

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.

1	Contribution of single-cell omics to microbial ecology
2	Authors: Mauger, S. ^{1,2} , Monard, C. ¹ , Thion, C. ² , Vandenkoornhuyse, P. ¹
3	¹ Université de Rennes 1, CNRS, UMR 6553 ECOBIO, 35042 Rennes (France)
4	² Cellenion, 60 Avenue Rockefeller, 69008 Lyon (France)
5	Corresponding authors : Solène Mauger (solene.mauger@univ-rennes1.fr); Philippe
6	Vandenkoornhuyse (philippe.vandenkoornhuyse@univ-rennes1.fr)
7	
8	Keywords: Microbiota; single cell; omics; population; scales
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	

19 Highlights

20 Microbes are involved in many ecosystems but remain understudied mostly due to 21 technical limitations.

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

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

26 Single microbial cell omics enable to assess and explore the dynamics of genetic 27 changes from individuals to higher levels of ecological complexity.

28

29 Abstract

Microorganisms play key roles in various ecosystems but many of their functions and 30 interactions remain undefined. To investigate the ecological relevance of microbial 31 32 communities, new molecular tools are being developed. Among them, single-cell 33 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 34 access to a bigger range of ecological scales (from individual to community) than 35 culture-based approaches and meta-omics, single-cell omics can contribute to 36 37 microorganisms genomes and functions identification but also to the testing of concepts in ecology. Here, we discuss the contribution of single-cell omics to 38 39 possible breakthroughs in concepts and knowledge on microbial ecosystems and 40 eco-evolutionary processes.

41 Main Text

42 Ecological scales

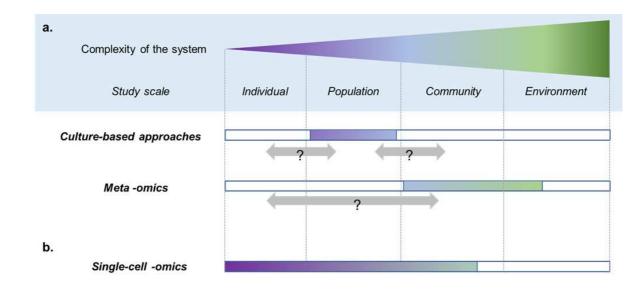
43 Interactions between organisms take place at all organizational levels, from molecules to communities, within or between species, and shape ecosystem 44 dynamics. Ecological interactions are difficult to understand due to the number of 45 biotic and abiotic parameters involved. Assembling knowledge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

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 defined entities. Interactions and diversity at the populational level, i.e. between 60 individual cells of the same population (Box 1), is still obscure since it is not often 61 62 analyzed in environmental microbiology. Given the natural mutation rate in bacteria (~10⁻⁷ substitution per nucleotide, e.g. [5]), even a single colony contains genetic 63 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).

71 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 72 73 remains a frontier in microbial ecology. Beyond the technical limits that microbiology is facing, the information gathered from culture-based studies or from natural 74 75 ecosystems can be difficult to interpret (Figure 1a). Laboratory experiments attempt to reproduce optimal ecological conditions for microbes by selecting among the 76 many biotic and environmental parameters [6] in order to understand specific 77 78 processes such as trait tradeoffs [9], interactions between strains [10], the production 79 of metabolites [11] or genome evolution [12]. Extrapolating observations obtained in 80 vitro, at restricted scales, to higher ecological scales such as natural communities 81 and ecosystems requires particular attention. Conversely, observations made from environmental samples, including microbial community composition, diversity or 82 global functions are less specific and represent an average of the microbial 83 84 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 85 coupled with global function and diversity of a community with meta-omics). 86 However, our ignorance of intricacies and interactions of ecological scales with one 87 another still jeopardize the assembly of the resulting information to answer specific 88 ecological questions in environmental microbiology. Moreover, the uncertainty of 89

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.



93

94 Figure 1: Gradient of ecological complexity, study scales and associated95 approaches.

96 Although other higher levels of ecology can be used, individual-populationcommunity-environment scales describe much of the subject of ecology. a) The 97 culture-based approaches aim at studying microbial populations or a very limited 98 number of strains, therefore the outcoming data cannot be fully informative about 99 100 higher levels of ecology. On the contrary, meta-omics approaches cover from the 101 community to environmental scales of microbial ecology and does not provide 102 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 103 104 cover the scales from the individual microbe to the community from the same environmental sample which allows to connect the outcoming information of each 105 106 ecological scale.

108 Specific tools for specific ecological questions

109 Like in any new field of exploration, ecological patterns within microbial communities are firstly observed and described but poorly understood [13], testifying to the 110 111 enormous lack of knowledge concerning microbes [14]. The use of DNA- and RNAbased methods to study natural microbial communities have demonstrated the 112 existence of a prodigious wealth of microorganisms which remained unsuspected 113 114 some years ago (e.g. [15]). Among meta-omics techniques, metagenomics and metatranscriptomics are the most widely used methods to explore microbiota. These 115 116 techniques enabled a breakthrough in our understanding of microbial phylogenetic 117 relationships [16], species diversity and abundance [17], metabolic abilities [18] and functional diversification [19]. The development and use of metagenome-118 assembled genomes (MAGs) led to discoveries that advanced our understanding 119 120 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 example, Crits-Christoph et al. [20] investigated genetic variation within populations 122 123 of highly abundant soil bacteria by studying MAGs and observed spatial differentiation of alleles. However, inferring population features from meta-omics 124 data remains limited, especially when genomes are inaccurately or incompletely 125 reconstructed from short sequence fragments (Figure 2, Key Figure). The use of 126 127 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

131 populations. In this case, the approach would necessarily conceal a considerable proportion of molecular diversity [22]. In addition, the molecular biology and 132 133 bioinformatics methods used in meta-omics approaches are varied and based on 134 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 135 136 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 137 138 regulation. Meta-omics approaches proved useful in describing communities using 139 large-scale sampling and make it possible to answer questions related to community 140 composition and its associated global functions but not to fully understand the 141 mechanisms underlying these patterns. Nevertheless, bioinformatics research has 142 developed algorithms aiming to identify genetic variations in microbial populations: vertically and horizontally inherited genes can be differentiated, and from population-143 144 specific sweeps, single nucleotide polymorphisms can be detected (e.g. DiscoSNP; 145 PopCOGenT) [4, 25].

To complement meta-omics data, modelling approaches are used to explore 146 microbial interactions, fluxes of metabolites, and to reconstruct ecological networks 147 in complex microbiomes [26-29]. These approaches provide a possible explanation 148 and scenarios of interactions in natural communities, but sometimes end in 149 150 contradictions with culture-based experiments [1]. Indeed, models can predict a 151 certain kind of interaction (e.g. cross-feeding) that is not verified or proven wrong in 152 an experimental setup due to the oversight of key parameters such as growing 153 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 154 155 relationships. Culture-dependent approaches may help reach the population level in

156 simple community compositions through controlled and simplified laboratory-scale 157 experiments [30] and can be effective for testing patterns observed in meta-omics studies and deconstructing mechanistic hypotheses (e.g. interactions, metabolism) 158 159 (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 160 161 observed processes. Many studies aim at unravelling microbiota diversity, primarily in plants, soil, water, and animal bodies, but few explain associated community 162 assembly and evolutionary mechanisms [32-34]. In this context, microbiologists and 163 164 ecologists are searching for other technical possibilities or approaches such as single-cell omics to complete the knowledge provided by current methods. 165

166 The alternative scope of single -cell omics

Single-cell whole genome sequencing (scWGS) and single-cell transcriptomics 167 (scRNA-seq) were first developed for eukaryotic cells and used in cancer research, 168 169 revealing both intra-population genetic diversity and heterogeneous genome 170 expression. As in cancer research where differentiation in space of the genome expression among cell has been observed [35], a pioneer paper on Pseudomonas 171 172 aeruginosa biofilm using a fluorescent based approach (i.e. parallel sequential fluorescence in situ hybridization (par-seqFISH)), revealed a differentiation in space 173 and time of cell-expression [36]. Because spatial single-cell microbial approaches 174 could allow the understanding of the drivers and mechanisms leading to the self-175 176 organization of these microbial structures, new developments are expected to 177 expand in health science and many other fields of microbial ecology research. Single-cell approaches are promising candidates for microbial studies, as they 178

provide a complementary view to metagenomics and metatranscriptomics that havedifferent strengths, but also weaknesses (Figure 2).

181 Single-cell omics technologies require additional steps to prepare a sample for sequencing as compared to meta-omics techniques, especially with regards to cell 182 isolation, for which different technical options are available [37,38]. Once the cells 183 184 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 used approach for bacteria [39]). 188

189 Single-cell approaches enable accurate access to genomic and transcriptomic 190 information for each cell, so that the assembled cell information is highly representative of the original population (Figure 2). This enables the identification of 191 192 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 the individual gene expression profiles of genetically close cells may differ. What is 196 more, some cell isolation tools such as automated image-based isolation devices 197 198 (CellenONE Cellenion; icell8 Takara) make it possible to select cells based on their integrity, their physiology and/or functional markers, and to minimize contamination 199 200 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.

203 A seminal paper on single cell microbial genome analysis was published in 2005 [40] 204 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 205 206 omics have improved our understanding of intraspecific diversification and metabolic capacities at a limited scale [41, 42]. Assessing the true individual cell gene 207 208 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 209 210 microbiome by linking the different ecological scales (Figure 1b). Indeed, single-cell -211 omics enable access to additional information than in culture-based and meta-omics 212 studies (the single individual ecological scale) whilst also, from the environmental 213 sample, giving information on the population and community interactions. To a 214 broader extent, this will enable better access to eco-evolutionary pressures and evolutionary processes within microbial communities. 215

216 Applications of single-cell omics in microbial ecology

217 Single-cell omics provide information at the cell level by changing the camera angle 218 when studying environmental communities and can contribute to microbiology and 219 ecology at many levels by exploring microbial diversity or microbial interactions.

220 The tremendous diversity of single-microbe genomes.

Observations of microbiota can complement/validate the diversity observed by metaomics on fungi [43], human samples [8], and marine viruses [44], and resolve cryptic bacterial species, which currently mainly rely on cultivable strains [45-47]. Singlemicrobe -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

227 possible new branches. Isolating cells from environmental samples can also cast 228 light on rare organisms that might be obscured in millions of genome fragments 229 using meta-omics. These rare organisms are considered to play key roles in 230 community dynamics since they over-proportionally contribute to the functions of the microbial community in fluctuating environments [49,50]. While most studies that 231 232 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 233 234 a high degree of uniqueness and limited clonality in the analyzed samples of sea 235 water and providing evidence for the ecological roles of uncultured microbial groups. Single-cell genomics, by looking at individual genomes instead of core genomes 236 237 from meta-omics, from natural microbial communities represent an unprecedented 238 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 239 and why to formulate hypotheses in ecology and better understand processes 240 241 involving microbes.

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

243 Single-cell genomics and transcriptomics approaches therefore help to answer the questions "What are these microbes?" and "What are they doing (or capable of 244 doing)?" but also to understand why and how observed patterns happen. Linking 245 environmental and community parameters to individual gene expression and 246 247 bacterial interactions enables a mechanistic understanding of underlying biotic and 248 abiotic conditions to patterns. This represents an opportunity to explore multiple ecological theories and hypotheses, notably on interactions of microbes at many 249 250 levels: within the community, with external microbes (i.e. viruses) and with their host.

251 One of the hottest topics in microbial ecology is the link between diversity and 252 function, including the productivity of the ecosystem that relies on the niche 253 **partitioning** theory. This hypothesis states that species coexistence is enabled by 254 species specialization in different available resources (or combinations of resources), thereby reducing interspecific competition [53] but likely modifying the microbial 255 256 population structure [54]. Specialization in specific resources raises many questions concerning microbial interactions through the exchange of metabolites [55], loss of 257 258 traits [56] and genome reduction [57]. The Black Queen Hypothesis (BQH), one of 259 the ecological theories that conceptualized this phenomenon, states that microbial community assembly and complexity are at least partially determined by functional 260 261 dependencies resulting from gene loss(es). Testing this hypothesis requires using 262 environmental samples to evaluate functional redundancy within communities, the expression and distribution of the functions between interacting (micro-)organisms 263 264 and the impact of genotypic interactions on these functions. At the community level, 265 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 266 most of the meta-omics tools used so far. Eco-evolutionary processes can so far be 267 explicitly assessed only through ecological models [58] or from dedicated in-vitro 268 experiments [42] that require deciding which gene and which organisms to look at. 269 270 Single cell approaches can help to investigate such hypotheses and other theories in 271 ecology and highlight patterns of interactions that shape microbial communities.

272 A breach to viral-host ecology.

273 The diversity of microbial communities is also influenced by the viral infections to 274 which they are exposed, which is particularly difficult to evaluate in nature [59]. A

275 recent study assessed viral infections in the ocean thanks to single cell genomics by 276 identifying virus sequences in uncultured protist cells [44]. Such interactions are widespread but generally missing in current microbial ecology analyses. Among 277 278 other roles, viruses are known to (1) control microbial community dynamics and drive microbial-host evolution [60,61] and (2) impact ecosystem changes and 279 280 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 281 282 prevalence of prophage sequences and possible lateral gene transfers [63,64] which would reveal preferential association of specific viruses with specific bacteria. 283

284 Single-cell insights into bigger scale interactions.

285 Generally speaking, single-cell omics can highlight diverse levels of interaction in 286 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 287 composition of human microbiota, very little is known about interactions between the 288 289 microbes and with their host [65]. The single cell study of host-pathogen interactions 290 paves the way for understanding infectious processes through microbiota dynamics, 291 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 292 [67]. Among other explanations, this apparent higher stochasticity may be the result 293 of drastic pathogen-induced changes in the habitat (i.e. transitory state in the 294 295 microbiota dynamics), modification of a component of the microbiota caused by 296 genetic change(s) and/or a functional modification expressing a modified phenotype 297 that leads to disequilibrium in the microbiota community. With the aim of 298 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 299 300 analyses of genomes and transcriptomes would help better define the characteristics of dysbiosis (i.e. dysbiosis mechanisms). Using single cell genomics makes it 301 302 possible to address hypotheses related to changes in bacterial populations, while microbial single-cell transcriptomics may be more appropriate to decrypt the 303 304 functions of microbes, metabolic abilities, and cellular states [68]. These two strategies are necessary to understand how microbial interactions occur within 305 306 communities as well as their possible impact on the ecosystem [69,51].

307

308 Microbial single-cell omics are not solely expected to better understand functional 309 interactions but also the underlying evolutionary processes. Microbial single-cell omics should also promote a shift in standpoint from observation to interpretation 310 and also offer new opportunities to test macroecological theories on microbes. The 311 312 development of microbial single-cell omics will have high impacts on our 313 understanding of microbial communities in many environments. For instance (i) in 314 freshwater and marine ecosystems, to define the interaction of bacteria and 315 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 316 317 nutrient and water uptake and protection against pathogens [71], (iii) in plant, human 318 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 319 320 omics applied to microorganisms is limited by technical limits.

321

322 Limits of single-cell omics on microbes

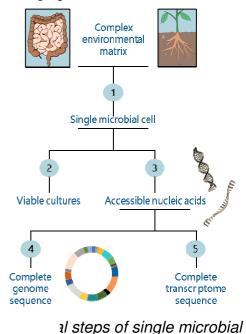
323 The current limited number of cells studied in the published papers questions the 324 representativeness of the analyses. Considering the number of microbial cells 325 contained in a given environmental sample, one can wonder how many cells need to 326 be isolated to cover the diversity of a sample, from hundreds [41], to thousands [51]. The limited number of analyzed cells is mostly due to technical problems and the 327 328 cost of such experiments. It has to be emphasized that the current use of single-cell omics for microbes have to solve many technical obstacles (Box 2), reviewed in 329 330 [72,73] such as cell isolation, lysis, and a biased amplification step. The structure of the microbial cell wall is complex, and unlike the animal cells, they do not break 331 easily. The diversity of cell wall composition across phylogeny and physiological 332 333 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 inhibiting enzymatic reactions downstream. Different protocols have been used in 335 recent microbial single-cell based studies, using either heat, temperature shocks, 336 337 sonication, enzymes, detergents, or combinations of these [74,75]. In addition to technical issues, one can wonder how to be certain that cells are isolated and lysed 338 equally and not preferentially depending on their physical/physiological status. The 339 340 amplification step, usually made via MDA, has been reported to be imprecise concerning the genome amplification uniformity, even though it presents a better 341 genome coverage than other approaches (i.e MALBAC [39]). This amplification step 342 is highly relevant, as it was suggested to be the cause of incomplete reconstruction 343 of single amplified genomes (SAGs) [76], although solutions are being developed 344 [77]. As the price of library preparation represents most of the cost of these new 345 single- cell omics for microorganisms, reduction of reaction volumes in 346 "nanolibraries" should be very cost-effective (Box 2). The probability of contamination 347

348 decreases with the miniaturization of the reagent volumes [78] and associated 349 robotics. Working in nanovolumes seems to be a convenient solution to solve 350 multiple problems, however this introduces new volume-related challenges such as 351 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 singlecell 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).

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

Box 2: Single microbial cell -omics approaches are tech		
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 + 父	
2. Culturomics	experiments	
Maximum viability/cultivability	Gentle cell handling, liquid dispensing	
Choice of culture conditions	High throughput media screening + 𝔍	
Assessment of monoclonality (= culture purity)	Microscopy, targeted sequencing	
3. Microbial cell lysi	s/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 + 父	
4. Single-microbe WGS		



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 + 父	
Cost reduction	Cell barcoding for multiplexing + $igvee$	
Bioinformatics	Dedicated tools for cell demultiplexing, monoclonality test	
5. Single-microbe RNA-seq		
Same as WGS (up to 100 fg RNA prokaryotic cell ⁻¹)	𝔍, 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	

360 **Table i.** Padlocks and possible keys offered by single microbial cell –omics 361 approaches.

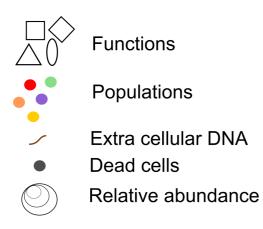
Crucial, but solvable, issues (Table i) should be addressed at each step of single microbe approaches (Fig. i). The reduction of reaction volumes down to "nanovolumes" (see \textcircled symbols), is often the most sensible solution, limiting contamination probability and allowing high throughput and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

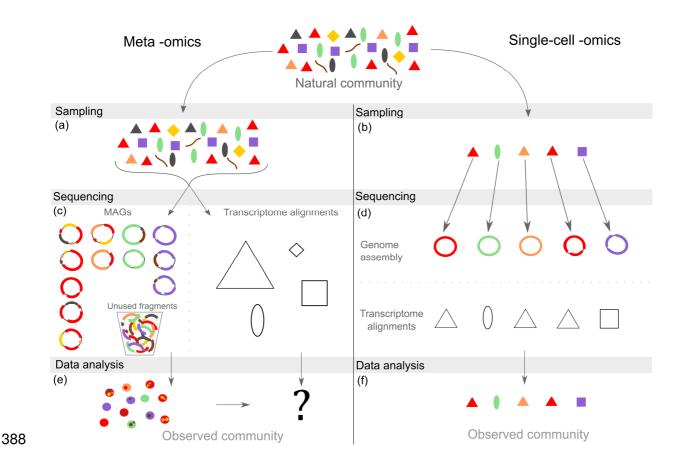
366 Concluding Remarks

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 374 and eco-evolutionary boundaries through the enabling of mechanistic characterization of deterministic populations and community assembly processes. 375 Currently, the use of metagenomics and single cell genomics in the same study 376 377 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 378 379 and/or scRNAseg [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 380 single-cell approaches. Approaches that will allow more accurate assessment of 381 microbial genome diversity and genome functioning within complex microbiota, are 382 impatiently awaited. Still, the future of microbial single-cell omics will likely fuel a new 383 384 perception of the microorganisms world.

385 Acknowledgements

- We would like to thank LABCOM 'Microscale' for funding.
- 387 Key figure





389 Outstanding Questions

390 To what extent can a single technical approach realistically reflect complex natural 391 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?

395 What criteria should be used to determine the scale of sampling in natural 396 environments?

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

How can we make the application of proteomics, metabolomics and multi-omicsapproaches 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**

419 In ecology, populations are individuals belonging to the same species, living in 420 the same environment, although the definition varies with different viewpoints [81]. Microbial populations represent a unit of diversity and selection. Within 421 422 these populations, diversity can be either genetic or phenotypic. The diversity within a population to some extent buffers an environmental stress because 423 424 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 425 426 also be induced by the stress. This organizational level is therefore a key to 427 understanding genetic structure and haplotype fitness, the dynamics of ecological interactions, including associations of microbial species, symbioses, host-428 429 pathogen interactions, and ecosystem functioning, resilience, and stability. For 430 instance, resistance to antibiotics can vary within populations [83] and the virulence pathogens can vary across subpopulations [84]. Genetic diversity and 431 432 ecological features such as niches can vary between lineages [85], so that 433 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 434 must thus be tuned down to the individuals that comprise the population in a 435 given sample [84]. Otherwise, applying the current bacterial species concept to 436 make population level inferences may lead to false or partial interpretation of 437 438 ecological phenomena.

439 **Reference list**

440 1- Coyte, K.Z., and Rakoff-Nahoum, S. (2019) Understanding competition
441 and cooperation within the mammalian gut microbiome. *Curr. Biol.* 29(11),
442 538–544. DOI: https://doi.org/10.1016/j.cub.2019.04.017

443	2-	Pacheco, A.R., and Segrè, D. (2019) A multidimensional perspective on
444		microbial interactions. FEMS Microbiol. Lett. 366(11), 1–11. DOI:
445		https://doi.org/10.1093/femsle/fnz125
446	3-	Dini-Andreote, F. et al. (2020) Towards meaningful scales in ecosystem
447		microbiome research. Environ. Microbiol. 23(1), 1-4. DOI:
448		https://doi.org/10.1111/1462-2920.15276
449	4-	VanInsberghe, D. et al. (2020) How can microbial population genomics
450		inform community ecology? Phil. Trans. R. Soc. B. 375(1798), 20190253.
451		DOI: https://dx.doi.org/10.1098%2Frstb.2019.0253
452	5-	Westra, E.R. et al. (2017) Mechanisms and consequences of diversity-
453		generating immune strategies. Nat Rev Immunol, 17(11), 719-728. DOI:
454		https://doi.org/10.1038/nri.2017.78
455	6-	García-García, N. et al. (2019) Microdiversity ensures the maintenance of
456		functional microbial communities under changing environmental
457		conditions. ISME J. 13(12), 2969-2983. DOI :
458		https://doi.org/10.1038/s41396-019-0487-8
459	7-	Niccum, B.A. et al. (2020) Strain-level diversity impacts cheese rind
460		microbiome assembly and function. <i>mSystems</i> 5(3) e00149-20. DOI:
461		https://doi.org/10.1128/msystems.00149-20
462	8-	Marcy, Y. et al. (2007) Dissecting biological "dark matter" with single-cell
463		genetic analysis of rare and uncultivated TM7 microbes from the human
464		mouth. PNAS 104(29), 11889–11894. DOI:
465		https://doi.org/10.1073/pnas.0704662104

466	9- Moreno-Gámez, S. et al. (2020) Wide lag time distributions break a trade-
467	off between reproduction and survival in bacteria. PNAS 117(31), 18729-
468	18736. DOI: https://doi.org/10.1073/pnas.2003331117
469	10-Ratzke, C. et al. (2020) Strength of species interactions determines
470	biodiversity and stability in microbial communities. Nat. Ecol. Evol. 4(3),
471	376-383. DOI: <u>https://doi.org/10.1038/s41559-020-1099-4</u>
472	11-Rütting, T. et al. (2021) The contribution of ammonia-oxidizing archaea
473	and bacteria to gross nitrification under different substrate availability. Soil
474	Biol. Biochem. 160, 108353. DOI:
475	https://doi.org/10.1016/j.soilbio.2021.108353
476	12-Hall, J. P. et al. (2018) Competitive species interactions constrain abiotic
477	adaptation in a bacterial soil community. Evol. Lett. 2(6), 580-589. DOI:
478	https://doi.org/10.1002/evl3.83
479	13-Tripathi, A. et al. (2018) Are microbiome studies ready for hypothesis-
480	driven research? Curr. Opin. Microbiol. 44, 61–69. DOI:
481	https://doi.org/10.1016/j.mib.2018.07.002
482	14-Lloyd, K.G. et al. (2018) Phylogenetically novel uncultured microbial cells
483	dominate earth microbiomes. <i>mSystems</i> 3(5) e00055-18. DOI:
484	https://doi.org/10.1128/msystems.00055-18
485	15-Hug, L.A. et al. (2016) A new view of the tree of life. Nat. Microbiol. 1(5),
486	16048. DOI: https://doi.org/10.1038/nmicrobiol.2016.48
487	16-Alteio, L.V. et al. (2020) Complementary metagenomic approaches
488	improve reconstruction of microbial diversity in a forest
489	soil. <i>mSystems</i> , <i>5</i> (2), e00768-19. DOI:
490	https://doi.org/10.1128/mSystems.00768-19

491	17-Daims, H. et al. (2015) Complete nitrification by Nitrospira bacteria. Nature
492	528(7583), 504. DOI: https://doi.org/10.1038/nature16461
493	18-Cernava, T. et al. (2017) Deciphering functional diversification within the
494	lichen microbiota by meta-omics. <i>Microbiome</i> 5(1), 82. DOI:
495	https://doi.org/10.1186/s40168-017-0303-5
496	19- Parks, D.H. et al. (2017) Recovery of nearly 8,000 metagenome-
497	assembled genomes substantially expands the tree of life. Nat. Microbiol.
498	2(11), 1533-1542. DOI: <u>https://doi.org/10.1038/s41564-017-0012-7</u>
499	20-Crits-Christoph, A. et al. (2020) Soil bacterial populations are shaped by
500	recombination and gene-specific selection across a grassland meadow.
501	ISME J. 14(7), 1834-1846. DOI: https://doi.org/10.1038/s41396-020-0655-
502	X
503	21-Nayfach, S. et al. (2019) New insights from uncultivated genomes of the
504	global human gut microbiome. Nature 568(7753), 505-510. DOI:
505	https://doi.org/10.1038/s41586-019-1058-x
506	22-Nelson, W.C. et al. (2020) Biases in genome reconstruction from
507	metagenomic data. PeerJ 8, e10119. DOI:
508	https://doi.org/10.7717/peerj.10119
509	23-Li, L. et al. (2015) Comparing viral metagenomics methods using a highly
510	multiplexed human viral pathogens reagent. J. Virol. Methods, 213, 139-
511	146. DOI: https://doi.org/10.1016/j.jviromet.2014.12.002
512	24-Mande, S.S. et al. (2012) Classification of metagenomic sequences:
513	methods and challenges. Brief. Bioinform. 13(6), 669-681. DOI:
514	https://doi.org/10.1093/bib/bbs054

515	25-Uricaru, R. et al. (2015) Reference-free detection of isolated
516	SNPs. Nucleic Acids Res. 43(2), e11-e11. DOI:
517	https://doi.org/10.1093/nar/gku1187
518	26-Antoniewicz, M.R. (2020) A guide to deciphering microbial interactions and
519	metabolic fluxes in microbiome communities. Curr. Opin. Biotechnol. 64,
520	230–237. DOI: https://doi.org/10.1016/j.copbio.2020.07.001
521	27-Belcour, A. et al. (2020) Metage2Metabo, microbiota-scale metabolic
522	complementarity for the identification of key species. <i>eLife 9</i> , e61968. DOI:
523	https://doi.org/10.7554/eLife.61968
524	28-Gorter, F.A. et al. (2020) Understanding the evolution of interspecies
525	interactions in microbial communities. Philos. Trans. R. Soc. Lond., B, Biol.
526	Sci. 375(1798), 20190256. DOI: https://doi.org/10.1098/rstb.2019.0256
527	29-Perez-Garcia, O. et al. (2016) Metabolic network modeling of microbial
528	interactions in natural and engineered environmental systems. Front.
529	Microbiol. 7, 673. DOI: https://doi.org/10.3389/fmicb.2016.00673
530	30-Morin, M. et al. (2018) Changes in the genetic requirements for microbial
531	interactions with increasing community complexity. eLife 7, e37072. DOI:
532	https://doi.org/10.7554/eLife.37072
533	31-Beaumont, H. J. E. et al. (2009) Experimental evolution of bet hedging.
534	Nature 462(7269), 90–93. DOI: https://doi.org/10.1038/nature08504
535	32-Kiers, E.T. et al. (2011) Reciprocal rewards stabilize cooperation in the
536	mycorrhizal symbiosis. Science 333(6044), 880–882. DOI:
537	https://doi.org/10.1126/science.1208473

538	33-Werner, G.D., and Kiers, E.T. (2015) Partner selection in the mycorrhizal
539	mutualism. New Phytol. 205(4), 1437-1442. DOI:
540	https://doi.org/10.1111/nph.13113
541	34-Amin, S.A. et al. (2015) Interaction and signaling between a cosmopolitan
542	phytoplankton and associated bacteria. Nature 522(7554), 98–101. DOI:
543	https://doi.org/10.1038/nature14488
544	35-Alon, S. et al. (2021) Expansion sequencing: Spatially precise in situ
545	transcriptomics in intact biological systems. Science 371(6528). DOI:
546	10.1126/science.aax2656
547	36-Dar, D. et al. (2021) Spatial transcriptomics of planktonic and sessile
548	bacterial populations at single-cell resolution. Science 373(6556). DOI:
549	10.1126/science.abi4882
550	37-Gawad, C. et al. (2016) Single-cell genome sequencing: current state of
551	the science. Nat. Rev. Genet. 17(3), 175. DOI:
552	https://doi.org/10.1038/nrg.2015.16
553	38-Woyke, T. et al. (2017) The trajectory of microbial single-cell sequencing.
554	Nat. Methods 14(11), 1045. DOI: https://doi.org/10.1038/nmeth.4469
555	39-Chen, Z. et al. (2017) Tools for genomic and transcriptomic analysis of
556	microbes at single-cell level. Front Microbiol. 8, 1831. DOI:
557	https://doi.org/10.3389/fmicb.2017.01831
558	40-Raghunathan, A. et al. (2005) Genomic DNA amplification from a single
559	bacterium. Appl. Environ. Microbiol. 71(6), 3342-3347. DOI:
560	https://doi.org/10.1128/AEM.71.6.3342-3347.2005

