



# Expanding the Diversity of Bacterioplankton Isolates and Modeling Isolation Efficacy with Large-Scale Dilution-to-Extinction Cultivation

Michael W. Henson,<sup>a\*</sup> V. Celeste Lanclos,<sup>a</sup> David M. Pitre,<sup>b</sup> Jessica Lee Weckhorst,<sup>b\*</sup> Anna M. Lucchesi,<sup>b</sup> Chuankai Cheng,<sup>a</sup> Ben Temperton,<sup>c</sup> J. Cameron Thrash<sup>a</sup>

<sup>a</sup>Department of Biological Sciences, University of Southern California, Los Angeles, California, USA

<sup>b</sup>Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA

<sup>c</sup>School of Biosciences, University of Exeter, Exeter, United Kingdom

**ABSTRACT** Cultivated bacterioplankton representatives from diverse lineages and locations are essential for microbiology, but the large majority of taxa either remain uncultivated or lack isolates from diverse geographic locales. We paired large-scale dilution-to-extinction (DTE) cultivation with microbial community analysis and modeling to expand the phylogenetic and geographic diversity of cultivated bacterioplankton and to evaluate DTE cultivation success. Here, we report results from 17 DTE experiments totaling 7,820 individual incubations over 3 years, yielding 328 repeatably transferable isolates. Comparison of isolates to microbial community data for source waters indicated that we successfully isolated 5% of the observed bacterioplankton community throughout the study; 43% and 26% of our isolates matched operational taxonomic units and amplicon single-nucleotide variants, respectively, within the top 50 most abundant taxa. Isolates included those from previously uncultivated clades such as SAR11 LD12 and *Actinobacteria* acIV, as well as geographically novel members from other ecologically important groups like SAR11 subclade IIIa, SAR116, and others, providing isolates in eight putatively new genera and seven putatively new species. Using a newly developed DTE cultivation model, we evaluated taxon viability by comparing relative abundance with cultivation success. The model (i) revealed the minimum attempts required for successful isolation of taxa amenable to growth on our media and (ii) identified possible subpopulation viability variation in abundant taxa such as SAR11 that likely impacts cultivation success. By incorporating viability in experimental design, we can now statistically constrain the effort necessary for successful cultivation of specific taxa on a defined medium.

**IMPORTANCE** Even before the coining of the term “great plate count anomaly” in the 1980s, scientists had noted the discrepancy between the number of microorganisms observed under the microscope and the number of colonies that grew on traditional agar media. New cultivation approaches have reduced this disparity, resulting in the isolation of some of the “most wanted” bacterial lineages. Nevertheless, the vast majority of microorganisms remain uncultured, hampering progress toward answering fundamental biological questions about many important microorganisms. Furthermore, few studies have evaluated the underlying factors influencing cultivation success, limiting our ability to improve cultivation efficacy. Our work details the use of dilution-to-extinction (DTE) cultivation to expand the phylogenetic and geographic diversity of available axenic cultures. We also provide a new model of the DTE approach that uses cultivation results and natural abundance information to predict taxon-specific viability and iteratively constrain DTE experimental design to improve cultivation success.

**Citation** Henson MW, Lanclos VC, Pitre DM, Weckhorst JL, Lucchesi AM, Cheng C, Temperton B, Thrash JC. 2020. Expanding the diversity of bacterioplankton isolates and modeling isolation efficacy with large-scale dilution-to-extinction cultivation. *Appl Environ Microbiol* 86:e00943-20. <https://doi.org/10.1128/AEM.00943-20>.

**Editor** Robert M. Kelly, North Carolina State University

**Copyright** © 2020 Henson et al. 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 Ben Temperton, [b.temperton@exeter.ac.uk](mailto:b.temperton@exeter.ac.uk), or J. Cameron Thrash, [thrash@usc.edu](mailto:thrash@usc.edu).

\* Present address: Michael W. Henson, Department of Geophysical Sciences, University of Chicago, Chicago, Illinois, USA; Jessica Lee Weckhorst, Quantitative and Computational Biosciences Program, Baylor College of Medicine, Houston, Texas, USA.

**Received** 21 April 2020

**Accepted** 13 June 2020

**Accepted manuscript posted online** 19 June 2020

**Published** 18 August 2020

**KEYWORDS** dilution to extinction, cultivation, bacterioplankton, LSUCC, microbial ecology, coastal microbiology

Axenic cultures of environmentally important microorganisms are critical for fundamental microbiological investigation, including generating physiological information about environmental tolerances, determining organism-specific metabolic and growth rates, testing hypotheses generated from *in situ* omics observations, and experimentally examining microbial interactions. Research using important microbial isolates has been critical to a number of discoveries, such as defining microorganisms involved in surface ocean methane saturation (1–3), the role of proteorhodopsin in maintaining cellular functions during states of carbon starvation (4, 5), the complete nitrification of ammonia within a single organism (6), and identifying novel metabolites and antibiotics (7–10). However, the vast majority of taxa remain uncultivated (11–13), restricting valuable experimentation on such topics as genes of unknown function, the role of analogous gene substitutions in overcoming auxotrophy, and the multifaceted interactions occurring in the environment inferred from sequence data (11, 14–16).

The quest to bring new microorganisms into culture and the recognition that traditional agar plate-based approaches have limited success (17–19) have compelled numerous methodological advances spanning a wide variety of techniques, such as diffusion chambers, microdroplet encapsulation, and slow acclimatization of cells to artificial media (20–25). Dilution-to-extinction (DTE) cultivation using sterile seawater as the medium has also proven highly successful for isolating bacterioplankton (26–32). Pioneered by Don Button and colleagues for the cultivation of oligotrophic bacteria, this method essentially preisolates organisms after serial dilution by separating individual or small groups of cells into their own incubation vessel (32, 33). This prevents slow-growing, obligately oligotrophic bacterioplankton from being outcompeted by faster-growing organisms, as would occur in enrichment-based isolation methods like those that would target aerobic heterotrophs. DTE is also a practical method for taxa that cannot grow on solid media. Natural seawater media provide these taxa with the same chemical surroundings from which they are collected, reducing the burden of anticipating all the relevant compounds required for growth (33).

Improvements to DTE cultivation in multiple labs have increased the number of inoculated wells and decreased the time needed to detect growth (26, 28, 34), thereby earning the moniker “high-throughput culturing” (26, 28). We (35) and others (30) have also adapted DTE culturing by incorporating artificial media in place of natural seawater media to successfully isolate abundant bacterioplankton. Thus far, DTE culturing has led to isolation of many important marine and freshwater groups, such as marine SAR11 *Alphaproteobacteria* (28, 29, 34–36), the freshwater SAR11 LD12 clade (29), SUP05/Arctic96BD-19 *Gammaproteobacteria* (37–39), OM43 *Betaproteobacteria* (26, 27, 31, 40, 41), HIMB11-type *Roseobacter* spp. (35, 42), numerous so-called oligotrophic marine *Gammaproteobacteria* (43), and *acl Actinobacteria* (44).

Despite the success of DTE cultivation, many taxa continue to elude domestication (11–13, 16). Explanations include a lack of required nutrients or growth factors in media (20, 45–49) and biological phenomena such as dormancy and/or phenotypic heterogeneity within populations (47, 48, 50–56). However, there have been few studies empirically examining the factors underlying isolation success in DTE cultivation experiments (34, 57, 58), restricting our ability to determine the relative importance of methodological versus biological influences on cultivation reliability for any given organism. Moreover, even for those taxa that we have successfully cultivated, in many cases we lack geographically diverse strains, restricting comparisons of the phenotypic and genomic diversity that may influence taxon-specific cultivability.

We undertook a 3-year cultivation effort in the coastal northern Gulf of Mexico (nGOM), from which we lack representatives of many common bacterioplankton groups, to provide new model organisms for investigating microbial function, ecology, biogeography, and evolution. Simultaneously, we paired our cultivation efforts with

**TABLE 1** Cultivation statistics, including whole-community viability estimates

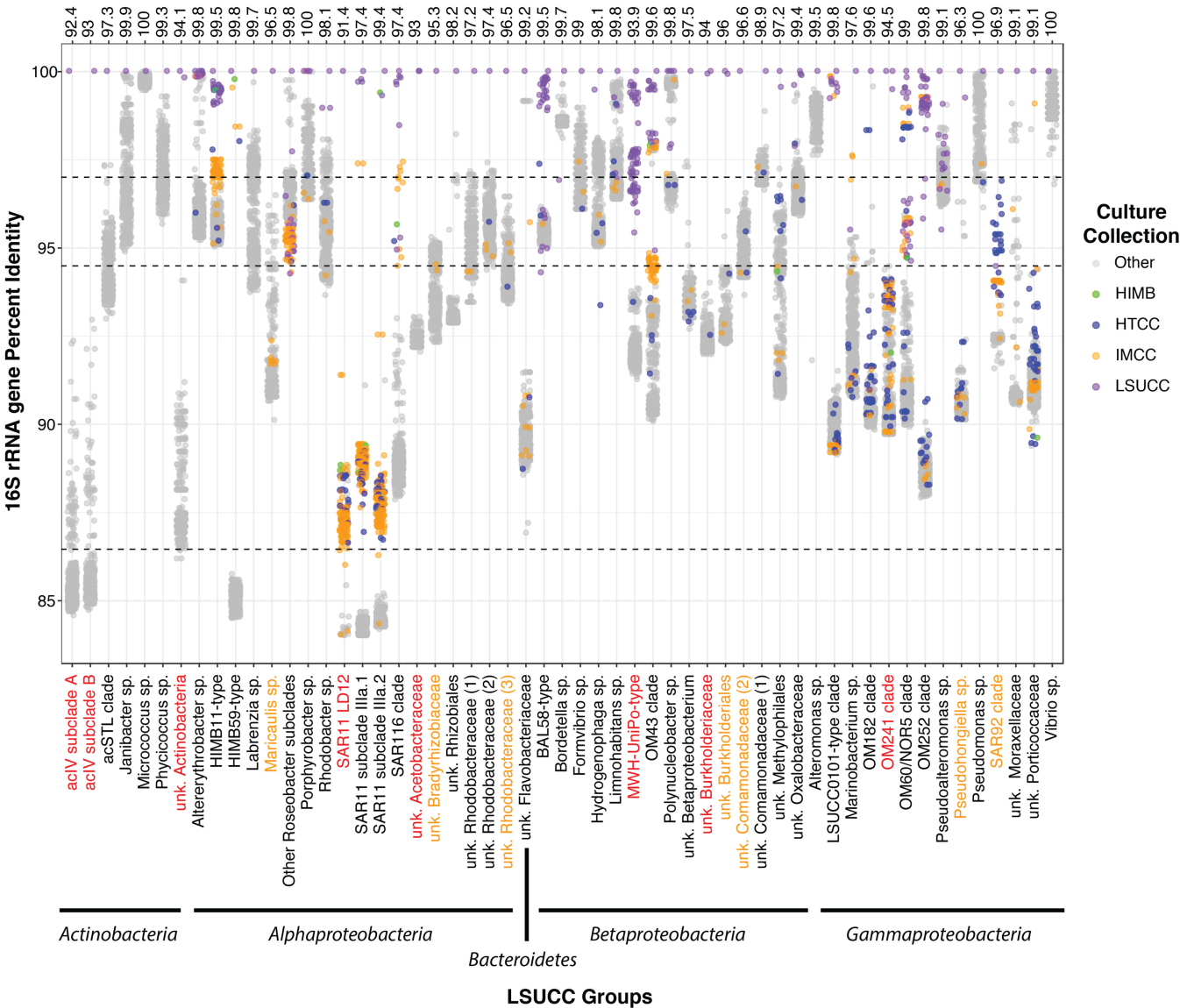
Site	Date	<i>n</i>	<i>z</i>	<i>p</i>	$\lambda$	<i>V</i> <sup>a</sup> (ASE)	CV	Estimated no. of wells with 1 cell if <i>V</i> = 1 <sup>b</sup>	<i>V</i> <sub>est</sub> <sup>c</sup>	<i>In situ</i> salinity	Medium salinity	Medium	Reference
CJ	September 2014	460	15	0.033	1.27	2.6 (0.67)	0.259	164 (144–185)	1.5–4.2	24.6	34.8	JWAMPFe	35
ARD	November 2014	460	1	0.002	1.5	0.1 (0.15)	1.451	154 (134–174)	0.1–0.7	1.72	34.8	JW1	35
JLB	January 2015	460	61	0.133	1.96	7.3 (0.93)	0.127	127 (109–146)	5.6–9.2	26.0	34.8	JW1	35
FWC <sup>d</sup>	March 2015	460	301	0.654	2	53.1 (3.2)	0.06	125 (106–143)	47.1–59.7	5.39	5.79	JW4	35
LKB	June 2015	460	15	0.033	1.8	1.8 (0.48)	0.266	137 (118–156)	1.1–3.0	2.87	5.79	JW4	35
Tbon2	August 2015	460	41	0.089	1.56	6.0 (0.93)	0.156	151 (132–171)	4.3–8.1	14.2	11.6	JW3	35
CJ2	October 2015	460	61	0.133	2	7.1 (0.91)	0.128	125 (106–143)	5.6–9.1	22.2	23.2	JW2	35
FWC2 <sup>d</sup>	April 2016	460	403	0.876	2	104.4 (6.2)	0.059	125 (106–143)	>92.3	20.9	23.2	JW2	This study
ARD2c	June 2016	460	7	0.015	2	0.8 (0.29)	0.362	125 (106–143)	0.3–1.5	0.18	1.45	JW5	This study
JLB2c <sup>d</sup>	May 2016	460	103	0.224	2	12.7 (1.25)	0.099	125 (106–143)	10.3–15.4	6.89	5.79	JW4	This study
LKB2	July 2016	460	39	0.085	2	4.4 (0.71)	0.161	125 (106–143)	3.2–6.0	2.39	1.45	JW5	This study
Tbon3	July 2016	460	78	0.17	2	9.3 (1.05)	0.113	125 (106–143)	7.4–11.5	17.7	34.8	MWH1	This study
CJ3	September 2016	460	69	0.15	2	8.1 (0.98)	0.121	125 (106–143)	6.4–10.2	23.7	23.2	MWH2	This study
FWC3	November 2016	460	27	0.059	2	3.0 (0.58)	0.194	125 (106–143)	2.0–4.4	18.0	23.2	MWH2	This study
ARD3	December 2016	460	58	0.126	2	6.7 (0.89)	0.132	125 (106–143)	5.2–8.6	3.72	1.45	MWH5	This study
JLB3 <sup>d</sup>	January 2017	460	146	0.317	2	19.1 (1.59)	0.083	125 (106–143)	16.1–22.4	12.4	11.6	MWH3	This study
LKB3	February 2017	460	38	0.083	2	4.3 (0.70)	0.163	125 (106–143)	3.1–5.8	3.55	1.45	MWH5	This study

<sup>a</sup>Viability according to equation 1 (see text). ASE, asymptotic standard error.<sup>b</sup>Bootstrapped median, with 95% CI in parentheses. Based on 9,999 bootstraps.<sup>c</sup>Minimum–maximum 95% CI. Based on 9,999 bootstraps tested at viability increments of 0.1%.<sup>d</sup>Experiments where a subset of positive wells were transferred. FWC2 shows the advantage of our method over equation 1 for extreme values.

16S rRNA gene amplicon analyses to compare cultivation results with the microbial communities in the source waters. We have previously reported on the success of our artificial media in obtaining abundant taxa over the course of the first seven experiments from this campaign (35). Here, we expand our report to include cultivation results from a total of 17 experiments, and we update the classic viability calculations of Button et al. (33) with a new model to estimate the viability of individual taxa using relative-abundance information. New isolates belonged to cultivated groups in eight putatively novel genera and seven putatively novel species in previously cultivated genera and expanded cultured geographic representation for many important clades like SAR11. Additionally, using model-based predictions, we identified possible taxon-specific viability variation that can influence cultivation success. By incorporating these new viability estimates into the model, our method facilitates statistically informed experimental design for targeting individual taxa, thereby reducing uncertainty for future culturing work (59).

## RESULTS

**General cultivation campaign results.** We conducted a total of 17 DTE cultivation experiments to isolate bacterioplankton (sub-2.7- $\mu$ m fraction), with paired microbial community characterization of source waters (0.22- to 2.7- $\mu$ m fraction), from six coastal Louisiana sites over a 3-year period (see Table S1 in the supplemental material, available at <https://doi.org/10.6084/m9.figshare.12142113>). We inoculated 7,820 distinct cultivation wells (all experiments) with an estimated 1 to 3 cells · well<sup>-1</sup> using overlapping suites of artificial seawater media, JW (years 1 and 2 [35]) and MWH (year 3), designed to match the natural environment (Table 1). The MWH suite of media was modified from the JW media to additionally include choline, glycerol, glycine betaine, cyanate, dimethyl sulfoxide (DMSO), dimethylsulfoniopropionate (DMSP), thiosulfate, and orthophosphate (Table S1). These compounds have been identified as important metabolites and osmolytes for marine and freshwater microorganisms and were absent in the first iteration (JW) of our media (60–66). A total of 1,463 wells were positive (>10<sup>4</sup> cells · ml<sup>-1</sup>), and 738 of these were transferred to 125-ml polycarbonate flasks. For four experiments (FWC, FWC2, JLB2, and JLB3) we transferred only a subset of positives (48/301, 60/403, 60/103, and 60/146) because the number of isolates exceeded our ability to maintain and identify them at that time (Table 1). The subset of positive wells for these four experiments was selected using flow cytometry signatures usually

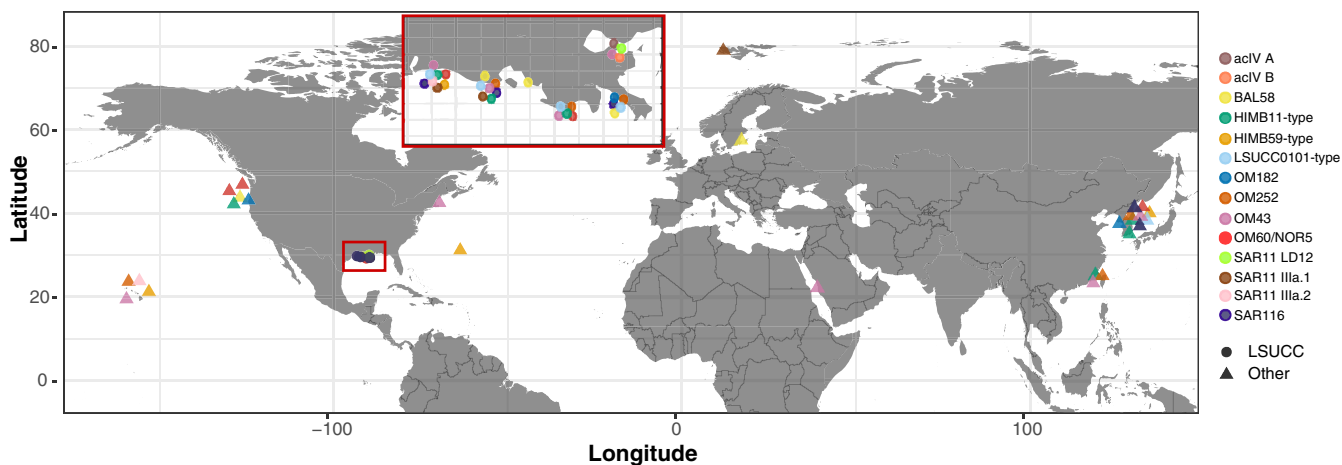


**FIG 1** Percent identity of LSUCC isolate 16S rRNA genes compared with those from other isolates in NCBI (“Other,” gray dots) or from the DTE culture collections IMCC (gold dots), HTCC (blue dots), and HIMB (green dots). Each dot represents a pairwise 16S rRNA gene comparison (via BLASTn). x-axis categories are groups designated according to  $\geq 94\%$  sequence identity and phylogenetic placement (see Fig. S1 to S5 in the supplemental material). Above the graph is the 16S rRNA gene sequence percent identity to the closest non-LSUCC isolate within a column. Groups colored in red are those where LSUCC isolates represent putatively novel genera, whereas orange indicates putatively novel species.

indicative of smaller oligotrophic cells like SAR11 strain HTCC1062 (49) using our settings. Of the 738 wells from which we transferred cells across all experiments, 328 of these yielded repeatedly transferable isolates that we deemed pure cultures based on 16S rRNA gene PCR and Sanger sequencing.

**Phylogenetic and geographic novelty of our isolates.** The 328 isolates belonged to three phyla: *Proteobacteria* ( $n = 319$ ), *Actinobacteria* ( $n = 8$ ), and *Bacteroidetes* ( $n = 1$ ) (see Fig. S1 to S5 in the supplemental material). We placed these isolates into 55 groups based on their positions within 16S rRNA gene phylogenetic trees (Fig. S1 to S5) and as a result of having  $\geq 94\%$  16S rRNA gene sequence identity to other isolates. We applied a nomenclature to each group based on previous 16S rRNA gene database designations and/or other cultured representatives (Fig. 1; Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]). Isolates represented eight putatively novel genera with  $\leq 94.5\%$  16S rRNA gene identity to a previously cultured representative: the *Actinobac-*

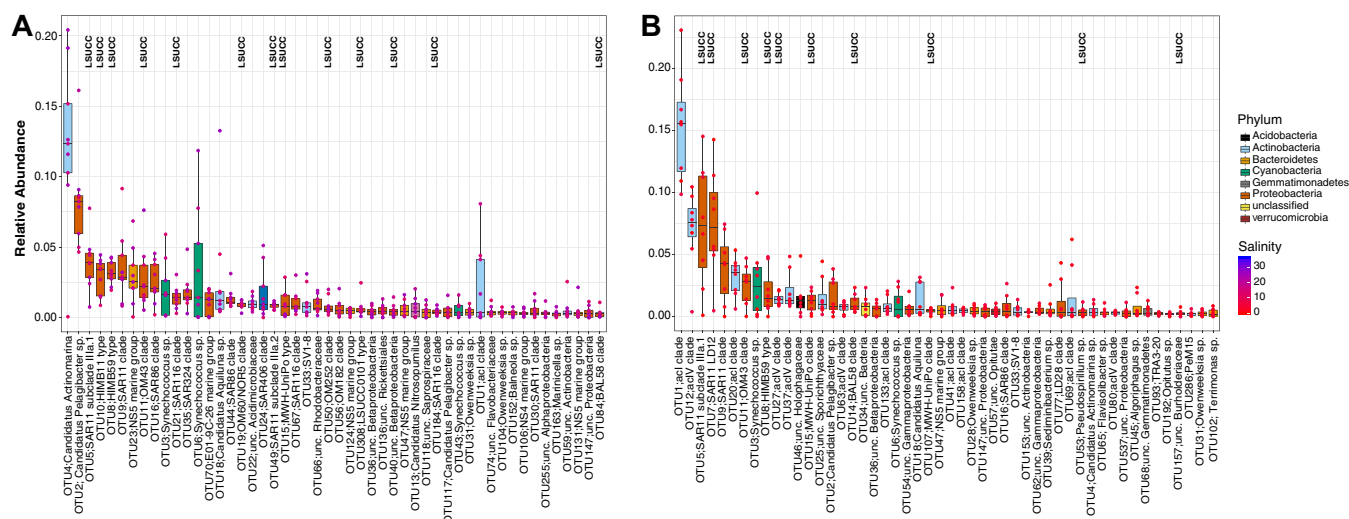




**FIG 2** A global map of the isolation locations of isolates from selected important aquatic bacterioplankton clades. All depicted taxa were isolated from surface water (<1 to 20 m), or the depth of the sample was not reported (see Table S1 in the supplemental material [available at <https://doi.org/10.6084/m9.figshare.12142113>] for details). Circles represent LSUCC isolates, while triangles are non-LSUCC isolates. Inset, a zoomed view of the coastal Louisiana region where LSUCC bacterioplankton originated. The map was made with the R package ggplot2 using the command map\_data (see <https://github.com/thrash-lab/Cultivar-novelty-map>).

teria acIV subclades A and B and one other unnamed *Actinobacteria* group, an undescribed *Acetobacteraceae* clade (*Alphaproteobacteria*), the freshwater SAR11 LD12 (“*Candidatus Fonsibacter ubiquis*” [29]), the MWH-UniPo and an unnamed *Burkholderiaceae* clade (*Betaproteobacteria*), and the OM241 *Gammaproteobacteria* (Fig. 1; Table S1). Seven additional putatively novel species in other genera were also isolated (between 94.6 and 96.9% 16S rRNA gene sequence identity) in unnamed *Comamonadaceae* and *Burkholderiales* clades (*Betaproteobacteria*), the SAR92 clade and *Pseudohongiella* genus (*Gammaproteobacteria*), and unnamed *Rhodobacteraceae* and *Bradyrhizobiaceae* clades, as well as *Maricaulis* spp. (*Alphaproteobacteria*) (Fig. 1). Louisiana State University Culture Collection (LSUCC) isolates belonging to the groups BAL58 *Betaproteobacteria* (Fig. S4), OM252 *Gammaproteobacteria*, HIMB59 *Alphaproteobacteria*, and what we designated the LSUCC0101-type *Gammaproteobacteria* (Fig. S5) had close 16S rRNA gene matches to other isolates at the species level; however, none of those previously cultivated organisms have been formally described (Fig. 1). The OM252, BAL58, and MWH-UniPo clades were the most frequently cultivated, with 124 of our 328 isolates belonging to these three groups (Table S1). In total, 73 and 10 of the 328 isolates belonged in putatively novel genera and novel species in previously cultivated genera, respectively. We estimated that at least 310 of these isolates were geographically novel, being the first of their type cultivated from the nGOM (Fig. 2). This included isolates from cosmopolitan groups like SAR11 subclade IIIa, OM43 *Betaproteobacteria*, SAR116, and HIMB11-type “*Roseobacter*” spp. Cultivars from *Vibrio* sp. and *Alteromonas* sp. were the only two groups with close relatives (species level) isolated from the GOM.

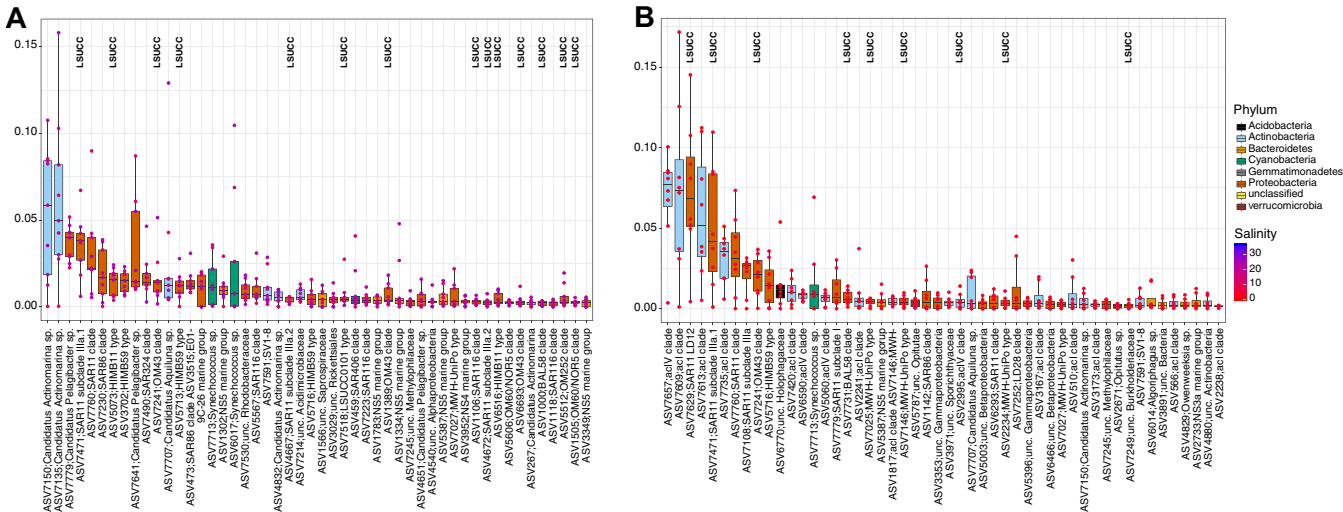
**Natural abundance of isolates.** We matched LSUCC isolate 16S rRNA gene sequences with both operational taxonomic units (OTUs) and amplicon single-nucleotide variants (ASVs) from bacterioplankton communities to assess the relative abundances of our isolates in the coastal nGOM waters that served as inocula. While OTUs provide a broad group-level designation (97% sequence identity), this approach can artificially combine multiple ecologically distinct taxa (67). Due to higher stringency for defining a matching 16S rRNA gene, ASVs can increase the confidence that our isolates represent environmentally relevant organisms (68, 69). However, while many abundant oligotrophic bacterioplankton clades, such as SAR11 (29, 70), OM43 (40, 41), SAR116 (71), and *Sphingomonas* spp. (72), have a single copy of the rRNA gene operon, other taxa can have multiple rRNA gene copies (70, 73), complicating ASV analyses. Since we could not *a priori* rule out multiple rRNA gene operons for novel groups with no genome-sequenced representatives, we used both OTU and ASV approaches.



**FIG 3** Rank abundances of the 50 most abundant OTUs from all sites based on median relative abundance at salinities greater than 12 (A) and less than 7 (B). The boxes indicate the interquartile range (IQR) of the data, with vertical lines indicating the upper and lower extremes according to  $1.5 \times \text{IQR}$ . Horizontal lines within each box indicate the median. The data points comprising the distribution are plotted on top of the boxplots. The shade of the dot represents the salinity at the sample site (red to blue, lower to higher), while the color of the box indicates broad taxonomic identity. LSUCC labels indicate OTUs with at least one cultivated representative.

In total, we obtained at least one isolate from 40 of the 777 OTUs and 71 of the 1,323 ASVs observed throughout the 3-year data set. Forty-three percent and 26% of LSUCC isolates matched the top 50 most abundant OTUs (median relative abundances for all sites, from 8.1 to 0.11% [see Fig. S6A in the supplemental material]) and ASVs (mean relative abundances for all sites, from 3.8 to 0.11% [Fig. S6B]), respectively, across all sites and samples. Microbial communities from all collected samples clustered into two groups corresponding to those inhabiting salinities below 7 and above 12, and salinity was the primary environmental driver distinguishing community beta diversity (OTU,  $R^2 = 0.88$ ,  $P = 0.001$ ; ASV,  $R^2 = 0.89$ ,  $P = 0.001$ ). As part of the cultivation strategy after the first five experiments, we used a suite of five media differing by salinity and matched the experiment with the medium that most closely resembled the salinity at the sample site. Consequently, our isolates matched abundant environmental groups from both high- and low-salinity regimes. At salinities above 12, LSUCC isolates matched 13 and 14 of the 50 most abundant OTUs and ASVs, respectively (Fig. 3A and 4A; Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]). These taxa included the abundant SAR11 subclade IIIa.1, HIMB59, HIMB11-type “*Roseobacter*,” and SAR116 *Alphaproteobacteria*; the OM43 *Betaproteobacteria*; and the OM182 and LSUCC0101-type *Gammaproteobacteria*. At salinities below seven, 10 and 9 of the 50 most abundant OTUs and ASVs, respectively, were represented by LSUCC isolates, including one of the most abundant taxa in both cluster sets, SAR11 LD12 (Fig. 3B and 4B). Some taxa, such as SAR11 IIIa.1 and OM43, were among the top 15 most abundant taxa in both salinity regimes (Fig. 3 and 4; Table S1), suggesting a euryhaline lifestyle. In fact, our cultured SAR11 IIIa.1 ASV7471 was the most abundant ASV in the aggregate data set (Fig. S6).

Overall, this effort isolated taxa representing 18 and 12 of the top 50 most abundant OTUs and ASVs, respectively (Table 2; Fig. S6). When categorized by different median relative abundance categories of  $>1\%$ ,  $0.1\%$  to  $1\%$ , and  $<0.1\%$ , isolate OTUs were distributed across those categories at 15%, 20%, and 27%, and isolate ASVs were distributed accordingly at 4%, 26%, and 37% (Table 2). Isolates with median relative abundances of  $<0.1\%$ , such as *Pseudohongiella* spp., *Rhodobacter* spp., and *Bordetella* spp., would canonically fall within the rare biosphere (74) (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]). A number of isolates did not match any identified OTUs or ASVs (38% and 33% of LSUCC isolates compared to available



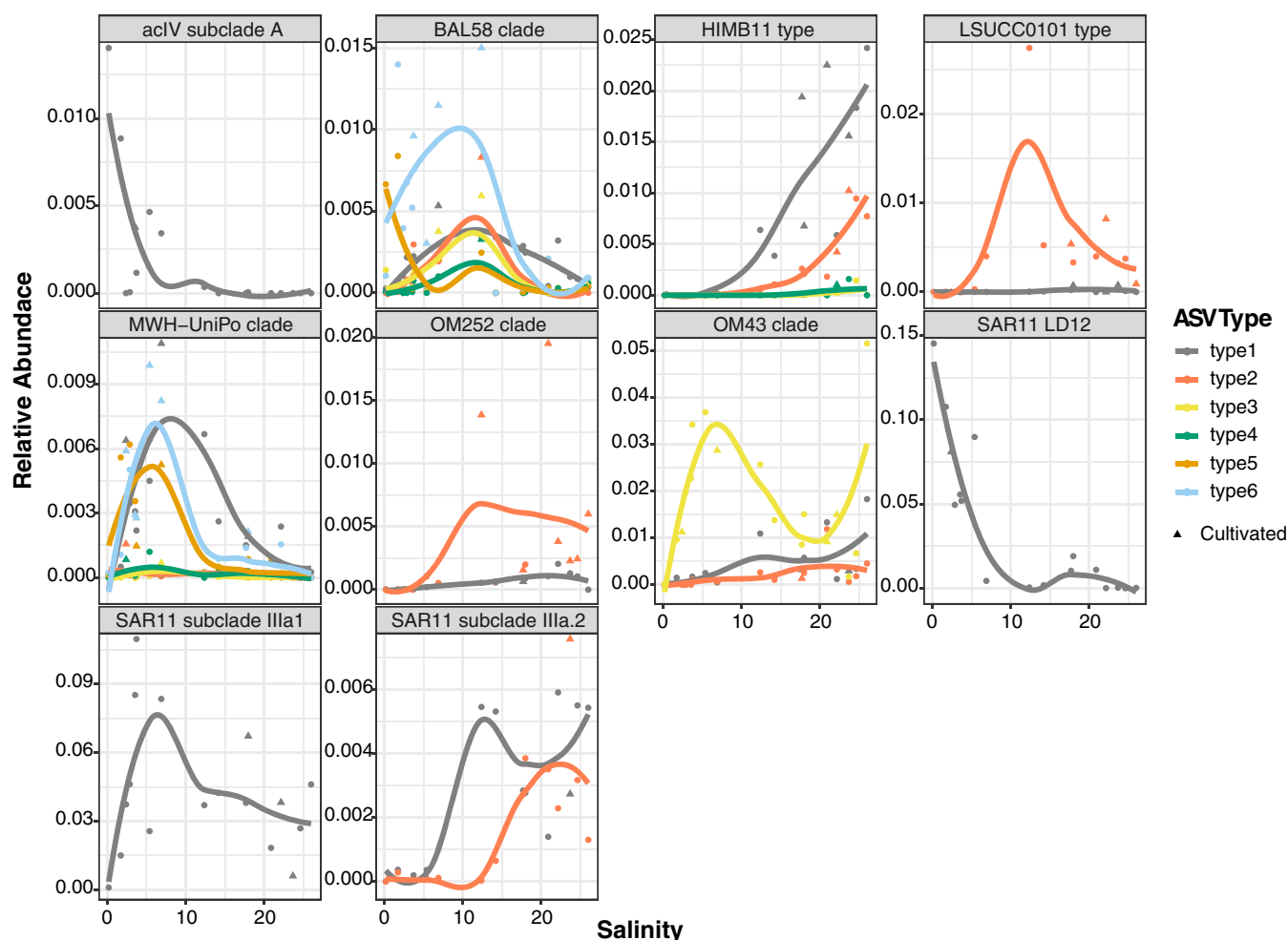
**FIG 4** Rank abundances of the 50 most abundant ASVs from all sites based on median relative abundance at salinities greater than 12 (A) and less than 7 (B). The boxes indicate the interquartile range (IQR) of the data, with vertical lines indicating the upper and lower extremes according to  $1.5 \times \text{IQR}$ . Horizontal lines within each box indicate the median. The data points comprising the distribution are plotted on top of the boxplots. The shade of the dot represents the salinity at the sample site (red to blue, lower to higher), while the color of the box indicates broad taxonomic identity. LSUCC labels indicate ASVs with at least one cultivated representative.

OTUs and ASVs, respectively), either because their matching OTUs/ASVs were below our thresholds for inclusion (at least two reads from at least two sites) or because they were below the detection limit from our sequencing effort (Table 2). Thus, 43% and 30% of our isolates belonged to OTUs and ASVs, respectively, with median relative abundances of  $>0.1\%$ .

**Modeling DTE cultivation.** An enigma that became immediately apparent through a review of our data was the absence of an obvious relationship between the abundance of a given taxon in the inoculum and the frequency of obtaining an isolate of the same type from a DTE cultivation experiment (see Fig. S7 and S8 in the supplemental material). For example, although we could culture SAR11 LD12 over a range of medium conditions (29) and the matching ASV had relative abundances of  $>5\%$  in six of our 17 experiments (Fig. 5), we isolated only one representative (LSUCC0530). In an ideal DTE cultivation experiment where cells are randomly subsampled from a Poisson-distributed population, if the medium is sufficient for a given microorganism's growth, then the number of isolates should correlate with that microorganism's abundance in the inoculum. However, a qualitative examination of several abundant taxa that grew in our media, some of which we cultured on multiple occasions, revealed no clear pattern between abundance and isolation success (Fig. 5). Considering that medium composition was sufficient for cultivation of these organisms on at least some occasions, we hypothesized that cultivation frequency may reflect differences in the capacity for growth within populations of a given taxon. Thus, we decided to model cultivation frequency in relationship to estimated abundances in a way that could generate estimates of cellular viability, defined herein as meaning "presently able to grow in defined medium," as opposed to a broader definition equating viability with being alive more generally, since we evaluated only growth capacity in this study. We

**TABLE 2** Median relative abundances of cultured OTUs and ASVs across all samples

Parameter	No. in top 50 ranks (no. of isolates)	No. (%) of isolates			
		With <i>r</i> :			Not detected
		$>1\%$	$1\%–0.1\%$	$<0.1\%$	
OTU	18 (140)	50 (15)	90 (27)	65 (20)	123 (38)
ASV	12 (84)	13 (4)	84 (26)	122 (37)	109 (33)



**FIG 5** Relative abundance of ASVs within key taxonomic groups compared with salinity. ASV types are colored independently, and triangle points indicate experiments for which at least one isolate was obtained. Nonlinear regression lines are provided as a visual aid for abundance trends.

hoped that modeling might also help us inform experimental design and make DTE cultivation efforts more predictable (59).

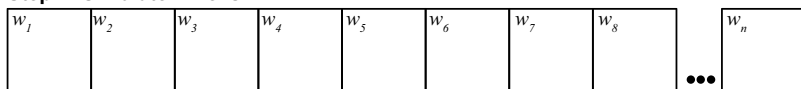
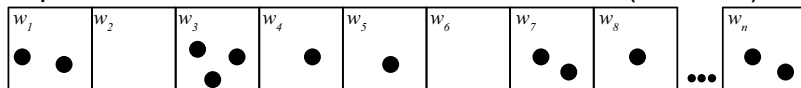
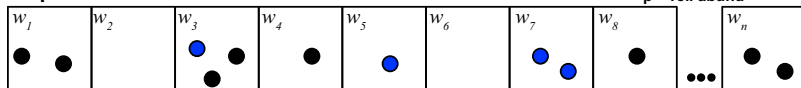
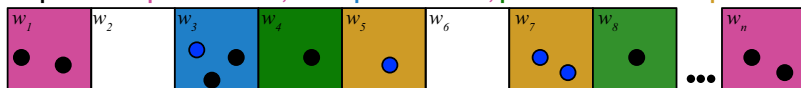
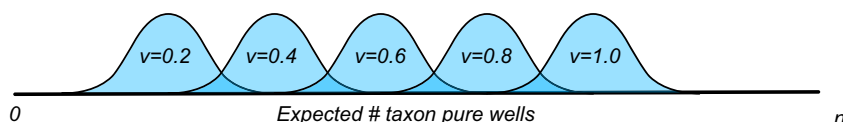
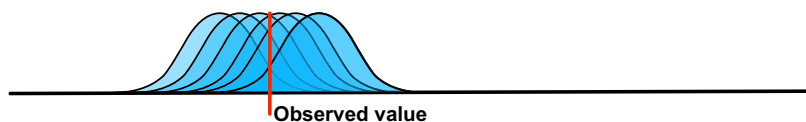
Previously, Don Button and colleagues developed a statistical model for viability ( $V$ ) of cells in the entire population for a DTE experiment (33):

$$V = -\frac{\ln(1-p)}{\lambda} \quad (1)$$

Where  $p$  is the proportion of wells or tubes,  $n$ , with growth  $z$  ( $p = z/n$ ) and  $\lambda$  is the estimated number of cells inoculated per well (the authors used  $X$  originally). The equation uses a Poisson distribution to account for the variability in cell distribution within the inoculum and therefore the variability in the number of wells or tubes receiving the expected number of cells. We and others have used this equation in the past (26, 28, 35) to evaluate the efficacy of our cultivation experiments in the context of commonly cited numbers for cultivability using agar plate-based methods (13, 17, 75).

While equation 1 was effective for its intended purpose, it has a number of drawbacks that limit its utility for taxon-specific application: (i) if  $p = 1$ , i.e., all wells are positive, then the equation is invalid; (ii) At high values of  $p$  and low values of  $\lambda$ , estimates of  $V$  can exceed 100% (Table 1); (iii) accuracy of viability, calculated by the asymptotic standard error (ASE) or the coefficient of variation (CV), was shown to be nonuniform across a range of  $\lambda$ , with greatest accuracy when true viability was  $\sim 10\%$



**Step 1: Simulate  $n$  wells****Step 2: Simulate inoculation of wells from Poisson distribution ( $\lambda$ =inoculum)****Step 3: Simulate likelihood of taxon from Binomial distribution ( $n$  = # cells in well,  $p$  = rel. abund)****Step 4: Count positive wells, taxon positive wells, pure wells and taxon pure wells****Step 5: For taxon pure wells, simulate likelihood of viability from Binomial distribution ( $n$  = # cells in well,  $p$  = viability)****Step 6: Count wells where viable cells  $\geq 1$** **Step 7: Bootstrap steps 1-6  $k$  times at different levels of viability,  $0 \leq v \leq 1$** **Step 8: Identify min, max values of  $v$  where # observed wells falls within bootstrapped 95% CI****FIG 6** Graphical depiction of the viability model.

(33), and thus, low viability, low values of  $\lambda$ , and small values of  $n$  were found to produce unreliable results; (iv) if  $p = 0$ , i.e., no positive wells are observed, estimates of viability that could produce 0 positive wells cannot be calculated; and (v) Button's original model assumes that a well will produce a pure culture only if the inoculated well contains one cell, but in contrast, in low-diversity samples, samples dominated by a single taxon, or experiments evaluating viability from axenic cultures across different media, a limitation that only wells with single cells are axenic will underestimate the expected number of pure wells.

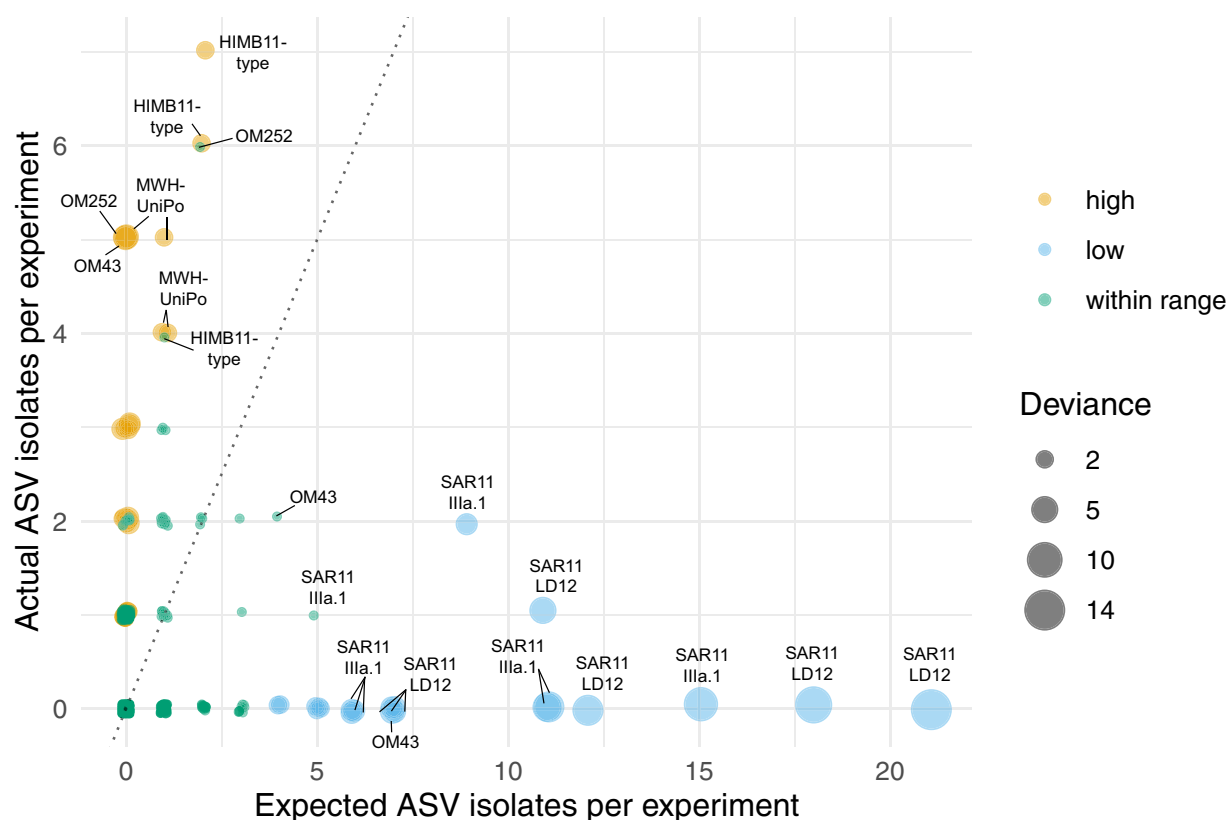
To overcome these limitations, we developed a Monte Carlo simulation model that facilitates the incorporation of relative-abundance data from complementary community profiling data (e.g., 16S rRNA gene amplicons) to calculate the likelihood of positive wells, pure wells, and viability at a taxon-specific level, based on the observed number of wells for which we obtained an isolate of a particular taxon (Fig. 6). By employing a Monte Carlo approach, our model is robust across all values of  $p$  and  $n$  with uniform prediction accuracy, and we can estimate the accuracy of our prediction within 95% confidence intervals (CI). Furthermore, the width of 95% CI boundaries of viability, as well as the expected number of positive and pure wells, is entirely controllable and

dependent only on available computational capacity for bootstrapping (i.e., these can be improved with more bootstrapping, but at greater computational cost). When zero positive wells are observed experimentally, our approach enables estimation of a maximum viability that could explain such an observation by identifying the range of variability values for which zero resides within the bootstrapped 95% CI. Finally, the ability to calculate the viability of the entire community, as in equation 1, is retained simply by estimating viability using a relative abundance of one.

We compared our model to that of Button et al. for evaluating viability from whole-community experimental results, similarly to previous reports (26, 28, 35) (Table 1). Our viability estimates ( $V_{\text{est}}$ ) generally agreed with those using equation 1, but we have now provided 95% CI to depict the maximum and minimum viabilities that would match the returned positive-well distribution, as well as maximum and minimum values for the number of wells that ought to have contained a single cell. Maximum  $V_{\text{est}}$  ranged from 1.1% to >92.3% depending on the experiment, with a median  $V_{\text{est}}$  across all our experiments of 8.6% (Table 1). In one case, the extremely high value (FWC2) was better handled by our model than by equation 1, because it did not lead to a viability estimation greater than 100%. FWC and FWC2 represent  $V_{\text{est}}$  outliers compared with the entire data set (maximums of 59.7% and >92.3%, respectively) (Table 1). We believe these high numbers most likely resulted from underestimating the number of cells inoculated into each well (because of the use of microscopy, the presence of clumped cells, or possible pipette error [described in reference 35]), thus increasing the estimated viability.

**Isolate-specific viability estimates.** Our new model also facilitates taxon-specific viability estimates. Cultivation efficacy was evaluated for 71 cultured taxa matching ASVs within our detection limits (219 isolates) across 17 sites (1,207 pairwise combinations) by comparing the number of observed pure wells to those predicted by the Monte Carlo simulation using 9,999 bootstraps, 460 wells per experiment, and an assumption that all cells were viable (i.e.,  $V = 100\%$ ). In total, for 1,158 out of 1,207 pairwise combinations (95.9%), the observed number of pure wells fell within the 95% CI of data simulated at matching relative abundance and inoculum size, suggesting that these two parameters alone could explain the observed cultivation success for most taxa (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]). A total of 1,059 out of these 1,158 combinations (91%) recorded zero observed wells, but with a maximum relative abundance of 2.8% within these combinations, a score of zero fell within predicted 95% CI of simulations with 460 wells. Sensitivity analysis showed that with 460 wells per experiment, an observation of zero pure wells falls below the 95% CI's lower bound (and is thus significantly depleted to enable viability to be estimated) for taxa with relative abundances of 2.3%, 2.9%, and 4.5% for inoculum sizes of one, two, and three cells per well, respectively (see Fig. S9 in the supplemental material). In fact, modeling DTE experiments from 92 wells to 9,200 wells per experiment showed that for taxa comprising ~1% of a microbial community 1,104 wells (or 12 plates at 92 wells per plate), 1,380 wells (15 plates), and 2,576 wells (28 plates) were required to be statistically likely to recover at least one positive, pure well using inocula of one, two, or three cells per well, respectively, with  $V = 100\%$  (Fig. S9).

A small but taxonomically relevant minority (49 out of 1,207) of pairwise combinations had a number of observed pure wells that fell outside the simulated 95% CI with  $V = 100\%$  (Fig. 7). Of these, 28 had either one, two, or three more observed pure wells than the upper 95% CI (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]), suggesting cultivability higher than expected based purely a model capturing the interaction between a Poisson-distributed inoculum and a binomially distributed relative abundance, with  $V = 100\%$ . However, the deviance from the expected number of positive wells for those above the 95% CI was limited to three or fewer wells, meaning that we obtained only 1 to 3 more isolates than expected (Table S1). Conversely, those organisms that we isolated less frequently than expected showed greater deviance. 21 out of the 49 outliers had lower-than-expected cultivability (Fig. 7).



**FIG 7** Actual versus expected numbers of isolates. Each point represents the actual number of isolates for every ASV/experiment pair compared to the expected number of isolates based on our model assuming 100% viability. Colors represent the relationship to the model predictions: green, isolates within the 95% CI for expected frequency; orange, actual isolates > maximum 95% CI for expected isolates; blue, actual isolates < minimum 95% CI for expected isolates. Circle size is proportional to the deviation of the number of actual isolates from the maximum (for orange) or minimum (for blue) 95% CI for expected isolates. The dotted line is the 1:1 ratio. Notable taxa on the extremities of the actual and expected values are labeled. All data points are provided in Table S1 in the supplemental material (available at <https://doi.org/10.6084/m9.figshare.12142113>).

These taxa had relative abundances ranging from 2.7% to 14.5% but recorded only 0, 1, or 2 isolates. In the most extreme case, ASV7629 (SAR11 LD12) at site ARD2c comprised 14.5% of the community but recorded no observed pure wells, compared to expected number of 13 to 30 isolates (95% CI) predicted by the Monte Carlo simulation. All the examples of taxa that were isolated less frequently than expected given the assumption of  $V = 100\%$  belonged to either SAR11 LD12, SAR11 IIIa.1, or one particular OM43 ASV (7241) (Fig. 6).

We used our model to calculate estimated viability ( $V_{\text{est}}$ ) for these organisms based on their cultivation frequency at sites where the assumption of  $V = 100\%$  appeared to be violated (Table 3). Using the extreme example of SAR11 LD12 ASV7629 at site ARD2c, simulations across a range of  $V$  indicated that a result of zero positive wells fell within 95% of simulated values when the associated taxon  $V_{\text{est}}$  was  $\leq 15\%$ . When considering all anomalous cultivation results, LD12 had estimated maximum viabilities that ranged up to 55% (Table 3). OM43 (ASV7241) estimated maximum viabilities ranged from 52 to 80%, depending on the site, and similarly, SAR11 IIIa.1 ranged between 22 and 82% maximum viability (Table 3).

## DISCUSSION

This work paired 17 DTE cultivation experiments with cultivation-independent assessments of microbial community structure in source waters to evaluate cultivation efficacy. We generated 328 new bacterial isolates representing 40 of the 777 OTUs and 71 of the 1,323 ASVs observed across all samples from which we inoculated DTE experiments. Stated another way, we successfully cultivated 5% of the total 3-year

**TABLE 3** Estimated viabilities for taxa cultivated less frequently than expected

ASV	Group	Site	$r^a$	$n$	$z$	$\lambda$	Estimated no. of wells with 1 cell if $V = 1^b$	$V_{est}^c$
7241	OM43	ARD3	0.03	460	0	2	4 (1–9)	0.1–80
7241	OM43	FWC <sup>d</sup>	0.04	460	0	2	5 (1–9)	0.1–77
7241	OM43	JLB	0.05	460	0	1.96	7 (2–12)	0.1–52
7471	SAR11 IIIa.1	ARD3	0.11	460	0	2	15 (8–23)	0.1–22
7471	SAR11 IIIa.1	CJ	0.03	460	0	1.27	4 (1–9)	0.1–82
7471	SAR11 IIIa.1	FWC3	0.07	460	2	2	9 (4–15)	2.5–80
7471	SAR11 IIIa.1	JLB	0.05	460	0	1.96	6 (2–11)	0.1–59
7471	SAR11 IIIa.1	JLB2c <sup>d</sup>	0.08	460	0	2	11 (5–18)	0.1–31
7471	SAR11 IIIa.1	JLB3 <sup>d</sup>	0.04	460	0	2	5 (1–9)	0.1–74
7471	SAR11 IIIa.1	LKB	0.05	460	0	1.8	6 (2–12)	0.1–55
7471	SAR11 IIIa.1	LKB2	0.04	460	0	2	5 (1–9)	0.1–77
7471	SAR11 IIIa.1	LKB3	0.09	460	0	2	11 (6–19)	0.1–30
7471	SAR11 IIIa.1	TBON2	0.04	460	0	1.56	6 (2–12)	0.1–56
7471	SAR11 IIIa.1	TBON3	0.04	460	0	2	5 (1–10)	0.1–73
7629	SAR11 LD12	ARD	0.11	460	0	1.5	18 (10–27)	0.1–20
7629	SAR11 LD12	ARD2c	0.15	460	0	2	21 (13–30)	0.1–15
7629	SAR11 LD12	ARD3	0.05	460	0	2	7 (2–12)	0.1–53
7629	SAR11 LD12	FWC <sup>d</sup>	0.09	460	0	2	12 (6–19)	0.1–28
7629	SAR11 LD12	LKB	0.05	460	0	1.8	7 (2–13)	0.1–51
7629	SAR11 LD12	LKB2	0.08	460	1	2	11 (5–18)	0.3–49
7629	SAR11 LD12	LKB3	0.06	460	0	2	7 (3–13)	0.1–48

<sup>a</sup>Fractional relative abundance.<sup>b</sup>Bootstrapped median, with 95% CI in parentheses. Based on 9,999 bootstraps.<sup>c</sup>Minimum–maximum 95% CI based on cultivation results. Based on 9,999 bootstraps tested at viability increments of 0.1%.<sup>d</sup>Experiments where a subset of positive wells were transferred.

bacterioplankton community observed via either OTU or ASV analyses. A large fraction of our isolates (43% of cultured OTUs and 30% of cultured ASVs) represented taxa present at median relative abundances of >0.1%, with 15% and 4% of cultured OTUs and ASVs, respectively, at median abundances of >1%. One hundred forty of our isolates matched the top 50 most abundant OTUs, and 84 isolates matched the top 50 most abundant ASVs.

This campaign led to the first isolations of the abundant SAR11 LD12 and *Actinobacteria* acIV, the second isolate of the HIMB59 *Alphaproteobacteria*, and new genera within the *Acetobacteraceae*, *Burkholderiaceae*, OM241 and LSUCC0101-type *Gammaproteobacteria*, and MWH-UniPo *Betaproteobacteria*, thereby demonstrating again that continued DTE experimentation leads to isolation of previously uncultured organisms with value for aquatic microbiology. We have also added a considerable collection of isolates to previously cultured groups such as OM252 *Gammaproteobacteria*, BAL58 *Betaproteobacteria*, and HIMB11-type “*Roseobacter*” spp., and the majority of our isolates represent the first versions of these types of taxa from the Gulf of Mexico, adding comparative biogeographic value to these cultures.

Our viability model improved upon the statistical equation developed by Button and colleagues (33) to extend viability estimates to individual taxa within a mixed community and provide 95% CI to constrain those estimates. We cultured several groups of organisms abundant enough to evaluate viability with 460 wells (Fig. 7; see Fig. S9 in the supplemental material). The fact that these organisms were successfully cultured at least once meant that we could reasonably assume that the medium was sufficient for growth.

Some taxa were cultivated more frequently than expected (Fig. 7). We explore two possible explanations for this phenomenon: errors in quantification and variation in microbial cell organization. Any systematic error that led to underestimating the abundance of an organism would have correspondingly resulted in our underestimating the number of wells in which we would expect to find a pure culture of that organism. Such underestimations could come from primer biases associated with amplicon sequencing (76–78), but we do not know if those protocols specifically underestimate the OM252, MWH-UniPo, and HIMB11-type taxa cultured more fre-

quently than expected (Fig. 7). However, due the low number of expected isolates in these groups and the small deviances in actual isolates from those expected numbers (within 1 to 3 isolates compared to expected values), the biases inherent in the relative abundance estimations for these taxa were probably small. Furthermore, one of the microorganisms isolated more frequently than expected matched the OM43 ASV1389 (see Fig. S6 in the supplemental material), whereas another OM43 ASV (7241) was cultivated less frequently than expected (see below), meaning that if primer bias were the cause of this discrepancy, it would have to be operating differently on very closely related organisms.

One possible biological explanation for why some isolates might have been cultured more frequently than expected is clumped cells. If cells of any given taxon in nature grew in small clusters, then the number of cells we added to a well would have been greater than expected based on a Poisson distribution. Furthermore, the model assumes that each cell is independent and that the composition of a subset of cells is only a function of the relative abundance of the taxon in the community. Within a cluster of cells, this assumption is violated, as the probability of cells being from the same taxon is higher. Thus, the model will underestimate the probability of a well being pure and therefore underestimate the number of pure wells likely to be observed within an experiment, leading to a greater number of isolates than expected. Future microscopy work could examine whether microorganisms such as OM252 and MWH-UniPo form small clusters *in situ* and/or in pure culture and whether this phenomenon may be different for different ASVs of OM43, or if clumping may be a transient phenotype.

We also identified three taxa—SAR11 LD12, SAR11 subclade IIIa.1, and the aforementioned OM43 ASV7241—that were isolated much less frequently than expected based on their abundances (Fig. 7; Table 3). This could mean that our assumption of  $V = 100\%$  was incorrect or that, in contrast to the taxa that were cultured more frequently than expected (see above), our methods had biases that overestimated the abundance of these organisms, thereby overinflating the expected number of isolates. We used the modified 515/806RB primers, which have been shown to be much more accurate in quantifying SAR11 compared to fluorescent *in situ* hybridization (FISH) than the original 515/806 primers (within  $6\% \pm 4\%$  [standard deviation]), and this protocol almost always underestimates SAR11 abundance (76). This suggests that our expected number of isolates may have actually been underestimated and our cultivation success poorer than we measured, and therefore we may be overestimating viability for the SAR11 taxa in this study. Other sources of systematic error that might impinge on successful transfers, and thereby reduce our recovery, include sensitivity to pipette tip and/or flask material. However, the fact that these taxa were sometimes successfully isolated means that if these mechanisms were impacting successful transfers, then their activity was less than 100% efficient, which implies variations in subpopulation vulnerability that would be very similar conceptually to variations in subpopulation viability.

Another possible source of error that could have resulted in lower-than-expected numbers of isolates was the subset of experiments for which we did not transfer all positive wells due to limitations in available personnel time (Tables 1 and 3). However, our selection criteria for the subset of wells to transfer was based on flow cytometric signatures that would have encompassed small cells like SAR11 (see Results), and in any case, there were many examples of lower-than-expected recovery from other experiments where we transferred all positive wells (Table 3). Thus, we believe that these four experiments were unlikely to contribute major errors biasing our estimates of viability for SAR11 LD12, SAR11 IIIa.1, and other small cells like OM43.

If we instead explore biological reasons for the lower-than-expected numbers of positive wells in DTE experiments, a plausible explanation supported by the literature is simply that a large fraction of the population is in some state of inactivity or at least not actively dividing (79). Studies using uptake of a variety of radiolabeled carbon and sulfur sources have demonstrated that substantial fractions of SAR11 cells may be inactive, depending on the population (80–83). SAR11 cells in the northwest Atlantic and Mediterranean showed variable uptake of labeled leucine (30 to 50% [80, 81] and



~25 to 55% [83, 84]) and amino acids (34 to 61% [80, 81] and 34 to 66% [80, 81]). Taken in reverse, this means that up to 75% of the SAR11 population may be dormant at any given time. In another study focused on brackish communities, fewer than 10% of SAR11 LD12 cells took up labeled leucine and/or thymidine (82). While this was likely not the ideal habitat for LD12 based on salinities above six (29, 82), this study supports the others described above that show substantial proportions of inactive SAR11 cells, the fraction of which may depend on environmental conditions and other unknown factors. Biorthogonal noncanonical amino acid tagging (BONCAT) showed a similar trend for SAR11 (85). These results also match general data indicating prevalent inactivity among aquatic bacterioplankton (79, 86–88). Although labeled uptake methods do not directly measure rates of cell division, the incorporation of these compounds requires active DNA replication or translation, which represent an even more fundamental level of activity than cell division (89).

Why might selection favor high percentages of subpopulation dormancy? One possibility is as an effective defense mechanism against abundant viruses. Viruses infecting SAR11 have been shown to be extremely abundant in both marine (90, 91) and freshwater (92) systems. Indeed, the paradox of high viral abundances and high host abundances in SAR11 has led to a refining of negative density-dependent selection through Lotka-Volterra predator-prey dynamics (93) to include heterogeneous susceptibility at the strain level (94, 95) and positive density-dependent selection through intraspecific proliferation of defense mechanisms (96). Activity of lytic viruses infecting SAR11 *in situ* demonstrated that phages infecting SAR11 have lower ratios of viral transcripts to host cells than in other abundant taxa and that observed abrupt changes in these ratios suggest coexistence of several SAR11 strains with different life strategies and phage susceptibilities (97). Phenotypic stochasticity of phage receptor expression has been shown to maintain a small proportion of phage-insensitive hosts within a population, enabling coexistence of predator and prey without extinction (98). Phages adsorb to a vast array of receptor proteins on their hosts, with many well-characterized receptors (e.g., OmpC, TonB, BtuB, and LamB) associated with nutrient uptake or osmoregulation (99). Selection therefore favors phenotypes that limit receptor expression, with an associated fitness cost, particularly in nutrient-limited environments.

However, an alternative mechanism is possible if a population of cells comprised a small number of susceptible cells and a large number of either resistant or dormant cells where presentation of receptor proteins is retained. The majority of host-virus encounters would occur with resistant or dormant cells and would constrain viral propagation through inefficient or failed infection, effectively acting as a sink for infectious particles. Prevalent lysogeny in SAR11 populations would provide a mechanism for establishing resistant cells via superinfection immunity (100, 101), where integration of a temperate phage prevents infection by other closely related viruses. There is growing evidence that many viruses infecting SAR11 are temperate (102, 103) and that reversion to virulence can be triggered through nutrient limitation (103), in contrast to other systems where lysogeny is favored in nutrient-poor conditions (104). Viral infection may also trigger host dormancy, lowering cellular metabolism to minimize energy requirements under nutrient-limited conditions (105). Such cells would be selected against during cultivation experiments, potentially explaining the rarity of SAR11 isolate genomes found to contain prophages. Dormancy and/or lysogeny would also enable long-term costability between abundant phages and their hosts (106) and resolve the apparent paradox of high host and virus abundances (101).

Detailed measurements of dormancy in SAR11 and what types of cellular functions become inactivated are part of our ongoing work. In the meantime, it is prudent to examine the implications of a substantial proportion of nondividing cells for our understanding of basic growth dynamics. Studies attempting to measure SAR11 growth rates in nature have yielded a wide range of results, ranging from 0.03 to 1.8 day<sup>-1</sup> (70, 80, 83, 107–109). These span wider growth rates than observed for axenic cultures of SAR11 (0.4 to 1.2 day<sup>-1</sup>), but isolate-specific growth ranges within that

spread are much more constrained (29, 36, 49, 110, 111). Conversion factors for determining production from [ $^3\text{H}$ ]leucine incorporation (112) are accurate for at least some Ia subclade members of SAR11 (113), so variations in growth rate estimates from microradiography experiments likely have other explanations. It is possible that different strains of SAR11 simply have variations in growth rate not captured by existing isolates. Another, not mutually exclusive, possibility is that the differences in *in situ* growth rate estimates also reflect variations in the proportion of actively dividing cells within the population. A simple model of cell division with binary fission where only a subset of cells divide and nondividing cells persist, rather than die, can still yield logarithmic growth curves (see Fig. S10 in the supplemental material) like those observed for SAR11 in pure culture (29, 49, 114). However, this subpopulation variability means that the division rate for the subset of cells that are actively dividing is much higher than calculated when assuming 100% dividing cells in the population. Based on our estimated viability for SAR11 LD12 of 15 to 55%, to obtain our previously calculated maximum division rate ( $0.5 \text{ day}^{-1}$ ) for the whole culture (29), the per-cell division rate for only a subpopulation would span 2.48 to  $0.79 \text{ day}^{-1}$  (Fig. S10 and supplemental text). Verifying the proportion of SAR11 cells actively dividing in a culture may be challenging. Time-lapse microscopy (115) offers an elegant solution if SAR11 can be maintained for the requisite time periods for accurate measurements in a microfluidic device.

In addition to identifying taxa whose isolation success suggested deviations from biological assumptions of single planktonic cells with 100% viability, the model also revealed the limitations of DTE cultivation in assessing viability depending on relative abundance (Fig. S9). We cannot ascertain whether any given taxon may violate an assumption of  $V = 100\%$  unless we have enough wells to demonstrate that it grew in fewer wells than expected. For example, taxa at 1% of the microbial community require more than 1,000 wells before the lack of a cultured organism represents a significant negative event, rather than a taxon simply lacking sufficient abundance to ensure inclusion in a well within 95% CI. In our 460-well experiments, we could not resolve whether taxa may have had viabilities below 100% if they were less than 3% of the community for any given experiment (Fig. S9). Modeling DTE experiments showed that for experiments targeting rare taxa, lower inoculum sizes are favored where a selective medium for enrichment is either unknown or undesirable. The exponential increase in the number of required wells with respect to the inoculum size is a function of a pure well requiring all cells within it to belong to the same taxon, assuming all cells are equally and optimally viable.

By providing taxon-specific predictions of viability from cultivation data, our model now facilitates an iterative process to improve experimental design and make cultivation more reliable. First, we use the cultivation success rates to determine for which taxa the assumption of 100% viability was violated. Second, we use the model to estimate viability for those organisms. Third, we use the viability and relative-abundance data to determine, within 95% CI, the appropriate number of inoculation attempts required to isolate a new version of that organism. Using SAR11 LD12 as an example, given a relative abundance of 10% and a viability of 15%, 800 DTE wells should yield four pure, positive wells (95% CI, 1 to 8). This means that, for microorganisms that we know successfully grow in our media, we can now statistically constrain the appropriate number of wells required to culture a given taxon again. For organisms that were not abundant enough to estimate viability using the model, we can use a conservative viability assumption (e.g., 50% [86]) with which to base our cultivation strategy, thereby still reducing uncertainty about the experimental effort necessary to reisolate one of these microorganisms.

**Conclusions.** This work has provided hundreds of new cultures for microbiological research, many among the most abundant members of the nGOM coastal bacterioplankton community. It also provides another demonstration of the effectiveness of sustained cultivation efforts for bringing previously uncultivated strains into culture.

Our modeled cultivation results have generated compelling evidence for low viability within subpopulations of SAR11 LD12 and Illa.1, as well as OM43 *Betaproteobacteria*. The prevalence of and controls on dormancy in these clades deserve further study. We anticipate that future work with larger DTE experiments will yield similar viability data about other groups of taxa with lower abundance, highlighting a valuable diagnostic application of DTE cultivation/modeling beyond the primary role in isolating new microorganisms. The integration of cultivation results, natural-abundance data from inoculum communities, and DTE modeling represents an important step forward in quantifying the risk associated with DTE efforts to isolate valuable taxa from new sources or repeating isolation from the same locations. We hope variations of this approach will be incorporated into wider community efforts to invest in culturing the uncultured.

## MATERIALS AND METHODS

**Sampling.** Surface water samples were collected at six different sites once a year for 3 years, except for Terrebonne Bay, which was collected twice. The sites sampled were Lake Borgne (LKB) (Shell Beach, LA), Bay Pomme d'Or (JLB) (Buras, LA), Terrebonne Bay (TBON) (Cocodrie, LA), Atchafalaya River Delta (ARD) (Franklin, LA), Freshwater City (FWC) (Kaplan, LA), and Calcasieu Jetties (CJ) (Cameron, LA) (latitude and longitude coordinates are provided in Table S1 in the supplemental material [available at <https://doi.org/10.6084/m9.figshare.12142113>]). Water collection for biogeochemical and biological analysis followed the protocol described previously (35). Briefly, we collected surface water in a sterile, acid-washed polycarbonate bottle. Duplicate 120-ml water samples were filtered serially through 2.7- $\mu$ m Whatman GF/D (GE Healthcare, Little Chalfont, UK) and 0.22- $\mu$ m Sterivex (Millipore, MA, USA) filters and placed on ice until being transferred to  $-20^{\circ}\text{C}$  in the laboratory (maximum of 3 h on ice). The University of Washington Marine Chemistry Laboratory analyzed duplicate subsamples of 50 ml 0.22- $\mu$ m-filtered water collected in sterile 50-ml Falcon tubes (VWR, PA, USA) for concentrations of  $\text{SiO}_4$ ,  $\text{PO}_4^{3-}$ ,  $\text{NH}_4^+$ ,  $\text{NO}_3^-$ , and  $\text{NO}_2^-$ . Samples for cell counts were filtered through a 2.7- $\mu$ m GF/D filter, fixed with 10% formaldehyde, and stored on ice until enumeration (maximum of 3 h). Temperature, salinity, pH, and dissolved oxygen were measured using a handheld YSI 556 multiprobe system (YSI Inc., OH, USA). All metadata are available in Table S1.

**DTE culturing and propagation.** Isolation, propagation, and identification of isolates were completed as previously reported (29, 35, 116). A subsample of 2.7- $\mu$ m-filtered surface water was stained with  $1 \times$  SYBR green (Lonza, Basel, Switzerland) using a repeat pipettor and disposal tip (Gilson, WI, USA) and enumerated using a Guava EasyCyte 5HT HPL flow cytometer (Millipore, MA, USA) as described previously (116). After serial dilution to a predicted 1 to 3 cells  $\cdot \mu\text{l}^{-1}$ , 2  $\mu\text{l}$  water was inoculated into five 2-ml 96-well polytetrafluoroethylene (PTFE) plates (Radleys, Essex, UK) containing 1.7 ml artificial seawater medium (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]) using a 20- $\mu\text{l}$  multi-channel pipette (Gilson, WI, USA) to achieve an estimated 1 to 3 cells  $\cdot \text{well}^{-1}$  (Table 1). The salinity of the medium was chosen to match *in situ* salinity after experiment JLB (January 2015) (Tables 1 and S1). After year two, a second generation of media, designated MWH, was designed to incorporate additional important osmolytes, reduced sulfur compounds, and other constituents (Tables 1 and S1) potentially necessary for *in vitro* growth of uncultivated clades (49, 117–123). The four corner wells of each plate were left uninoculated as negative controls for every experiment. Plates were covered using sterile, PTFE-coated silicon 96-well plate mats (Thermo Scientific, MA, USA). Cultures were incubated at *in situ* temperatures (Table S1) in the dark for 3 to 6 weeks and evaluated for positive growth ( $>10^4$  cells  $\cdot \text{ml}^{-1}$ ) by flow cytometry. Two hundred microliters from positive wells was transferred using a 200- $\mu\text{l}$  single-channel pipette (Gilson, WI, USA) to duplicate 125-ml polycarbonate flasks (Corning, NY, USA) containing 50 ml of medium (29, 35, 116). At FWC, FWC2, JLB2c, and JLB3, not all positive wells were transferred because of the large number of positive wells. At each site, 48/301, 60/403, 60/103, and 60/146 of the positive wells were transferred, respectively, selected using flow cytometry signatures with  $<10^2$  green fluorescence and  $<10^2$  side scatter, which maximized our chances of isolating small microorganisms that encompass many of the most abundant and most wanted taxa, like SAR11, using our settings (116).

**Culture identification.** Cultures reaching  $\geq 1 \times 10^5$  cells  $\cdot \text{ml}^{-1}$  had 35 ml of the 50-ml volume filtered for identification via 16S rRNA gene PCR onto 25-mm, 0.22- $\mu$ m polycarbonate filters (Millipore, MA, USA). DNA extractions were performed using the MoBio PowerWater DNA kit (Qiagen, MA, USA) following the manufacturer's instructions and eluted in sterile water. The 16S rRNA gene was amplified as previously reported by Henson et al. (35) and sequenced at the Michigan State University Research Technology Support Facility Genomics Core. Evaluation of Sanger sequence quality was performed with 4Peaks (v. 1.7.1) (<https://nucleobytes.com/4peaks/index.html>), and forward and reverse complement sequences (converted via [http://www.bioinformatics.org/sms/rev\\_comp.html](http://www.bioinformatics.org/sms/rev_comp.html)) were assembled where overlap was sufficient using the CAP3 web server (<http://droua.prabi.fr/software/cap3>).

**Community iTag sequencing, OTUs, and single-nucleotide variants.** Sequentially filtered (2.7  $\mu\text{m}$ , 0.22  $\mu\text{m}$ ) duplicate samples were extracted and analyzed using our previously reported protocols and settings (35, 124). We sequenced the 2.7- to 0.22- $\mu\text{m}$  fraction for this study because this fraction corresponded with the  $<2.7\text{-}\mu\text{m}$  communities that were used for the DTE experiments. To avoid batch sequencing effects, DNA from the first seven collections reported previously (35) was resequenced with the additional samples from this study (FWC2 and after) (Table 1). We targeted the 16S rRNA gene V4

region with the 515F-806RB primer set (which corrects for poor amplification of taxa like SAR11) (76, 77) using Illumina MiSeq 2× 250-bp paired-end sequencing at Argonne National Laboratories, resulting in 2,343,106 raw reads for the 2.7- to 0.22- $\mu$ m fraction. Using mothur v1.33.3 (125), we clustered 16S rRNA gene amplicons into distinctive operational taxonomic units (OTUs) with a 0.03 dissimilarity threshold (OTU<sub>0.03</sub>) and classified them according to the Silva v119 database (126, 127). After these steps, 55,256 distinct OTU<sub>0.03</sub> remained. We also used minimum entropy decomposition (MED) to partition reads into fine-scale amplicon single-nucleotide variants (ASVs) (68). Reads were first analyzed using mothur as described above up to the screen.seqs() command. The cleaned-reads fasta file was converted to MED-compatible headers with the “mothur2oligo” tool renamer.pl from the functions in MicrobeMiseq (<https://github.com/DenefLab/MicrobeMiseq>) (128) using the fasta output from screen.seqs() and the mothur group file. These curated reads were analyzed using MED (v. 2.1) with the flags -M 60, and -d 1. MED resulted in 2,813 refined ASVs. ASVs were classified in mothur using classify.seqs(), the Silva v119 database, and a cutoff bootstrap value of 80% (129). After classification, we removed ASVs identified as “chloroplast,” “mitochondria,” or “unknown” from the data set.

**Community analyses.** OTU (OTU<sub>0.03</sub>) and ASV abundances were analyzed within the R statistical environment v.3.2.1 (130) following previously published protocols (29, 35, 124). Using the package PhyloSeq (131), OTUs and ASVs were rarefied using the command rarefy\_even\_depth(), and OTUs/ASVs without at least two reads in four of the 34 samples (2 sites; ~11%) were removed. This cutoff was used to remove potentially spurious OTUs/ASVs resulting from sequencing errors. Our modified PhyloSeq script is available on our GitHub repository, <https://github.com/thrash-lab/Modified-Phyloseq>. After filtering, the data sets contained 777 unique OTUs and 1,323 unique ASVs (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>]). For site-specific community comparisons, beta diversity between sites was examined using Bray-Curtis distances via ordination with nonmetric multidimensional scaling (NMDS) (Table S1). The nutrient data were normalized using the R function “scale,” which subtracts the mean and divides by the standard deviation for each column. The influence of the transformed environmental parameters on beta diversity was calculated in R with the envfit function. Relative abundances of an OTU or ASV from each sample were calculated as the number of reads over the sum of all the reads in that sample. The relative abundance was then averaged between biological duplicates for a given OTU or ASV. To determine the best-matching OTU or ASV for a given LSUCC isolate, the OTU representative fasta file, provided by mothur using get.oturep(), and the ASV fasta file were used to create a BLAST database (makeblastdb) against which the LSUCC isolate 16S rRNA genes could be searched via blastn (BLAST v. 2.2.26) (“OTU\_ASVrep\_db”; available as supplemental information at <https://doi.org/10.6084/m9.figshare.12142098>). We designated an LSUCC isolate 16S rRNA gene match with an OTU or ASV sequence based on  $\geq 97\%$  or  $\geq 99\%$  sequence identity, respectively, as well as a  $\geq 247$ -bp alignment.

**16S rRNA gene phylogeny.** Taxa in the *Alpha*-, *Beta*-, and *Gammaproteobacteria* phylogenies from reference 35 served as the backbone for the trees in the current work. For places in these trees with poor representation near isolate sequences, additional taxa were selected by searching the 16S rRNA genes of LSUCC isolates against the NCBI nucleotide database online with BLASTn (132) and selecting a variable number of best hits. The *Bacteroidetes* and *Actinobacteria* trees were composed entirely of nonredundant top 100 to 300 MegaBLAST hits to a local version of the NCBI nucleotide database, accessed August 2018. Sequences were aligned with MUSCLE v3.6 (133) using default settings and culled with TrimAl v1.4.rev22 (134) using the -automated1 flag, and the final alignment was inferred with IQ-TREE v1.6.11 (135) with default settings and -bb 1000 for ultrafast bootstrapping (136). Tips were edited with the nw\_rename script within Newick Utilities v1.6 (137), and trees were visualized with Archaeopteryx (138). Fasta files for these trees and the naming keys are available as supplemental information at <https://doi.org/10.6084/m9.figshare.12142098>.

**Assessment of isolate novelty.** We quantified taxonomic novelty using BLASTn of our isolate 16S rRNA genes to those of other known isolates collected in three databases: (i) The NCBI nucleotide database (accessed August 2018) (NCBIdb), (ii) a custom database comprised of sequences from DTE experiments in other labs (DTEdb), and (iii) a database containing all of our isolate 16S rRNA genes (LSUCCdb). The DTEdb and LSUCCdb fasta files are available as supplemental information at <https://doi.org/10.6084/m9.figshare.12142098>. We compared our isolate sequences to these databases as follows.

(i) All representative sequences were searched against the nucleotide database using BLASTn (BLAST+v. 2.7.1) with the flags -perc\_identity 84, -evalue 1E-6, -task blastn, -outfmt “6 qseqid sseqid pident length qlen mismatch evalbitscore sscinames sbblastnames stitle,” and -negative\_glist to remove uncultured and environmental sequences. The negative GI list was obtained by searching “environmental samples[organism] OR metagenomes[orgn]” in the NCBI nucleotide database (accessed 12 September 2018), and hits were downloaded in GI list format. This negative GI list is available as supplemental information at <https://doi.org/10.6084/m9.figshare.12142098>. The resultant hits from the NCBIdb search were further manually curated to remove sequences classified as single-cell genomes, clones, duplicates, and previously deposited LSUCC isolates.

(ii) We observed that many known HTCC, IMCC, and HIMB isolates that had previously been described as matching our clades (see Fig. S1 to S5 in the supplemental material) were missing from the resultant lists of nucleotide hits, so we extracted isolate accession numbers from numerous DTE experiments (26–28, 31, 34, 37, 44, 139, 140) from the nucleotide database via blastdbcmd and generated a separate DTEdb using makeblastdb. Duplicate accession numbers found in the NCBIdb were removed. The same BLASTn settings as for the first step were used to search our isolate sequences against DTEdb. Any match that fell below the lowest percent identity hit to the NCBIdb was removed from the DTEdb search, since the match would not have been present in the first NCBIdb search.

(iii) Finally, using the same BLASTn settings, we compared all pairwise identities of our 328 LSUCC isolate 16S rRNA gene sequences via the LSUCCdb.

The output from these searches is available in Table S1 (<https://doi.org/10.6084/m9.figshare.12142113>) under the “taxonomic novelty” tab.

We placed our LSUCC isolates into 55 taxonomic groups based on sharing  $\geq 94\%$  identity and/or their occurrence in monophyletic groups within our 16S rRNA gene trees (Fig. S1 to S5) (see above). For visualization purposes, in groups with multiple isolates we used our chronologically first cultivated isolate as the representative sequence for blastn searches, and these are the top point (100% identity to itself) in each group column of Fig. 1. Sequences from the other DTE culture collections were labeled with the corresponding collection name, while all other hits were labeled “other.”

Geographic novelty was assessed by manually screening the accession numbers from hits to LSUCC isolates with  $\geq 99\%$  16S rRNA gene sequence identity for the latitude and longitude from a connected publication or location name (e.g., source, country, or site) in the NCBI description. LSUCC isolates in the *Janibacter* sp., *Micrococcus* sp., *Altererythrobacter* sp., *Pseudomonas* sp., and *Phycoccus* sp. groups (16 total isolates) were not assessed because of missing isolation source information and no traceable publication. Isolation locations were plotted for a subset of important taxa (Table S1 [available at <https://doi.org/10.6084/m9.figshare.12142113>], “Map\_cultivars” tab) using the “LSUCC\_cultivar\_map.R” available at our GitHub repository, <https://github.com/thrash-lab/Cultivar-novelty-map>.

**Modeling DTE cultivation via Monte Carlo simulations.** We developed a model using Monte Carlo simulation to estimate the median number of positive and pure wells (and associated 95% confidence intervals [CI]) expected from a DTE experiment for a given taxon at different inoculum sizes ( $\lambda$ ), relative abundances ( $r$ ), and viability ( $V$ ) (Fig. 5). For each bootstrap, the number of cells added to each well was simulated using a Poisson distribution at a mean inoculum size of  $\lambda$  cells per well across  $n$  wells. The number of cells added to each well that belonged to a specific taxon was then estimated using a binomial distribution where the number of trials was set as the number of cells in a well and the probability of a cell belonging to a specific taxon,  $r$ , was the relative abundance of its representative ASV in the community analysis. Wells that contained at least one cell of a specific taxon were designated “positive.” Wells in which all the cells belonged to a specific taxon were designated “pure.” Finally, the influence of taxon-specific viability on recovery of pure wells was simulated using a second binomial distribution, where the number of cells within a pure well was used as the number of trials and the probability of growth was a viability score ranging from 0 to 1. For each simulation, 9,999 bootstraps were performed. Code for the model and all simulations is available at our GitHub repository, [https://github.com/thrash-lab/viability\\_test](https://github.com/thrash-lab/viability_test).

**Actual versus expected number of isolates.** For each taxon in each DTE experiment, the Monte Carlo simulation was used to evaluate whether the number of recovered pure wells for each taxon was within 95% CI of simulated estimates, assuming optimum growth conditions (i.e.,  $V = 100\%$ ). For each of 9,999 bootstraps, 460 wells were simulated with the inoculum size used for the experiment and the relative abundance of the ASV. For taxa where the number of expected wells fell outside the 95% CI of the model, a deviance score was calculated as the difference between the actual number of wells observed and median of the simulated data set. The results of this output are presented in Table S1 (available at <https://doi.org/10.6084/m9.figshare.12142113>) under the “Expected vs actual” tab, and the R script for visualizing this output (Fig. 7) is available at our GitHub repository, <https://github.com/thrash-lab/EvsA-visualization>.

**Estimating viability in underrepresented taxa.** For taxa where the observed number of positive wells was lower than the 95% CI lower limit within a given experiment, and because our analysis was restricted to only those organisms for which our medium was sufficient for growth at least once, the deviance was assumed to be a function of a viability term,  $V$  (ranging from 0 to 1), associated with suboptimal growth conditions, dormancy, persister cells, etc. To estimate a value of viability for a given taxon within a particular experiment, the Monte Carlo simulation was run using an experiment-appropriate inoculum size, relative abundance, and number of wells (460 for each experiment). Taxon-specific viability was tested across a range of decreasing values from 99% to 1% until such time as the observed number of pure wells for a given taxon fell between the 95% CI bounds of the simulated data. At this point, the viability value is the maximum viability of the taxon that enables the observed number of pure wells for a given taxon to be explained by the model. The results of this output are presented in Table 3 and Table S1 (available at <https://doi.org/10.6084/m9.figshare.12142113>) under the “Expected vs actual” tab.

**Likelihood of recovering taxa at different relative abundances.** To estimate the number of wells required in a DTE experiment to have a significant chance of recovering a taxon with a relative abundance of  $r$ , assuming optimum growth conditions ( $V = 100\%$ ), the Monte Carlo model was used to simulate experiments from 92 wells to 9,200 wells per experiment across a range of relative abundances from 0% to 100% in 0.1% increments and a range of inoculum sizes (cells per well of 1, 1.27, 1.96, 2, 3, 4, and 5). Each experiment was bootstrapped 999 times, and the number of bootstraps in which the lower bound of the 95% CI was  $\geq 1$  was recorded.

**Data availability.** All iTag sequences are available at the Sequence Read Archive with accession numbers [SRR6235382](https://www.ncbi.nlm.nih.gov/sra/SRR6235382) to [SRR6235415](https://www.ncbi.nlm.nih.gov/sra/SRR6235415) (29). PCR-generated 16S rRNA gene sequences from this study are accessible in NCBI GenBank under the accession numbers [MK603525](https://www.ncbi.nlm.nih.gov/nuclink/MK603525) to [MK603769](https://www.ncbi.nlm.nih.gov/nuclink/MK603769). Previously generated 16S rRNA gene sequences are accessible in NCBI GenBank under the accession numbers [KU382357](https://www.ncbi.nlm.nih.gov/nuclink/KU382357) to [KU382438](https://www.ncbi.nlm.nih.gov/nuclink/KU382438) (35). Table S1 is available at <https://doi.org/10.6084/m9.figshare.12142113>.



## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 1.7 MB.

## ACKNOWLEDGMENTS

We thank Nancy Rabalais for her comments and edits on a very early draft of the manuscript. Portions of this research were conducted with high-performance computing resources provided by Louisiana State University (<http://www.hpc.lsu.edu>) and the University of Southern California (<https://hpcc.usc.edu/>).

This work was supported by the Department of Biological Sciences at Louisiana State University, a Louisiana Board of Regents grant to J.C.T. [LEQSF(2014-17)-RD-A-06], a National Academies of Science, Engineering, and Medicine Gulf Research Program Early Career Fellowship to J.C.T., and the Dornsife College of Letters, Arts and Sciences at the University of Southern California. B.T. was partially funded by NERC award NE/R010935/1 and by the Simons Foundation BIOS-SCOPE program.

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

M.W.H. led sample collections, cultivation experiments, nucleic acid extraction, amplicon sequencing, and analyses; V.C.L., D.M.P., J.L.W., and A.M.L. supported sample collections, cultivation experiments, and nucleic acid extractions; M.W.H., and J.C.T. conducted cultivation comparisons and phylogenetic analyses; B.T. developed the viability model and led the statistical analyses; C.C. derived the cell-specific growth rate equations incorporating viability; J.C.T. designed the study and assisted with sample collections and model refinement; M.W.H., B.T., and J.C.T. led manuscript writing; and all authors contributed to and reviewed the text.

## REFERENCES

1. Metcalf WW, Griffin BM, Cicchillo RM, Gao J, Janga SC, Cooke HA, Circello BT, Evans BS, Martens-Habben W, Stahl DA, van der Donk WA. 2012. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science* 337:1104–1107. <https://doi.org/10.1126/science.1219875>.
2. Carini P, White AE, Campbell EO, Giovannoni SJ. 2014. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nat Commun* 5:4346. <https://doi.org/10.1038/ncomms5346>.
3. Karl DM, Beversdorf L, Björkman KM, Church MJ, Martinez A, Delong EF. 2008. Aerobic production of methane in the sea. *Nat Geosci* 1:473–478. <https://doi.org/10.1038/ngeo234>.
4. Steindler L, Schwalbach MS, Smith DP, Chan F, Giovannoni SJ. 2011. Energy starved *Candidatus Pelagibacter ubique* substitutes light-mediated ATP production for endogenous carbon respiration. *PLoS One* 6:e19725. <https://doi.org/10.1371/journal.pone.0019725>.
5. Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T, Neutze R, Pedrós-Alió C, Pinhassi J. 2007. Light stimulates growth of proteorhodopsin-containing marine *Flavobacteria*. *Nature* 445: 210–213. <https://doi.org/10.1038/nature05381>.
6. Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard RH, von Bergen M, Rattei T, Bendinger B, Nielsen PH, Wagner M. 2015. Complete nitrification by *Nitrospira* bacteria. *Nature* 528:504–509. <https://doi.org/10.1038/nature16461>.
7. Lam KS. 2006. Discovery of novel metabolites from marine actinomycetes. *Curr Opin Microbiol* 9:245–251. <https://doi.org/10.1016/j.mib.2006.03.004>.
8. Jensen PR, Williams PG, Oh D-C, Zeigler L, Fenical W. 2007. Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. *Appl Environ Microbiol* 73:1146–1152. <https://doi.org/10.1128/AEM.01891-06>.
9. Nett M, Erol Ö, Kehraus S, Köck M, Krick A, Eguereva E, Neu E, König GM. 2006. Siphonazole, an unusual metabolite from *Herpetosiphon* sp. *Angew Chem Int Ed Engl* 45:3863–3867. <https://doi.org/10.1002/anie.200504525>.
10. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K. 2015. A new antibiotic kills pathogens without detectable resistance. *Nature* 517:455–459. <https://doi.org/10.1038/nature14098>.
11. 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>.
12. Hug LA. 2018. Sizing up the uncultured microbial majority. *mSystems* 3:e00185-18. <https://doi.org/10.1128/mSystems.00185-18>.
13. Steen AD, Crits-Christoph A, Carini P, DeAngelis KM, Fierer N, Lloyd KG, Thrash JC. 2019. High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J* 13:3126–3130. <https://doi.org/10.1038/s41396-019-0484-y>.
14. Overmann J, Abt B, Sikorski J. 2017. Present and future of culturing bacteria. *Annu Rev Microbiol* 71:711–730. <https://doi.org/10.1146/annurev-micro-090816-093449>.
15. Rappé MS. 2013. Stabilizing the foundation of the house that 'omics builds: the evolving value of cultured isolates to marine microbiology. *Curr Opin Microbiol* 16:618–624. <https://doi.org/10.1016/j.mib.2013.09.009>.
16. Carini P. 2019. A “cultural” renaissance: genomics breathes new life into an old craft. *mSystems* 4:e00092-19. <https://doi.org/10.1128/mSystems.00092-19>.
17. Staley JT, Konopka A. 1985. Measurement of in situ activities of non-photosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* 39:321–346. <https://doi.org/10.1146/annurev.mi.39.100185.001541>.
18. Amann RL, Ludwig W, Schleifer KH. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59:143–169. <https://doi.org/10.1128/MMBR.59.1.143-169.1995>.
19. Hahn MW, Koll U, Schmidt J. 2019. Isolation and cultivation of bacteria, p 313–351. In Hurst CJ (ed), *The structure and function of aquatic microbial communities*. Springer International Publishing, Cham, Switzerland.
20. Kaeberlein T, Lewis K, Epstein SS. 2002. Isolating “uncultivable” micro-

- organisms in pure culture in a simulated natural environment. *Science* 296:1127–1129. <https://doi.org/10.1126/science.1070633>.
21. Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M. 2002. Cultivating the uncultured. *Proc Natl Acad Sci U S A* 99:15681–15686. <https://doi.org/10.1073/pnas.252630999>.
  22. Steinert G, Whitfield S, Taylor MW, Thoms C, Schupp PJ. 2014. Application of diffusion growth chambers for the cultivation of marine sponge-associated bacteria. *Mar Biotechnol* (NY) 16:594–603. <https://doi.org/10.1007/s10126-014-9575-y>.
  23. Nichols D, Cahoon N, Trakhtenberg EM, Pham L, Mehta A, Belanger A, Kanigan T, Lewis K, Epstein SS. 2010. Use of Ichip for high-throughput in situ cultivation of “uncultivable” microbial Species. *Appl Environ Microbiol* 76:2445–2450. <https://doi.org/10.1128/AEM.01754-09>.
  24. Tandogan N, Abadian PN, Epstein S, Aoi Y, Goluch ED. 2014. Isolation of microorganisms using sub-micrometer constrictions. *PLoS One* 9:e101429. <https://doi.org/10.1371/journal.pone.0101429>.
  25. Hahn MW, Stadler P, Wu QL, Pöckl M. 2004. The filtration–acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *J Microbiol Methods* 57:379–390. <https://doi.org/10.1016/j.jmimet.2004.02.004>.
  26. Connon SA, Giovannoni SJ. 2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol* 68:3878–3885. <https://doi.org/10.1128/aem.68.8.3878-3885.2002>.
  27. Yang SJ, Kang I, Cho JC. 2016. Expansion of cultured bacterial diversity by large-scale dilution-to-extinction culturing from a single seawater sample. *Microb Ecol* 71:29–43. <https://doi.org/10.1007/s00248-015-0695-3>.
  28. Stingl U, Tripp HJ, Giovannoni SJ. 2007. Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *ISME J* 1:361–371. <https://doi.org/10.1038/ismej.2007.49>.
  29. Henson MW, Lanclos VC, Faircloth BC, Thrash JC. 2018. Cultivation and genomics of the first freshwater SAR11 (LD12) isolate. *ISME J* 12:1846–1860. <https://doi.org/10.1038/s41396-018-0092-2>.
  30. Hahnke RL, Bennis CM, Fuchs BM, Mann AJ, Rhie E, Teeling H, Amann R, Harder J. 2015. Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environ Microbiol* 17:3515–3526. <https://doi.org/10.1111/1462-2920.12479>.
  31. Sosa OA, Gifford SM, Repeta DJ, DeLong EF. 2015. High molecular weight dissolved organic matter enrichment selects for methylotrophs in dilution to extinction cultures. *ISME J* 9:2725–2739. <https://doi.org/10.1038/ismej.2015.68>.
  32. Schut F, de Vries EJ, Gottschal JC, Robertson BR, Harder W, Prins RA, Button DK. 1993. Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl Environ Microbiol* 59:2150–2160. <https://doi.org/10.1128/AEM.59.7.2150-2160.1993>.
  33. Button DK, Schut F, Quang P, Martin R, Robertson BR. 1993. Viability and isolation of marine bacteria by dilution culture: theory, procedures, and initial results. *Appl Environ Microbiol* 59:881–891. <https://doi.org/10.1128/AEM.59.3.881-891.1993>.
  34. Song J, Oh HM, Cho JC. 2009. Improved culturability of SAR11 strains in dilution-to-extinction culturing from the East Sea, West Pacific Ocean. *FEMS Microbiol Lett* 295:141–147. <https://doi.org/10.1111/j.1574-6968.2009.01623.x>.
  35. Henson MW, Pitre DM, Weckhorst JL, Lanclos VC, Webber AT, Thrash JC. 2016. Artificial seawater media facilitate cultivating members of the microbial majority from the Gulf of Mexico. *mSphere* 1:e00028-16. <https://doi.org/10.1128/mSphere.00124-16>.
  36. Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418:630–633. <https://doi.org/10.1038/nature00917>.
  37. Marshall KT, Morris RM. 2013. Isolation of an aerobic sulfur oxidizer from the SUP05/Arctic96BD-19 clade. *ISME J* 7:452–455. <https://doi.org/10.1038/ismej.2012.78>.
  38. Shah V, Chang BX, Morris RM. 2017. Cultivation of a chemoautotroph from the SUP05 clade of marine bacteria that produces nitrite and consumes ammonium. *ISME J* 11:263–271. <https://doi.org/10.1038/ismej.2016.87>.
  39. Spietz RL, Lundeen RA, Zhao X, Nicastro D, Ingalls AE, Morris RM. 2019. Heterotrophic carbon metabolism and energy acquisition in *Candidatus Thiomargarita singularis* strain PS1, a member of the SUP05 clade of marine Gammaproteobacteria. *Environ Microbiol* 21:2391–2401. <https://doi.org/10.1111/1462-2920.14623>.
  40. Huggett MJ, Hayakawa DH, Rappé MS. 2012. Genome sequence of strain HIMB624, a cultured representative from the OM43 clade of marine Betaproteobacteria. *Stand Genomic Sci* 6:11–20. <https://doi.org/10.4056/signs.2305090>.
  41. Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho J-C, Oh H-M, Kitner JB, Vergin KL, Rappé MS. 2008. The small genome of an abundant coastal ocean methylotroph. *Environ Microbiol* 10:1771–1782. <https://doi.org/10.1111/j.1462-2920.2008.01598.x>.
  42. Durham BP, Grote J, Whittaker KA, Bender SJ, Luo H, Grim SL, Brown JM, Casey JR, Dron A, Florez-Leiva L, Krupke A, Luria CM, Mine AH, Nigro OD, Pather S, Talarmin A, Wear EK, Weber TS, Wilson JM, Church MJ, DeLong EF, Karl DM, Steward GF, Eppley JM, Kyrpides NC, Schuster S, Rappé MS. 2014. Draft genome sequence of marine alphaproteobacterial strain HIMB11, the first cultivated representative of a unique lineage within the Roseobacter clade possessing an unusually small genome. *Stand Genomic Sci* 9:632–645. <https://doi.org/10.4056/signs.4998989>.
  43. Cho JC, Giovannoni SJ. 2004. Cultivation and growth characteristics of a diverse group of oligotrophic marine gammaproteobacteria. *Appl Environ Microbiol* 70:432–440. <https://doi.org/10.1128/AEM.70.1.432-440.2004>.
  44. Kim S, Kang I, Seo J-H, Cho J-C. 2019. Culturing the ubiquitous freshwater actinobacterial acd lineage by supplying a biochemical “helper” catalase. *ISME J* 13:2252–2263. <https://doi.org/10.1038/s41396-019-0432-x>.
  45. Tripp HJ. 2013. The unique metabolism of SAR11 aquatic bacteria. *J Microbiol* 51:147–153. <https://doi.org/10.1007/s12275-013-2671-2>.
  46. Nichols D, Lewis K, Orjala J, Mo S, Ortenberg R, O'Connor P, Zhao C, Vouras P, Kaeberlein T, Epstein SS. 2008. Short peptide induces an “uncultivable” microorganism to grow in vitro. *Appl Environ Microbiol* 74:4889–4897. <https://doi.org/10.1128/AEM.00393-08>.
  47. Stewart EJ. 2012. Growing unculturable bacteria. *J Bacteriol* 194:4151–4160. <https://doi.org/10.1128/JB.00345-12>.
  48. Kell DB, Young M. 2000. Bacterial dormancy and culturability: the role of autocrine growth factors. *Curr Opin Microbiol* 3:238–243. [https://doi.org/10.1016/s1369-5274\(00\)00082-5](https://doi.org/10.1016/s1369-5274(00)00082-5).
  49. Carini P, Steindler L, Beszteri S, Giovannoni SJ. 2013. Nutrient requirements for growth of the extreme oligotroph “*Candidatus Pelagibacter ubique*” HTCC1062 on a defined medium. *ISME J* 7:592–602. <https://doi.org/10.1038/ismej.2012.122>.
  50. Epstein SS. 2013. The phenomenon of microbial uncultivability. *Curr Opin Microbiol* 16:636–642. <https://doi.org/10.1016/j.mib.2013.08.003>.
  51. Shah D, Zhang Z, Khodursky A, Kaldalu N, Kurg K, Lewis K. 2006. Persisters: a distinct physiological state of *E. coli*. *BMC Microbiol* 6:53. <https://doi.org/10.1186/1471-2180-6-53>.
  52. Kell D, Potgieter M, Pretorius E. 2015. Individuality, phenotypic differentiation, dormancy and “persistence” in culturable bacterial systems: commonalities shared by environmental, laboratory, and clinical microbiology. *F1000Res* 4:179. <https://doi.org/10.12688/f1000research.6709.2>.
  53. Lennon JT, Jones SE. 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 9:119–130. <https://doi.org/10.1038/nrmicro2504>.
  54. Grassi L, Di Luca M, Maisetta G, Rinaldi AC, Esin S, Trampuz A, Batoni G. 2017. Generation of persister cells of *Pseudomonas aeruginosa* and *Staphylococcus aureus* by chemical treatment and evaluation of their susceptibility to membrane-targeting agents. *Front Microbiol* 8:1917. <https://doi.org/10.3389/fmicb.2017.01917>.
  55. Bergkessel M, Basta DW, Newman DK. 2016. The physiology of growth arrest: uniting molecular and environmental microbiology. *Nat Rev Microbiol* 14:549–562. <https://doi.org/10.1038/nrmicro.2016.107>.
  56. Kussell E, Leibler S. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309:2075–2078. <https://doi.org/10.1126/science.1114383>.
  57. Kurm V, van der Putten WH, Hol WHG. 2019. Cultivation-success of rare soil bacteria is not influenced by incubation time and growth medium. *PLoS One* 14:e0210073. <https://doi.org/10.1371/journal.pone.0210073>.
  58. D’Onofrio A, Crawford JM, Stewart EJ, Witt K, Gavrish E, Epstein S, Clardy J, Lewis K. 2010. Siderophores from neighboring organisms promote the growth of uncultured bacteria. *Chem Biol* 17:254–264. <https://doi.org/10.1016/j.chembiol.2010.02.010>.
  59. Thrash JC. 2019. Culturing the uncultured: risk versus reward. *mSystems* 4:e00130-19. <https://doi.org/10.1128/mSystems.00130-19>.

60. Noell SE, Giovannoni SJ. 2019. SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter. *Environ Microbiol* 21: 2559–2575. <https://doi.org/10.1111/1462-2920.14649>.
61. Thume K, Gebser B, Chen L, Meyer N, Kieber DJ, Pohnert G. 2018. The metabolite dimethylsulfoxonium propionate extends the marine organosulfur cycle. *Nature* 563:412–415. <https://doi.org/10.1038/s41586-018-0675-0>.
62. Asher EC, Dacey JWH, Stukel M, Long MC, Tortell PD. 2017. Processes driving seasonal variability in DMS, DMSP, and DMSO concentrations and turnover in coastal Antarctic waters. *Limnol Oceanogr* 62:104–124. <https://doi.org/10.1002/lno.10379>.
63. Mizuno CM, Rodríguez-Valera F, Ghai R. 2015. Genomes of planktonic Acidimicrobiales: widening horizons for marine Actinobacteria by metagenomics. *mBio* 6:e02083-14. <https://doi.org/10.1128/mBio.02083-14>.
64. Lee J, Kwon KK, Lim S-I, Song J, Choi AR, Yang S-H, Jung K-H, Lee J-H, Kang SG, Oh H-M, Cho JC. 2019. Isolation, cultivation, and genome analysis of proteorhodopsin-containing SAR116-clade strain *Candidatus Puniceispirillum marinum* IMCC1322. *J Microbiol* 57:676–687. <https://doi.org/10.1007/s12275-019-9001-2>.
65. Kitzinger K, Padilla CC, Marchant HK, Hach PF, Herbold CW, Kidane AT, Könneke M, Littmann S, Mooshammer M, Niggemann J, Petrov S, Richter A, Stewart FJ, Wagner M, Kuypers MMM, Bristow LA. 2019. Cyanate and urea are substrates for nitrification by Thaumarchaeota in the marine environment. *Nat Microbiol* 4:234–243. <https://doi.org/10.1038/s41564-018-0316-2>.
66. Kamennaya NA, Post AF. 2011. Characterization of cyanate metabolism in marine *Synechococcus* and *Prochlorococcus* spp. *Appl Environ Microbiol* 77:291–301. <https://doi.org/10.1128/AEM.01272-10>.
67. Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, Sogin ML. 2013. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 4:1111–1119. <https://doi.org/10.1111/2041-210X.12114>.
68. Eren AM, Borisy GG, Huse SM, Mark Welch JL. 2014. Oligotyping analysis of the human oral microbiome. *Proc Natl Acad Sci U S A* 111: E2875–E2884. <https://doi.org/10.1073/pnas.1409644111>.
69. Needham DM, Fuhrman JA. 2016. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat Microbiol* 1:16005. <https://doi.org/10.1038/nmicrobiol.2016.5>.
70. Campbell BJ, Yu L, Heidelberg JF, Kirchman DL. 2011. Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci U S A* 108: 12776–12781. <https://doi.org/10.1073/pnas.1101405108>.
71. Oh HM, Kwon KK, Kang I, Kang SG, Lee JH, Kim SJ, Cho JC. 2010. Complete genome sequence of “*Candidatus puniceispirillum marinum*” IMCC1322, a representative of the SAR116 clade in the Alphaproteobacteria. *J Bacteriol* 192:3240–3241. <https://doi.org/10.1128/JB.00347-10>.
72. Fegatella F, Lim J, Kjelleberg S, Cavicchioli R. 1998. Implications of rRNA operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium *Sphingomonas* sp. strain RB2256. *Appl Environ Microbiol* 64:4433–4438. <https://doi.org/10.1128/AEM.64.11.4433-4438.1998>.
73. Acinas SG, Marcelino LA, Klepac-Ceraj V, Polz MF. 2004. Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* 186:2629–2635. <https://doi.org/10.1128/JB.186.9.2629-2635.2004>.
74. Pedrós-Alió C. 2012. The rare bacterial biosphere. *Annu Rev Mar Sci* 4:449–466. <https://doi.org/10.1146/annurev-marine-120710-100948>.
75. Martiny AC. 2019. High proportions of bacteria are culturable across major biomes. *ISME J* 13:2125–2128. <https://doi.org/10.1038/s41396-019-0410-3>.
76. Apprill A, McNally S, Parsons R, Weber L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb Ecol* 75:129–137. <https://doi.org/10.3354/ame01753>.
77. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA, Apprill A, Knight R. 2016. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* 1:e00009-15. <https://doi.org/10.1128/mSystems.00009-15>.
78. Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* 18:1403–1414. <https://doi.org/10.1111/1462-2920.13023>.
79. del Giorgio PA, Gasol JM. 2008. Physiological structure and single-cell activity in marine bacterioplankton, p 243–298. In Kirchman DL (ed), *Microbial ecology of the oceans*, 2nd ed. Wiley, New York, NY.
80. Malmstrom RR, Kiene RP, Cottrell MT, Kirchman DL. 2004. Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic