Systematic and Applied Microbiology 38 (2015) 231–236

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Review

The changing landscape of microbial biodiversity exploration and its implications for systematics

Brian P. Hedlund[a,b,*], Jeremy A. Dodsworth[c], James T. Staley[d]

[a] School of Life Sciences, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA
[b] Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, NV 89154, USA
[c] Department of Biology, California State University, San Bernardino, CA 92407, USA
[d] Department of Microbiology, University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Keywords:
Candidatus taxonomy
Candidate phyla
"Microbial dark matter" metagenomics
Single-cell genomics

ABSTRACT

A vast diversity of *Bacteria* and *Archaea* exists in nature that has evaded axenic culture. Advancements in single-cell genomics, metagenomics, and molecular microbial ecology approaches provide ever-improving insight into the biology of this so-called "microbial dark matter"; however, due to the *International Code of Nomenclature of Prokaryotes*, yet-uncultivated microorganisms are not accommodated in formal taxonomy regardless of the quantity or quality of data. Meanwhile, efforts to calibrate the existing taxonomy with phylogenetic anchors and genomic data are increasingly robust. The current climate provides an exciting opportunity to leverage rapidly expanding single-cell genomics and metagenomics datasets to improve the taxonomy of *Bacteria* and *Archaea*. However, this opportunity must be weighted carefully in light of the strengths and limitations of these approaches. We propose to expand the definition of the *Candidatus* taxonomy to include taxa, from the phylum level to the species level, that are described genomically, particularly when genomic work is coupled with advanced molecular ecology approaches to probe metabolic functions in situ. This system would preserve the rigor and value of traditional microbial systematics while enabling growth of a provisional taxonomic structure to facilitate communication about "dark" lineages on the tree of life.

Current framework for systematics and its strengths and weaknesses

Currently, editors and reviewers of manuscripts describing new taxa of *Bacteria* and *Archaea* often mandate a polyphasic study of axenic cultures, including physiological, chemotaxonomic, morphological, and genetic comparisons to existing related taxa [63]. The precise combination of experiments and the degree of dissimilarity to justify new taxa is highly taxon-dependent, drawing on historical precedent and best practices defined by the International Committee on the Systematics of Prokaryotes (ICSP) and its subcommittees. The strength of the polyphasic approach and the focus on axenic cultures is rooted in its reproducibility and the large amount of descriptive data it provides, which is an invaluable resource for the greater microbiology community, including ecologists, physiologists, molecular biologists, and clinicians.

These strengths notwithstanding, the rigor demanded by the current system combined with the difficulties in cultivating many microorganisms and a dearth in funding for these efforts limits the efficacy of the current taxonomic framework [64]. The funding climate for microbial systematics is especially poor in Europe and North America. For example, the Systematics and Biodiversity Science Cluster at the U.S. National Science Foundation, the only U.S. federal program dedicated to systematics, currently allocates only ~2% of their projects and ~2% of the program's funding to microbiology projects (Table S1). With almost no funding available, it is not surprising that the rate of new taxon descriptions from these continents has stagnated, with Asia taking over the lion's share of publications [42,61]. There are currently ~12,000 validly described species of *Bacteria* and *Archaea* [42] and it is estimated that it will take >1000 years to describe the remaining species, given the current rate of ~600–700 new species descriptions per year [42,51]. If one considers that the current taxonomic descriptions are heavily biased toward taxa that are relatively easy to cultivate [27], the prognosis is even less rosy. On the other hand, a more optimistic view is that advancements, both methodological and ideological, will continue to push biodiversity exploration and systematics forward at an ever-accelerating pace. One such

* Corresponding author at: School of Life Sciences, University of Nevada, Las Vegas, 4004, 4505 Maryland Parkway, Las Vegas, NV 89154-4004, USA. Tel.: +1 702 895 0809.
E-mail address: brian.hedlund@unlv.edu (B.P. Hedlund).

http://dx.doi.org/10.1016/j.syapm.2015.03.003
0723-2020/

ideological advancement could come through the development of 'minimal standards' for systematics [57,64], which would likely deprioritize extensive phenotypic and chemotaxonomic characterizations for individual species descriptions and better leverage more objective genomic criteria and other high-throughput technologies [51,57,62].

It is also useful to note that our poor understanding of microbial diversity is not only at the species level, but it also extends deep into the tree of life. Common estimates place the number of yet-uncultivated bacterial phyla at ~50–100 [3,26]; however, a recent effort using conservative criteria estimated ~1300 yet-uncultivated phylum-level clades [73]. The diversity of yet-uncultivated *Archaea* is also high.

The changing landscape of microbial biodiversity exploration

Progress exploring the microbial world outside of formal systematics has blossomed in recent years, fueled in large part by rapid advancements in DNA sequencing technology, which dropped in cost of over four orders of magnitude over the last decade [6]. Two approaches, single-cell genomics and metagenomics, provide distinct but highly complementary types of datasets (Fig. 1). Although both produce data that are inferior to full genomes from microbial isolates, they have greatly accelerated genomic exploration. For example, the Genomic Encyclopedia of *Bacteria* and *Archaea* (GEBA) project, designed to maximize genomic exploration of microbial isolates, doubled genomic novelty, as measured by phylogenetic diversity per genome [72]. Yet, the project's successor, GEBA-MDM (MDM, microbial dark matter) nearly quintupled genomic novelty by using single-cell genomics to explore habitats rich in underexplored microbial diversity [49]. The current iteration, GEBA-MDM II, inclusive of both single-cell genomics and functional exploration, promises future advancement of "microbial dark matter" biology.