561	41-Engel, P. et al. (2014) Hidden diversity in honey bee gut symbionts
562	detected by single-cell genomics. PLoS Genet. 10(9), e1004596. DOI:
563	https://doi.org/10.1371/journal.pgen.1004596
564	42-Hennon, G.M. et al. (2018) The impact of elevated CO2 on
565	Prochlorococcus and microbial interactions with 'helper' bacterium
566	Alteromonas. ISME J. 12(2), 520–531. DOI:
567	https://doi.org/10.1038/ismej.2017.189
568	43-Ahrendt, S.R. et al. (2018) Leveraging single-cell genomics to expand the
569	fungal tree of life. Nat. Microbiol. 3(12), 1417-1428. DOI:
570	https://doi.org/10.1038/s41564-018-0261-0
571	44-Castillo, Y.M. et al. (2019) Assessing the viral content of uncultured
572	picoeukaryotes in the global ocean by single cell genomics. Mol. Ecol.
573	28(18), 4272–4289. DOI: <u>https://doi.org/10.1111/mec.15210</u>
574	45-Hahn, M.W. et al. (2016) Complete ecological isolation and cryptic
575	diversity in Polynucleobacter bacteria not resolved by 16S rRNA gene
576	sequences. ISME J. 10(7), 1642–1655. DOI:
577	https://doi.org/10.1038/ismej.2015.237
578	46-Hoetzinger, M. et al. (2019) Polynucleobacter paneuropaeus sp. nov.,
579	characterized by six strains isolated from freshwater lakes located along a
580	3000 km north-south cross-section across Europe. Int. J. Syst. Evol.
581	Microbiol. 69(1), 203–213. DOI: https://doi.org/10.1099/ijsem.0.003130
582	47-Pitt, A. et al. (2019) Aquírufa antheringensis gen nov., sp nov. and
583	Aquirufa nivaíisiívae sp nov., representing a new genus of widespread
584	freshwater bacteria. Int. J. Syst. Evol. Microbiol. 69(9), 2739–2749. DOI:
585	https://doi.org/10.1099/ijsem.0.003554

586	48-Bernard, G. et al. (2018) Microbial dark matter investigations: how
587	microbial studies transform biological knowledge and empirically sketch a
588	logic of scientific discovery. Genome Biol. Evol. 10(3), 707-715. DOI:
589	https://doi.org/10.1093/gbe/evy031
590	49-Shade, A. et al. (2014) Conditionally rare taxa disproportionately contribute
591	to temporal changes in microbial diversity. <i>MBio</i> 5(4), 1371–1385. DOI:
592	https://doi.org/10.1128/mBio.01371-14
593	50-Xu, Q. et al. (2021) Rare Bacterial Assembly in Soils Is Mainly Driven by
594	Deterministic Processes. Microb. Ecol. 1–14. DOI:
595	https://doi.org/10.1007/s00248-021-01741-8
596	51-Pachiadaki, M.G. et al. (2019) Charting the complexity of the marine
597	microbiome through single-cell genomics. Cell 179(7), 1623-1635.e11.
598	DOI: https://doi.org/10.1016/j.cell.2019.11.017
599	52-Prosser, J.I. (2020) Putting science back into microbial ecology: a question
600	of approach. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 375(1798),
601	20190240. DOI: https://doi.org/10.1098/rstb.2019.024
602	53-Mas, A. et al. (2020) Reflections on the predictability of evolution: towards
603	a conceptual framework. iScience, 23(11), 101736. DOI:
604	https://doi.org/10.1016/j.isci.2020.101736
605	54-Mas, A. et al. (2016) Beyond the Black Queen Hypothesis. ISME J, 10(9),
606	2085–2091. DOI : https://doi.org/10.1038/ismej.2016.22
607	55-Zelezniak, A. et al. (2015) Metabolic dependencies drive species co-
608	occurrence in diverse microbial communities. PNAS 112(20), 6449–6454.
609	DOI: https://doi.org/10.1073/pnas.1421834112

610	56- Ellers, J. et al. (2012) Ecological interactions drive evolutionary loss of
611	traits. Ecol. Lett. 15(10), 1071–1082. DOI: https://doi.org/10.1111/j.1461-
612	<u>0248.2012.01830.x</u>
613	57-Morris, J.J. et al. (2012) The Black Queen Hypothesis: evolution of
614	dependencies through adaptive gene loss. MBio, 3(2) e00036-12. DOI:
615	10.1128/mBio.00036-12
616	58-Wang, M. et al. (2020) Selfishness driving reductive evolution shapes
617	interdependent patterns in spatially structured microbial communities.
618	ISME J. 15(5), 1387-1401. DOI: https://doi.org/10.1038/s41396-020-
619	<u>00858-x</u>
620	59-Weitz, J.S., and Wilhelm, S.W. (2012) Ocean viruses and their effects on
621	microbial communities and biogeochemical cycles. F1000 Biol. Rep. 4(1),
622	17. DOI: https://doi.org/10.3410/B4-17
623	60-Frada, M. et al. (2008) The "Cheshire Cat" escape strategy of the
624	coccolithophore Emiliania huxleyi in response to viral infection. PNAS
625	105(41), 15944–15949. DOI: <u>https://doi.org/10.1073/pnas.0807707105</u>
626	61-Scanlan, P.D., and Buckling, A. (2012) Co-evolution with lytic phage
627	selects for the mucoid phenotype of <i>Pseudomonas fluorescens</i> SBW25.
628	ISME J. 6(6), 1148–1158. DOI: https://doi.org/10.1038/ismej.2011.174
629	62-Rousk, J., and Bengtson, P. (2014) Microbial regulation of global
630	biogeochemical cycles. Front. microbiol. 5, 103. DOI:
631	https://doi.org/10.3389/fmicb.2014.00103
632	63-Roux, S. et al. (2014) Ecology and evolution of viruses infecting
633	uncultivated SUP05 bacteria as revealed by single-cell-and meta-
634	genomics. <i>eLife</i> , <i>3</i> , e03125. DOI: <u>https://doi.org/10.7554/eLife.03125</u>

635	64-Labonté, J.M. et al. (2015) Single-cell genomics-based analysis of virus-
636	host interactions in marine surface bacterioplankton. ISME J. 9(11), 2386-
637	2399. DOI: https://doi.org/10.1038/ismej.2015.48
638	65-Hassani, M.A. et al. (2018) Microbial interactions within the plant holobiont.
639	Microbiome 6(1)58. DOI: https://doi.org/10.1186/s40168-018-0445-0
640	66-Mills, E. and Avraham, R. (2017) Breaking the population barrier by single
641	cell analysis: one host against one pathogen. Curr. Opin. Microbiol. 36, 69-
642	75. DOI: https://doi.org/10.1016/j.mib.2017.01.005
643	67-Zaneveld, J.R. et al. (2017) Stress and stability: Applying the Anna
644	Karenina principle to animal microbiomes. Nat. Microbiol. 2(9),1-8. DOI:
645	https://doi.org/10.1038/nmicrobiol.2017.121
646	68-Hatzenpichler, R. et al. (2020) Next-generation physiology approaches to
647	study microbiome function at single cell level. Nat Rev Microbiol. 18(4),
648	241-256. DOI: https://doi.org/10.1038/s41579-020-0323-1
649	69-McNulty, R. et al. (2021) Droplet-based single cell RNA sequencing of
650	bacteria identifies known and previously unseen cellular states. BioRxiv
651	2021.03.10.434868. DOI: https://doi.org/10.1101/2021.03.10.434868
652	70-Landa, M. et al. (2016) Shifts in bacterial community composition
653	associated with increased carbon cycling in a mosaic of phytoplankton
654	blooms. <i>ISME J</i> . 10(1), 39–50. DOI:
655	https://doi.org/10.1038/ismej.2015.105
656	71-Vandenkoornhuyse, P. et al. (2015) The importance of the microbiome of
657	the plant holobiont. New Phytol. 206(4), 1196-1206. DOI:
658	https://doi.org/10.1111/nph.13312

659	72-Kaster, A.K., and Sobol, M.S. (2020) Microbial single-cell omics: the crux
660	of the matter. Appl. Microbiol. Biotechnol. 104(19),8209-8220. DOI:
661	https://doi.org/10.1007/s00253-020-10844-0
662	73-Clingenpeel, S. et al. (2015) Reconstructing each cell's genome within
663	complex microbial communities - dream or reality? Front. Microbiol. 5, 771.
664	DOI: https://doi.org/10.3389/fmicb.2014.00771
665	74-He, J. et al. (2016) Improved lysis of single bacterial cells by a modified
666	alkaline-thermal shock procedure. Biotechniques 60(3), 129-135. DOI:
667	10.2144/000114389
668	75-Liu, Y. et al. (2018) The development of an effective bacterial single-cell
669	lysis method suitable for whole genome amplification in microfluidic
670	platforms. <i>Micromachines 9</i> (8), 367. DOI:
671	https://doi.org/10.3390/mi9080367
672	76-Alneberg, J. et al. (2018) Genomes from uncultivated prokaryotes: a
673	comparison of metagenome-assembled and single-amplified genomes.
674	<i>Microbiome 6</i> (1), 1-14. DOI: <u>https://doi.org/10.1186/s40168-018-0550-0</u>
675	77-Kogawa, M. et al. (2018) Obtaining high-quality draft genomes from
676	uncultured microbes by cleaning and co-assembly of single-cell amplified
677	genomes. <i>Sci Rep. 8</i> (1), 1-11. DOI: <u>https://doi.org/10.1038/s41598-018-</u>
678	<u>20384-3</u>
679	78-Blainey, P.C. (2013) The future is now: single-cell genomics of bacteria
680	and archaea. FEMS Microbiol. Rev. 37(3), 407-427. DOI:
681	https://doi.org/10.1111/1574-6976.12015
682	79-Probst, A.J. et al. (2018) Differential depth distribution of microbial function
683	and putative symbionts through sediment-hosted aquifers in the deep

r
ond.,
rical
е
DOI:
in
C-
Vat.
2
ity
-755.
N

707 Figure 2 – Key Figure: Microbial communities observed through meta-

708 omics versus single-cell omics

709 The key steps of meta-omics (left) and single-cell omics (right) approaches are 710 shown, resulting in contrasting representations of a natural microbial community. 711 Different genomes are represented by different colors, red, orange and yellow show genomes of close relatives, i.e. intra-population genomic variants. Dead 712 713 cells are shown in grey, and extracellular or host DNA in brown. Different 714 functions are represented by different symbols (triangles, squares, diamonds, or 715 ovals). (a) After meta-omics sampling, the cell proportions are maintained but 716 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 717 718 meta-omics, the unit sequenced is the complete extracted sample. Metagenome 719 assembled genomes (MAGs) are partial and include chimeras, i.e. unreal 720 collages of closely related genomes, dead cells, and extracellular material. Meta-721 transcriptome analysis yields averaged relative abundances (represented by the size of the symbol) of functions within the sample. (d) By contrast, in single-cell 722 723 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 724 representative of the composition of the whole community but not of the 725 726 associated genes and functions. (f) With genome and transcriptome information 727 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 728 729 functions are more likely to be detected.