

# Discovery and cultivation of prokaryotic taxa in the age of metagenomics and artificial intelligence

Diego Javier Jiménez<sup>1,\*</sup>, Ramona Marasco<sup>1</sup>, Júnia Schultz<sup>1</sup>, Carlos Andrés Díaz Rodríguez<sup>2,3</sup>, Juan Nogales<sup>4</sup>, Luis Miguel Rodríguez-R<sup>5</sup>, Jörg Overmann<sup>6,7</sup>, Alexandre Soares Rosado<sup>1,\*</sup>

<sup>1</sup>Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), 23955-6900, Thuwal, Kingdom of Saudi Arabia

<sup>2</sup>Católica Biomedical Research Centre, Medicine Faculty, Universidade Católica Portuguesa, 2780-156, Oeiras, Portugal

<sup>3</sup>Gulbenkian Institute for Molecular Medicine, 2780-156, Oeiras, Portugal

<sup>4</sup>Department of Systems Biology, Centro Nacional de Biotecnología, CSIC, C/Darwin n° 3. 28048, Madrid, Spain

<sup>5</sup>Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H9220, Aalborg, Denmark

<sup>6</sup>Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, and Institute for Microbiology, Faculty of Life Sciences, Technical University of Braunschweig, 38124, Braunschweig, Germany

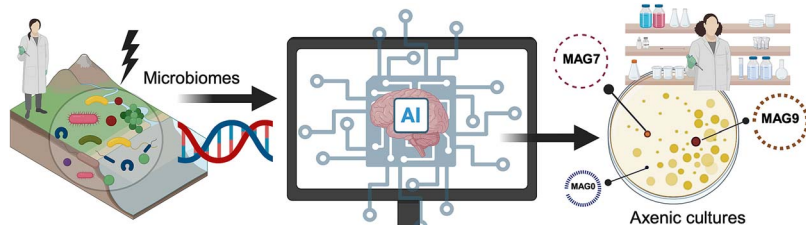
<sup>7</sup>Current affiliation: Bavarian State Collections of Natural History, Munich, Germany; and Faculty of Biology, Ludwig-Maximilians-Universität München, 80638, Munich, Germany

\*Corresponding authors: Diego Javier Jiménez, Microbial EcoGenomics and Biotechnology Laboratory (MEGB Lab), Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), 23955-6900, Thuwal, Kingdom of Saudi Arabia. E-mail: [diego.jimenezavella@kaust.edu.sa](mailto:diego.jimenezavella@kaust.edu.sa); Alexandre Soares Rosado, Microbial EcoGenomics and Biotechnology Laboratory (MEGB Lab), Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), 23955-6900, Thuwal, Kingdom of Saudi Arabia. E-mail: [alexandre.rosado@kaust.edu.sa](mailto:alexandre.rosado@kaust.edu.sa).

## Abstract

Despite advances in sequencing, microbial genomics, and cultivation techniques, the vast majority of prokaryotic species remain uncultured, which is a persistent bottleneck in microbiology and microbial ecology. This perspective outlines a conceptual framework to improve the transition from genome-resolved metagenomics to the targeted isolation of yet-uncultured prokaryotic taxa. The proposed framework integrates the induced reshaping of microbiomes, genome-based inferences of physiological and phenotypic traits, culture media design, and targeted culturomics, enabling hypothesis-driven cultivation. In addition, this manuscript addresses the critical limitations in the field, including the sequence-to-function gap, and emphasizes the synergistic potential of experimental microbiology, microbial ecology, metagenomics, and artificial intelligence-based predictions to enhance rational and actionable roadmaps for discovering and cultivating novel prokaryotic lineages.

## Graphical Abstract



BioRender.com was employed to create the figure

**Keywords** culturomics, isolation, genotype-phenotype inference, genome-scale metabolic models, machine learning, metagenome-assembled genomes, microbiome perturbation

**Received:** 14 December 2025. **Revised:** 26 January 2026. **Accepted:** 29 January 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of the International Society for Microbial Ecology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Advances in genome-resolved metagenomics (see Table 1 for a glossary) has improved the systematics of prokaryotes and expanded the discovery of novel taxa across a wide range of ecosystems [1–4]. However, our understanding of the functions of newly identified genomes and how to isolate the corresponding microorganisms is still limited. The exact fraction of prokaryotic diversity belonging to yet-uncultured species remains debated [5–7], as does the estimation of the total microbial diversity on this planet [8, 9]. It is widely recognized that far more than 90% of known prokaryotic species have never been brought into culture [4, 10]. Even among species for which genome sequences are available, recent estimates suggest that >75% remain uncultivated [11, 12]. In fact, from the trillions of predicted prokaryotic species inhabiting Earth [13, 14], currently <22 000 have been formally described (<https://lpsn.dsmz.de/text/numbers>) of which ~20 000 have been deposited as pure cultures in the German Collection of Microorganisms and Cell Cultures (DSMZ), one of the world's largest culture collections. Most of these bacterial strains belong to three phyla: *Actinomycetota*, *Pseudomonadota*, and *Bacillota* (<https://bacdiv.dsmz.de/>).

Whereas genome sequencing provides essential insights into the genetic potential of prokaryotes, cultivation remains indispensable for experimentally exploring their physiology, metabolism, and ecological roles that cannot be fully inferred from sequence data alone [15–17]. This is crucial for lineages influencing key ecosystem processes, taxa with few or no cultivated representatives, and species relevant to human and environmental health or hold potential for biotechnological applications [15, 18]. For example, members of the widely distributed phyla *Patescibacteriota* and *Sysuimicrobiota* (formerly phylum CSP1-3) are among the most desirable prokaryotic lineages for cultivation because of their ecological significance and persistent uncultivability [19–21].

The urgency of safeguarding microbial diversity through cultivation efforts has intensified in the face of accelerating climate change, land degradation, and habitat fragmentation [22, 23]. These environmental pressures might threaten prokaryotic diversity, risking the loss of species before they can be identified, studied, and harnessed. Monitoring and preserving microbial diversity must therefore be discussed in future conference of the parties of the United Nations Convention on Biological Diversity (Philippe Sansonetti; <https://www.pasteur.fr/en/research-journal/news/challenge-and-importance-preserving-microbial-biodiversity>). Such discussions have paved the way to formally integrate microorganisms and microbiomes in global conservation and biodiversity policy scenarios, now also endorsed by the Microbial Conservation Specialist Group of the International Union for Conservation of Nature [24].

For decades, prokaryotic species have been isolated by plating microbial cells (usually from serial dilutions) on agar-based culture media. These traditional methods are often biased towards isolating aerobic, mesophilic, heterotrophic, and fast-growth species [15, 25]. The isolation of many prokaryotic taxa using traditional methods is challenging, primarily due to the disruption of microbial interactions and obligate metabolic interdependencies (e.g. vitamins produced by accompanying populations) [26, 27]. Moreover, the nutritional needs and physicochemical tolerances of yet-uncultivated prokaryotic taxa are highly specific and often unknown [15, 28].

To improve the isolation of prokaryotic species, diverse strategies have been used, such as the dilution-to-extinction method, *in situ*

cultivation devices, microfluidics, cell sorting, and culturomics [15, 25, 28–30]. Examples of ingenious approaches that have yielded cultures of previously uncultivated taxa include modifications in the isolation media (e.g. using low-nutrient concentrations, growth substrate combinations, and supplementation with cAMP), cocultivation in dialysis vessels, physical enrichments via chemotaxis [31], and diffusion chamber-based approach to mimic natural environment [32, 33]. Over the past three decades, cultivation techniques have improved, enabling the isolation of a multitude of novel prokaryotic taxa. These efforts have increased the number of newly described species from a mere 100 per year (in the 1980s to 1990s) to ~1000 species per year in the last decade (<https://lpsn.dsmz.de/statistics/figure/15>) [34]. Yet, according to the SeqCode registry (<https://registry.seqco.de>), 55 phylum names and 21 673 species names (without synonym) of cultured prokaryotes are valid under the International Code of Nomenclature of Prokaryotes (as of December 2025).

In the age of metagenomics, innovative strategies (e.g. reverse genomics) have been pushing the boundaries of possibilities for isolating target prokaryotic species [35, 36]. (Meta)genome information has provided insights into the metabolic capabilities and growth requirements of specific prokaryotic taxa, guiding the development of new cultivation endeavours [16, 36–38]. However, the systematic integration of physiological profiling, genotype–phenotype inference, and genome-scale metabolic models (known as GEMs or GSMMs), derived from high-quality metagenome-assembled genomes (MAGs), has yet to be fully incorporated into targeted strategies for microbial isolation. In this regard, it has been argued that advancing cultivation efforts requires the integration of microbiome-derived data with classical microbiological methods [39].

Recently, machine learning (ML) (see definition in Table 1) has been employed to analyse the relationships between prokaryotic genotypes and phenotypes based on high-quality genomic and phenotypic data [40]. The reconstruction of ML-powered GEMs represents a substantial opportunity to improve the robustness and quality of traditional GEMs by providing more complete fluxes of metabolites and accurate environmental and metabolic context, including information about potential auxotrophies, metabolic interdependencies, substrate preferences, and niche constraints. Integration of ML, GEMs, and omics data can provide a more comprehensive link between genotype and phenotype. This joint information could offer clues for the appropriate design of cultivation strategies, including tailored culture media formulations and growth conditions [41, 42]. Although genome-based design of culture media to isolate representatives of novel prokaryotic species is an established approach for certain species-poor microbial communities [16, 43, 44], it has not been systematically applied using MAGs representing novel taxa.

Currently, the challenges of isolation of novel or yet-unculturable taxa are technical and conceptual, and several questions must be addressed in this topic. For example: (i) How can microbial communities be reshaped in favour of the enrichment of novel taxa? (ii) How can phenotypic and physiological traits be predicted from genomes and genes with unknown functions? (iii) How can genomes (including MAGs) be systematically used to improve target isolation strategies?

### From microbiome reshaping to axenic cultures: an integrative framework

This paper outlines an integrated conceptual framework, merging the controlled restructuring of microbial communities—via

**Table 1** Key terms and concepts in this perspective.

Term	Definition
Genome-resolved metagenomics	Metagenomic sequences are assembled and binned into genomes that can be employed in microbial community analyses
Cultivation	The laboratory process of growing microorganisms under controlled conditions, such as temperature, pH, and oxygen level, in culture media (sometimes as part of a community reduced in diversity) to promote cell division, recover viable cells, and isolate pure (axenic) cultures (when possible)
Isolation	Procedures aimed at separating a single microbial cell from a mixed community to obtain an axenic culture, involving streak plating, serial dilution, selective or differential media, and tailored incubation conditions that favour the target organism. Successful isolation yields genetically uniform populations comprising clonal descendants of a single, separated cell
Culturomics	High-throughput strategy of growth trials in which cells from microbial communities are cultivated in various types of media and under different physicochemical conditions in parallel; the resulting colonies are typically subjected to DNA sequencing for rapid assignment to taxa
Dilution-to-extinction	Microbial communities are serially diluted to reduce their diversity, ideally to a single cell
Novel species	Defined as taxa represented by genomes (typically MAGs) with <95% of average nucleotide identity (ANI) or <70% similarity regarding digital DNA–DNA hybridization to any members of known taxa
Reverse genomics	Cultivation strategy using antibodies targeting predicted cell surface proteins inferred from genomes (or MAGs) to capture and isolate specific microorganism types
Auxotrophies	Inability of any microbe to generate specific chemical constituents needed for growth. This inability is common in environmental microbes
Perturbation	A disturbance or shift in abiotic or biotic factors, such as temperature, nutrient level, chemical exposure, or species introduction, that disrupts the structure, function, or dynamics of a microbial community
Rare biosphere	Prokaryotic species with low relative abundance or low prevalence <i>in situ</i> . For practical purposes, any species with <0.1% of relative abundance is considered rare
Flux balance analysis (FBA)	This computational and mathematical approach to analysing GEMs relies on fluxes of metabolites via network models, which is useful to predict and maximize cell growth rates
Machine learning (ML)	ML is a subfield of artificial intelligence (AI) that uses algorithms and statistical models enabling machines to build systems and to learn from data, optimizing their performance over time on a specific goal or task

perturbation—with genome-resolved metagenomic analysis (including predictions of phenotypic traits and GEM development), artificial intelligence (AI)-guided culture media design, and experimental validation. This framework reimagines the standard isolation workflow by positioning microbiome perturbations or enrichment cultures, not just as an experimental condition, but as a deliberate, selective ecological filter. Perturbation experiments, enrichment culture methods, and genome-guided isolation have been refined in separate contexts, whether to probe the microbial community assembly in ecological studies [45] or to cultivate taxa with specific metabolic activities [38]. However, they have never been united in a single workflow. This perspective is not intended to ‘reinvent the wheel’; instead, it proposes combining these established approaches into a coordinated, iterative roadmap to facilitate the systematic discovery and targeted isolation of novel or yet-uncultured prokaryotic taxa. In this context, by addressing the challenges and highlighting innovation opportunities, this manuscript aims to lay the groundwork for a systematic and predictive science of prokaryotic cultivation and isolation based on improved and interpretable genome information coupled with AI-driven tools such as ML. The next sections explain and discuss the topics concerning the conceptual framework (Fig. 1).

### Improving the chance of retrieving novel taxa via microbiome reshaping

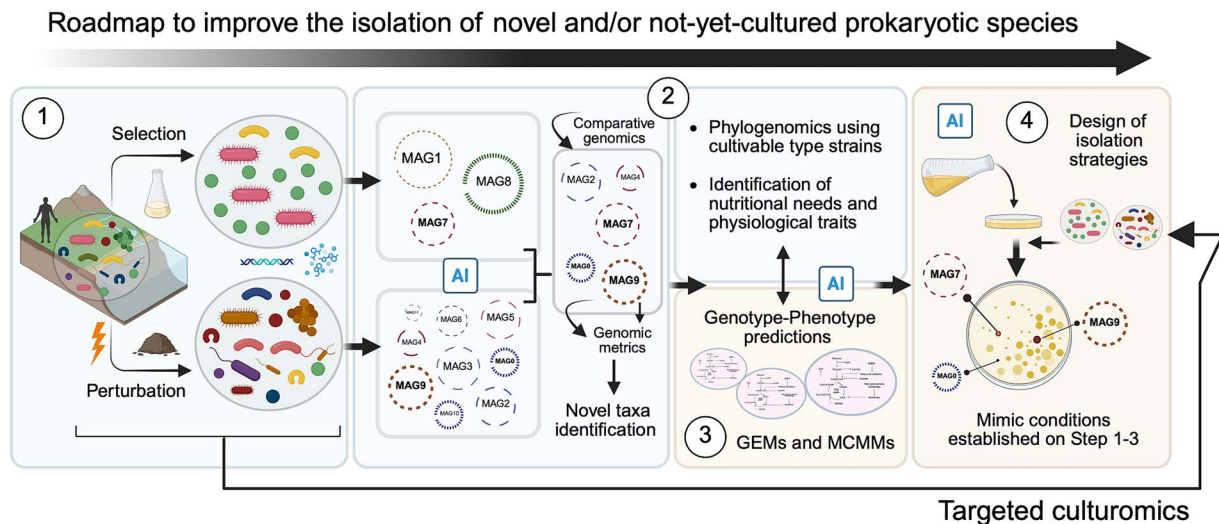
Several prokaryotic species are found in natural ecosystems in low abundance (i.e. rare biosphere) or in a dormant stage [15, 46–48]. Normally, low-abundance species are challenging to characterize via metagenome sequencing owing to the high sequencing depth required to detect them [49]. Ecological studies have proven that under

changed, perturbed, or selective conditions, some rare or dormant species may become highly active, outcompeting dominant species, increasing in abundance, and contributing to community functioning [50–52]. The induced selection of these prokaryotic taxa may improve their detectability using metagenomics, increasing the chances of informing successful isolation strategies.

In this context, microbiome reshaping aims to increase the odds of detecting rare, novel, or divergent prokaryotic lineages. Two primary approaches can be employed to restructure microbial communities: (i) perturbation experiments and (ii) enrichment cultures. Commonly, these two approaches select for specific functions, but the combination of factors or dynamic shifts in physicochemical conditions could produce random changes in the structure of the original microbiomes. Phenotype or function-based enrichment can be applied in instances where no phylogenetic markers exist to track specific prokaryotic taxa in the reshaped microbial communities.

#### (i) Perturbation experiments

Acute or chronic perturbation experiments have long been employed to examine microbial dynamics, stochastic/deterministic shifts, and the resilience or resistance of microbial communities under stress conditions [53, 54]. These studies often apply biotic (e.g. introduction of invading populations) or abiotic stressors (e.g. nutrient pulses, desiccation–rewetting cycles, pH or temperature changes) directly to cells in the environmental matrix (e.g. soil or water) to mimic natural or anthropogenic disturbances [55–59]. Although microbial communities often display predictable successional patterns following a disturbance, such as transient declines in diversity followed by reassembly [45, 54], these shifts can also uncover cryptic community



**Figure 1** Integrative conceptual framework encompassing four steps. (1) Reshaping microbial communities via *ex situ* perturbation or targeted enrichment experiments can increase the chances of detecting hidden prokaryotic diversity by altering interspecies dynamics and triggering shifts in the microbial community. (2) Genome-resolved metagenomics and the analysis of high-quality metagenome-assembled genomes (MAGs) can provide information about the novelty of the detected prokaryotic taxa (e.g. comparative genomics) and their metabolic potential and predictive physiology. (3) Artificial intelligence (AI)-based genotype-phenotype predictions, genome-scale metabolic model (GEM) reconstruction, and microbial community-scale metabolic models (MCMMs) using MAGs (detected in Step 2) can enhance the understanding of the nutritional needs and metabolic interdependencies of species, yielding crucial information on the potential design of cultivation and isolation strategies for target prokaryotic taxa. During Steps 2 and 3, ML-based tools improve MAG reconstruction, gene function annotation in MAGs, and the accuracy of GEMs and MCMMs (see Fig. 2). (4) Based on Steps 1 to 3, targeted isolation strategies (i.e. targeted culturomics) can be established and tested, including AI-driven combinatorial synthetic culture media design. At this stage, information on physicochemical conditions applied in perturbation or enrichment culture assays is input into an AI-predictive analysis to enhance the isolation of specific taxa. BioRender.com was employed to create the figure.

members previously undetectable. Such community reshuffling can enhance the likelihood of detecting novel prokaryotic taxa, especially when stressors select for traits that are underrepresented in the original microbial community, including tolerance of extreme conditions and metabolic flexibility [50, 60]. This process occurs because perturbations can alter the ecological balance or the steady state of microbial communities via competitive hierarchy disruption, resource monopolization, and niche partitioning [60, 61]. Natural ecosystems rarely settle into permanent climaxes; instead, they are dynamic and punctuated by successional turnovers over weeks, months, seasonal and diel cycles, and epicyclic swings in dominant and rare taxa [62]. Dominant taxa can successfully occupy environmental niches, suppressing the proliferation of slow-growing species via competitive exclusion. However, environmental fluctuations—here through applied perturbation—can disrupt existing microbial hierarchies by selectively disadvantaging the dominant taxa, allowing rare taxa to increase in relative abundance and exploit newly available resources or vacated niches [60, 63].

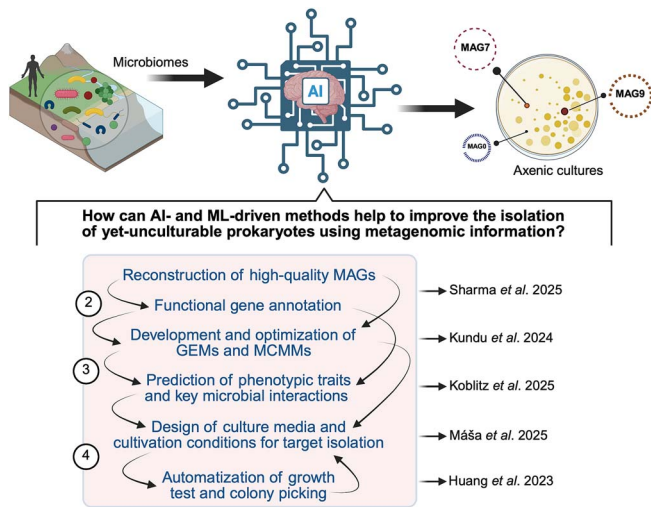
Thus, induced perturbation of microbiomes could be an approach to uncover hidden microbial diversity for elusive lineages, which can then be enriched and likely further isolated. To the best of the authors' knowledge, very few studies have applied controlled *ex situ* perturbations to environmental samples, aiming to enrich rare and/or novel yet-uncultivated prokaryotic taxa [50], compared with the more widespread application of enrichment strategies that often rely on adding a selective substrate to favour specific metabolic traits over successive culture cycles.

### (ii) Enrichment cultures

In enrichment cultures, a small amount of the natural sample or derived microbial cells is inoculated in selective liquid media which

then favour the growth of the most adapted microorganisms able to respond to the stimulus, i.e. nutrient based or adaptative [64–66]. In these systems, strong selection pressure is imposed where the niche space and resource availability are temporarily restructured, allowing otherwise excluded taxa to expand [67]. Enrichment cultures have been instrumental in cultivating difficult-to-grow microbes, particularly those involved in specialized metabolic transformations, such as methanogenesis [27], acetogenesis [68], anoxygenic photosynthesis [69], or polysaccharide [38] and plastic [70] degradation.

These controlled systems have allowed the enrichment and cultivation of many unknown and novel prokaryotic taxa. For instance, MAGs belonging to three novel bacterial genera (i.e. *Pristimantibacillus*, *Andeanibacterium*, and *Mangrovimarina*) have been detected in polymer-amendment liquid enrichment cultures using tropical soil as the inoculum source [58, 71, 72]. However, enriching these novel taxa is only part of the journey; they must be brought into a pure culture to harness these organisms fully. For instance, enrichment cultures coupled with a dilution-to-extinction approach have been employed to isolate a known plastic-degrading microorganism, *Piscinibacter sakaiensis* (formerly *Ideonella sakaiensis*) [73]. A liquid enrichment culture has some advantages because it can detect species that have low growth rates, cannot grow on a solid surface, or are inhibited by agar constituents. In addition, the low microbial complexity in these systems, compared to environmental microbiomes, can facilitate the reconstruction of high-quality MAGs and accurate microbial community-scale metabolic models (MCMMs). Therefore, enrichment cultures should not be considered passive pre-cultivation steps but active ecological tools functioning as targeted, scalable perturbations that can reshape the community structure in favour of rare and/or novel prokaryotic species.



**Figure 2** Schematic representation of the proposed steps for the target isolation of yet-uncultivable microbial species detected by genome-resolved metagenomics. Most of the processes within the pink square (i.e. from MAG reconstruction to identification of axenic cultures containing the target MAG) can be optimized by using AI-driven methods. In particular, the reconstruction and optimization (via gap filling) of genome-scale metabolic models (GEMs and MCMs) and prediction of physiological and phenotypic traits are processes where ML-driven methods have been implemented. Arrows within the square represent input for each step and processes. [BioRender.com](https://www.biorender.com) was employed to create the figure.

### Genome-resolved metagenomics: starting point for new cultivation endeavours

Advances in metagenomics and computational tools have enabled the analysis of thousands of novel MAGs [2], offering gateways to the predominantly uncharted microbial diversity. In this regard, employing long-read DNA sequencing technologies (e.g. PacBio-HiFi and Oxford nanopore) [74, 75] and hybrid assemblies combining short- and long-read sequencing data [76] offers advantages for reconstructing high-quality (>90% completeness and <5% contamination) [77] and low-fragmented MAGs (e.g. <100 contigs) [78]. When combined with high sequencing depth and advanced bioinformatic analysis integrated into ML-based approaches [79] (Fig. 2) and supported by high-quality databases, these hybrid strategies greatly improve the detection of low-abundance taxa and characterization of intraspecies diversity based on MAGs [80, 81]. However, a considerable fraction of prokaryotic diversity with low-to-intermediate abundances still escapes genome-resolved metagenomics, affecting the characterization of microbiomes with high diversity [82, 83]. Although researchers have demonstrated a nanopore-based selective sequencing method that increases the genome coverage of rare prokaryotic taxa [84], the reconstruction of high-quality MAGs of the taxonomically broad, low-abundance species remains a challenging task. Cell-sorting techniques coupled with reconstruction of single-amplified genomes have partially solved this problem.

Curated and up-to-date databases, different genomic metrics, and comparative phylogenomic analyses are necessary to determine the uniqueness of a specific set of high-quality MAGs. Recently, thresholds and standards have been proposed for genomes if used for taxonomy, nomenclature, and delineating novel prokaryotic species [85, 86]. Phylogenomic comparisons can provide data

regarding the circumscription of a novel taxon, identifying related prokaryotic species with isolated representatives (e.g. type strains). This information can be applied to guide the initial isolation attempts for targeted species by applying the cultivation conditions for their closely related isolated representatives, similar as is reported in the Known Media Database [87].

In this regard, the *BacDive* and *MediaDive* databases [88, 89] report comprehensive data on the physiological traits and cultivation conditions of ~100 000 prokaryotic strains and nearly all type strains. Providing information on the strain level is essential because prokaryotic physiology, metabolic functions, and ecological dynamics are conserved at different phylogenetic depths, sometimes even varying between subpopulations of the same species [90, 91]. Moreover, a significant amount of genetic variation exists within a single prokaryotic species [92–94], making cultivation efforts more challenging. For example, various strains of the same species may harbour divergent sets of biosynthetic pathways, nutrient requirements, or stress response mechanisms, affecting their ability to grow under laboratory conditions. Neglecting this intraspecies diversity can oversimplify microbial diversity, masking strain-specific functions and potentially reducing the predictive power of the GEMs employed to guide cultivation efforts.

### Functional gene annotation of MAGs as a keystone for target cultivation strategies

In addition to genome characterization using various features (e.g. ANI values, GC content, size, and number of protein-coding genes), a robust functional gene annotation is crucial to decipher the predictive metabolism of detected prokaryotic species. Currently, gene functions are primarily predicted based on their sequence similarity to previously characterized proteins deposited in databases. The use and integration of databases (e.g. KEGG, InterPro, MEROPS, Gene Ontology, CAZy, and Pfam), relaxed sequence homology thresholds (e.g. a minimum of 40% amino acid identity and a minimum of 70% sequence coverage), and genomic context [95, 96] can expand the gene annotation space, offering a comprehensive functional profiling of prokaryotic-derived genomes. Recently, a new method was developed to improve the prediction of functions in human gut microbiomes by leveraging expression data, gene proximity, sequence similarity, and a ML-based classifier system [97] (Fig. 2).

Highly relaxed protein similarity thresholds could spur errors in the assignment of potential functions. Thus, deep homology detection approaches for annotating gene catalogues are also recommended [98]. The coupling of these methods to unveil metabolic profiling and identify relevant ecological traits, such as stress tolerance [99], carbon and energy sources [100], growth rates [101], optimal growth temperatures [102], and amino acid biosynthesis pathways [103], provides useful physiological information of the target species, offering context for the design of cultivation strategies. Recently, the metabolic traits and physiological information retrieved from MAGs have offered important clues about carbon and nitrogen substrates for the isolation of the following: a fastidious crude-oil degrader from oil-contaminated seashores (*Macondimonas*; [104]), thermophilic spirochetes from a deep terrestrial subsurface biosphere (*Longinema*; [37]), a phenanthrene degrader from petroleum-contaminated soil (*Achromobacter*; [105]), an anaerobic arsenic-methylating bacterium from soil (*Paraclostridium*; [106]), and a hyperthermophilic polysaccharide degrader from hot spring sediments (*Fervidibacter*; [38]). A recent study has

identified metabolic traits and pathways associated with culturability, highlighting the functional novelty found in yet-uncultured microbial species [107]. In addition, a novel genome-based computational model using amino acid frequencies was developed to predict oxygen tolerance, optimum temperature, salinity, and pH for a range of prokaryotic groups [108].

A major challenge in genome-assisted cultivation is the sequence-to-function gap because many prokaryotic genomes contain genes with unknown or poorly annotated functions that complicate the interpretation of genome-derived data (see Box 1), especially when attempting to predict the entire metabolic potential, physiology, or phenotype of a microorganism [109–112]. This limitation can impair the development and curation of GEMs and the translation of genomic information into actionable cultivation strategies. High-throughput experimental platforms (e.g. gene synthesis or CRISPR/Cas-guided gene mutations coupled with microfluidic phenotypic screening), AI-driven predictions, and accurate databases are still necessary to bridge this sequence-to-function gap [109, 110, 113]. Emerging AI-based approaches are powerful tools in protein discovery and should be widely adopted to enhance the accuracy of the current predictions of protein domain annotations. For instance, structural features extracted from amino acid sequences using prediction software, such as AlphaFold [114], could further improve gene annotation. Moreover, large language models (LLMs) have demonstrated significant superiority over state-of-the-art approaches based on protein family classification, reducing prediction errors [40, 115, 116]. Further, a flux balance analysis (FBA; Table 1) to scrutinize GEMs may serve as a framework to identify and correct candidate genes whose functions might be considered unknown during annotation but are likely to support growth, a process typically known as gap filling [117].

### Box 1. Unknown functions of MAGs

Diverse factors can hinder the complete functional analysis of metagenomes: (i) Metagenome sequencing commonly omits a considerable fraction of the rare biosphere [111]. (ii) Gene and MAG catalogues are developed after an assembly process, and unassembled sequences are often discarded [94]. (iii) Gene annotation approaches provide vague and nonspecific functional information (e.g. transporters, transcriptional regulators, or hydrolases). In addition, many proteins (40% to 60%) derived from metagenomic gene catalogues or MAGs are categorized simply as hypothetical proteins [112]. This set of genes without functional annotation can be denominated as the ‘microbial gray matter’. This term was proposed by Dr Pedro J. Torres in a post where he aimed to clarify the concept of microbial dark matter that has been indiscriminately used in literature to refer to the unknown fraction of metagenomes and unseen not-yet-culturable microbial species ([https://medium.com/@PedroJTorres\\_/microbial-dark-and-gray-matter-exploring-the-unknown-67fc9ac0c712](https://medium.com/@PedroJTorres_/microbial-dark-and-gray-matter-exploring-the-unknown-67fc9ac0c712)). Four recent studies have computationally analysed unknown functions of MAGs found in diverse ecosystems, helping improve their functional assignments (e.g. using protein structural similarities and genomic context) and discovering new protein families [96, 97, 112, 114]. In addition, some funded research programmes (<https://www.darpa.mil/research/programs/duf>) and collaborative international networks (<https://dark.metagenomics.eu/who.html>) have been established to address this topic. There are still open questions regarding the bottlenecks of gene function

annotation and how to address the unknown functions of MAGs with high-throughput experimental validations. In one example, a study of mutants derived from 32 bacterial strains growing under various conditions measured gene fitness, proposing functions for poorly annotated enzymes and transporters [118]. The science is still far from filling the sequence-to-function gap in complex microbiomes, considering the relevance of the environmental context in genotype–phenotype relationships. However, computational biology is moving forward fast, and it is predicted that AI-based tools can help prioritize protein families to validate their functions experimentally.

## MAG-based metabolic modelling and ML-aided phenotype inference

Cultivating prokaryotes by using MAG information requires translating the functional potential into testable hypotheses about physiology and growth. To this end, reconstruction of GEMs, MCMs, and ML-based phenotype predictions can be used as complementary approaches.

### (i) Mechanistic modelling from MAGs

An accurate reconstruction and appropriate analysis of GEMs and MCMs can help integrate and contextualize the functional profiles of MAGs across diverse environmental scenarios. These predictive metabolic models enable a better understanding of the physiology, nutritional needs, and substrate preferences of target prokaryotic species under certain growth conditions [41, 119]. In microbial communities, MCMs can provide information about how small molecules, such as amino acids and vitamins, in the extracellular matrix affect the cell growth of specific members, detecting potential cross-feeding events and auxotrophies [120, 121]. For example, MAG-based metabolic models have been used to predict possible metabolic exchange events and interactions in hydrothermal vents and hot spring-derived microbiomes [122, 123].

The predictive power of the resulting GEMs and MCMs from metagenomic data is still limited due to intrinsic (e.g. low completeness of MAGs, drawbacks in gene annotation, and gaps in many metabolic pathways) [124] and extrinsic factors (e.g. lack of physicochemical context in microbial growth). Beyond improving the functional gene annotation, integrating transcriptomic, proteomic, and metabolomic data is recommended to overcome these challenges, providing insight into the overall context of the sampled microbial community, including viability, metabolism, expression levels, adaptations, and lifestyle of target-specific members [125]. This integration can be achieved by using Metabolic-Informed Neural Network that utilizes omics data to predict metabolic fluxes under different conditions [126]. Additionally, emerging ML-driven strategies contribute to refining, structuring, optimizing, and integrating data, enhancing GEMs and MCMs via gap-filling support (e.g. AMMEDEUS and Boost-GAPFILL) and improved FBA feasibility spaces [127] (Fig. 2).

### (ii) ML for phenotype predictions

Phenotypic traits can also be inferred using ML with training sets based on the available high-quality genome and well-curated phenotypic (e.g. motility, spore formation, and salinity tolerance) and physicochemical environmental data (e.g. temperature, pH, and oxygen). These approaches can link phenotypic traits with protein families, also

reducing the sequence-to-function gap in prokaryotic genomes [40]. ML-driven strategies have been used to predict optimal growth temperatures [128] and phenotypic traits [116] and to functionally classify MAGs [129] (Fig. 2). Nonetheless, these ML-based models are trained with known microbial taxa; they depend on draft genomes and partial annotation of genes and often lack the inclusion of unconventional or new metabolic pathways, and main assumptions are based on fixed environmental conditions without biological/ecological context (e.g. microbial interactions). These factors could generate low-quality predictions and false-positive associations [116, 130]; however, ML-based models still constitute a powerful pathway for phenotype predictions.

### (iii) Integration of mechanistic models and ML-powered inferences

The above-described strategies stem from fundamentally distinct paradigms: mechanistic modelling (e.g. GEMs) seeks to elucidate biological processes through detailed physicochemical representations yet faces scalability challenges in complex microbial systems; conversely, ML-driven phenotypic inferences can predict outcomes in high dimensional contexts without requiring mechanistic insight, albeit at the cost of large, high-quality training datasets. In other words, whereas GEMs enable metabolic predictions under defined nutritional and/or genetic constraints, ML-based approaches can refine these predictions by capturing phenotypic responses to environmental factors not explicitly represented in GEMs.

Process-informed neural networks (PINNs) represent a family of hybrid frameworks that integrate prior mechanistic knowledge—expressed as differential equations or constraint-based models—into the architecture or training of neural networks. Unlike purely data-driven models, PINNs are trained not only to fit data but also to satisfy biophysical, biochemical, and ecological constraints such as mass balances, reaction stoichiometry, and thermodynamic feasibility [126, 131, 132]. These hybrid modelling frameworks can be used to predict optimal media compositions, pH, temperature, or other physicochemical parameters that maximize microbial growth. For example, closely related physics-informed variants have modelled population dynamics by embedding environmental variables directly in the loss function, achieving accurate predictions for monocultures and microbial communities [133]. Therefore, incorporating enzyme constrained metabolic models can capture proteome allocation limits and catalytic capacities, yielding more realistic predictions of substrate utilization hierarchies and aiding the rational selection of crucial nutrients for cultivation testing [134].

Taken together, combining mechanistic metabolic knowledge with ML-driven approaches within an iterative design–build–test–learn framework is poised to deliver reliable genotype-to-phenotype inferences from MAGs and to improve targeted culturomics. In practice, hybrid GEMs/MCMs, accounting for the influence of environmental factors, might predict cultivation conditions for target prokaryotic taxa, reducing the resources, time, and labour required for extensive testing. These hybrid modelling paradigms have the potential to provide mechanistic information that can be sensitive to extrinsic variables, making more accurate predictions [135, 136].

### Transforming genomic clues into an *ad hoc* isolation strategy

Reshaping microbiomes may provide improved access to hidden prokaryotic diversity, but a significant bottleneck remains because

numerous species continue to escape isolation in axenic culture. Traditional isolation methods may be unsuitable for some prokaryotic taxa if they fail to replicate the specific conditions of their natural habitats, including the presence of key growth factors [25]. With this goal, various culture media have been designed to re-create the nutritional milieu occurring in the ecosystem (e.g. soil-, seawater-, or host-extract agar) [137, 138]. Efforts in designing new solid or liquid synthetic culture media are vital to improve the isolation of target prokaryotic species [37, 139]. The nutrient conditions found in a hot spring have recently been employed as the basis for the design of semi-solid culture media to target the isolation of chemolithotrophic iron-oxidizing bacteria [140]. Although agar-agar has been used for ~150 years as a gelling agent for solid media, gellan gum has helped isolate a representative of the lineage WOR-3 from hydrothermal vents [141]. This factor and the bias associated with solid culture media must be considered in any targeted isolation strategy. Moreover, buffered culture media might restrict the growth of certain species types and certain buffers have long been known to inhibit bacterial growth; thus, adapting or even avoiding buffer solutions could be a critical choice for cultivating target taxa [142, 143].

In the future, AI-based models could help to conduct a proper and optimized synthetic culture media design [42, 130]. Although these approaches have not been extensively evaluated for the target cultivation of prokaryotic species, they could reduce the scope and rapidly screen thousands of predictive nutrient formulations, increasing the likelihood of a successful isolation [144, 145]. Recently, an AI-based tool was developed to learn the amino acid requirements for two oral streptococci [146]. In addition, rule-based classifiers and LLMs have been used to predict culture media preferences using microbial traits beyond genomic features or phylogeny [147]. These pioneering studies could accelerate the AI-guided design of culture media (Fig. 2). As mentioned, the metabolic data retrieved from MAG-derived genes, MAG-based metabolic models, and MMCMs can be applied (Fig. 1, see Steps 2 and 3; Fig. 2) to facilitate identifying unknown essential nutritional requirements, potential auxotrophies, cofactor dependencies, and environmental tolerances of specific microbes. This process can be fuelled by using complementary tools, such as DRAM [148], gapseq, and CarveMe [120], which perform functional annotation and GEM reconstruction. Such diverse data can guide the design of AI-assisted culture media and adjust cultivation parameters, such as carbon and nitrogen sources, vitamins, metals, pH, redox potential, temperature, or oxygen levels, which specifically cater to the physiological needs of target taxa [37, 41]. Active learning and Bayesian optimization frameworks can iteratively refine media composition using experimental growth feedback, accelerating the identification of optimized culture conditions [149, 150].

In this context, isolation of target prokaryotic species can be tested via a specific culturomic approach, where cultures can be analysed using Polymerase Chain Reaction (PCR) with specific probes [151] or using mass spectrometry [152] to detect the target prokaryotic taxa (Fig. 1, see Step 4). Recently, culturomic-based metagenomics has enriched specific prokaryotic species in human gut microbiomes [30] and has captured some of the unknown microbial diversity in desert soil [153]. Current culturomic strategies and growth media optimization efforts often rely on empirical methods. Thus, genome-guided culturomics transform a trial-and-error process into a hypothesis-driven one, improving the odds of isolating target species that have not yet been cultured. However, this *ad hoc* strategy must be supported by sophisticated technologies, heavy infrastructure, new methods,

complementary approaches (see Box 2), and special cultivation workflows, especially for anaerobic microorganisms [125, 154].

### Box 2. Complementary innovations to improve the isolation of prokaryotes

Identifying biosynthetic gaps or substrate dependencies from MAGs offers information on the requirements for cocultivation, conditioned media, or externally supplemented growth factors. Thus, culture media targeting specific microorganisms can be combined with other methods to avoid disrupting complex networks and metabolic interactions commonly occurring in the ecosystems. In cocultivation, syntrophic partners provide essential metabolites to enhance the growth of auxotrophic taxa [155]. For example, an approach called the sandwich agar plate was developed, aiming to coculture various populations throughout the isolation [156]. For liquid cultures, cocultivation technology is well established, using vessels in which partners are separated by dialysis membranes [157]. An anaerobic bioreactor-mediated enrichment allows the cocultivation of a pure syntrophic community, recovering a target archaeal taxon (*Promethearchaeum syntrophicum*; [158]). In the future, single-cell and CRISPR technology [159] could improve the isolation of specific target species from complex microbial communities. For example, taxon-specific probes derived from target MAGs can enable targeted cell sorting for cultivation and isolation in a selective culture media designed for its growth [160, 161]. Another interesting approach includes labelling cells, detection based on the physiological state or response to environmental changes, and isolation mediated by cell-sorting techniques [162]. In addition, AI-based models coupled with robotics can improve the detection of microcolonies (Fig. 2) that cannot be detected by the human eye, increasing the chance of isolation of the target species [163, 164].

### Concluding remarks

By integrating various approaches, methods, models, and tools, the proposed framework represents a multidisciplinary gateway to discovering and isolating novel or yet-uncultivated target prokaryotic taxa. In this context, perturbation and enrichment culture experiments offer powerful means to reshape microbial communities under controlled conditions, promoting the potential of discovering taxa that are otherwise difficult to detect in natural systems. Advancing these wet-laboratory strategies can permit access to previously elusive microbial lineages, enhance opportunities for cultivation, and broaden the knowledge of microbial diversity and ecosystem functioning.

Although all efforts to characterize genomes, genome-based models, and MMCMs seem complex, they can generate comprehensive and useful information regarding the metabolic and physiological profiles of target prokaryotic species, offering a foundation to refine culture media formulations and design innovative isolation strategies. However, due to the drawbacks and barriers in MAG reconstruction and gene function annotation, gaps in metabolic pathways, and limited representation in databases, this approach is still far from a systematic application. Experimentally annotating the vast number of uncharacterized genes is an expensive, time-consuming, and daunting task and a significant hurdle facing genome-driven cultivation. A better understanding of protein functions,

physiological features, metabolic pathways, and nutritional needs based on MAGs can provide the basis to isolate thousands of unknown prokaryotes. These obstacles are becoming increasingly surmountable with the convergence of advanced bioinformatics and emerging AI-based tools to assist in functional annotation [165] and synthetic media designs, defining a new era of genome-informed microbiology.