Advancements have also taken place to reframe microbial systematics in the genomics age. Central to this work is the calibration of genomic comparisons to existing microbial species. This body of work has already been quite decisive, with microbial species boundaries corresponding to 95–96% average nucleotide identity (ANI), ~10 Karlin genomic signature, and 70% digital DNA–DNA hybridization [48,51,62]. With this robust calibration in hand, along with robust pipelines for phylogenomics analyses to guide higher-order taxonomy [49,72], a fantastic opportunity exists to expand the taxonomic structure to include single-cell genomics and/or metagenomics datasets.

This expansion has already begun on an ad hoc basis but it is currently poorly coordinated with the systematics community. At this time, significant genomic data have been recovered, assembled, and annotated from >35 candidate microbial phyla, >25 of which have been given informal names in the peer-reviewed literature (Table 1). A few of these names are listed in the List of Prokaryotic Names with Standing in Nomenclature (LPSN) as *Candidatus* taxa [44]; however, most of them cannot readily be identified microscopically in their natural environment (i.e., morphologically or by fluorescence in situ hybridization) so they do not meet the current criteria for *Candidatus* status [37,38]. The future of these candidate phyla within microbial systematics is uncertain on several grounds. First, under the current system of microbial systematics, no taxonomic proposals can be accepted by the ICSP without well-described axenic cultures. Second, the *International Code of Nomenclature of Prokaryotes* does not cover taxonomy above the rank of class, so, in fact, no phylum or domain names are approved by the ICSP [16,29,59]. Instead, Bergey's Manual Trust has become the *de facto* governing body for the higher-order taxonomy. In this vein, it is useful to note that *Bergey's Manual* is not specifically endorsed by the ICSP to serve in this role, and Bergey's Manual Trust has only acted on decadal timeframes to revise phylum-level taxonomy. This may change because *Bergey's Manual of Systematics of Archaea and Bacteria* is currently moving to an online platform that provides a venue for more frequent updates. Finally, the primary literature also includes debates on whether several proposed candidate phyla deserve that rank even informally or whether they are equivalent to lower taxonomic ranks, such as classes [21,54]. This problem is particularly acute among *Archaea* [15,16,19].

Single-cell genomics and metagenomics: strengths and limitations

Single-cell genomics and metagenomics both provide data that are invaluable for biodiversity exploration. These approaches could further be leveraged to guide expansion of a provisional taxonomy; however, each approach has strengths and weaknesses that should be considered carefully (Table 2). Single-cell genomics requires the separation of individual cells from a population, cell lysis, whole genome amplification (WGA), DNA sequencing, and bioinformatics analysis. The details of different workflows for single-cell genomics

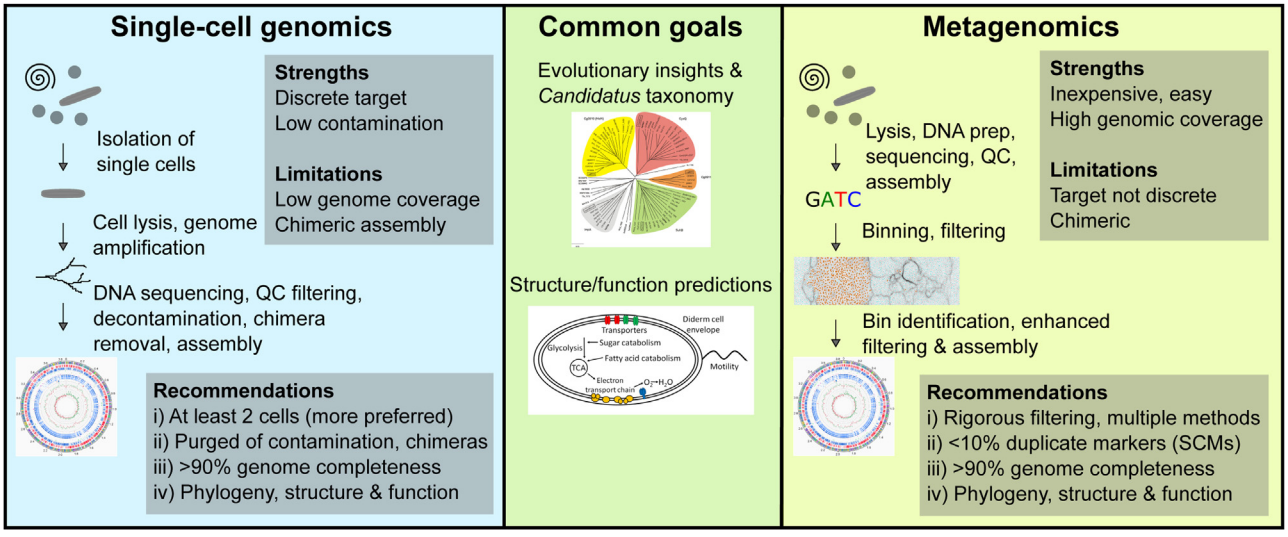

Fig. 1. Summary of workflows, strengths and limitations, and recommendations for single-cell genomics and metagenomics datasets for *Candidatus* taxonomy.

Table 1
Candidate phyla with one or more members having partial or complete genomic coverage.

