Contribution of single-cell omics to microbial ecology

S Mauger, C Monard, C Thion, P Vandenkoornhuyse

► To cite this version:

S Mauger, C Monard, C Thion, P Vandenkoornhuyse. Contribution of single-cell omics to microbial ecology. Trends in Ecology & Evolution, 2022, 37 (1), pp.67-78. 10.1016/j.tree.2021.09.002 . hal-03414494

HAL Id: hal-03414494

<https://hal.science/hal-03414494>

Submitted on 15 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Contribution of single-cell omics to microbial ecology

Authors: Mauger, S.^{1,2}, Monard, C.¹, Thion, C.², Vandenkoornhuyse, P.¹

¹ Université de Rennes 1, CNRS, UMR 6553 ECOBIO, 35042 Rennes (France)

² Cellenion, 60 Avenue Rockefeller, 69008 Lyon (France)

Corresponding authors : Solène Mauger (solene.mauger@univ-rennes1.fr); Philippe Vandenkoornhuyse (philippe.vandenkoornhuyse@univ-rennes1.fr)

Keywords: Microbiota; single cell; omics; population; scales

Highlights

Microbes are involved in many ecosystems but remain understudied mostly due to technical limitations.

Applying -omics technologies to microbes involve changes in the scales of ecological studies.

Single-cell omics offer the opportunity to study microbes at a finer scale than meta-omics tools.

Single microbial cell omics enable to assess and explore the dynamics of genetic changes from individuals to higher levels of ecological complexity.

Abstract

Microorganisms play key roles in various ecosystems but many of their functions and interactions remain undefined. To investigate the ecological relevance of microbial communities, new molecular tools are being developed. Among them, single-cell omics assessing genetic diversity at the population and community level and linking each individual cell to its functions is gaining interest in microbial ecology. By giving access to a bigger range of ecological scales (from individual to community) than culture-based approaches and meta-omics, single-cell omics can contribute to microorganisms genomes and functions identification but also to the testing of concepts in ecology. Here, we discuss the contribution of single-cell omics to possible breakthroughs in concepts and knowledge on microbial ecosystems and eco-evolutionary processes.

41 **Main Text**

42 **Ecological scales**

43 Interactions between organisms take place at all organizational levels, from
44 molecules to communities, within or between species, and shape ecosystem
45 dynamics. Ecological interactions are difficult to understand due to the number of
46 biotic and abiotic parameters involved. Assembling knowledge at various ecological
47 scales and from different standpoints is therefore crucial in the study of ecological
48 and evolutionary processes. This is particularly true in the case of microbes, which
49 individuals can be seen as metabolic units involved in complex metabolic networks
50 at much higher ecological scales (e.g. [1,2]). Therefore, accessing genetic and
51 metabolic information of microbes is a necessary step to understand ecosystem
52 functioning. In microbial metabolic units, “small” changes in genomes and metabolic
53 pathways may have significant impact on the microbial community organization and
54 hence on ecosystems. Deciphering large ecological scale processes therefore
55 requires observation of fine ecological scale (i.e. at the individual cell level), which is
56 the biggest challenge of current environmental microbiology [3] (Figure 1a).

57 A fundamental level of organization in ecology is the species. However, due to gene
58 flow between cells that increases with ecological overlap and genetic similarity [4],
59 microbial species concept and thus, **populations** (see Glossary), are not clearly
60 defined entities. Interactions and diversity at the populational level, i.e. between
61 individual cells of the same population (Box 1), is still obscure since it is not often
62 analyzed in environmental microbiology. Given the natural mutation rate in bacteria
63 ($\sim 10^{-7}$ substitution per nucleotide, e.g. [5]), even a single colony contains genetic
64 variations, i.e. variants within a cell population. The population has been suggested

to be more relevant than the species level for microbes [6,7] and species usually contain genetically divergent microorganisms. Considering the hierarchical levels of ecology, populations are keys to assessing genetic structure within species and over time, changes therein. They thereby provide insights into eco-evolutionary processes and advance our understanding of microbiota composition dynamics (Box 1).

Our understanding of the microbial world and its ecological roles is still very limited [8]. Understanding the functions played by microbial cells in a complex community remains a frontier in microbial ecology. Beyond the technical limits that microbiology is facing, the information gathered from culture-based studies or from natural ecosystems can be difficult to interpret (Figure 1a). Laboratory experiments attempt to reproduce optimal ecological conditions for microbes by selecting among the many biotic and environmental parameters [6] in order to understand specific processes such as trait tradeoffs [9], interactions between strains [10], the production of metabolites [11] or genome evolution [12]. Extrapolating observations obtained in vitro, at restricted scales, to higher ecological scales such as natural communities and ecosystems requires particular attention. Conversely, observations made from environmental samples, including microbial community composition, diversity or global functions are less specific and represent an average of the microbial community. Ideally, we want to get the most information out of each level of approach (i.e precise interactions and genetic dynamics from culture-based studies coupled with global function and diversity of a community with meta-omics). However, our ignorance of intricacies and interactions of ecological scales with one another still jeopardize the assembly of the resulting information to answer specific ecological questions in environmental microbiology. Moreover, the uncertainty of

community composition and the complexity of microbial interactions [2] make it even more difficult to target specific scientific hypotheses on natural communities and to choose the appropriate tools.

Figure 1: Gradient of ecological complexity, study scales and associated approaches.

Although other higher levels of ecology can be used, individual-population-community-environment scales describe much of the subject of ecology. a) The culture-based approaches aim at studying microbial populations or a very limited number of strains, therefore the outcoming data cannot be fully informative about higher levels of ecology. On the contrary, meta-omics approaches cover from the community to environmental scales of microbial ecology and does not provide finer information on the ecological gradient. In both cases, the individual scale is unattainable while being at the basis of ecological processes. b) Single-cell omics cover the scales from the individual microbe to the community from the same environmental sample which allows to connect the outcoming information of each ecological scale.