Emerging AI methods (e.g. AI-assisted gene annotation, AI-based predictions of physiological traits, AI-improved GEMs, and AI-guided media formulation) can generate vast data and new hypotheses to test the cultivation of prokaryotes. For example, an announced collaboration between the Align Foundation and American Type Culture Collection can train and validate AI-based models that predict microbial physiology from genetic information, helping to predict cultivation conditions for numerous species ([align-announces-new-collaboration-with-atcc-to-bring-ai-ready-microbial-datasets-to-life](#)).

Fostering synergy between microbiome scientists, computational and system biologists, and microbiologists would facilitate and accelerate the high-throughput isolation of novel or yet-uncultivated prokaryotic taxa, uncovering ecological roles and biotechnological potential. Finally, we pose provocative and forward-looking questions: (i) Are researchers sequencing a prokaryotic world that is increasingly challenging to cultivate, isolate, or fully understand? (ii) Without the availability of microbes as pure cultures, can researchers understand microbial life in depth or harness it for real-world applications? These questions are aimed at inspiring discussion, collaboration, and innovation in the scientific community.

### Author contributions

D.J.J.: Conceived the idea, framework design, figures, tables, and box development, and wrote the original draft and final version of the manuscript. R.M.: Conceptualization, framework design, participated in writing, discussion, and editing the manuscript. J.S.: Participated in writing, discussion, and editing the manuscript. C.A.D.R.: Participated in writing, discussion, and editing the manuscript. J.N.: Participated in writing, discussion, and editing the manuscript. L.M.R.: Participated in writing, discussion, and editing the manuscript. J.O.: Participated in writing, discussion, and editing the manuscript. A.S.R.: Funding acquisition, participated in writing, revision, and editing the manuscript.

### Conflicts of interest

None declared.

### Funding

We thank King Abdullah University of Science and Technology (KAUST) for the financial support under baseline BAS/1/1096-01-01 linked to A.S.R. The work by J.N. was supported by the European Union under grant agreement numbers 101081782 (deCYpher) and 101036768 (PROMISEANG) and project PID2022-139247OB-I00 (Rob3D), funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033) and the European Union through NextGenerationEU/PRTR 'A way to make Europe'.

### Data availability

No datasets were generated or analysed during the current study.

## References

- Hedlund B, Dodsworth J, Staley J. The changing landscape of microbial biodiversity exploration and its implications for systematics. *Syst Appl Microbiol* 2015;**38**:231–6. <https://doi.org/10.1016/j.syam.2015.03.003>
- Nayfach S, Roux S, Seshadri R *et al*. A genomic catalog of Earth's microbiomes. *Nat Biotechnol* 2021;**39**:499–509. <https://doi.org/10.1038/s41587-020-0718-6>
- Jiménez D, Rosado AS. SeqCode in the golden age of prokaryotic systematics. *ISME J* 2024;**18**:wrae109. <https://doi.org/10.1093/ismejo/wrae109>
- Wu D, Seshadri R, Kyrpides N *et al*. A metagenomic perspective on the microbial prokaryotic genome census. *Sci Adv* 2025;**11**:eadq2166. <https://doi.org/10.1126/sciadv.adq2166>
- Lloyd K, Steen A, Ladau J *et al*. Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 2018;**3**:e00055–18. <https://doi.org/10.1128/msystems.00055-18>
- Martiny AC. High proportions of bacteria are culturable across major biomes. *ISME J* 2019;**13**:2125–8. <https://doi.org/10.1038/s41396-019-0410-3>
- Steen A, Crits-Christoph A, Carini P *et al*. High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J* 2019;**13**:3126–30. <https://doi.org/10.1038/s41396-019-0484-y>
- Schloss P, Girard R, Martin T *et al*. Status of the archaeal and bacterial census: an update. *mBio* 2016;**7**:e00201–16. <https://doi.org/10.1128/mBio.00201-16>
- Louca S, Mazel F, Doebeli M *et al*. A census-based estimate of Earth's bacterial and archaeal diversity. *PLoS Biol* 2019;**17**:e3000106. <https://doi.org/10.1371/journal.pbio.3000106>
- Martiny A. The '1% culturability paradigm' needs to be carefully defined. *ISME J* 2020;**14**:10–1. <https://doi.org/10.1038/s41396-019-0507-8>
- Parks D, Chuvochina M, Rinke C *et al*. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;**50**:D785–94. <https://doi.org/10.1093/nar/gkab776>
- Venter S, Rodriguez-R L, Chuvochina M *et al*. Options and considerations for validation of prokaryotic names under the SeqCode. *Syst Appl Microbiol* 2024;**47**:126554. <https://doi.org/10.1016/j.syam.2024.126554>
- Curtis T, Sloan W, Scannell J. Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci* 2002;**99**:10494–9. <https://doi.org/10.1073/pnas.142680199>
- Locey K, Lennon J. Scaling laws predict global microbial diversity. *Proc Natl Acad Sci* 2016;**113**:5970–5. <https://doi.org/10.1073/pnas.1521291113>
- Lewis W, Tahon G, Geesink P *et al*. Innovations to culturing the uncultured microbial majority. *Nat Rev Microbiol* 2021;**19**:225–40. <https://doi.org/10.1038/s41579-020-00458-8>
- Gutleben J, Chaib De Mares M, van Elsas J *et al*. The multi-omics promise in context: from sequence to microbial isolate. *Crit Rev Microbiol* 2018;**44**:212–29. <https://doi.org/10.1080/1040841X.2017.1332003>
- Overmann J, Huang S, Nübel U *et al*. Relevance of phenotypic information for the taxonomy of not-yet-cultured microorganisms. *Syst Appl Microbiol* 2019;**42**:22–9. <https://doi.org/10.1016/j.syam.2018.08.009>
- Carini P. A “cultural” renaissance: genomics breathes new life into an old craft. *mSystems* 2019;**4**:e00092–19. <https://doi.org/10.1128/mSystems.00092-19>
- Simon S, Aschmann V, Behrendt A *et al*. Earth's most needed uncultivated aquatic prokaryotes. *Water Res* 2025;**273**:122928. <https://doi.org/10.1016/j.watres.2024.122928>
- Zhao R, Farag I, Jørgensen S *et al*. Occurrence, diversity, and genomes of “*Candidatus* Patescibacteria” along the early diagenesis of marine sediments. *Appl Environ Microbiol* 2022;**88**:e0140922–2. <https://doi.org/10.1128/aem.01409-22>
- Liu L, Lian Z, Lv A *et al*. Insights into chemoautotrophic traits of a prevalent bacterial phylum CSP1-3, herein *Sysuimicrobiota*. *Natl Sci Rev* 2024;**11**:nwae378. <https://doi.org/10.1093/nsr/nwae378>
- Cavicchioli R, Ripple W, Timmis K *et al*. Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Microbiol* 2019;**17**:569–86. <https://doi.org/10.1038/s41579-019-0222-5>
- Dominguez-Bello M, Steiger D, Fankhauser M *et al*. The microbiota vault initiative: safeguarding Earth's microbial heritage for future generations. *Nat Commun* 2025;**16**:5373. <https://doi.org/10.1038/s41467-025-61008-5>
- Gilbert J, Peixoto R, Scholz A *et al*. Launching the IUCN microbial conservation specialist group as a global safeguard for microbial biodiversity. *Nat Microbiol* 2025;**10**:2359–60. <https://doi.org/10.1038/s41564-025-02113-5>
- Schultz J, Modolon F, Peixoto R *et al*. Shedding light on the composition of extreme microbial dark matter: alternative approaches for culturing extremophiles. *Front Microbiol* 2023;**14**:1167718. <https://doi.org/10.3389/fmicb.2023.1167718>
- Hessler T, Huddy R, Sachdeva R *et al*. Vitamin interdependencies predicted by metagenomics-informed network analyses and validated in microbial community microcosms. *Nat Commun* 2023;**14**:4768. <https://doi.org/10.1038/s41467-023-40360-4>
- Laso-Pérez R, Wegener G, Knittel K *et al*. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature* 2016;**539**:396–401. <https://doi.org/10.1038/nature20152>
- Lagier J, Dubourg G, Million M *et al*. Culturing the human microbiota and culturomics. *Nat Rev Microbiol* 2018;**16**:540–50. <https://doi.org/10.1038/s41579-018-0041-0>
- Yang Z, Lian Z, Liu L *et al*. Cultivation strategies for prokaryotes from extreme environments. *iMeta* 2023;**2**:e123. <https://doi.org/10.1002/imt2.123>
- Armetta J, Li S, Vaaben T *et al*. Metagenome-guided culturomics for the targeted enrichment of gut microbes. *Nat Commun* 2025;**16**:663. <https://doi.org/10.1038/s41467-024-55668-y>
- Overmann J, Abt B, Sikorski J. Present and future of culturing bacteria. *Ann Rev Microbiol* 2017;**71**:711–30. <https://doi.org/10.1146/annurev-micro-090816-093449>
- Sefrji F, Marasco R, Michoud G *et al*. *Kaustia mangrovi* gen. nov., sp. nov. isolated from Red Sea mangrove sediments belongs to the recently proposed Parvibaculaceae family within the order Rhizobiales. *Int J Syst Evol Microbiol* 2021;**71**:004806. <https://doi.org/10.1099/ijsem.0.004806>
- Sefrji F, Michoud G, Marasco R *et al*. *Mangrovivirga cuniculi* gen. nov., sp. nov., a moderately halophilic bacterium isolated from bioturbated Red Sea mangrove sediment, and proposal of the novel family Mangrovivirgaceae fam. nov. *Int J Syst Evol Microbiol* 2021;**71**:004866. <https://doi.org/10.1099/ijsem.0.004866>
- Parte A, Sardà Carbasse J, Meier-Kolthoff J *et al*. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020;**70**:5607–12. <https://doi.org/10.1099/ijsem.0.004332>
- Cross K, Campbell J, Balachandran M *et al*. Targeted isolation and cultivation of uncultivated bacteria by reverse genomics.