Candidate phylum	Domain	Genomic datasets	References
Acetothermia (OP1)	Bacteria	MG, SAG	[49,60]
Aerophobetes (CD12)	Bacteria	SAG	[49]
Aminicenantes (OP8)	Bacteria	SAG	[49]
Atribacteria (OP9/JS1)	Bacteria	MG, SAG	[11,49]
BD1-5	Bacteria	MG	[70]
Berkelbacteria (ACD58)	Bacteria	MG	[71]
BRC1	Bacteria	SAG	[49]
Calescamantes (EM19)	Bacteria	SAG	[49]
Cloacimonetes (WWE1)	Bacteria	MG, SAG	[45,49]
EM3	Bacteria	SAG	[49]
Fervidibacteria (OctSpA1-106)	Bacteria	SAG	[49]
Gracilibacteria (GN02)	Bacteria	SAG	[49]
Hydrogenogenetes (NKB19)	Bacteria	SAG	[49]
Latescibacteria (WS3)	Bacteria	SAG	[49]
Marinimicrobia (SAR406)	Bacteria	SAG	[49]
Melainabacteria	Bacteria	MG	[9]
Microgenomates (OP11)	Bacteria	MG, SAG	[70,74]
NC10	Bacteria	MG	[14]
Omnitrophica (OP3)	Bacteria	MG, SAG	[18,49]
Parcobacteria (OD1)	Bacteria	MG, SAG	[49,70]
PER	Bacteria	MG, SAG	[49,70]
Poribacteria	Bacteria	SAG	[53]
SBR1093	Bacteria	MG	[66]
SR1	Bacteria	MG	[26]
Tectomicrobia	Bacteria	MG, SAG	[67]
TM6	Bacteria	SAG	[36]
Saccharibacteria (TM7)	Bacteria	MG, SAG	[26,33,46]
WS1	Bacteria	SAG	[49]
WWE3	Bacteria	MG	[20,71]
Aenigmarchaeota (DSEG)	Archaea	SAG	[49]
Aigarchaeota (pSL4; HWCG-I)	Archaea	MG, SAG	[41,49]
Diapherotrites (pMC2A384)	Archaea	SAG	[49]
Korarchaeota	Archaea	MG	[13]
MCG	Archaea	MG	[31]
Nanoarchaeota	Archaea	MG	[23]
Nanohaloarchaeota	Archaea	MG, SAG	[17,39]
Parvarchaeota (ARMAN)	Archaea	MG, SAG	[2,49]

Abbreviations: MG, metagenome-derived; SAG, single-cell amplified genome.

are provided elsewhere [5,28,30,50,55]. An important strength of single-cell genomics is that the target is discrete, providing an ideological parallel to the clonal structure of axenic cultures used in classical microbial systematics; however, DNA contamination is a significant concern. A recent study showed that protocols that minimize handling result in very low levels of contamination (0.035% [8]); however, bulk tube amplifications were inconsistent, often yielding very high amounts of contamination. As a result of contamination risk, among other factors, very few laboratories are equipped for single-cell genomics and those that do spend considerable energy to minimize contamination [55,68]. Bioinformatic approaches are also useful in removing contaminating sequences, including nucleotide word frequency filtering [11], comparison to similar datasets (e.g., similar single-call amplified genomes (SAGs) or metagenomes [11]), comparison to databases of common contaminants [69], and screening for single-copy conserved markers (SCMs) that are present in multiple copies [49]. Ultimately, if rigorous procedures are used to limit and filter contaminating DNA, the risk for significant contamination in SAG datasets is low.

Another limitation of single-cell genomics is that the genomic coverage of individual SAG datasets is limited. For example, in a recent study describing 201 SAGs, individual SAGs ranged in estimated genome completeness from 4% to 100%, with a mean of 40%. (Genome completeness is typically assessed as a percentage of SCMs [2,49].) Low genome completeness is a result of amplification bias inherent in all current methods for WGA. Although 20% genomic coverage is suitable for ANI comparisons to delimit microbial species [48] and for stable placement of taxa in phylogenomics analyses [49], the value of individual SAG datasets for prediction of phenotypic features is limited. This problem has been solved by analyzing several SAGs within a comparative framework (e.g., to identify core genomic features) or by coassembling them into a composite genome [11,49]. For example, recent studies have used either a 95% [11] or 97% [49] ANI boundary to guide decisions on whether to coassemble individual SAGs. Both of these approaches can lead to great improvements in genomic coverage and have other benefits, such as providing a mechanism to filter chimeric artifacts of MDA, leading to better assemblies [11,34]. Although composite SAG genomes violate the ideological simplicity of working with a discrete target, they provide a wealth of data to explore the phylogenetics, evolution, and possible structural and functional features of yet-uncultivated microorganisms. Comparative analysis of individual SAG datasets is, of course, limited by the completeness of the SAGs themselves; however, this approach is similar to the best practice in microbial systematics to compare many distinct isolates of a taxon to describe both core and accessory phenotypic properties [63].

Metagenomics is easier to implement than single-cell genomics and can provide a route to obtain deep genomic coverage of abundant taxa (Table 1). The data are typically assembled and then assigned to organisms. In cases where closely related genomes exist from microbial isolates, the extraction of related genomic fragments based on sequence homology can be relatively simple. In the absence of closely related genomes, various pipelines exist to filter and "bin" DNA fragments according to genomic characteristics, homology, and read depth [10,12,24,32,52,56], which can then be assigned phylogenetically based on 16S rRNA genes or other phylogenetic anchors. As with composite SAG assemblies, carefully filtered metagenome bins can provide a wealth of information on coarse taxonomic levels, even providing information on accessory

Table 2
Summary of features of different types of genomic datasets relevant to microbial systematics.

