



Sizing Up the Uncultured Microbial Majority

 Laura A. Hug^a

^aDepartment of Biology, University of Waterloo, Waterloo, Ontario, Canada

ABSTRACT Predicting the total number of microbial cells on Earth and exploring the full diversity of life are fundamental research concepts that have undergone paradigm shifts in the genomic era. In this issue, Lloyd and colleagues (K. G. Lloyd, A. D. Steen, J. L. Ladau, J. Yin, and L. Crosby, *mSystems* 3:e00055-18, <https://doi.org/10.1128/mSystems.00055-18>, 2018) present results that combine these two concepts by estimating the total diversity of all cells from Earth's environments. Leveraging publicly available amplicon, metagenomic, and metatranscriptomic datasets, they determined that nearly all environments are dominated by uncultured lineages, with the exception of humans and human-associated habitats. They define a new concept: phylogenetically diverse noncultured cells (PDNC). Unlike viable but nonculturable cells (VBNC), PDNC are microorganisms for which traditional isolation techniques may never succeed. Lloyd et al. estimate that the majority of microorganisms in Earth's ecosystems may be PDNC and conclude that culture-independent methods combined with innovative culturing techniques may be required to understand the ecology and physiology of these abundant and divergent microorganisms.

KEYWORDS amplicon sequencing, metagenomics, metatranscriptomics, microbial diversity

Attempts to characterize microbial life on Earth typically fall into two research categories. Either they seek to quantify the total number of microbial cells on the planet (see, e.g., references 1, 2, and 3) or they strive to describe the entirety of microbial diversity from phylogenetic and taxonomic perspectives (4–6). New work from Lloyd and colleagues (7), published in this issue, attempts to connect these two categories by defining the diversity of all microbial cells from available sequence data derived from Earth's diverse environments. Their results suggest that the majority of microorganisms detected within nearly all samples affiliate with divergent and uncultured lineages, despite these microbial cells being relatively abundant and likely important to ecosystem functioning within their respective source environments (7).

Lloyd and colleagues mined 16S rRNA genes, metagenomes, and metatranscriptomes from publicly available databases. Simultaneously, they counted and classified microorganisms within the selected habitats by comparing environmentally derived 16S rRNA gene sequence reads to full-length reference sequences from cultivated microorganisms in the SILVA database (8, 9). Their analyses showed that the vast majority of Earth's environments are dominated by taxa from uncultured lineages (Fig. 1). Only human and human-associated habitats (e.g., indoor and outdoor aerosols, drinking water, building materials) showed the reverse pattern, with the majority of detected microorganisms belonging to genera with cultured representatives (Fig. 1). It is striking that this global dominance of uncultured microbial taxa is consistent for systems that directly impact human activities, including livestock, agricultural soils, and engineered bioreactors. Microbial communities in systems with obvious economic importance remain uncultured to a similar extent as those in extreme environments such as hydrothermal vents. These findings highlight ongoing research bias toward

Published 25 September 2018


Citation Hug LA. 2018. Sizing up the uncultured microbial majority. *mSystems* 3:e00185-18. <https://doi.org/10.1128/mSystems.00185-18>.

Copyright © 2018 Hug. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to laura.hug@uwaterloo.ca.

For the article discussed, see <https://doi.org/10.1128/mSystems.00055-18>.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

 Sizing up the uncultured microbial majority - work by @archaearama et al. defines the phylogenetically diverse non-cultured cell (PDNC).

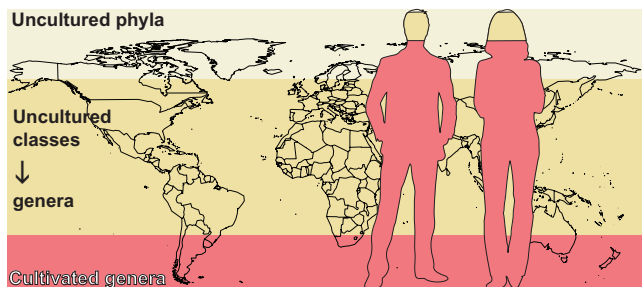


FIG 1 Graphical representation of data adapted from the work by Lloyd et al. (7) showing the proportion of uncultured cells from Earth's habitats, with an inset demonstrating that culturability is highest for humans and human-associated environments. Based on the conservative simplification of a single 16S rRNA gene per genome, Lloyd and colleagues estimate that 19% of Earth's microbial cells are from cultivated lineages (pink). The remaining 81% are cells from uncultured genera or higher levels of taxa, with 25% from uncultured phyla (cream). In contrast, microbial populations from humans are overwhelmingly from cultured genera, with only ~1% of cells belonging to uncultured phyla. (Image generated from the R package "maps" and royalty-free clipart.)

human systems and human health in microbiology research but may also reflect the comparative ease of mimicking human-associated habitats in a laboratory.

The comparison between the amplicon sequencing and metagenomic data sets identified important differences between these techniques, as well as probable biases in our current understanding of microbial diversity. As expected, amplicon sequencing data underestimated the proportion of phylogenetically divergent microorganisms compared to total community sequencing methods, with approximately 35% of amplicon sequences classified to the genus level with cultured members compared to only 15% of metagenome-derived 16S rRNA genes. This discrepancy is likely due to traditional primer design approaches favoring cultured microorganisms in the resulting amplicon sequence data. Archaea contributed a higher proportion of sequences from highly divergent lineages than bacteria in both amplicon and meta-omic data sets. From the 16S rRNA gene amplicon data, 44% of archaeal sequences were assigned to uncultured phyla compared to 17% of bacterial sequences. Our understanding of the domain Archaea is undergoing rapid expansion as culture-independent methods allow access to new environments and low-abundance microorganisms (10, 11). The results reported by Lloyd et al. indicate that the lower diversity of sequences from archaea than from bacteria in current genome databases may be, in part, a consequence of sampling bias (4).

Lloyd and colleagues made data handling decisions that impact their estimates. First, they used 16S rRNA gene sequence identity thresholds of 86% and higher to define phylum boundaries (12), which is a conservative choice compared to the thresholds of 75% to 83% and higher that have recently, and historically, been used for defining novel phyla identified through culture-independent methods (13–15). Were a lower threshold for phyla used (e.g., 80%), the proportion of detected taxa assigned to novel uncultured phyla might be much lower. This becomes a question of semantics, however, as these microorganisms would still be highly divergent and uncultured even if placed within a phylum with distantly related cultured relatives. Second, metatranscriptomic data are used to provide estimates of microbial abundance, despite the rRNA subtraction methods undoubtedly changing the abundance distribution of the organisms present (see, e.g., reference 16). Third, read depth is used to provide estimates of microbial abundance from metagenomes. This is problematic because unassembled reads of 16S rRNA genes can map spuriously to conserved regions of the 16S rRNA gene on assembled scaffolds. Most 16S rRNA genes in assembled metagenomes are on short scaffolds, with insufficient additional genomic information for an accurate average read depth to be calculated. They do not overstep these limitations in their analyses. Instead, they apply metatranscriptomic data primarily to confirm microbial viability in the environments surveyed, and they provide read depth-based

estimates of abundance side by side with the count-based data for metagenomes. With these concerns in mind, their estimates based on metagenome-derived 16S rRNA gene counts likely suffer the least bias. From all analyses, it is clear that the dominance of abundant and highly divergent microbial cells from the majority of Earth's environments is a robust finding regardless of the precise taxonomic rankings or exact numbers.

The key innovative concept introduced in this paper addresses the issue of culturability. A meta-analysis of cultivation trials from diverse environments yielded a median of 0.5% cultured cells—a new lower bound that updates the oft-cited 1% associated with the “Great Plate Count Anomaly” (17). Lloyd et al. predict that innovations in culture techniques will be successful for cultivation of viable but nonculturable cells (VBNC) (18). The hypothesis proposed by Lloyd et al. is that new cultivation methods do not succeed for the vast majority of microbial representatives from more divergent taxa. These divergent lineages may comprise as much as 95% of taxa in some environments and may require nonaxenic growth conditions to survive in a laboratory environment. Lloyd et al. coin the name and define the concept of phylogenetically divergent uncultured cells (PDNC), characterized here as representatives from the order level or higher with no cultured representatives. The PDNC represent lineages whose metabolisms may preclude growth as pure cultures, including obligate syntrophs, extreme oligotrophs, and cells with very low growth rates. In contrast to VBNC, which are dormant cells awaiting the correct conditions for growth, the PDNC may require fundamentally different techniques to be brought into culture. Lloyd et al. suggest the PDNC and VBNC can be separated by taxonomic novelty, with organisms novel at the order level or higher belonging to the PDNC. It seems unlikely this suggested boundary will hold in the future, as several studies have already reported cultivation of highly novel organisms using relatively routine methods (see, e.g., references 19 and 20). PDNC will instead be applied to those cells resisting cultivation efforts or requiring innovative new protocols to be brought into the laboratory environment.

The PDNC is an important concept with respect to analyses of environmental microbial communities, and members of the PDNC represent the current grand challenge for cultivation assays. Lloyd and colleagues conclude that culture-independent techniques and mixed enrichment cultures may be the best avenues for studying the physiology and ecology of these abundant and phylogenetically divergent cells. Their report helps refocus attention on the challenge of cultivating microorganisms because, as they state, “uncultured microbes, often from deeply phylogenetically divergent groups, dominate nonhuman environments on Earth, and their undiscovered physiologies may matter for Earth systems.”

REFERENCES

- Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* 95:6578–6583. <https://doi.org/10.1073/pnas.95.12.6578>.
- Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. 2012. Global distribution of microbial abundance and biomass in seafloor sediment. *Proc Natl Acad Sci U S A* 109:16213–16216. <https://doi.org/10.1073/pnas.1203849109>.
- Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proc Natl Acad Sci U S A* 115:6506–6511. <https://doi.org/10.1073/pnas.1711842115>.
- Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hemsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF. 2016. A new view of the tree of life. *Nat Microbiol* 1:16048. <https://doi.org/10.1038/nmicrobiol.2016.48>.
- Hinchliff CE, Smith SA, Allman JF, Burleigh JG, Chaudhary R, Coghill LM, Crandall KA, Deng J, Drew BT, Gazis R, Gude K, Hibbett DS, Katz LA, Laughinghouse HD, McTavish EJ, Midford PE, Owen CL, Ree RH, Rees JA, Soltis DE, Williams T, Cranston KA. 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc Natl Acad Sci U S A* 112:12764–12769. <https://doi.org/10.1073/pnas.1423041112>.
- Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 27 August 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* <https://doi.org/10.1038/nbt.4229>.
- Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. 2018. Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *mSystems* 3:e00055-18. <https://doi.org/10.1128/mSystems.00055-18>.
- Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188–7196. <https://doi.org/10.1093/nar/gkm864>.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and Web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.
- Spang A, Caceres EF, Etema TJG. 2017. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* 357:eaaf3883. <https://doi.org/10.1126/science.aaf3883>.
- Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, Frischkorn KR, Tringe SG, Singh A, Markillie LM, Taylor RC, Williams KH, Banfield JF. 2015. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* 25:690–701. <https://doi.org/10.1016/j.cub.2015.01.014>.

12. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12:635–645. <https://doi.org/10.1038/nrmicro3330>.
13. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* 523: 208–211. <https://doi.org/10.1038/nature14486>.
14. Hugenholtz P, Pitulle C, Hershberger KL, Pace NR. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol* 180: 366–376.
15. Becraft ED, Woyke T, Jarett J, Ivanova N, Godoy-Vitorino F, Poulton N, Brown JM, Brown J, Lau MCY, Onstott T, Eisen JA, Moser D, Stepanauskas R. 2017. Rokubacteria: genomic giants among the uncultured bacterial phyla. *Front Microbiol* 8:2264. <https://doi.org/10.3389/fmicb.2017.02264>.
16. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG, Wang Z, Chen F, Lindquist EA, Sorek R, Hugenholtz P. 2010. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Methods* 7:807–812. <https://doi.org/10.1038/nmeth.1507>.
17. Staley JT, Konopka A. 1985. Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* 39:321–346. <https://doi.org/10.1146/annurev.mi.39.100185.001541>.
18. Xu H-S, Roberts N, Singleton FL, Attwell RW, Grimes DJ, Colwell RR. 1982. Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment. *Microb Ecol* 8:313–323. <https://doi.org/10.1007/BF02010671>.
19. Sait M, Hugenholtz P, Janssen PH. 2002. Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environ Microbiol* 4:654–666. <https://doi.org/10.1046/j.1462-2920.2002.00352.x>.
20. Joseph SJ, Hugenholtz P, Sangwan P, Osborne CA, Janssen PH. 2003. Laboratory cultivation of widespread and previously uncultured soil bacteria. *Appl Environ Microbiol* 69:7210–7215. <https://doi.org/10.1128/AEM.69.12.7210-7215.2003>.