107

108 **Specific tools for specific ecological questions**

109 Like in any new field of exploration, ecological patterns within microbial communities
110 are firstly observed and described but poorly understood [13], testifying to the
111 enormous lack of knowledge concerning microbes [14]. The use of DNA- and RNA-
112 based methods to study natural microbial communities have demonstrated the
113 existence of a prodigious wealth of microorganisms which remained unsuspected
114 some years ago (e.g. [15]). Among **meta-omics** techniques, metagenomics and
115 metatranscriptomics are the most widely used methods to explore microbiota. These
116 techniques enabled a breakthrough in our understanding of microbial phylogenetic
117 relationships [16], species diversity and abundance [17], metabolic abilities [18] and
118 functional diversification [19]. The development and use of **metagenome-**
119 **assembled genomes (MAGs)** led to discoveries that advanced our understanding
120 of bacterial life and modified our perception of the tree of life [15,19]. Some studies
121 attempted to reconstitute population level genomes from metagenomes. For
122 example, Crits-Christoph et al. [20] investigated genetic variation within populations
123 of highly abundant soil bacteria by studying MAGs and observed spatial
124 differentiation of alleles. However, inferring population features from meta-omics
125 data remains limited, especially when genomes are inaccurately or incompletely
126 reconstructed from short sequence fragments (Figure 2, Key Figure). The use of
127 MAGs becomes challenging when microbial richness and diversity within a
128 community are high and taxa are phylogenetically close [21]. During genome
129 assembly, stitching of fragments from different individual genomes and/or
130 contaminant DNA can occur, creating chimeras that are irrelevant for the study of

populations. In this case, the approach would necessarily conceal a considerable proportion of molecular diversity [22]. In addition, the molecular biology and bioinformatics methods used in meta-omics approaches are varied and based on different criteria and assumptions in the absence of a consensus, leading to contrasting results and interpretations [23,24]. Overall, it might be difficult to directly link the detected functions to their original microbial cell from meta-omics data, thereby limiting the identification of signaling pathways and trade-offs in gene regulation. Meta-omics approaches proved useful in describing communities using large-scale sampling and make it possible to answer questions related to community composition and its associated global functions but not to fully understand the mechanisms underlying these patterns. Nevertheless, bioinformatics research has developed algorithms aiming to identify genetic variations in microbial populations: vertically and horizontally inherited genes can be differentiated, and from population-specific sweeps, single nucleotide polymorphisms can be detected (e.g. DiscoSNP; PopCOGenT) [4, 25].

To complement meta-omics data, modelling approaches are used to explore microbial interactions, fluxes of metabolites, and to reconstruct ecological networks in complex microbiomes [26-29]. These approaches provide a possible explanation and scenarios of interactions in natural communities, but sometimes end in contradictions with culture-based experiments [1]. Indeed, models can predict a certain kind of interaction (e.g. cross-feeding) that is not verified or proven wrong in an experimental setup due to the oversight of key parameters such as growing conditions, space and (very often) time [1]. They also rely on the co-occurrence of phenomena, which is more associated with correlations than cause-consequences relationships. Culture-dependent approaches may help reach the population level in

simple community compositions through controlled and simplified laboratory-scale experiments [30] and can be effective for testing patterns observed in meta-omics studies and deconstructing mechanistic hypotheses (e.g. interactions, metabolism) (e.g. [31]). The lack of information on individual cells and mostly on populations (i.e. both functions and phylogeny) using existing methods limits our understanding of observed processes. Many studies aim at unravelling microbiota diversity, primarily in plants, soil, water, and animal bodies, but few explain associated community assembly and evolutionary mechanisms [32-34]. In this context, microbiologists and ecologists are searching for other technical possibilities or approaches such as single-cell omics to complete the knowledge provided by current methods.

The alternative scope of single -cell omics

Single-cell whole genome sequencing (scWGS) and single-cell transcriptomics (scRNA-seq) were first developed for eukaryotic cells and used in cancer research, revealing both intra-population genetic diversity and heterogeneous genome expression. As in cancer research where differentiation in space of the genome expression among cell has been observed [35], a pioneer paper on *Pseudomonas aeruginosa* biofilm using a fluorescent based approach (i.e. parallel sequential fluorescence in situ hybridization (par-seqFISH)), revealed a differentiation in space and time of cell-expression [36]. Because spatial single-cell microbial approaches could allow the understanding of the drivers and mechanisms leading to the self-organization of these microbial structures, new developments are expected to expand in health science and many other fields of microbial ecology research. Single-cell approaches are promising candidates for microbial studies, as they

179 provide a complementary view to metagenomics and metatranscriptomics that have
180 different strengths, but also weaknesses (Figure 2).

181 Single-cell omics technologies require additional steps to prepare a sample for
182 sequencing as compared to meta-omics techniques, especially with regards to cell
183 isolation, for which different technical options are available [37,38]. Once the cells
184 lysed, DNA and RNA content from a single cell is in the femtogram scale for
185 bacteria, i.e. 1,000-fold less than in animal cells. Preparing the sequencing library,
186 which typically requires ng ranges of material, will need an ultra-efficient prior
187 amplification step (e.g. Multiple Displacement Amplification (MDA), the most widely
188 used approach for bacteria [39]).

189 Single-cell approaches enable accurate access to genomic and transcriptomic
190 information for each cell, so that the assembled cell information is highly
191 representative of the original population (Figure 2). This enables the identification of
192 heterogeneity in gene assemblage, gene expression and metabolic pathways
193 between cells. Single-cell transcriptomic and genomic information provide a link
194 between phylogeny and functional traits and reveal the physiological status of an
195 individual cell, at the time of sampling. This is particularly important considering that
196 the individual gene expression profiles of genetically close cells may differ. What is
197 more, some cell isolation tools such as automated image-based isolation devices
198 (CellenONE Cellenion; icell8 Takara) make it possible to select cells based on their
199 integrity, their physiology and/or functional markers, and to minimize contamination
200 by the host or extracellular DNA. This is very promising for microbiology to, for
201 instance, select active cells in the studied sample at the time of sampling and reveal
202 which of them are taking part in the community productivity.

A seminal paper on single cell microbial genome analysis was published in 2005 [40] and paved the way for further improvement of single-cell omics, notably on the amplification method (here MDA) and lysis buffer. Recent studies using single-cell omics have improved our understanding of intraspecific diversification and metabolic capacities at a limited scale [41, 42]. Assessing the true individual cell gene assemblage and expression using single-cell omics will make it possible to study the hitherto unexplored microbial population level and the functioning of a given microbiome by linking the different ecological scales (Figure 1b). Indeed, single-cell -omics enable access to additional information than in culture-based and meta-omics studies (the single individual ecological scale) whilst also, from the environmental sample, giving information on the population and community interactions. To a broader extent, this will enable better access to eco-evolutionary pressures and evolutionary processes within microbial communities.

Applications of single-cell omics in microbial ecology

Single-cell omics provide information at the cell level by changing the camera angle when studying environmental communities and can contribute to microbiology and ecology at many levels by exploring microbial diversity or microbial interactions.

The tremendous diversity of single-microbe genomes.

Observations of microbiota can complement/validate the diversity observed by meta-omics on fungi [43], human samples [8], and marine viruses [44], and resolve cryptic bacterial species, which currently mainly rely on cultivable strains [45-47]. Single-microbe -omics therefore contribute to the microbial inventory which is still in its infancy in many ecosystems [48]. The use of single-cell based approaches could demonstrate the existence of the discovered sequence-based lineages and discover

possible new branches. Isolating cells from environmental samples can also cast light on rare organisms that might be obscured in millions of genome fragments using meta-omics. These rare organisms are considered to play key roles in community dynamics since they over-proportionally contribute to the functions of the microbial community in fluctuating environments [49,50]. While most studies that apply this approach use a limited number of cells, a recent survey of the marine microbiome recovered no less than 12,000 genomes from single cells [51], revealing a high degree of uniqueness and limited clonality in the analyzed samples of sea water and providing evidence for the ecological roles of uncultured microbial groups. Single-cell genomics, by looking at individual genomes instead of core genomes from meta-omics, from natural microbial communities represent an unprecedented opportunity to complete the identification and classification of microbes. This is a key step sometimes missing in environmental microbiology [52]: knowing what to look at and why to formulate hypotheses in ecology and better understand processes involving microbes.

Ecological and evolutionary hypothesis testing using single-cell -omics.

Single-cell genomics and transcriptomics approaches therefore help to answer the questions “What are these microbes?” and “What are they doing (or capable of doing)?” but also to understand why and how observed patterns happen. Linking environmental and community parameters to individual gene expression and bacterial interactions enables a mechanistic understanding of underlying biotic and abiotic conditions to patterns. This represents an opportunity to explore multiple ecological theories and hypotheses, notably on interactions of microbes at many levels: within the community, with external microbes (i.e. viruses) and with their host.

One of the hottest topics in microbial ecology is the link between diversity and function, including the productivity of the ecosystem that relies on the **niche partitioning** theory. This hypothesis states that species coexistence is enabled by species specialization in different available resources (or combinations of resources), thereby reducing interspecific competition [53] but likely modifying the microbial population structure [54]. Specialization in specific resources raises many questions concerning microbial interactions through the exchange of metabolites [55], loss of traits [56] and genome reduction [57]. The Black Queen Hypothesis (BQH), one of the ecological theories that conceptualized this phenomenon, states that microbial community assembly and complexity are at least partially determined by functional dependencies resulting from gene loss(es). Testing this hypothesis requires using environmental samples to evaluate functional redundancy within communities, the expression and distribution of the functions between interacting (micro-)organisms and the impact of genotypic interactions on these functions. At the community level, it is impossible to access this information through meta-omics since BQH evolution is supposed to also occur at the population level [53,54] which is undetectable from most of the meta-omics tools used so far. Eco-evolutionary processes can so far be explicitly assessed only through ecological models [58] or from dedicated *in-vitro* experiments [42] that require deciding which gene and which organisms to look at. Single cell approaches can help to investigate such hypotheses and other theories in ecology and highlight patterns of interactions that shape microbial communities.

A breach to viral-host ecology.

The diversity of microbial communities is also influenced by the viral infections to which they are exposed, which is particularly difficult to evaluate in nature [59]. A

recent study assessed viral infections in the ocean thanks to single cell genomics by identifying virus sequences in uncultured protist cells [44]. Such interactions are widespread but generally missing in current microbial ecology analyses. Among other roles, viruses are known to (1) control microbial community dynamics and drive microbial-host evolution [60,61] and (2) impact ecosystem changes and biogeochemical cycles [59,62]. Single-cell omics can help us understand the roles played by viruses in microbial populations by making it possible to assess both the prevalence of prophage sequences and possible lateral gene transfers [63,64] which would reveal preferential association of specific viruses with specific bacteria.

Single-cell insights into bigger scale interactions.

Generally speaking, single-cell omics can highlight diverse levels of interaction in natural habitats either between the microbes that comprise a microbiota or between the microbes/microbiota and their host. Despite the great number of studies on the composition of human microbiota, very little is known about interactions between the microbes and with their host [65]. The single cell study of host-pathogen interactions paves the way for understanding infectious processes through microbiota dynamics, metabolic capacities, and host resistance [66]. Dysbiosis is often shown to display a higher beta-diversity interpreted as a higher stochasticity in the microbiota assembly [67]. Among other explanations, this apparent higher stochasticity may be the result of drastic pathogen-induced changes in the habitat (i.e. transitory state in the microbiota dynamics), modification of a component of the microbiota caused by genetic change(s) and/or a functional modification expressing a modified phenotype that leads to disequilibrium in the microbiota community. With the aim of understanding how a disorder of the microbiota leads to disease or the reverse, i.e.

how a disease can modify the composition of a microbiota, single-cell microbiota analyses of genomes and transcriptomes would help better define the characteristics of dysbiosis (i.e. dysbiosis mechanisms). Using single cell genomics makes it possible to address hypotheses related to changes in bacterial populations, while microbial single-cell transcriptomics may be more appropriate to decrypt the functions of microbes, metabolic abilities, and cellular states [68]. These two strategies are necessary to understand how microbial interactions occur within communities as well as their possible impact on the ecosystem [69,51].

Microbial single-cell omics are not solely expected to better understand functional interactions but also the underlying evolutionary processes. Microbial single-cell omics should also promote a shift in standpoint from observation to interpretation and also offer new opportunities to test macroecological theories on microbes. The development of microbial single-cell omics will have high impacts on our understanding of microbial communities in many environments. For instance (i) in freshwater and marine ecosystems, to define the interaction of bacteria and phytoplankton through the exchange of metabolites, and to test links to blooms [70], (ii) in soils, to better assess the provision of services by plant microbiota including nutrient and water uptake and protection against pathogens [71], (iii) in plant, human and animal health, to better decipher how dysbiosis could be cause or consequence of a disease. However, it can't be ignored that the wide development of single-cell omics applied to microorganisms is limited by technical limits.

Limits of single-cell omics on microbes

323 The current limited number of cells studied in the published papers questions the
324 representativeness of the analyses. Considering the number of microbial cells
325 contained in a given environmental sample, one can wonder how many cells need to
326 be isolated to cover the diversity of a sample, from hundreds [41], to thousands [51].
327 The limited number of analyzed cells is mostly due to technical problems and the
328 cost of such experiments. It has to be emphasized that the current use of single-cell
329 omics for microbes have to solve many technical obstacles (Box 2), reviewed in
330 [72,73] such as cell isolation, lysis, and a biased amplification step. The structure of
331 the microbial cell wall is complex, and unlike the animal cells, they do not break
332 easily. The diversity of cell wall composition across phylogeny and physiological
333 status (peptidoglycan layer(s), spores, capsules, etc.) makes it challenging to find a
334 universal lysis method able to breach each cell without damaging its content or
335 inhibiting enzymatic reactions downstream. Different protocols have been used in
336 recent microbial single-cell based studies, using either heat, temperature shocks,
337 sonication, enzymes, detergents, or combinations of these [74,75]. In addition to
338 technical issues, one can wonder how to be certain that cells are isolated and lysed
339 equally and not preferentially depending on their physical/physiological status. The
340 amplification step, usually made via MDA, has been reported to be imprecise
341 concerning the genome amplification uniformity, even though it presents a better
342 genome coverage than other approaches (i.e MALBAC [39]). This amplification step
343 is highly relevant, as it was suggested to be the cause of incomplete reconstruction
344 of single amplified genomes (SAGs) [76], although solutions are being developed
345 [77]. As the price of library preparation represents most of the cost of these new
346 single- cell omics for microorganisms, reduction of reaction volumes in
347 “nanolibraries” should be very cost-effective (Box 2). The probability of contamination

decreases with the miniaturization of the reagent volumes [78] and associated robotics. Working in nanovolumes seems to be a convenient solution to solve multiple problems, however this introduces new volume-related challenges such as pipetting or sample purification.

For these reasons, single-cell omics have been sometimes used in combination with meta-omics to combine the possibility of fine scale analysis with high throughput [76,79]. It also represents a good opportunity to validate multiple aspects of single-cell omics: 1) the isolation and lysis universality, 2) the sample preparation (genome amplification and library preparation for sequencing), 3) the lack of contamination, and 4) the representativeness of the sample covered by single -cell omics (See outstanding questions).

Box 2: Single microbial cell -omics approaches are technically challenging.

Challenge	Possible solution
1. Isolation of single microbial cells	
Community representative sampling	High throughput
Cell isolation from complex matrices, such as soil, sediments, host tissue, feces, mucus...	Sonication, filtration, density gradient centrifugation
Exhaustive/targeted labelling/detection	Fluorescence, antibodies
Maintenance of axenic conditions	(As) clean (as possible) room + ⑤
2. Culturomics experiments	
Maximum viability/cultivability	Gentle cell handling, liquid dispensing
Choice of culture conditions	High throughput media screening + ⑤
Assessment of monoclonality (= culture purity)	Microscopy, targeted sequencing
3. Microbial cell lysis/permeabilization	
Efficiency across phylogeny	(Ultra-) sonication, thermal shock, heat, enzymolysis, detergents...
Preservation of DNA/RNA quality and quantity	Gentle procedure, avoid purification
Minimum contamination from reagents and prevention of subsequent steps	Physical rather than biological/chemical + ⑤
4. Single-microbe WGS

5 steps of single microbial approaches.

Super-efficient (100K- to 1 Mo-fold) amplification (1-10 fg DNA prokaryotic cell ⁻¹)	Ⓟ, molecular crowding, linear amplification (e.g. <i>in vitro</i> transcription)
Even/broad coverage, high fidelity, and no chimera creation	Minimum number of PCR cycles and/or primary template amplification (e.g. <i>in vitro</i> transcription)
Minimum contamination from reagents	Minimum reagent amount + Ⓟ
Cost reduction	Cell barcoding for multiplexing + Ⓟ
Bioinformatics	Dedicated tools for cell demultiplexing, monoclonality test...
5. Single-microbe RNA-seq	
Same as WGS (up to 100 fg RNA prokaryotic cell ⁻¹)	Ⓟ, molecular crowding
Amplification bias	Unique molecular identifier (UMI)
No poly-A tail on prokaryotic mRNA	RNA poly-A tailing, random priming, ribosomal RNA targeted depletion

Table i. Padlocks and possible keys offered by single microbial cell –omics approaches.

Crucial, but solvable, issues (Table i) should be addressed at each step of single microbe approaches (Fig. i). The reduction of reaction volumes down to “nanovolumes” (see Ⓟ symbols), is often the most sensible solution, limiting contamination probability and allowing high throughput and cost reduction.

Concluding Remarks

Soon, single-cell omics applied to microorganisms could become a gold standard in microbial ecology thanks to the knowledge produced by focusing on individual genomes and transcriptomes, and, possibly, individual proteomes and metabolomes. Today, technical problems prevent the testing of broad ecological hypotheses. Generalizing ecological single-cell studies on microbes requires the development of robust high throughput techniques with a high-cost effectiveness ratio (See Outstanding Questions). A knowledge upshot is expected in microbial interactions

and eco-evolutionary boundaries through the enabling of mechanistic characterization of deterministic populations and community assembly processes. Currently, the use of metagenomics and single cell genomics in the same study appears as the best solution, combining the strengths of the two approaches, i.e. (i) high throughput and alpha/beta diversity, and (ii) fine scale analysis by scWGS and/or scRNAseq [76,80]. Ideally, one would not over-interpret meta-omics data and rather use those to build hypotheses on mechanisms, which can be tested using single-cell approaches. Approaches that will allow more accurate assessment of microbial genome diversity and genome functioning within complex microbiota, are impatiently awaited. Still, the future of microbial single-cell omics will likely fuel a new perception of the microorganisms world.

Acknowledgements

We would like to thank LABCOM 'Microscale' for funding.

Key figure

Outstanding Questions

To what extent can a single technical approach realistically reflect complex natural processes?

How many cells need to be isolated from a natural environment to accurately represent the population and/or community from which they originate? And how can this be assessed?

What criteria should be used to determine the scale of sampling in natural environments?

How can we use information obtained by single-cell omics for a better understanding of ecosystem functioning?

399 How can we make the application of proteomics, metabolomics and multi-omics
400 approaches to single microbes more realistic?

401

402 **Glossary**

403 **Metagenome-assembled genome (MAG):** in-silico reconstruction of an artificial
404 microbial genome obtained from one or multiple binned metagenomes that represent
405 the core genome of the population

406 **Meta-omics:** group of molecular biology technologies, extensively used to access
407 unculturable organisms, by studying the bulk pool of biomolecules from
408 environmental samples, to reveal genomes (metagenomics), transcriptomes
409 (metatranscriptomics), proteomes (metaproteomics) and metabolites
410 (metabolomics).

411 **Niche complementarity/partitioning:** ecological concept describing how species
412 differential specialization in different combinations of resource uses and functions
413 allow them to coexist in the same environment.

414 **Population:** applied to bacteria in natural communities, individuals with identical or
415 different genomes from the same species gathered in a specific environment or
416 sample. An isolated microorganism culture also comprises a population.

417

418 **Box 1: Microbial population**

In ecology, populations are individuals belonging to the same species, living in the same environment, although the definition varies with different viewpoints [81]. Microbial populations represent a unit of diversity and selection. Within these populations, diversity can be either genetic or phenotypic. The diversity within a population to some extent buffers an environmental stress because existing variants are able to survive the stress and/or allow rapid phenotype switching (e.g. Bet-Hedging, [82,31]) but positive selection of new variants can also be induced by the stress. This organizational level is therefore a key to understanding genetic structure and haplotype fitness, the dynamics of ecological interactions, including associations of microbial species, symbioses, host-pathogen interactions, and ecosystem functioning, resilience, and stability. For instance, resistance to antibiotics can vary within populations [83] and the virulence pathogens can vary across subpopulations [84]. Genetic diversity and ecological features such as niches can vary between lineages [85], so that subpopulations are able to co-exists through niche diversity. To capture the total genetic and phenotypic diversity and get a holistic view of populations, the scale must thus be tuned down to the individuals that comprise the population in a given sample [84]. Otherwise, applying the current bacterial species concept to make population level inferences may lead to false or partial interpretation of ecological phenomena.

Reference list

- 1- Coyte, K.Z., and Rakoff-Nahoum, S. (2019) Understanding competition and cooperation within the mammalian gut microbiome. *Curr. Biol.* 29(11), 538–544. DOI: <https://doi.org/10.1016/j.cub.2019.04.017>

- 2- Pacheco, A.R., and Segrè, D. (2019) A multidimensional perspective on microbial interactions. *FEMS Microbiol. Lett.* 366(11), 1–11. DOI: <https://doi.org/10.1093/femsle/fnz125>
- 3- Dini-Andreote, F. *et al.* (2020) Towards meaningful scales in ecosystem microbiome research. *Environ. Microbiol.* 23(1), 1-4. DOI: <https://doi.org/10.1111/1462-2920.15276>
- 4- VanInsberghe, D. *et al.* (2020) How can microbial population genomics inform community ecology? *Phil. Trans. R. Soc. B.* 375(1798), 20190253. DOI: <https://dx.doi.org/10.1098%2Frstb.2019.0253>
- 5- Westra, E.R. *et al.* (2017) Mechanisms and consequences of diversity-generating immune strategies. *Nat Rev Immunol*, 17(11), 719-728. DOI: <https://doi.org/10.1038/nri.2017.78>
- 6- García-García, N. *et al.* (2019) Microdiversity ensures the maintenance of functional microbial communities under changing environmental conditions. *ISME J.* 13(12), 2969-2983. DOI : <https://doi.org/10.1038/s41396-019-0487-8>
- 7- Niccum, B.A. *et al.* (2020) Strain-level diversity impacts cheese rind microbiome assembly and function. *mSystems* 5(3) e00149-20. DOI: <https://doi.org/10.1128/msystems.00149-20>
- 8- Marcy, Y. *et al.* (2007) Dissecting biological “dark matter” with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *PNAS* 104(29), 11889–11894. DOI: <https://doi.org/10.1073/pnas.0704662104>

- 9- Moreno-Gómez, S. *et al.* (2020) Wide lag time distributions break a trade-off between reproduction and survival in bacteria. *PNAS* 117(31), 18729-18736. DOI: <https://doi.org/10.1073/pnas.2003331117>
- 10-Ratzke, C. *et al.* (2020) Strength of species interactions determines biodiversity and stability in microbial communities. *Nat. Ecol. Evol.* 4(3), 376-383. DOI: <https://doi.org/10.1038/s41559-020-1099-4>
- 11-Rütting, T. *et al.* (2021) The contribution of ammonia-oxidizing archaea and bacteria to gross nitrification under different substrate availability. *Soil Biol. Biochem.* 160, 108353. DOI: <https://doi.org/10.1016/j.soilbio.2021.108353>
- 12-Hall, J. P. *et al.* (2018) Competitive species interactions constrain abiotic adaptation in a bacterial soil community. *Evol. Lett.* 2(6), 580-589. DOI: <https://doi.org/10.1002/evl3.83>
- 13-Tripathi, A. *et al.* (2018) Are microbiome studies ready for hypothesis-driven research? *Curr. Opin. Microbiol.* 44, 61–69. DOI: <https://doi.org/10.1016/j.mib.2018.07.002>
- 14-Lloyd, K.G. *et al.* (2018) Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 3(5) e00055-18. DOI: <https://doi.org/10.1128/msystems.00055-18>
- 15-Hug, L.A. *et al.* (2016) A new view of the tree of life. *Nat. Microbiol.* 1(5), 16048. DOI: <https://doi.org/10.1038/nmicrobiol.2016.48>
- 16-Alteio, L.V. *et al.* (2020) Complementary metagenomic approaches improve reconstruction of microbial diversity in a forest soil. *mSystems*, 5(2), e00768-19. DOI: <https://doi.org/10.1128/mSystems.00768-19>

- 17-Daims, H. *et al.* (2015) Complete nitrification by *Nitrospira* bacteria. *Nature* 528(7583), 504. DOI: <https://doi.org/10.1038/nature16461>
- 18-Cernava, T. *et al.* (2017) Deciphering functional diversification within the lichen microbiota by meta-omics. *Microbiome* 5(1), 82. DOI: <https://doi.org/10.1186/s40168-017-0303-5>
- 19- Parks, D.H. *et al.* (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2(11), 1533-1542. DOI: <https://doi.org/10.1038/s41564-017-0012-7>
- 20-Crits-Christoph, A. *et al.* (2020) Soil bacterial populations are shaped by recombination and gene-specific selection across a grassland meadow. *ISME J.* 14(7), 1834-1846. DOI: <https://doi.org/10.1038/s41396-020-0655-x>
- 21-Nayfach, S. *et al.* (2019) New insights from uncultivated genomes of the global human gut microbiome. *Nature* 568(7753), 505-510. DOI: <https://doi.org/10.1038/s41586-019-1058-x>
- 22-Nelson, W.C. *et al.* (2020) Biases in genome reconstruction from metagenomic data. *PeerJ* 8, e10119. DOI: <https://doi.org/10.7717/peerj.10119>
- 23-Li, L. *et al.* (2015) Comparing viral metagenomics methods using a highly multiplexed human viral pathogens reagent. *J. Virol. Methods*, 213, 139-146. DOI: <https://doi.org/10.1016/j.jviromet.2014.12.002>
- 24-Mande, S.S. *et al.* (2012) Classification of metagenomic sequences: methods and challenges. *Brief. Bioinform.* 13(6), 669-681. DOI: <https://doi.org/10.1093/bib/bbs054>

- 25-Uricaru, R. *et al.* (2015) Reference-free detection of isolated
SNPs. *Nucleic Acids Res.* 43(2), e11-e11. DOI:
<https://doi.org/10.1093/nar/gku1187>
- 26-Antoniewicz, M.R. (2020) A guide to deciphering microbial interactions and
metabolic fluxes in microbiome communities. *Curr. Opin. Biotechnol.* 64,
230–237. DOI: <https://doi.org/10.1016/j.copbio.2020.07.001>
- 27-Belcour, A. *et al.* (2020) Metage2Metabo, microbiota-scale metabolic
complementarity for the identification of key species. *eLife* 9, e61968. DOI:
<https://doi.org/10.7554/eLife.61968>
- 28-Gorter, F.A. *et al.* (2020) Understanding the evolution of interspecies
interactions in microbial communities. *Philos. Trans. R. Soc. Lond., B, Biol.*
Sci. 375(1798), 20190256. DOI: <https://doi.org/10.1098/rstb.2019.0256>
- 29-Perez-Garcia, O. *et al.* (2016) Metabolic network modeling of microbial
interactions in natural and engineered environmental systems. *Front.*
Microbiol. 7, 673. DOI: <https://doi.org/10.3389/fmicb.2016.00673>
- 30-Morin, M. *et al.* (2018) Changes in the genetic requirements for microbial
interactions with increasing community complexity. *eLife* 7, e37072. DOI:
<https://doi.org/10.7554/eLife.37072>
- 31-Beaumont, H. J. E. *et al.* (2009) Experimental evolution of bet hedging.
Nature 462(7269), 90–93. DOI: <https://doi.org/10.1038/nature08504>
- 32-Kiers, E.T. *et al.* (2011) Reciprocal rewards stabilize cooperation in the
mycorrhizal symbiosis. *Science* 333(6044), 880–882. DOI:
<https://doi.org/10.1126/science.1208473>

- 33-Werner, G.D., and Kiers, E.T. (2015) Partner selection in the mycorrhizal mutualism. *New Phytol.* 205(4), 1437-1442. DOI: <https://doi.org/10.1111/nph.13113>
- 34-Amin, S.A. *et al.* (2015) Interaction and signaling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 522(7554), 98–101. DOI: <https://doi.org/10.1038/nature14488>
- 35-Alon, S. *et al.* (2021) Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. *Science* 371(6528). DOI: [10.1126/science.aax2656](https://doi.org/10.1126/science.aax2656)
- 36-Dar, D. *et al.* (2021) Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science* 373(6556). DOI: [10.1126/science.abi4882](https://doi.org/10.1126/science.abi4882)
- 37-Gawad, C. *et al.* (2016) Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* 17(3), 175. DOI: <https://doi.org/10.1038/nrg.2015.16>
- 38-Woyke, T. *et al.* (2017) The trajectory of microbial single-cell sequencing. *Nat. Methods* 14(11), 1045. DOI: <https://doi.org/10.1038/nmeth.4469>
- 39-Chen, Z. *et al.* (2017) Tools for genomic and transcriptomic analysis of microbes at single-cell level. *Front Microbiol.* 8, 1831. DOI: <https://doi.org/10.3389/fmicb.2017.01831>
- 40-Raghunathan, A. *et al.* (2005) Genomic DNA amplification from a single bacterium. *Appl. Environ. Microbiol.* 71(6), 3342–3347. DOI: <https://doi.org/10.1128/AEM.71.6.3342-3347.2005>

- 41-Engel, P. *et al.* (2014) Hidden diversity in honey bee gut symbionts detected by single-cell genomics. *PLoS Genet.* 10(9), e1004596. DOI: <https://doi.org/10.1371/journal.pgen.1004596>
- 42-Hennon, G.M. *et al.* (2018) The impact of elevated CO₂ on *Prochlorococcus* and microbial interactions with ‘helper’ bacterium *Alteromonas*. *ISME J.* 12(2), 520–531. DOI: <https://doi.org/10.1038/ismej.2017.189>
- 43-Ahrendt, S.R. *et al.* (2018) Leveraging single-cell genomics to expand the fungal tree of life. *Nat. Microbiol.* 3(12), 1417–1428. DOI: <https://doi.org/10.1038/s41564-018-0261-0>
- 44-Castillo, Y.M. *et al.* (2019) Assessing the viral content of uncultured picoeukaryotes in the global ocean by single cell genomics. *Mol. Ecol.* 28(18), 4272–4289. DOI: <https://doi.org/10.1111/mec.15210>
- 45-Hahn, M.W. *et al.* (2016) Complete ecological isolation and cryptic diversity in *Polynucleobacter* bacteria not resolved by 16S rRNA gene sequences. *ISME J.* 10(7), 1642–1655. DOI: <https://doi.org/10.1038/ismej.2015.237>
- 46-Hoetzing, M. *et al.* (2019) *Polynucleobacter paneuropaeus* sp. nov., characterized by six strains isolated from freshwater lakes located along a 3000 km north-south cross-section across Europe. *Int. J. Syst. Evol. Microbiol.* 69(1), 203–213. DOI: <https://doi.org/10.1099/ijsem.0.003130>
- 47-Pitt, A. *et al.* (2019) *Aquirufa antheringensis* gen nov., sp nov. and *Aquirufa nivaísiívae* sp nov., representing a new genus of widespread freshwater bacteria. *Int. J. Syst. Evol. Microbiol.* 69(9), 2739–2749. DOI: <https://doi.org/10.1099/ijsem.0.003554>

- 48-Bernard, G. *et al.* (2018) Microbial dark matter investigations: how microbial studies transform biological knowledge and empirically sketch a logic of scientific discovery. *Genome Biol. Evol.* 10(3), 707-715. DOI: <https://doi.org/10.1093/gbe/evy031>
- 49-Shade, A. *et al.* (2014) Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *MBio* 5(4), 1371–1385. DOI: <https://doi.org/10.1128/mBio.01371-14>
- 50-Xu, Q. *et al.* (2021) Rare Bacterial Assembly in Soils Is Mainly Driven by Deterministic Processes. *Microb. Ecol.* 1–14. DOI: <https://doi.org/10.1007/s00248-021-01741-8>
- 51-Pachiadaki, M.G. *et al.* (2019) Charting the complexity of the marine microbiome through single-cell genomics. *Cell* 179(7), 1623-1635.e11. DOI: <https://doi.org/10.1016/j.cell.2019.11.017>
- 52-Prosser, J.I. (2020) Putting science back into microbial ecology: a question of approach. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 375(1798), 20190240. DOI: <https://doi.org/10.1098/rstb.2019.024>
- 53-Mas, A. *et al.* (2020) Reflections on the predictability of evolution: towards a conceptual framework. *iScience*, 23(11), 101736. DOI: <https://doi.org/10.1016/j.isci.2020.101736>
- 54-Mas, A. *et al.* (2016) Beyond the Black Queen Hypothesis. *ISME J*, 10(9), 2085–2091. DOI : <https://doi.org/10.1038/ismej.2016.22>
- 55-Zelezniak, A. *et al.* (2015) Metabolic dependencies drive species co-occurrence in diverse microbial communities. *PNAS* 112(20), 6449–6454. DOI: <https://doi.org/10.1073/pnas.1421834112>

- 56- Ellers, J. *et al.* (2012) Ecological interactions drive evolutionary loss of traits. *Ecol. Lett.* 15(10), 1071–1082. DOI: <https://doi.org/10.1111/j.1461-0248.2012.01830.x>
- 57-Morris, J.J. *et al.* (2012) The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio*, 3(2) e00036-12. DOI: [10.1128/mBio.00036-12](https://doi.org/10.1128/mBio.00036-12)
- 58-Wang, M. *et al.* (2020) Selfishness driving reductive evolution shapes interdependent patterns in spatially structured microbial communities. *ISME J.* 15(5), 1387-1401. DOI: <https://doi.org/10.1038/s41396-020-00858-x>
- 59-Weitz, J.S., and Wilhelm, S.W. (2012) Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biol. Rep.* 4(1), 17. DOI: <https://doi.org/10.3410/B4-17>
- 60-Frada, M. *et al.* (2008) The “Cheshire Cat” escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. *PNAS* 105(41), 15944–15949. DOI: <https://doi.org/10.1073/pnas.0807707105>
- 61-Scanlan, P.D., and Buckling, A. (2012) Co-evolution with lytic phage selects for the mucoid phenotype of *Pseudomonas fluorescens* SBW25. *ISME J.* 6(6), 1148–1158. DOI: <https://doi.org/10.1038/ismej.2011.174>
- 62-Rousk, J., and Bengtson, P. (2014) Microbial regulation of global biogeochemical cycles. *Front. microbiol.* 5, 103. DOI: <https://doi.org/10.3389/fmicb.2014.00103>
- 63-Roux, S. *et al.* (2014) Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell-and meta-genomics. *eLife*, 3, e03125. DOI: <https://doi.org/10.7554/eLife.03125>

- 64-Labonté, J.M. *et al.* (2015) Single-cell genomics-based analysis of virus–
host interactions in marine surface bacterioplankton. *ISME J.* 9(11), 2386–
2399. DOI: <https://doi.org/10.1038/ismej.2015.48>
- 65-Hassani, M.A. *et al.* (2018) Microbial interactions within the plant holobiont.
Microbiome 6(1)58. DOI: <https://doi.org/10.1186/s40168-018-0445-0>
- 66-Mills, E. and Avraham, R. (2017) Breaking the population barrier by single
cell analysis: one host against one pathogen. *Curr. Opin. Microbiol.* 36, 69–
75. DOI: <https://doi.org/10.1016/j.mib.2017.01.005>
- 67-Zaneveld, J.R. *et al.* (2017) Stress and stability: Applying the Anna
Karenina principle to animal microbiomes. *Nat. Microbiol.* 2(9),1–8. DOI:
<https://doi.org/10.1038/nmicrobiol.2017.121>
- 68-Hatzenpichler, R. *et al.* (2020) Next-generation physiology approaches to
study microbiome function at single cell level. *Nat Rev Microbiol.* 18(4),
241–256. DOI: <https://doi.org/10.1038/s41579-020-0323-1>
- 69-McNulty, R. *et al.* (2021) Droplet-based single cell RNA sequencing of
bacteria identifies known and previously unseen cellular states. *BioRxiv*
2021.03.10.434868. DOI: <https://doi.org/10.1101/2021.03.10.434868>
- 70-Landa, M. *et al.* (2016) Shifts in bacterial community composition
associated with increased carbon cycling in a mosaic of phytoplankton
blooms. *ISME J.* 10(1), 39–50. DOI:
<https://doi.org/10.1038/ismej.2015.105>
- 71-Vandenkoornhuyse, P. *et al.* (2015) The importance of the microbiome of
the plant holobiont. *New Phytol.* 206(4), 1196–1206. DOI:
<https://doi.org/10.1111/nph.13312>

- 72-Kaster, A.K., and Sobol, M.S. (2020) Microbial single-cell omics: the crux of the matter. *Appl. Microbiol. Biotechnol.* 104(19),8209–8220. DOI: <https://doi.org/10.1007/s00253-020-10844-0>
- 73-Clingenpeel, S. *et al.* (2015) Reconstructing each cell's genome within complex microbial communities - dream or reality? *Front. Microbiol.* 5, 771. DOI: <https://doi.org/10.3389/fmicb.2014.00771>
- 74-He, J. *et al.* (2016) Improved lysis of single bacterial cells by a modified alkaline-thermal shock procedure. *Biotechniques* 60(3), 129-135. DOI: [10.2144/000114389](https://doi.org/10.2144/000114389)
- 75-Liu, Y. *et al.* (2018) The development of an effective bacterial single-cell lysis method suitable for whole genome amplification in microfluidic platforms. *Micromachines* 9(8), 367. DOI: <https://doi.org/10.3390/mi9080367>
- 76-Alneberg, J. *et al.* (2018) Genomes from uncultivated prokaryotes: a comparison of metagenome-assembled and single-amplified genomes. *Microbiome* 6(1), 1-14. DOI: <https://doi.org/10.1186/s40168-018-0550-0>
- 77-Kogawa, M. *et al.* (2018) Obtaining high-quality draft genomes from uncultured microbes by cleaning and co-assembly of single-cell amplified genomes. *Sci Rep.* 8(1), 1-11. DOI: <https://doi.org/10.1038/s41598-018-20384-3>
- 78-Blainey, P.C. (2013) The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* 37(3), 407-427. DOI: <https://doi.org/10.1111/1574-6976.12015>
- 79-Probst, A.J. *et al.* (2018) Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep

terrestrial subsurface. *Nat. Microbiol.* (3), 328-336. DOI:

<https://doi.org/10.1038/s41564-017-0098-y>

80-Santoro, A.E. *et al.* (2019) Contributions of single-cell genomics to our understanding of planktonic marine archaea. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 374(1786), 20190096. DOI:

<https://doi.org/10.1098/rstb.2019.0096>

81-Waples, R.S. and Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* 15(6), 1419–1439. DOI:

<https://doi.org/10.1111/j.1365-294X.2006.02890.x>

82-Veening, J.-W. *et al.* (2008) Bistability, Epigenetics, and Bet-Hedging in Bacteria. *Annu. Rev. Microbiol.* 62(1), 193–210. DOI:

<https://doi.org/10.1146/annurev.micro.62.081307.163002>

83-Chua, S.L. *et al.* (2016) Selective labelling and eradication of antibiotic-Tolerant bacterial populations in *Pseudomonas aeruginosa* biofilms. *Nat. Commun.* 7(1), 1–11. DOI : <https://doi.org/10.1038/ncomms10750>

84-Davis, K.M., and Isberg, R.R. (2016) Defining heterogeneity within bacterial populations via single cell approaches. *BioEssays* 38(8), 782–790. DOI: <https://doi.org/10.1002/bies.201500121>

85-Garcia, S.L. *et al.* (2018) Contrasting patterns of genome-level diversity across distinct co-occurring bacterial populations. *ISME J.* 12(3), 742–755. DOI : <https://doi.org/10.1038/s41396-017-0001-0>

Figure 2 – Key Figure: Microbial communities observed through meta-omics versus single-cell omics

The key steps of meta-omics (left) and single-cell omics (right) approaches are shown, resulting in contrasting representations of a natural microbial community. Different genomes are represented by different colors, red, orange and yellow show genomes of close relatives, i.e. intra-population genomic variants. Dead cells are shown in grey, and extracellular or host DNA in brown. Different functions are represented by different symbols (triangles, squares, diamonds, or ovals). (a) After meta-omics sampling, the cell proportions are maintained but transient DNA and dead cells are not filtered. (b) In single-cell omics, a smaller proportion of the community is sampled, and dead cells can be excluded. (c) In meta-omics, the unit sequenced is the complete extracted sample. Metagenome assembled genomes (MAGs) are partial and include chimeras, i.e. unreal collages of closely related genomes, dead cells, and extracellular material. Meta-transcriptome analysis yields averaged relative abundances (represented by the size of the symbol) of functions within the sample. (d) By contrast, in single-cell omics, each cell is a sequenced unit and can be associated with its genome and/or transcriptome. (e) The community observed through meta-omics is representative of the composition of the whole community but not of the associated genes and functions. (f) With genome and transcriptome information from single cells, the observed community is under-sampled but is closer to the natural community: if the sampling scale is appropriate, rare populations and functions are more likely to be detected.