- Nat Biotechnol* 2019;**37**:1314–21. <https://doi.org/10.1038/s41587-019-0260-6>
36. Liu S, Moon C, Zheng N *et al*. Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome* 2022;**10**:76. <https://doi.org/10.1186/s40168-022-01272-5>
  37. Karnachuk O, Lukina A, Kadnikov V *et al*. Targeted isolation based on metagenome-assembled genomes reveals a phylogenetically distinct group of thermophilic spirochetes from deep biosphere. *Environ Microbiol* 2021;**23**:3585–98. <https://doi.org/10.1111/1462-2920.15218>
  38. Nou N, Covington J, Lai D *et al*. Genome-guided isolation of the hyperthermophilic aerobe *Fervidibacter sacchari* reveals conserved polysaccharide metabolism in the *Armatimonadota*. *Nat Commun* 2024;**15**:9534. <https://doi.org/10.1038/s41467-024-53784-3>
  39. Radlinski L, Bäuml A. Microbiome science needs more microbiologists. *Nat Microbiol* 2025;**10**:263–4. <https://doi.org/10.1038/s41564-024-01922-4>
  40. Koblit J, Reimer L, Pukall R *et al*. Predicting bacterial phenotypic traits through improved machine learning using high-quality, curated datasets. *Commun Biol* 2025;**8**:897. <https://doi.org/10.1038/s42003-025-08313-3>
  41. Tejera N, Crossman L, Pearson B *et al*. Genome-scale metabolic model driven design of a defined medium for *Campylobacter jejuni* M1cam. *Front Microbiol* 2020;**11**:1072. <https://doi.org/10.3389/fmicb.2020.01072>
  42. Sipkema D. Improving the odds: artificial intelligence and the great plate count anomaly. *Microb Biotechnol* 2024;**17**:e70004. <https://doi.org/10.1111/1751-7915.70004>
  43. Tyson G, Lo I, Baker B *et al*. Genome-directed isolation of the key nitrogen fixer *Leptospirillum ferrodiazotrophum* sp. nov. from an acidophilic microbial community. *Appl Environ Microbiol* 2005;**71**:6319–24. <https://doi.org/10.1128/AEM.71.10.6319-6324.2005>
  44. Pope P, Smith W, Denman S *et al*. Isolation of *Succinivibrionaceae* implicated in low methane emissions from Tamar wallabies. *Science* 2011;**333**:646–8. <https://doi.org/10.1126/science.1205760>
  45. Potts L, Douglas A, Perez CL *et al*. Chronic environmental perturbation influences microbial community assembly patterns. *Environ Sci Technol* 2022;**56**:2300–11. <https://doi.org/10.1021/acs.est.1c05106>
  46. Pedrós-Alió C. The rare bacterial biosphere. *Annu Rev Mar Sci* 2012;**4**:449–66. <https://doi.org/10.1146/annurev-marine-120710-100948>
  47. Pascoal F, Branco P, Torgo L *et al*. Definition of the microbial rare biosphere through unsupervised machine learning. *Commun Biol* 2025;**8**:544. <https://doi.org/10.1038/s42003-025-07912-4>
  48. McDonald M, Owusu-Ansah C, Ellenbogen J *et al*. What is microbial dormancy? *Trends Microbiol* 2024;**32**:142–50. <https://doi.org/10.1016/j.tim.2023.08.006>
  49. Lindner B, Gerhardt K, Feistel D *et al*. A user's guide to the bioinformatic analysis of shotgun metagenomic sequence data for bacterial pathogen detection. *Int J Food Microbiol* 2024;**410**:110488. <https://doi.org/10.1016/j.ijfoodmicro.2023.110488>
  50. Delmont T, Eren A, Maccario L *et al*. Reconstructing rare soil microbial genomes using in situ enrichments and metagenomics. *Front Microbiol* 2015;**6**:358. <https://doi.org/10.3389/fmicb.2015.00358>
  51. Jia Z, Lipus D, Burckhardt O *et al*. Enrichment of rare methanogenic archaea shows their important ecological role in natural high-CO<sub>2</sub> terrestrial subsurface environments. *Front Microbiol* 2023;**14**:1105259. <https://doi.org/10.3389/fmicb.2023.1105259>
  52. Heritier-Robbins P, Karthikeyan S, Hatt J *et al*. Beach sand oil spills select for generalist microbial populations. *ISME J* 2021;**15**:3418–22. <https://doi.org/10.1038/s41396-021-01017-6>
  53. Cordero I, Leizeaga A, Hicks L *et al*. High intensity perturbations induce an abrupt shift in soil microbial state. *ISME J* 2023;**17**:2190–9. <https://doi.org/10.1038/s41396-023-01512-y>
  54. Knight C, Nicolitch O, Griffiths R *et al*. Soil microbiomes show consistent and predictable responses to extreme events. *Nature* 2024;**636**:690–6. <https://doi.org/10.1038/s41586-024-08185-3>
  55. Ossowicki A, Raaijmakers J, Garbeva P. Disentangling soil microbiome functions by perturbation. *Environ Microbiol Rep* 2021;**13**:582–90. <https://doi.org/10.1111/1758-2229.12989>
  56. Liu X, Salles J. Drivers and consequences of microbial community coalescence. *ISME J* 2024;**18**:wrae179. <https://doi.org/10.1093/ismejo/wrae179>
  57. Jurburg S, Nunes I, Brejnrod A *et al*. Legacy effects on the recovery of soil bacterial communities from extreme temperature perturbation. *Front Microbiol* 2017;**8**:1832. <https://doi.org/10.3389/fmicb.2017.01832>
  58. Jiménez D, Chaparro D, Sierra F *et al*. Engineering the mangrove soil microbiome for selection of polyethylene terephthalate-transforming bacterial consortia. *Trends Biotechnol* 2025;**43**:162–83. <https://doi.org/10.1016/j.tibtech.2024.08.013>
  59. Lee K, Liu S, Crocker K *et al*. Functional regimes define soil microbiome response to environmental change. *Nature* 2025;**644**:1028–38. <https://doi.org/10.1038/s41586-025-09264-9>
  60. Kurm V, Geisen S, Gera HW. A low proportion of rare bacterial taxa responds to abiotic changes compared with dominant taxa. *Environ Microbiol* 2019;**21**:750–8. <https://doi.org/10.1111/1462-2920.14492>
  61. Liang Y, Xiao X, Nuccio E *et al*. Differentiation strategies of soil rare and abundant microbial taxa in response to changing climatic regimes. *Environ Microbiol* 2020;**22**:1327–40. <https://doi.org/10.1111/1462-2920.14945>
  62. Richter-Heitmann T, Hofner B, Krahe F *et al*. Stochastic dispersal rather than deterministic selection explains the spatio-temporal distribution of soil bacteria in a temperate grassland. *Front Microbiol* 2020;**11**:1391. <https://doi.org/10.3389/fmicb.2020.01391>
  63. Du S, Dini-Andreote F, Zhang N *et al*. Divergent co-occurrence patterns and assembly processes structure the abundant and rare bacterial communities in a salt marsh ecosystem. *Appl Environ Microbiol* 2020;**86**:e00322–0. <https://doi.org/10.1128/AEM.00322-20>
  64. Díaz-García L, Huang S, Spröer C *et al*. Dilution-to-stimulation/extinction method: a combination enrichment strategy to develop a minimal and versatile lignocellulolytic bacterial consortium. *Appl Environ Microbiol* 2021;**87**:e02427–0. <https://doi.org/10.1128/AEM.02427-20>
  65. Jiménez DJ, Sanchez A, Dini-Andreote F. Engineering microbiomes to transform plastics. *Trends Biotechnol* 2024;**42**:265–8. <https://doi.org/10.1016/j.tibtech.2023.09.011>
  66. Hu H, Wang M, Huang Y *et al*. Guided by the principles of microbiome engineering: accomplishments and perspectives for environmental use. *mLife* 2022;**1**:382–98. <https://doi.org/10.1002/mlf2.12043>
  67. Jiménez D, Dini-Andreote F, DeAngelis K *et al*. Ecological insights into the dynamics of plant biomass-degrading microbial consortia. *Trends Microbiol* 2017;**25**:788–96. <https://doi.org/10.1016/j.tim.2017.05.012>
  68. Leadbetter J, Schmidt T, Graber J *et al*. Acetogenesis from H<sub>2</sub> plus CO<sub>2</sub> by spirochetes from termite guts. *Science* 1999;**283**:686–9. <https://doi.org/10.1126/science.283.5402.686>

69. Methner A, Kuzyk S, Petersen J *et al.* *Thiorhodovibrio frisius* and *Trv. Litoralis* spp. nov., two novel members from a clade of fastidious purple sulfur bacteria that exhibit unique red-shifted light-harvesting capabilities. *Microorganisms* 2023;**11**:2394. <https://doi.org/10.3390/microorganisms11102394>
70. Meyer-Cifuentes I, Werner J, Jehmlich N *et al.* Synergistic biodegradation of aromatic-aliphatic copolyester plastic by a marine microbial consortium. *Nat Commun* 2020;**11**:5790. <https://doi.org/10.1038/s41467-020-19583-2>
71. Díaz Rodríguez C, Díaz-García L, Bunk B *et al.* Novel bacterial taxa in a minimal lignocellulolytic consortium and their potential for lignin and plastics transformation. *ISME Commun* 2022;**2**:89. <https://doi.org/10.1038/s43705-022-00176-7>
72. Díaz-García L, Chuvochina M, Feuerriegel G *et al.* Andean soil-derived lignocellulolytic bacterial consortium as a source of novel taxa and putative plastic-active enzymes. *Syst Appl Microbiol* 2024;**47**:126485. <https://doi.org/10.1016/j.syapm.2023.126485>
73. Yoshida S, Hiraga K, Takehana T *et al.* A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* 2016;**351**:1196–9. <https://doi.org/10.1126/science.aad6359>
74. Albertsen M. Long-read metagenomics paves the way toward a complete microbial tree of life. *Nat Methods* 2023;**20**:30–1. <https://doi.org/10.1038/s41592-022-01726-6>
75. Sereika M, Mussig A, Jiang C *et al.* Genome-resolved long-read sequencing expands known microbial diversity across terrestrial habitats. *Nat Microbiol* 2025;**10**:2018–30. <https://doi.org/10.1038/s41564-025-02062-z>
76. Orellana L, Krüger K, Sidhu C *et al.* Comparing genomes recovered from time-series metagenomes using long- and short-read sequencing technologies. *Microbiome* 2023;**11**:105. <https://doi.org/10.1186/s40168-023-01557-3>
77. Bowers R, Kyrpides N, Stepanauskas R *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017;**35**:725–31. <https://doi.org/10.1038/nbt.3893>
78. Kim N, Ma J, Kim W *et al.* Genome-resolved metagenomics: a game changer for microbiome medicine. *Exp Mol Med* 2024;**56**:1501–12. <https://doi.org/10.1038/s12276-024-01262-7>
79. Sharma S, Narahari H, Raman K. Harnessing machine learning for metagenomic data analysis: trends and applications. *mSystems* 2025;**10**:e0164224. <https://doi.org/10.1128/msystems.01642-24>
80. Rodríguez-R L, Conrad R, Viver T *et al.* An ANI gap within bacterial species that advances the definitions of intra-species units. *mBio* 2024;**15**:e02696–23. <https://doi.org/10.1128/mbio.02696-23>
81. Zachariassen T, Russel J, Petersen C *et al.* MAGinator enables accurate profiling of de novo MAGs with strain-level phylogenies. *Nat Commun* 2024;**15**:5734. <https://doi.org/10.1038/s41467-024-49958-8>
82. Anthony W, Allison S, Broderick C *et al.* From soil to sequence: filling the critical gap in genome-resolved metagenomics is essential to the future of soil microbial ecology. *Environ Microbiome* 2024;**19**:56. <https://doi.org/10.1186/s40793-024-00599-w>
83. Rodríguez del Río Á, Scheu S, Rillig MC. Soil microbial responses to multiple global change factors as assessed by metagenomics. *Nat Commun* 2025;**16**:5058. <https://doi.org/10.1038/s41467-025-60390-4>
84. Sun Y, Cheng Z, Li X *et al.* Genome enrichment of rare and unknown species from complicated microbiomes by nanopore selective sequencing. *Genome Res* 2023;**33**:612–21. <https://doi.org/10.1101/gr.277266.122>
85. Riesco R, Trujillo M. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2024;**74**:006300. <https://doi.org/10.1099/ijsem.0.006300>
86. Hedlund B, Chuvochina M, Hugenholtz P *et al.* SeqCode: a nomenclatural code for prokaryotes described from sequence data. *Nat Microbiol* 2022;**7**:1702–8. <https://doi.org/10.1038/s41564-022-01214-9>
87. Oberhardt M, Zarecki R, Gronow S *et al.* Harnessing the landscape of microbial culture media to predict new organism–media pairings. *Nat Commun* 2015;**6**:8493. <https://doi.org/10.1038/ncomms9493>
88. Schober I, Koblitz J, Sardà CJ *et al.* Bac dive in 2025: the core database for prokaryotic strain data. *Nucleic Acids Res* 2025;**53**:D748–56. <https://doi.org/10.1093/nar/gkae959>
89. Koblitz J, Halama P, Spring S *et al.* Media dive: the expert-curated cultivation media database. *Nucleic Acids Res* 2023;**51**:D1531–8. <https://doi.org/10.1093/nar/gkac803>
90. Jaspers E, Overmann J. Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysologies. *Appl Environ Microbiol* 2004;**70**:4831–9. <https://doi.org/10.1128/AEM.70.8.4831-4839.2004>
91. Martiny A, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 2013;**7**:830–8. <https://doi.org/10.1038/ismej.2012.160>
92. Viver T, Conrad R, Rodríguez-R L *et al.* Towards estimating the number of strains that make up a natural bacterial population. *Nat Commun* 2024;**15**:544. <https://doi.org/10.1038/s41467-023-44622-z>
93. Tong X, Luo D, Leung M *et al.* Diverse and specialized metabolic capabilities of microbes in oligotrophic built environments. *Microbiome* 2024;**12**:198. <https://doi.org/10.1186/s40168-024-01926-6>
94. Pinto Y, Bhatt A. Sequencing-based analysis of microbiomes. *Nat Rev Genet* 2024;**25**:829–45. <https://doi.org/10.1038/s41576-024-00746-6>
95. Jha N, Kravitz J, West-Roberts J *et al.* Gaia: an AI-enabled genomic context-aware platform for protein sequence annotation. *Sci Adv* 2025;**11**:eadv5109. <https://doi.org/10.1126/sciadv.adv5109>
96. Rodríguez del Río Á, Giner-Lamia J, Cantalapiedra C *et al.* Functional and evolutionary significance of unknown genes from uncultivated taxa. *Nature* 2024;**626**:377–84. <https://doi.org/10.1038/s41586-023-06955-z>
97. Zhang Y, Bhosle A, Bae S *et al.* Predicting functions of uncharacterized gene products from microbial communities. *Nat Biotechnol* 2025. <https://doi.org/10.1038/s41587-025-02813-7>
98. Coelho L, Alves R, del Río Á *et al.* Towards the biogeography of prokaryotic genes. *Nature* 2022;**601**:252–6. <https://doi.org/10.1038/s41586-021-04233-4>
99. Karaoz U, Brodie EL. microTrait: a toolset for a trait-based representation of microbial genomes. *Front Bioinform* 2022;**2**:918853. <https://doi.org/10.3389/fbinf.2022.918853>
100. Weimann A, Mooren K, Frank J *et al.* From genomes to phenotypes: TraitR, the microbial trait analyzer. *mSystems* 2016;**1**:e00101–16. <https://doi.org/10.1128/mSystems.00101-16>
101. Xu L, Zakem E, Weissman J. Improved maximum growth rate prediction from microbial genomes by integrating phylogenetic information. *Nat Commun* 2025;**16**:4256. <https://doi.org/10.1038/s41467-025-59558-9>
102. Sauer D, Wang D. Predicting the optimal growth temperatures of prokaryotes using only genome derived features. *Bioinformatics* 2019;**35**:3224–31. <https://doi.org/10.1093/bioinformatics/btz059>

103. Price M, Deutschbauer A, Arkin A. GapMind: automated annotation of amino acid biosynthesis. *mSystems* 2020;**5**:e00291–20. <https://doi.org/10.1128/msystems.00291-20>
104. Karthikeyan S, Rodriguez-R L, Heritier-Robbins P *et al*. “*Candidatus Macondimonas diazotrophica*”, a novel gammaproteobacterial genus dominating crude-oil-contaminated coastal sediments. *ISME J* 2019;**13**:2129–34. <https://doi.org/10.1038/s41396-019-0400-5>
105. Li J, Luo C, Cai X *et al*. Cultivation and characterization of functional-yet-uncultivable phenanthrene degraders by stable-isotope-probing and metagenomic-binning directed cultivation (SIP-MDC). *Environ Int* 2024;**185**:108555. <https://doi.org/10.1016/j.envint.2024.108555>
106. Viacava K, Qiao J, Janowczyk A *et al*. Meta-omics-aided isolation of an elusive anaerobic arsenic-methylating soil bacterium. *ISME J* 2022;**16**:1740–9. <https://doi.org/10.1038/s41396-022-01220-z>
107. Oduwole I, Babjac A, Royalty T *et al*. Functional genomic signatures predict microbial culturability across the tree of life. *bioRxiv* 2025. <https://doi.org/10.1101/2025.08.18.670795>
108. Barnum T, Crits-Christoph A, Molla M *et al*. Predicting microbial growth conditions from amino acid composition. *bioRxiv* 2024. <https://doi.org/10.1101/2024.03.22.586313>
109. Maranga M, Szczerbiak P, Bezshapkin V *et al*. Comprehensive functional annotation of metagenomes and microbial genomes using a deep learning-based method. *mSystems* 2023;**8**:e01178–22. <https://doi.org/10.1128/msystems.01178-22>
110. Rappsilber J. A dive into the unknown. *Trends Genet* 2024;**40**:15–6. <https://doi.org/10.1016/j.tig.2023.10.011>
111. Thomas A, Segata N. Multiple levels of the unknown in microbiome research. *BMC Biol* 2019;**17**:48. <https://doi.org/10.1186/s12915-019-0667-z>
112. Vanni C, Schechter M, Acinas S *et al*. Unifying the known and unknown microbial coding sequence space. *Elife* 2022;**11**:e67667. <https://doi.org/10.7554/eLife.67667>
113. Jaffe M, Dziulko A, Smith J *et al*. Improved discovery of genetic interactions using CRISPRiSeq across multiple environments. *Genome Res* 2019;**29**:668–81. <https://doi.org/10.1101/gr.246603.118>
114. Pavlopoulos G, Baltoumas F, Liu S *et al*. Unraveling the functional dark matter through global metagenomics. *Nature* 2023;**622**:594–602. <https://doi.org/10.1038/s41586-023-06583-7>
115. Xiang W, Xiong Z, Chen H *et al*. FAPM: functional annotation of proteins using multimodal models beyond structural modeling. *Bioinformatics* 2024;**40**:btae680. <https://doi.org/10.1093/bioinformatics/btae680>
116. Vitale R, Bugnon L, Fenoy E *et al*. Evaluating large language models for annotating proteins. *Brief Bioinform* 2024;**25**:bbae177. <https://doi.org/10.1093/bib/bbae177>
117. Bautista E, Zinski J, Szczepanek S *et al*. Semi-automated curation of metabolic models via flux balance analysis: a case study with *Mycoplasma gallisepticum*. *PLoS Comput Biol* 2013;**9**:e1003208. <https://doi.org/10.1371/journal.pcbi.1003208>
118. Price M, Wetmore K, Waters R *et al*. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* 2018;**557**:503–9. <https://doi.org/10.1038/s41586-018-0124-0>
119. Mol V, Bennett M, Sánchez B *et al*. Genome-scale metabolic modeling of *P. thermoglucosidasius* NCIMB 11955 reveals metabolic bottlenecks in anaerobic metabolism. *Metab Eng* 2021;**65**:123–34. <https://doi.org/10.1016/j.ymben.2021.03.002>
120. Tarzi C, Zampieri G, Sullivan N *et al*. Emerging methods for genome-scale metabolic modeling of microbial communities. *Trends Endocrinol Metab* 2024;**35**:533–48. <https://doi.org/10.1016/j.tem.2024.02.018>
121. Quinn-Bohmann N, Carr A, Diener C *et al*. Moving from genome-scale to community-scale metabolic models for the human gut microbiome. *Nat Microbiol* 2025;**10**:1055–66. <https://doi.org/10.1038/s41564-025-01972-2>
122. Kuppa Baskaran D, Umale S, Zhou Z *et al*. Metagenome-based metabolic modelling predicts unique microbial interactions in deep-sea hydrothermal plume microbiomes. *ISME Commun* 2023;**3**:42. <https://doi.org/10.1038/s43705-023-00242-8>
123. Peng X, Wang S, Wang M *et al*. Metabolic interdependencies in thermophilic communities are revealed using co-occurrence and complementarity networks. *Nat Commun* 2024;**15**:8166. <https://doi.org/10.1038/s41467-024-52532-x>
124. Zorrilla F, Buric F, Patil K *et al*. metaGEM: reconstruction of genome scale metabolic models directly from metagenomes. *Nucleic Acids Res* 2021;**49**:e126. <https://doi.org/10.1093/nar/gkab815>
125. Clavel T, Faber F, Groussin M *et al*. Enabling next-generation anaerobic cultivation through biotechnology to advance functional microbiome research. *Nat Biotechnol* 2025;**43**:878–88. <https://doi.org/10.1038/s41587-025-02660-6>
126. Tazza G, Moro F, Ruggeri D *et al*. MINN: a metabolic-informed neural network for integrating omics data into genome-scale metabolic modeling. *Comput Struct Biotechnol J* 2025;**27**:3609–17. <https://doi.org/10.1016/j.csbj.2025.08.004>
127. Kundu P, Beura S, Mondal S *et al*. Machine learning for the advancement of genome-scale metabolic modeling. *Biotechnol Adv* 2024;**74**:108400. <https://doi.org/10.1016/j.biotechadv.2024.108400>
128. Liu H, Zhu G, Chen L *et al*. Machine learning prediction of bacterial optimal growth temperature from protein domain signatures reveals thermoadaptation mechanisms. *BMC Genomics* 2025;**26**:957. <https://doi.org/10.1186/s12864-025-12162-z>
129. Bizzotto E, Fraulini S, Zampieri G *et al*. MICROPHERRET: MICROBIAL PHENotypic tRait Classifier using machine lEarning techniques. *Environ Microbiome* 2024;**19**:58. <https://doi.org/10.1186/s40793-024-00600-6>
130. James T, Williamson B, Tino P *et al*. Whole-genome phenotype prediction with machine learning: open problems in bacterial genomics. *Bioinformatics* 2025;**41**:btaf206. <https://doi.org/10.1093/bioinformatics/btaf206>
131. Lagergren J, Nardini J, Baker R *et al*. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLoS Comput Biol* 2020;**16**:e1008462. <https://doi.org/10.1371/journal.pcbi.1008462>
132. Wesselkamp M, Moser N, Kalweit M *et al*. Process-informed neural networks: a hybrid modelling approach to improve predictive performance and inference of neural networks in ecology and beyond. *Ecol Lett* 2024;**27**:e70012. <https://doi.org/10.1111/ele.70012>
133. Cuomo S, De Rosa M, Piccialli F *et al*. A numerical approach for soil microbiota growth prediction through physics-informed neural network. *Appl Numer Math* 2025;**207**:97–110. <https://doi.org/10.1016/j.apnum.2024.08.025>
134. Chen Y, Gustafsson J, Tafur RA *et al*. Reconstruction, simulation and analysis of enzyme-constrained metabolic models using GECKO toolbox 3.0. *Nat Protoc* 2024;**19**:629–67. <https://doi.org/10.1038/s41596-023-00931-7>
135. Faure L, Mollet B, Liebermeister W *et al*. A neural-mechanistic hybrid approach improving the predictive power of genome-scale metabolic models. *Nat Commun* 2023;**14**:4669. <https://doi.org/10.1038/s41467-023-40380-0>
136. Morrissey J, Barberi G, Strain B *et al*. NEXT-FBA: a hybrid stoichiometric/data-driven approach to improve intracellular flux

- predictions. *Metab Eng* 2025;**91**:130–44. <https://doi.org/10.1016/j.ymben.2025.03.010>
137. Schultz J, Modolon F, Rosado A *et al.* Methods and strategies to uncover coral-associated microbial dark matter. *mSystems* 2022;**7**:e0036722–2. <https://doi.org/10.1128/msystems.00367-22>
  138. Sarhan M, Mourad E, Nemr R *et al.* An inoculum-dependent culturing strategy (IDC) for the cultivation of environmental microbiomes and the isolation of novel endophytic Actinobacteria. *J Antibiot (Tokyo)* 2020;**73**:66–71. <https://doi.org/10.1038/s41429-019-0226-4>
  139. Tao X, Huang W, Pan L *et al.* Optimizing ex vivo culture conditions to study human gut microbiome. *ISME Commun* 2023;**3**:38. <https://doi.org/10.1038/s43705-023-00245-5>
  140. Uchijima T, Kato S, Tanimoto K *et al.* Custom-made medium approach for effective enrichment and isolation of chemolithotrophic iron-oxidizing bacteria. *FEMS Microbiol Ecol* 2025;**101**:fiaf051. <https://doi.org/10.1093/femsec/fiaf051>
  141. Mori K, Hidaka K, Tamazawa S *et al.* Isolation and characterization of a bacterium affiliated with the hitherto uncultured candidate phylum WOR-3 from a deep-sea hydrothermal fluid. *Appl Environ Microbiol* 2025;**91**:e0018825. <https://doi.org/10.1128/aem.00188-25>
  142. Prakash O, Waghmare U, Chauhan A *et al.* Optimizing experimental conditions: the role of buffered environments in microbial isolation, physiological studies, and taxonomic characterization. *Appl Environ Microbiol* 2025;**91**:e0172824. <https://doi.org/10.1128/aem.01728-24>
  143. Bartscht K, Cypionka H, Overmann J. Evaluation of cell activity and of methods for the cultivation of bacteria from a natural lake community. *FEMS Microbiol Ecol* 1999;**28**:249–59. [https://doi.org/10.1016/S0168-6496\(98\)00112-3](https://doi.org/10.1016/S0168-6496(98)00112-3)
  144. Hashizume T, Ozawa Y, Ying B. Employing active learning in the optimization of culture medium for mammalian cells. *NPJ Syst Biol Appl* 2023;**9**:20. <https://doi.org/10.1038/s41540-023-00284-7>
  145. Nikkha A, Rohani A, Zarei M *et al.* Toward sustainable culture media: using artificial intelligence to optimize reduced-serum formulations for cultivated meat. *Sci Total Environ* 2023;**894**:164988. <https://doi.org/10.1016/j.scitotenv.2023.164988>
  146. Dama A, Kim K, Leyva D *et al.* BacterAI maps microbial metabolism without prior knowledge. *Nat Microbiol* 2023;**8**:1018–25. <https://doi.org/10.1038/s41564-023-01376-0>
  147. Máša P, Kliegr T, Joachimiak MP. Explainable rule-based prediction of cultivation media for microbes. *Comput Struct Biotechnol J* 2025;**27**:5194–206. <https://doi.org/10.1016/j.csbj.2025.10.014>
  148. Shaffer M, Borton M, McGivern B *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* 2020;**48**:8883–900. <https://doi.org/10.1093/nar/gkaa621>
  149. Zhang S, Aida H, Ying B. Employing active learning in medium optimization for selective bacterial growth. *Appl Microbiol* 2023;**3**:1355–69. <https://doi.org/10.3390/applmicrobiol3040091>
  150. Lapiere F, Mattaliano P, Raith D *et al.* Multi-cycle high-throughput growth media optimization using batch Bayesian optimization. *J Chem Technol Biotechnol* 2025;**100**:1571–83. <https://doi.org/10.1002/jctb.7860>
  151. Kothe C, Renault P. Metagenomic driven isolation of poorly culturable species in food. *Food Microbiol* 2025;**129**:104722. <https://doi.org/10.1016/j.fm.2025.104722>
  152. Wang F, Yu L, Ren Y *et al.* An optimized culturomics strategy for isolation of human milk microbiota. *Front Microbiol* 2024;**15**:1272062. <https://doi.org/10.3389/fmicb.2024.1272062>
  153. Li S, Lian W, Han J *et al.* Capturing the microbial dark matter in desert soils using culturomics-based metagenomics and high-resolution analysis. *NPJ Biofilms Microbiomes* 2023;**9**:67. <https://doi.org/10.1038/s41522-023-00439-8>
  154. Sallet H, Calvo M, Titus M *et al.* High-throughput cultivation and isolation of environmental anaerobes using selectively permeable hydrogel capsules. *ISME Commun* 2025;**5**:ycaf117. <https://doi.org/10.1093/ismeco/ycaf117>
  155. Marmann A, Aly A, Lin W *et al.* Co-cultivation—a powerful emerging tool for enhancing the chemical diversity of microorganisms. *Mar Drugs* 2014;**12**:1043–65. <https://doi.org/10.3390/md12021043>
  156. Zhang J, Liang Q, Mu D *et al.* Cultivating the uncultured: harnessing the “sandwich agar plate” approach to isolate heme-dependent bacteria from marine sediment. *mLife* 2024;**3**:143–55. <https://doi.org/10.1002/mlf2.12093>
  157. Moutinho T, Panagides J, Biggs M *et al.* Novel co-culture plate enables growth dynamic-based assessment of contact-independent microbial interactions. *PLoS One* 2017;**12**:e0182163. <https://doi.org/10.1371/journal.pone.0182163>
  158. Imachi H, Nobu M, Miyazaki M *et al.* Cultivation of previously uncultured microorganisms with a continuous-flow down-flow hanging sponge (DHS) bioreactor, using a syntrophic archaeon culture obtained from deep marine sediment as a case study. *Nat Protoc* 2022;**17**:2784–814. <https://doi.org/10.1038/s41596-022-00735-1>
  159. Rottinghaus A, Vo S, Moon T. Computational design of CRISPR guide RNAs to enable strain-specific control of microbial consortia. *Proc Natl Acad Sci USA* 2023;**120**:e2213154120. <https://doi.org/10.1073/pnas.2213154120>
  160. Ma L, Kim J, Hatzenpichler R *et al.* Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project’s Most Wanted taxa. *Proc Natl Acad Sci USA* 2014;**111**:9768–73. <https://doi.org/10.1073/pnas.1404753111>
  161. Bellais S, Nehlich M, Ania M *et al.* Species-targeted sorting and cultivation of commensal bacteria from the gut microbiome using flow cytometry under anaerobic conditions. *Microbiome* 2022;**10**:24. <https://doi.org/10.1186/s40168-021-01206-7>
  162. Mulay S, Vishnivetskaya T, Hochanadel L *et al.* BONCAT-live for isolation and cultivation of active environmental bacteria. *mBio* 2025;**16**:e02389–25. <https://doi.org/10.1128/mbio.02389-25>
  163. Huang Y, Sheth R, Zhao S *et al.* High-throughput microbial culturomics using automation and machine learning. *Nat Biotechnol* 2023;**41**:1424–33. <https://doi.org/10.1038/s41587-023-01674-2>
  164. Jacot D, Gizha S, Orny C *et al.* Development and evaluation of an artificial intelligence for bacterial growth monitoring in clinical bacteriology. *J Clin Microbiol* 2024;**62**:e0165123–3. <https://doi.org/10.1128/jcm.01651-23>
  165. Pálsson B, Lee S, Kim G. Approaches for accelerating microbial gene function discovery using artificial intelligence. *Nat Microbiol* 2026;**11**:350–58. <https://doi.org/10.1038/s41564-025-02214-1>