samples relevant to health and industrial applications are lacking. They are essential for fundamental research on the biology of biofilms but also for the relevant testing and validation of novel interventions. There is a real need to develop *in vitro* biofilm biobanks, including polymicrobial communities that retain their physicochemical properties upon storage. There is also the need to create in vitro biofilm biobanks made using standardised methods that can be accessible to the research community to test novel interventions in a reproducible manner, reducing experimental variability between laboratories. A challenge in this area is the development of preservation methods for these biofilms that can retain the architecture of the biofilms as well as the genetic and metabolic status. For natural biofilms it is paramount to store metadata on the environmental conditions the biofilms were exposed to upon collection including antibiotics, temperature etc. Such biobanks also require an underpinning digital infrastructure to make them accessible and discoverable. Leveraging existing infrastructure within national capabilities and understanding the requirements of the biofilm community to advance biobanking technology should offer significant return to the biofilm field.

The National Biofilms Innovation Centre will be generating a roadmap for the development of a fully accessible national biofilm biobank which can support the development of diagnostics and interventions, with an initial focus on health but which can ultimately be translated across sectors. The Biofilm Biobank roadmap will include (i) preserved lab-based standard model biofilms (fully characterised single species and polymicrobial), and (ii) real-world biofilms, preserving the functionality and genetic diversity from their natural environment. Samples will be available for fundamental research, acceleration of product development and commercialisation. To establish this roadmap, NBIC will work with the UK Biological Resource Centre Network (UKBRCN), the UK Health Security Agency, and academic, industry and clinical partners (NIHR BRCs and UK Clinical Research Facility Network including NIHR Clinical Research Facilities). Work with the UKCRC Tissue Directory and Coordination Centre will support the development of a digital strategy to ensure the biobank is discoverable and accessible.

6.8 One Health Context

One Health is defined by the WHO as an integrated, unifying approach to 'balance and optimise the health of people, animals and the environment'. One of the main reasons to take this approach is to prevent and respond to global health threats, primarily from microbial hazards. This has a wide relevance, covering food and water safety, nutrition, the control of zoonoses, pollution management, and in combating antimicrobial resistance (AMR).

The WHO One Health Initiative aims to integrate work on different sectors (human, animal and environmental health) along with the Food and Agriculture Organization of the United Nations (FAO), the United Nations Environment Programme (UNEP) and the World Organisation for Animal Health (WOAH) as a One Health Quadripartite. A high-level expert panel (OHHLEP) was formed in May 2021 to specifically advise on One Health issues. At a global level, the Quadripartite promotes multi-sectoral approaches to reduce health threats at the human-animal-ecosystem interface.

The wide scope of relevance of the One Health approach has a requisite understanding of basic microbiology of microbes, with a fundamental requirement to enable access to banked strains and their genetic resources. This was exemplified in the SARS-CoV-2 outbreak (2019) where access to material was pivotal in determination of the zoonotic origins of the virus (ref[NH1]). The scale and impact of the SARS-

CoV-2 pandemic was unprecedented, affecting everyone globally yet we anticipate global pandemics to continue to arise, most likely from zoonotic origins, and influenced by the consequences of climatic change (Baker *et al*. 2022). Of equal or perhaps even greater importance is the so-called 'silent pandemic' of antimicrobial resistance (Murray *et al*. 2022), because it has been brought about almost entirely through anthropogenic activities and actions. The WHO has a global action plan on AMR (ref[NH2]), delegated to most countries as 5-year National Action Plans. However, our continued reliance on pharmaceutical, veterinary and pesticide antimicrobials for maintaining human, animal and plant health means control of AMR will continue to be a challenge. Furthermore, the multi-species and/or multi-environment nature of the transmission pathways of zoonotic and AMR threats place them centrally within the One Health agenda. The One Health agenda itself continues to evolve, with an increased appreciation of the function that soil health, plant health and biodiversity play, consolidating the role of microbiology.

DNA-sequence-based analysis is a key tool in molecular epidemiology, surveillance and detection of target microbes. Whole genome sequencing has become routine for clinical infectious diseases, allowing high-resolution analysis to inform on disease outbreaks, e.g. for Shigatoxigenic *Escherichia coli* (STEC) (Holmes *et al*. 2018). Access to historical samples and sequences extended the approach to identify emerging STEC threats from an alternative transmission pathway (Dallman *et al*. 2022). Similarly, high-resolution sequence analysis provides insight for crop health, for plant-associated microbes. This was exemplified with the extremely divergent species complex of *Pseudomonas fluorescens*, whose functions span from beneficial to virulent on plant hosts, making high-resolution identification crucial (Melnyk *et al*. 2019). These types of sequence-based studies are grounded in banked isolates that can be queried and re-investigated as needed.

As microbiology moves into the microbial community era , it opens new avenues for One Health centred investigations. Understanding of individual pathogens and of microbial hazards has been greatly enriched by taking a microbial community context, illuminating the significance of microbial ecological interactions within the given environment or host (Fierer 2017). Equally this approach has been recognised for some time to enhance surveillance and detection application (Miller *et al*. 2013). However, microbiome analyses have inherent challenges, not least for detection and identification of microbes in low abundance within a community, or determination of viability status. Therefore, to explore fully the power of microbiomes for One Health approaches, access to biobanked samples is vital for optimisation and validation.

[NH1]Webpage link if this is any good, otherwise can find a suitable paper citation https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part

[NH2]Another link https://www.who.int/publications/i/item/9789241509763

7. Proposition, Remit and Scope

We propose a model to establish an internationally-competitive Microbiome Biobank as a national capability. This will seek to provide a one health, integrated biobanking capacity in the UK to support both the academic and commercial microbiome community and enhance economic capability and the sustainability of the bioeconomy. This activity will also cover the vital area of standards development and provision.

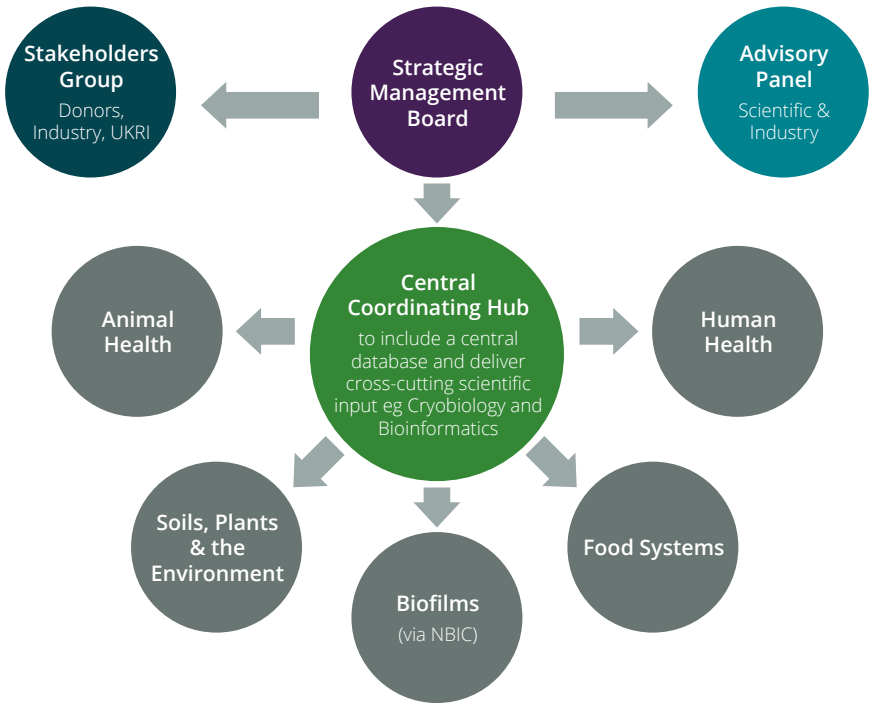
The model will

- Incorporate existing biobanking and culture collection infrastructure
- Encourage the exchange of data, technology and research expertise to facilitate improvements across different sectors
- Allow for common approaches in data handling, management and bioinformatic interpretation
- Allow for the development of metadata and sample collection standards
- Identify gaps in provision
- Seek to protect vital national assets with respect to biodiversity and human health
- Provide a strategy to ensure vital national assets are backed up
- Support research, education and industry

A key emphasis will be to link with other UKRI national capabilities, EU infrastructures such as ELIXIR, EBI-EMBL, BBMRI and MIRRI and International activities such as the World Data Centre for Microorganisms, the International Society for Biological and Environmental Repositories (ISBER) as well as international standards and regulatory organisations.

The biobanking regime will build on the model and practices of the UK Crop Microbiome Cryobank, ensuring links between sample provenance, the physical sample and its associated metadata, using an open access approach to both samples and data.

7.1 Hub and Spoke Model for Biobanking and Standards



The UK Microbiome Biobank
A hub and spoke model

Spokes will be representative of thematic areas.

It is envisaged that each thematic area will have its own advisory panel to ensure the needs of the respected research user community are met. A representative of each 'spoke panel' will ensure liaison with the UK Microbiome Biobank Steering Group. The steering group will consist of representatives from Innovate UK Business Connect, UKRI, academia and industry as well as participant organisations.

7.2 Strategic management

The Strategic Management Board will oversee the operation of the Biobank, with the prime objectives of ensuring:

i. spokes operate to the highest possible standard

ii. knowledge transfer occurs between the various spokes ensuring common approaches wherever possible

iii. infrastructure is available to meet the operating needs of hubs

iv. standardisation of data generation and format

v. responsive to the strategic needs of UKRI and other stakeholders

vi. linkage with international partners, institutes and governments

vii. adherence to ethical and legislative compliance e.g, the Nagoya Protocol of the CBD

viii. direction/management of accession policy in association with the spokes)

ix. translation of cryopreservation theory and application

i. identification of opportunities

The central coordinating hub will manage the central database, requests for samples through an appropriate website and taking into account IP, manage a 'biobank'-wide accession policy, ensure quality through development of standards and protocols and ensure enquiries are delivered to an appropriate spoke.

7.4 Data

Data will be handled via a database through the central coordinating hub based on a FAIR open access approach and with meta-data, linked to EBI's MGnify platform.

7.4.1 Common data infrastructure

A common data infrastructure to underpin the quality and reproducibility of all microbiome-based research for both academic and commercial applications is a desired outcome. This should encompass whether biobanks, culture collections or both can be further expanded to cover this area, while also considering current legislation – and adaptations of it, including data handling – quality and provenance and quality and standard operating protocols. This will require the identification of infrastructural overlaps in order to gauge what is missing and what is required within the UK, EU and beyond. Key international umbrella organisations such as the International Society for Biological and Environmental Repositories (ISBER), European and African Society for Biobanking, Word Federation for Culture Collections (WFCC) and Global Genome Biodiversity Network (GGBN) will need to be consulted to measure what data are accessible globally and whether these represent 'total' or restricted data sets.

Such a unified platform will support systemic analysis of biobank data at multiple levels simultaneously, in a way that balances respecting privacy and sharing data appropriately with other collaborators. Traditional database management systems, distributed computing systems, data lakes and file-based systems all exhibit constraints that limit the value to be gained from biobank data analysis. For example, they are often too slow to process heterogeneous scientific datasets, and a single powerful server could easily spend weeks computing the results of one large-scale phenome studies (e.g. GWAS or PheWAS). Furthermore, as systems grow in complexity, databases can easily become isolated and data sharing between systems is restricted.

ELIXIR is already building an infrastructure to support the development of (i) new rare-disease therapies, (ii) long-term storage and reuse of human biomedical data using the European Genome-phenome Archive (EGA) and (iii) Global Alliance for Genomics and Health (GA4GH) Beacons to discover genomics datasets. ELIXIR-UK aims to build links to the Farr Institute, Health Data Research UK (HDR UK), UK BioBank and UK Clinical Research Collaboration's (UKCRC) Tissue Directory and Coordination Centre (TDCC) as well as other national and internationally recognised projects in this area. ELIXIR also created the FAIR cookbook, along with researchers and data managers professionals. It is an online resource for the Life Sciences with recipes that help to make and keep data Findable, Accessible, Interoperable and Reusable (FAIR).

An example of how data can potentially be managed is summarised by the approach of ELIXIR/EMBL-EBI infrastructure and UK Biobank. This will be taken into consideration as a model for the Microbiome Biobank:

Genomic data from the 500,000 people participating in the UK Biobank initiative is distributed via the European Genome–phenome Archive (EGA), a resource developed jointly by the European Bioinformatics Institute (EMBL-EBI) and the Centre for Genomic Regulation (CRG).

- Distribution of the data via the EGA ensured long-term data security, accessibility, and sustainability, which will help researchers to better understand human disease.
- In its first few weeks of activity, more than 300 researchers across 139 institutes requested access to the genetic data from UK Biobank.
- By partnering with the EGA, the UK Biobank can make the data available through robust community-agreed standards and practices.
- It has seen computational biology research, with work spanning sequence analysis methods, multi-dimensional statistical analysis and data-driven biological discovery.
- In the future, the BioBank may encompass new datasets, such as data from fitness devices (Willettts et al. 2018). These can provide greater detail of an individual's health and lifestyle than a traditional survey.
- Handling this data resource calls for considerable IT infrastructure. The UK is always in need of growing our storage infrastructure, and our analysis infrastructure.
- Data are classified as 'very confidential', 'confidential', 'internal' and 'public'; confidential data can only

be moved to the public cloud with the express permission of the data controller.

- The organisation is pursuing a hybrid multicloud strategy, drawing on services from providers including Google, AWS, and the European Open Science Cloud. This ensures that the institute's cloud infrastructure is flexible and can support the diverse needs of the different teams based at the institute.

The EGA participates in the large funded projects euCanSHare [http://www.eucanshare.eu/], EUCANCan [https://eucancan.com/], and CINECA [https://edukad.etag.ee/project/4011?lang=en]. The CINECA project, will work with 18 organisations representing European, Canadian and African cohorts to develop and apply the necessary international infrastructure to responsibly share and analyse data based on existing cohorts' data, operating within existing consent and EU General Data Protection Regulation (GDPR) regulations. Genomics England in the UK, has now completed full genome sequencing for more than 100,000 participants, and has already demonstrated benefits by providing a diagnosis for one in four participants of the rare disease component of the initiative. No other national sequencing initiative has reached this scale, with most being currently at the stage of inception. We should learn from this example and be prepared to build the example in microbiome space as well.

7.4.2 Pan-European Data Harmonisation for Biobanks in ADOPT BBMRI-ERIC

To transform these data into a common representation, a uniform approach for data integration and harmonisation had to be developed. For example, the ADOPT Biomolecular Resources Research Infrastructure-European Research Infrastructure Consortium (BBMRI-ERIC) project kick-started BBMRI-ERIC by collecting colorectal cancer data from European biobanks and aims to facilitate access to such biological resources.

7.4.3 Data standards

There should be a minimal mechanism to address the appropriate data standards required for microbiome research, ensuring compatibility by bringing the best aspects of the many current data standards and approach together into open access universal standardised approach. The following are to be considered:

- European Committee for Standardisation (CEN) and International Standardisation Organisation (ISO) are such forums where experts are drafting standards related to biobanking and use of biological resources.
- Consultation with the relevant standards bodies in the alignment and inclusion of microbiome data and metadata including the Biodiversity Information Standards (TDWG) maintainers of the GGBN data standard, the Genomic Data Standards Consortium (GSC) and the World Federation for Culture Collections; URL: http://www.wfcc.info/).
- Best Practices Guidelines for human and microorganism biological resources were published by the International entities such as OECD, ISBER, NCI, IARC, UNESCO (http://portal.unesco.org/en/ev.php URL_ID=17720&URL_DO=DO_TOPIC&URL_SECTION=201.html), Human Genome Organisation (http://www.hugo-international.org/HUGO-CELS), CIOMS, Nuffield Council on bioethics, and European Society of Human Genetics.
- Many countries have specific legislation or national or regional guidelines for collecting and processing human tissues. On the other hand, there exist several quality assessments programs and standards: the Canadian Tissue Repository Network (CTRNet) Biobank Certification Program, the College of American Pathologist (CAP) and the OECI accreditation program (https://www.oeci.eu/Accreditation/Page.aspx?name=OECI_STANDARDS).

In France (NFS 96 900), Brazil, and UK (UK standard for biobanks), a national standard for biobanks has been implemented. It is estimated that there are about 90 guidelines for biobanks which were published by national or international entities. In addition, the European Union as well as other national funding agencies, supported projects aiming to define requirements and harmonise practices in biotechnologies, research activities, data storage and bioinformatics.

It is important to highlight that different raw data processing leads to different results that cannot be harmonised. For example, by using a different reference database for taxonomical classification, different "counts" for each taxa are obtained, sometimes even the taxa names are not compatible across databases (although researchers cannot be forced to use specific databases). Going beyond this, the clustering of

sequences should be also standardised (for the case of 16S rRNA gene data, OTUs and ASVs are not compatible). And finally, agreed ontologies should be used for genes and pathways annotation.

We will not only have DNA/RNA-based data but also protein data and metabolomics data. Current microbiome datasets are stored in different archives with different histories and different requirements for metadata. Thus, sometimes it is not even possible to have the same metadata set for the same sample if genomic, metabolomics, or proteomic data are stored. This emphasises the importance of unique identifiers and the urgent need for unifying principles to allow for easy discovery and interoperability. Information will also likely not be in the same data store, so it is important to keep these data associated to allow interoperability based on standards and infrastructures. Currently, there is little to no linkage, so this needs to be facilitated. Most importantly, it is the duty of individual microbiome researchers to actively and accurately record all data produced from their research. Ideally, it should be a condition of sample and metadata submission and publication of their research that standard formats are followed to ensure standardisation and reproducibility of their research. Although it must also be noted that not all metadata can be captured, efforts should be made to use good data entry methods – beyond entering of tabular data – to ease the adoption of metadata requirements, and to use smart technology to assist or pre-fill fields where appropriate.

7.4.4 Platform for data analysis

On top of the infrastructure, standards, harmonisation and interoperability, we need a platform ecosystem that can be linked to these data sets to navigate through the complex multi-omic data to enable data exploration, hypothesis generation and causal discoveries. Tools for genotype-phenotype exploration and ontology mappings are available. For example; Global Biobank Engine (GBE), is a web-based tool that enables exploration of the relationship between genotype and phenotype in biobank cohorts, such as the UK Biobank. GBE supports browsing for results from genome-wide association studies, phenome-wide association studies, gene-based tests and genetic correlation between phenotypes. ZOOMA, another tool for example, is an application for discovering optimal ontology mappings, developed by the Samples, Phenotypes and Ontologies Team at EBI. It can be used to automatically annotate "properties" (plain text, descriptive values about biological entities) with "semantic tags" (ontology classes). This means it can handle a wide variety of diverse datasets that contain metadata descriptions about species, anatomical components, cell types, drugs treatments and compounds, diseases and phenotypes and many others. These tools need to be developed further by utilising graph technology and network science for causal discoveries alongside implementing standards through the data submission, storage, onboarding and analysis through FAIRification recipes. ELIXIR FAIRification recipe for example focuses on (semi-)automated approaches and best practices to support the selection of the knowledge resources (i.e. ontologies, thesauri, etc.) that provide higher semantic coverage with respect to the description of the structure (i.e. metadata) and content of the dataset and enable the selection from these knowledge resources, of (sets of) concepts useful to describe the structure and content of the dataset.

7.5 Services to the research community

- Deposit of research materials– encouragement to/by UKRI / donors / funders with open access. This could be part of a specimen management plan for all publicly-funded research.
- Provide easy access to data and resources, including provision of samples for research and industry.
- Services to industry – safe deposit, patents, IP; Consultancy to industry, technology transfer.
- IP and compliance material-transfer agreements would be required. Material from UKRI funded projects to be made publicly available. Potential for public safe deposits.
- Open access to materials and data to appropriately qualified individuals
- Facilitate collaboration.
- Curation – identification, authentication and ordering.

References

Adamson AS. The Economics Burden of Atopic Dermatitis. *Advances in Experimental Medicine and Biology* 2017;1027:79-92. doi: 10.1007/978-3-319-64804-0_8. PMID: 29063433.

American Phytopathological Society Phytobiomes: A Roadmap for Research and Translation (American Phytopathological Society, St. Paul, MN. USA) 2016; Available at www.phytobiomes.org/Roadmap/Documents/PhytobiomesRoadmap.pdf Accessed 22 July, 2024.

Amos GCA, Logan A, Anwar S, Fritzsche M, Mate R, Bleazard T, Rijpkema S. Developing standards for the microbiome field. *Microbiome* 2020 Jun 26;8(1):98. doi: 10.1186/s40168-020-00856-3. PMID: 32591016; PMCID: PMC7320585.

Bahl MI, Bergström A, Rask-Licht T, Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis, FEMS Microbiology Letters, 2012;329(2):193–197, https://doi.org/10.1111/j.1574-6968.2012.02523.x

Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F *et al*. Infectious disease in an era of global change. *Nature Reviews Microbiology* 2022;20(4):193-205 doi: 10.1038/s41579-021-00639-z.

Botelho J, Machado V, Leira Y, Proença L, Chambrone L, Mendes JJ. Economic burden of periodontitis in the United States and Europe: An updated estimation. *Journal of Periodontology* 2022;93(3): 373-379. 10.1002/JPER.21-0111

British Beauty Council The Value of Beauty report 2019. https://britishbeautycouncil.com/wp-content/uploads/2020/05/Value_of_beauty_FINAL.pdf [Accessed on 22 July 2024]

Bycroft C, Freeman C, Petkova D *et al*. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–209. https://doi.org/10.1038/s41586-018-0579-z

Cámara, M, Green W, MacPhee CE *et al*. Economic significance of biofilms: a multidisciplinary and cross-sectoral challenge. npj Biofilms *Microbiomes* 2022;8: 42.. https://doi.org/10.1038/s41522-022-00306-y

Coxon CH, Longstaff C, Burns C Applying the science of measurement to biology: Why bother? *PLoS Biol* 2019;17(6): e3000338. https:// doi.org/10.1371/journal.pbio.3000338

Dallman TJ, Jalava K, Verlander NQ, Gally D, Jenkins C *et al*. Identification of domestic reservoirs and common exposures in an emerging lineage of Shiga toxin-producing *Escherichia col*i O157:H7 in England: a genomic epidemiological analysis. *The Lancet Microbe* 2022 doi: https://doi.org/10.1016/S2666-5247(22)00089-1

Department for Environment Food & Rural Affairs. Nagoya Protocol – UK processes for submitting Due Diligence Declarations, applying for Best Practice and registering a Collection in the UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/947137/nagoya-protocol-process-guide.pdf. Accessed on 22 July 2024

EFSA Journal (2009),7(9):1247. Scientific Opinion on the substantiation of health claims related to non-characterised microorganisms pursuant to Article 13(1) of Regulation (EC) No 1924/20061. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2009.1247

EFSA Journal (2018), 16(3): 5206y. Guidance on the characterisation of microorganisms used as feed additives or as production organisms. https://doi.org/10.2903/j.efsa.2018.5206

European Pharmacopoeia Commission. 3053E General monograph on Live Biotherapeutic Products. European Pharmacopoeia 9.7 (2019).

FAO/WHO, 2006. Food and Agriculture Organization / World Health Organization, 2006. Probiotics in food. Health and nutritional properties and guidelines for evaluation. *FAO Food and Nutrition paper 85*. ftp://ftp.fao.org/docrep/fao/009/a0512e/a0512e00.pdf

FDA. Early Clinical Trials With Live Biotherapeutic Products: Chemistry, Manufacturing and Control Information (2016).

Field D, Garrity G, Gray T *et al*. The minimum information about a genome sequence (MIGS) specification. *Nature Biotechnology* 2008;26:541–547. https://doi.org/10.1038/nbt1360

Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology* 2017;15(10):579-590 doi: 10.1038/nrmicro.2017.87.

Gevers, D. *et al*. The Human Microbiome Project: a community resource for the healthy human microbiome. *PLoS Biology* 2012;10: e1001377.

Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L. Validation of whole-genome sequencing for identification and characterization of Shiga Toxin-producing *Escherichia coli* to produce standardized data to enable data sharing. *Journal of Clinical Microbiology* 2018;56(3):e01388-01317 doi: doi:10.1128/JCM.01388-17.

Innovate UK KTN (2021) https://iuk.ktn-uk.org/programme/microbiome-landscape/ Accessed 2023.

KTN (2021) 'KTN's Microbiome Innovation Network's Microbiome Strategic Roadmap'. KTN. https://ktn-uk.org/wp- content/uploads/2021/02/ Accessed 2023.

Melnyk RA, Hossain SS, Haney CH. Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated *Pseudomonas*. *ISME Journal* 2019;13(6):1575-1588 doi: 10.1038/s41396-019-0372-5.

Miller RR, Montoya V, Gardy JL, Patrick DM, Tang P. Metagenomics for pathogen detection in public health. *Genome Medicine* 2013;5(9):81 doi: 10.1186/gm485.

Mirzayi C, Renson A, Genomic Standards Consortium, *et al*. Reporting guidelines for human microbiome research: the STORMS checklist. *Nature Medicine* 2021;27:1885–1892. https://doi.org/10.1038/s41591-021-01552-x

Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G *et al*. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* 2022 doi: https://doi.org/10.1016/S0140-6736(21)02724-0

Public Health England. National Dental Epidemiology Programme for England: oral health survey of 5-year-olds 2019. https://assets.publishing.service.gov.uk/media/5e721fdc86650c72713d2520/NDEP_for_England_OH_Survey_5yr_2019_v1.0.pdf (2020). Accessed on 22 July 2024.

REGULATION (EC) No 1924/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 December 2006 on nutrition and health claims made on foods. Official Journal of the European L12; 3–18. Corrigendum 18.1.2007. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:404:0009:0025.

Ryan MJ, Mauchline TH, Malone JG *et al*. The UK Crop Microbiome Cryobank: a utility and model for supporting Phytobiomes research. *CABI Agriculture and Bioscience* 2023;4:53. https://doi.org/10.1186/s43170-023-00190-2

Sergaki, C., Anwar, S., Fritzsche, M. *et al*. Developing whole cell standards for the microbiome field. *Microbiome* 10, 123 (2022). https://doi.org/10.1186/s40168-022-01313-z

Shayanthan A, Ordoñez PAC and Oresnik IJ (2022) The Role of Synthetic Microbial Communities (SynCom) in Sustainable Agriculture. *Frontiers in Agronomy* 2022;4:896307. doi: 10.3389/fagro.2022.896307

Smith D, Ryan, MJ, Buddie AG. *Managing Microorganisms*, Wallingford, Oxon, UK. CAB International (2023).

Sunagawa S, Acinas SG, Bork P, *et al*. Tara Oceans: towards global ocean ecosystems biology. *Nature Reviews Microbiology* 2020;18: 428–445. https://doi.org/10.1038/s41579-020-0364-5

ten Hoopen P, Finn RD, Ailo Bongo L, Corre E, Fosso B, Meyer F, Mitchell A, Pelletier E, Pesole G, Santamaria M, Peder Willassen N, Cochrane G. The metagenomic data life-cycle: standards and best practices, *GigaScience* 2017;6(8): gix047, https://doi.org/10.1093/gigascience/gix047

Thompson L, Sanders J, McDonald D, *et al*. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017;551, 457–463. https://doi.org/10.1038/nature24621

Tudela H, Claus SP, Saleh M. Next generation microbiome research: identification of keystone species in the metabolic regulation of host-gut microbiota interplay. *Frontiers in Cell and Developmental Biologyl* 2021 Sep

1;9:719072. doi: 10.3389/fcell.2021.719072. PMID: 34540837; PMCID: PMC8440917.

Vogel T, Simonet P, Jansson J, *et al*. TerraGenome: a consortium for the sequencing of a soil metagenome. *Nature Reviews Microbiology* 2009;7: 252. https://doi.org/10.1038/nrmicro2119

Weil C, Sugerman PB, Chodick G, Liang H, Wang H, Calimlim BM, Dorfman A, Shalev V, Ben Amitai D, Leshem YA. Epidemiology and Economic Burden of Atopic Dermatitis: Real-World Retrospective Data from a Large Nationwide Israeli Healthcare Provider Database. *Advances in Therapy* 2022;39(6):2502-2514. doi: 10.1007/s12325-022-02120-6. Epub 2022 Mar 29. Erratum in: *Advances in Therapy* 2022 Sep;39(9):4392-4393. doi: 10.1007/s12325-022-02272-5.

WFCC, 2022 https://www.wfcc.info/about_view (Accessed October 10, 2022).

WHO. A WHO collaborative study to evaluate the candidate 1st WHO International Reference Reagents for Gut Microbiome analysis by Next-Generation Sequencing WHO/BS/2022.2416: https://www.who.int/publications/m/item/who-bs-2022.2416

Willetts, M., Hollowell, S., Aslett, L. *et al*. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Scientific Reports* 2018;8: 7961. https://doi.org/10.1038/s41598-018-26174-1

Yilmaz P, Kottmann R, Field D *et al*. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology* 2011;29:415–420. https://doi.org/10.1038/nbt.1823

Yu Y, Wang W, Zhang F. The next generation fecal microbiota transplantation: to transplant bacteria or virome. *Advanced Science* 2023;10: 2301097. https://doi.org/10.1002/advs.202301097

Zhang Y-J, Hu H-W, Chen Q-L, Singh BK, Yan H, Chen D, He JZ. Transfer of antibiotic resistance from manure-amended soils to vegetable microbiomes. *Environment International* 2019;130:104912, ISSN 0160-4120, https://doi.org/10.1016/j.envint.2019.104912

Appendix

Current UK Biobanks and Culture Collections (Holding type)

Animal and Plant Health Agency (APHA), Department of Food and Rural Affairs (Animal and plant microbes)

CABI Genetic Resource Collection, CABI International (Fungi, Bacteria, Nematodes, Yeasts)

Culture Collection of Algae and Protozoa (CCAP), Scottish Association for Marine Science (Cyanobacteria, Algae, Protozoa)

Centre for Environment, Fisheries and Aquaculture Science (Cefas), Department of Food and Rural Affairs (Bacteria)

European Collection of Authenticated Cell Cultures (ECACC), UK Health Security Agency, (Cell lines, Stem cells)

Kew Scientific Collections, Royal Botanic Gardens Kew (Seed bank, Fungi)

National Collection of Industrial, Food and Marine Bacteria (NCIMB), Private Company (Bacteria)

National Collection of Type Cultures (NCTC), UK Health Security Agency (Bacteria)

National Collection of Pathogenic Fungi (NCPF), UK Health Security Agency (Fungi, Yeasts)

National Collection of Pathogenic Viruses (NCPV), UK Health Security Agency (Viruses)

National Collection of Plant Pathogenic Bacteria (NCPPB), Food and Environment Research Agency (Bacteria)

National Collection of Yeast Cultures (NCYC), Quadram Institute Biosciences (Yeast)

National Institute for Biological Standards and Controls (NIBSC) – UK Stem Cell Bank, Medicines and Healthcare products Regulatory Agency (Stem Cells)

Natural History Museum, (Various: animal cell lines from CRCCG Cambs UK, protozoan cultures)

UK Biocentre

Cancer Biobank

Rothamsted Soil Archive

Confederation of Cancer Biobanks:

Children's Cancer and Leukaemia Group (CCLG) Tissue Bank

Human Biomaterials Resource Centre

Wales Cancer Bank

Tayside Biorepository

Northern Ireland Biobank (NIB)

NHS Grampian Biorepository

University of Southampton Faculty of Medicine Tissue Bank

Leeds Multidisciplinary Research Tissue Bank

Arden Tissue Bank

Members of the UKCRC Tissue Directory and Coordination Centre (UKCRC TDCC)

National Stakeholders

NERC, BBSRC, MRC, DEFRA, DOH, BEIS, UKRI

Facilitated by the National Biofilms Innovation Centre:
biofilms.ac.uk

Published on behalf of the authors and designed by CABI:
www.cabi.org

For further information please contact
nbic@biofilms.ac.uk

or
Cryobank@cabi.org
with 'Biobank' in the subject heading.

To contact any of the contributors, please contact the lead author
m.ryan@cabi.org

National Biofilms Innovation Centre

CABI