	Isolate genomes		Environmental genomes		
	Complete genome	Draft genome[a]	SAG[a]	SAG coassmbly[b]	Metagenome assembly[c]
Coverage	100%	≥95%	4–100% (40%)	to >96%	to 100%
Source discrete	Yes	Yes	Yes	Yes	No
Source clonal	Yes	Yes	Yes	No	No
Appropriate for ANI[e]	Yes	Yes	Yes[e]	Somewhat	Somewhat
Available for polyphasic analysis	Yes	Yes	No	No	No
Value for biodiversity exploration	Moderate	Moderate	High	High	High
Value guiding systematics	High	High	Moderate	Moderate	Moderate

[a] Coverage data from Rinke et al. [49], representing range and mean (parentheses). 100% genomic coverage has been reported for an organism with high genome copy number [69].
[b] Highest genome completeness estimated based on recovery of single-copy conserved markers [49].
[c] Highest literature values [13,41].
[e] 20% genome completeness is sufficient for robust calculation of DDH or ANI [48].

genomic elements (i.e., low read depth genes or biochemical pathways); however, these genomic assemblies should be viewed as chimeras whose heterogeneity depends on the population structure of the target taxon and the rigor used to assign, filter, and assemble the data.

Methods to test functions of yet-uncultivated microorganisms

In our experience, it is common for members of the microbial systematics community, among other biologists, to express pessimism about what can be learned by single-cell genomics and metagenomics or, more generally, by cultivation-independent studies of yet-uncultivated microorganisms. We share these frustrations about the difficulties inherent in studying microorganisms in nature, but subscribe to a more optimistic viewpoint. A variety of approaches exist to investigate the activities of microorganisms in microcosms or mixed cultures. Metatranscriptomics and metaproteomics can reveal information on the expression genes in natural samples [22,47]. A classic approach that has gone through both evolutionary and revolutionary improvements involves the identification of radioactive or stable isotopes in target cells labeled by fluorescence in situ hybridization (FISH) [4,40,65]. Complementary high-throughput approaches are being developed constantly, including workflows to identify stable isotope labels incorporated into 16S rRNA for identification on phylogenetic microarrays [35] or to identify isotopically labeled peptides in metaproteomes, allowing species-level assignment of assimilatory activities [25,43]. Additionally, activity-based probes can reveal the functions of microorganisms in natural environments by labeling membrane transport proteins, even enabling the physical isolation and study of these organisms [7].

Outlook

Single-cell genomics and metagenomics provide windows into the biology of yet-uncultivated microorganisms and are particularly powerful when combined synergistically and with cultivation-independent approaches to interrogate the activities of microorganisms in nature. These data are potential resources for the systematics community; however, they are inferior to classical microbial systematics in several important ways: (1) they fail to deliver microbial isolates as a resource for the research community; (2) they are less reproducible in that sampling the same or similar habitats may or may not recover similar datasets; (3) they currently do not provide nearly as much information about the structure and function of microorganisms, although some progress is now being made in metabolic and biosynthetic pathway reconstruction [1].

These shortcomings notwithstanding, we believe there is value in a provisional taxonomic structure that can serve to promote order and facilitate communication about the "dark" branches of the tree of life. As such, we propose that *Candidatus* status be broadened to accommodate yet-uncultivated microorganisms for which complete or nearly complete and carefully curated genomic datasets exist, even if morphological identification and/or functional data are lacking. The literature proposing and endorsing a *Candidatus* system currently requires 'identification of a morphotype with a specific probe' as well as 'information concerning properties other than phylogenetic position', among other criteria [38]. These guidelines are well-founded, particularly in light of the occasional presence of highly divergent 16S rRNA genes in some taxa. However, these guidelines were conceived at a time when the 'rRNA approach' was just coming on line and we feel they should be revised in light of advancements in experimental and analytical tools available to microbiologists over the last twenty years. An expansion of the *Candidatus* concept to include single-cell or metagenomics datasets with or without expression or functional data would enable significant expansion of the tree of life (e.g., Table 1) while preserving the rigor intended for the *Candidatus* system. Of note, most *Candidatus* taxa currently listed in the LPSN are from microbial phyla that are well-represented in culture; >63% of *Candidatus* taxa belong to the *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, and *Actinobacteria*. We believe that, by adopting appropriate guidelines, these ideas would preserve the more successful elements of the current systematic practice while allowing the *Candidatus* taxonomic structure to expand within the "dark" branches on the tree of life for the mutual benefit of systematists and the microbiology community at large. We furthermore believe that the *Candidatus* nomenclature should be encouraged at all taxonomic levels, from phylum to species, to fully take advantage of a provisional taxonomic structure, rather than only at the genus level [38].

Proposal to the ICSP to revise guidelines for *Candidatus* status

We encourage the ICSP to discuss expansion of the *Candidatus* system to include microorganisms with complete or nearly complete and carefully curated genome datasets, but which have not necessarily been imaged in natural environments by FISH. (In this regard, it is noteworthy that microfluidics-based single-cell genomics allows microorganisms to be visualized by phase-contrast microscopy or FISH during sorting, thereby coupling morphology of the organisms and with genomic data [e.g., 11,28].) These recommendations are designed to mandate the clear delineation of distinct, species-level genomic datasets that are sufficiently comprehensive to predict key morphological, structural, and physiological traits. The specific parameters indicated below serve as initial suggestions that can be better defined by in depth empirical and in silico analyses, and should be rigorously debated within the microbiology community.

We recommend the following criteria for evaluating *Candidatus* proposals from single-cell genomic data: (i) data should be obtained by ≥2 single cells of a species (≥95% ANI [48,62]), with more cells being preferred, and should be distinct from genomes of other cultivated or *Candidatus* taxa (≤95% ANI); (ii) data should be screened and purged of contaminating DNA [11,50,68] and, where possible, chimeric artifacts of DNA amplification [11,34]; (iii) at least one SAG should have an estimated genomic coverage of ≥90%, or the combined data should represent ≥90% of the core genome of the proposed species (e.g., as assessed by accounting for SCMs in a SAG coassembly [e.g., 49]); (iv) the data should be used to infer characteristics that are commonly used to delineate taxa, including, but not limited to, phylogenetic relationships (e.g., 16S rRNA gene and multigene (i.e., phylogenomics) analyses), morphological and structural characteristics (e.g., cell morphology (e.g., MreB or crenactin), cell membrane structure (e.g., BamA/YaeT, OmpH, TolC, TonB, secretin [58]), and motility), and metabolic potential (e.g., electron donors and acceptors for respiration, fermentative capacity).

We recommend the following criteria for evaluating *Candidatus* proposals from metagenomic data: (i) data should be filtered rigorously using nucleotide word frequency, read depth filtering, and nucleic acid identity approaches (e.g., [10]); (ii) datasets should contain no more than 10% SCMs [49,72] present in more than one copy and have a monomodal % G+C content (anomalies in SCMs or % G+C content should carefully justified as representative of a single genome (e.g., through comparison with related isolate or SAG datasets)), and should be distinct from genomes of other cultivated or *Candidatus* taxa (≤95% ANI); (iii) data should cover ≥90% of the core genome (e.g., as assessed by accounting for SCMs in the metagenome [e.g., 49]); (iv) the data should be used to infer

characteristics that are commonly used to delineate taxa, including, but not limited to, phylogenetic relationships (e.g., 16S rRNA gene and multigene (i.e., phylogenomics) analyses), morphological and structural characteristics (e.g., cell morphology (e.g., MreB or crenactin), ultrastructure [58], and motility), and metabolic potential (e.g., electron donors and acceptors for respiration, fermentative capacity).

Acknowledgements

We thank Namritha Manoharan, Chrisabelle Cempron, Scott Thomas, Julienne Paraiso, Tim Alba, and two outstanding reviewers for help and editorial advice with this manuscript. This work was supported by U.S. National Science Foundation grant OISE 0968421 and NASA Exobiology grant EXO-NNX11AR78G. B.P.H. acknowledges generous support from Greg Fullmer through the UNLV Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.syapm.2015.03.003

References

[1] Amaral, G.R., Dias, G.M., Wellington-Oguri, M., Chimetto, L., Campeão, M.E., Thompson, F.L., Thompson, C.C. (2014) Genotype to phenotype: identification of diagnostic vibrio phenotypes using whole genome sequences. Int. J. Syst. Evol. Microbiol. 64, 357–365.
[2] Baker, B.J., Comolli, L.R., Dick, G.J., Hauser, L.J., Hyatt, D., Dill, B.D., Land, M.L., Verberkmoes, N.C., Hettich, R.L., Banfield, J.F. (2010) Enigmatic, ultrasmall, uncultivated archaea. Proc. Natl. Acad. Sci. U. S. A. 107, 8806–8811.
[3] Baker, B.J., Dick, G.J. (2013) Omic approaches in microbial ecology: charting the unknown. Microbe 8, 353–360.
[4] Behrens, S., Lösekann, T., Pett-Ridge, J., Weber, P.K., Ng, J.W.O., Stevenson, B.S., Hutcheon, I.D., Relman, D.A., Spormann, A.M. (2008) Linking phylogeny with metabolic activity of single microbial cells using FISH-NanoSIMS. Appl. Environ. Microbiol. 74, 3143–3150.
[5] Blainey, P.C., Quake, S.R. (2014) Dissecting genomic diversity, one cell at a time. Nat. Methods 11, 19–21.
[6] Carlson, R.H. 2011 Biology is Technology: The Promise, Peril, and New Business of Engineering Life, Harvard University Press, Cambridge, MA.
[7] Chauvigne-Hines, L.M., Anderson, L.N., Weaver, H.M., Brown, J.N., Koech, P.K., Nicora, C.D., Hofstad, B.A., Smith, R.D., Wilkins, M.J., Callister, S.J., Wright, A.T. (2012) Suite of activity-based probes for cellulose-degrading enzymes. J. Am. Chem. Soc. 134, 20521–20532.
[8] de Bourcy, C.F., De Vlaminck, I., Kanbar, J.N., Wang, J., Gawad, C., Quake, S.R. (2014) A quantitative comparison of single-cell whole genome amplification methods. PLOS ONE 9, e105585.
[9] Di Rienzi, S.C., Sharon, I., Wrighton, K.C., Koren, O., Hug, L.A., Thomas, B.C., Goodrich, J.K., Bell, J.T., Spector, T.D., Banfield, J.F., Ley, R.E. (2013) The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. eLife 2, e01102.
[10] Dick, G.J., Andersson, A.F., Baker, B.J., Simmons, S.L., Thomas, B.C., Yelton, A.P., Banfield, J.F. (2009) Community-wide analysis of microbial genome sequence signatures. Genome Biol. 10, R85.
[11] Dodsworth, J.A., Blainey, P.C., Murugapiran, S.K., Swingley, W.D., Ross, C.A., Tringe, S.G., Chai, P.S.G., Raymond, J., Quake, S.R., Hedlund, B.P. (2013) Single-cell and metagenomic analyses indicate a fermentative, saccharolytic lifestyle for members of the OP9 lineage. Nat. Commun. 4, 1854.
[12] Dröge, J., McHardy, A.C. (2012) Taxonomic binning of metagenome samples generated by next-generation sequencing technologies. Brief Bioinform. 13, 646–655.
[13] Elkins, J.G., Kunin, V., Anderson, I., Barry, K., Goltsman, E., Lapidus, A., Hedlund, B.P., Hugenholtz, P., Kyrpides, N., Graham, D., Keller, M., Wanner, G., Richardson, P., Stetter, K.O. (2008) A korarchaeal genome reveals insights into the evolution of archaea. Proc. Natl. Acad. Sci. U. S. A. 105, 8102–8107.
[14] Ettwig, K., Butler, M., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M., Schreiber, F., Dutilh, B., Zedelius, J., de Beer, D., Gloerich, J., Wessels, H., van Alen, T., Luesken, F., Wu, M., van de Pas-Schoonen, K., Op den Camp, H., Janssen-Megens, E., Francoijs, K., Stunnenberg, H. (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature 464, 543–548.
[15] Garrity, G.M., Oren, A. (2012) Response to Gribaldo and Brochier-Armanet: time for order in microbial systematics. Trends Microbiol. 20, 353–354.
[16] Garrity, G.M., Oren, A. (2013) Response to Sutcliffe et al.: regarding the international committee on systematics of prokaryotes. Trends Microbiol. 21, 53–55.
[17] Ghai, R., Pašić, L., Fernández, A.B., Martin-Cuadrado, A.B., Mizuno, C.M., McMahon, K.D., Papke, R.T., Stepanauskas, R., Rodriguez-Brito, B., Rohwer, F., Sánchez-Porro, C., Ventosa, A., Rodríguez-Valera, F. (2011) New abundant microbial groups in aquatic hypersaline environments. Sci. Rep. 1, 135.
[18] Glöckner, J., Kube, M., Shrestha, P.M., Weber, M., Glöckner, F.O., Reinhardt, R., Liesack, W. (2010) Phylogenetic diversity and metagenomics of candidate division OP3. Environ. Microbiol. 12, 1218–1229.
[19] Gribaldo, S., Brochier-Armanet, C. (2012) Time for order in microbial systematics. Trends Microbiol. 20, 209–210.
[20] Guermazi, S., Daegelen, P., Dauga, C., Rivière, D., Bouchez, T., Godon, J., Gyapay, G., Sghir, A., Pelletier, E., Weissenbach, J., Le Paslier, D. (2008) Discovery and characterization of a new bacterial candidate division by an anaerobic sludge digester metagenomic approach. Environ. Microbiol. 10, 2111–2123.
[21] Guy, L., Ettema, T.J. (2011) The archaeal 'TACK' superphylum and the origin of eukaryotes. Trends Microbiol. 19, 580–587.
[22] Handley, K.M., VerBerkmoes, N.C., Steefel, C.I., Williams, K.H., Sharon, I., Miller, C.S., Frischkorn, K.R., Chourey, K., Thomas, B.C., Shah, M.B., Long, P.E., Hettich, R.L., Banfield, J.F. (2013) Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community. ISME J. 7, 800–816.
[23] Huber, H., Hohn, M., Rachel, R., Fuchs, T., Wimmer, V., Stetter, K. (2002) A new phylum of archaea represented by a nanosized hyperthermophilic symbiont. Nature 417, 63.
[24] Imelfort, M., Parks, D., Woodcroft, B.J., Dennis, P., Hugenholtz, P., Tyson, G.W. (2014) GroopM: an automated tool for the recovery of population genomes from related metagenomes. PeerJ 2, e603.
[25] Jehmlich, N., Schmidt, F., Taubert, M., Seifert, J., Bastida, F., von Bergen, M., Richnow, H.H., Vogt, C. (2010) Protein-based stable isotope probing. Nat. Protoc. 5, 1957–1966.
[26] Kantor, R.S., Wrighton, K.C., Handley, K.M., Sharon, I., Hug, L.A., Castelle, C.J., Thomas, B.C., Banfield, J.F. (2013) Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. mBio 4, e00708–e00713.
[27] Klenk, H.P., Göker, M. (2010) En route to a genome-based classification of Archaea and Bacteria? Syst. Appl. Microbiol. 33 (June (4)), 175–182.
[28] Landry, Z.C., Giovanonni, S.J., Quake, S.R., Blainey, P.C. (2013) Optofluidic cell selection from complex microbial communities for single-genome analysis. Methods Enzymol. 531, 61–90.
[29] Lapage, S.P., Sneath, P.H.A., Lessel, E.F., Skerman, V.B.D., Seeliger, H.P.R., Clark, W.A. 1992 International Code of Nomenclature of Bacteria Bacteriological Code, 1990 Revision, ASM Press, Washington, DC.
[30] Lasken, R.S. (2012) Genomic sequencing of uncultured microorganisms from single cells. Nat. Rev. Microbiol. 10, 631–640.
[31] Lloyd, K.G., Schreiber, L., Petersen, D.G., Kjeldsen, K.U., Lever, M.A., Steen, A.D., Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., Schramm, A., Jørgensen, B.B. (2013) Predominant archaea in marine sediments degrade detrital proteins. Nature 496, 215–218.
[32] Mande, S.S., Mohammed, M.H., Ghosh, T.S. (2012) Classification of metagenomic sequences: methods and challenges. Brief Bioinform. 13, 669–681.
[33] Marcy, Y., Ouverney, C., Bik, E.M., Lösekann, T., Ivanova, N., Martin, H.G., Szeto, E., Platt, D., Hugenholtz, P., Relman, D.A., Quake, S.R. (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proc. Natl. Acad. Sci. U. S. A. 104, 11889–11894.
[34] Marshall, I.P.G., Blainey, P.C., Spormann, A.M., Quake, S.R. (2012) A single-cell genome for *Thiovulum* sp. Appl. Environ. Microbiol. 78, 8555–8563.
[35] Mayali, X., Weber, P.K., Brodie, E.L., Mabery, S., Hoeprich, P.D., Pett-Ridge, J. (2012) High-throughput isotopic analysis of RNA microarrays to quantify microbial resource use. ISME J. 6, 1210–1221.
[36] McLean, J.S., Lombardo, M.J., Badger, J.H., Edlund, A., Novotny, M., Yee-Greenbaum, J., Vyahhi, N., Hall, A.P., Yang, Y., Dupont, C.L., Ziegler, M.G., Chitsaz, H., Allen, A.E., Yooseph, S., Tesler, G., Pevzner, P.A., Friedman, R.M., Nealson, K.H., Venter, J.C., Lasken, R.S. (2013) Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc. Natl. Acad. Sci. U. S. A. 110, E2390–E2399.
[37] Murray, R.G.E., Scheifer, K.H. (1994) Taxonomic notes: a proposal for recording the properties of putative taxa of procaryotes. Int. J. Syst. Bacteriol. 44, 174–176.
[38] Murray, R.G.E., Stackebrandt, E. (1995) Taxonomic note: implementation of the provisional status candidatus for incompletely described procaryotes. Int. J. Syst. Bacteriol. 45, 186–187.
[39] Narasingarao, P., Podell, S., Ugalde, J.A., Brochier-Armanet, C., Emerson, J.B., Brocks, J.J., Heidelberg, K.B., Banfield, J.F., Allen, E.E. (2012) De novo metagenomic assembly reveals abundant novel major lineage of archaea in hypersaline microbial communities. ISME J. 6, 81–93.
[40] Neufeld, J.D., Murrell, J.C. (2007) Witnessing the last supper of uncultivated microbial cells with Raman-FISH. ISME J. 1, 269–270.
[41] Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., Chee, G.J., Hattori, M., Kanai, A., Atomi, H., Takai, K., Takami, H. (2011) Insights into the evolution of archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res. 39, 3204–3223.
[42] Oren, A., Garrity, G.M. (2014) Then and now: a systematic review of the systematics of prokaryotes in the last 80 years. Antonie Van Leeuwenhoek 106, 43–56.
[43] Pan, C., Fischer, C.R., Hyatt, D., Bowen, B.P., Hettich, R.L., Banfield, J.F. (2011) Quantitative tracking of isotope flows in proteomes of microbial communities. Mol. Cell. Proteomics 10, M110.006049.

[44] Parte, A.C. (2014) LPSN-list of prokaryotic names with standing in nomenclature. Nucleic Acids Res. 42, D613–D616.
[45] Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., Rivière, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G., Médigue, C., Weissenbach, J., Le Paslier, D. (2008) "*Candidatus* Cloacamonas Acidaminovorans": genome sequence reconstruction provides a first glimpse of a new bacterial division. J. Bacteriol. 190, 45.
[46] Podar, M., Abulencia, C.B., Walcher, M., Hutchison, D., Zengler, K., Garcia, J.A., Holland, T., Cotton, D., Hauser, L., Keller, M. (2007) Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl. Environ. Microbiol. 73, 3205–3214.
[47] Ram, R.J., Verberkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake, R.C., Shah, I.I., Hettich, M., Banfield, R.L.J.F. (2005) Community proteomics of a natural microbial biofilm. Science 308, 1915–1920.
[48] Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U. S. A. 106, 19126–191231.
[49] Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., Dodsworth, J.A., Hedlund, B.P., Tsiamis, G., Sievert, S.M., Liu, W.T., Eisen, J.A., Hallam, S.J., Kyrpides, N.C., Stepanauskas, R., Rubin, E.M., Hugenholtz, P., Woyke, T. (2013) Insights into the phylogeny and coding potential of microbial dark matter. Nature 499, 431–437.
[50] Rinke, C., Lee, J., Nath, N., Goudeau, D., Thompson, B., Poulton, N., Dmitrieff, E., Malmstrom, R., Stepanauskas, R., Woyke, T. (2014) Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. Nat. Protoc. 9, 1038–1048.
[51] Rosselló-Móra, R. (2012) Towards a taxonomy of bacteria and archaea based on interactive and cumulative data repositories. Environ. Microbiol. 14, 318–334.
[52] Scholz, M.B., Lo, C.C., Chain, P.S. (2012) Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Curr. Opin. Biotechnol. 23, 9–15.
[53] Siegl, A., Kamke, J., Hochmuth, T., Piel, J., Richter, M., Liang, C.G., Dandekar, T., Hentschel, U. (2011) Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges. ISME J. 5, 61–70.
[54] Spang, A., Martijn, J., Saw, J.H., Lind, A.E., Guy, L., Ettema, T.J. (2013) Close encounters of the third domain: the emerging genomic view of archaeal diversity and evolution. Archaea 2013, 202358.
[55] Stepanauskas, R. (2012) Single cell genomics: an individual look at microbes. Curr. Opin. Microbiol. 15, 613–620.
[56] Strous, M., Kraft, B., Bisdorf, R., Tegetmeyer, H.E. (2012) The binning of metagenomic contigs for microbial physiology of mixed cultures. Front. Microbiol. 3, 410.
[57] Sutcliffe, I.C., Trujillo, M.E., Goodfellow, M. (2012) A call to arms for systematists: revitalising the purpose and practises underpinning the description of novel microbial taxa. Antonie Van Leeuwenhoek 101, 13–20.
[58] Sutcliffe, I.C. (2011) Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. Environ. Microbiol. 13, 279–282.
[59] Sutcliffe, I.C., Trujillo, M.E., Whitman, W.B., Goodfellow, M. (2013) A call to action for the international committee on systematics of prokaryotes. Trends Microbiol. 21, 51–52.
[60] Takami, H., Noguchi, H., Takaki, Y., Uchiyama, I., Toyoda, A., Nishi, S., Chee, G.J., Arai, W., Nunoura, T., Itoh, T., Hattori, M., Takai, K. (2012) A deeply branching thermophilic bacterium with an ancient acetyl–CoA pathway dominates a subsurface ecosystem. PLoS ONE 7, e30559.
[61] Tamames, J., Rosselló-Móra, R. (2012) On the fitness of microbial taxonomy. Trends Microbiol. 20, 514–516.
[62] Thompson, C.C., Chimetto, L., Edwards, R.A., Swings, J., Stackebrandt, E., Thompson, F.L. (2013) Microbial genomic taxonomy. BMC Genomics 14, 913.
[63] Tindall, B.J., Rosselló-Móra, R., Busse, H.J., Ludwig, W., Kämpfer, P. (2010) Notes on the characterization of prokaryote strains for taxonomic purposes. Int. J. Syst. Evol. Microbiol. 60, 249–266.
[64] Vandamme, P., Peeters, C. (2014) Time to revisit polyphasic taxonomy. Antonie Van Leeuwenhoek 106, 57–65.
[65] Wagner, M., Nielsen, P.H., Loy, A., Nielsen, J.L., Daims, H. (2006) Linking microbial community structure with function: fluorescence in situ hybridization-microautoradiography and isotope arrays. Curr. Opin. Biotechnol. 17, 1–9.
[66] Wang, Z., Guo, F., Liu, L., Zhang, T. (2014) Evidence of carbon fixation pathway in a bacterium from candidate phylum SBR1093 revealed with genomic analysis. PLOS ONE 9, 1–9.
[67] Wilson, M., Mori, T., Rückert, C., Uria, A., Helf, M., Takada, K., Gernert, C., Steffens, U., Heycke, N., Schmitt, S., Rinke, C., Helfrich, E., Brachmann, A., Gurgui, C., Wakimoto, T., Kracht, M., Crüsemann, M., Hentschel, U., Abe, I., Matsunaga, S. (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. Nature 506, 58–62.
[68] Woyke, T., Sczyrba, A., Lee, J., Rinke, C., Tighe, D., Clingenpeel, S., Malmstrom, R., Stepanauskas, R., Cheng, J.F. (2011) Decontamination of MDA reagents for single cell whole genome amplification. PLoS ONE 6, e26161.
[69] Woyke, T., Tighe, D., Mavromatis, K., Clum, A., Copeland, A., Schackwitz, W., Lapidus, A., Wu, D., McCutcheon, J.P., McDonald, B.R., Moran, N.A., Bristow, J., Cheng, J.F. (2010) One bacterial cell, one complete genome. PLoS ONE 5, e10314.
[70] Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., Wilkins, M.J., Hettich, R.L., Lipton, M.S., Williams, K.H., Long, P.E., Banfield, J.F. (2012) Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. Science 337, 1661–1665.
[71] Wrighton, K.C., Castelle, C.J., Wilkins, M.J., Hug, L.A., Sharon, I., Thomas, B.C., Handley, K.M., Mullin, S.W., Nicora, C.D., Singh, A., Lipton, M.S., Long, P.E., Williams, K.H., Banfield, J.F. (2014) Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. ISME J. 8, 1452–1463.
[72] Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N.N., Kunin, V., Goodwin, L., Wu, M., Tindall, B.J., Hooper, S.D., Pati, A., Lykidis, A., Spring, S., Anderson, I.J., D'haeseleer, P., Zemla, A., Singer, M., Lapidus, A., Nolan, M., Copeland, A., Han, C., Chen, F., Cheng, J.F., Lucas, S., Kerfeld, C., Lang, E., Gronow, S., Chain, P., Bruce, D., Rubin, E.M., Kyrpides, N.C., Klenk, H.P., Eisen, J.A. (2009) A phylogeny-driven genomic encyclopaedia of bacteria and archaea. Nature 462, 1056–1060.
[73] Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., 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.
[74] Youssef, N.H., Blainey, P.C., Quake, S.R., Elshahed, M.S. (2011) Partial genome assembly for candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). Appl. Environ. Microbiol. 77, 7804–7814.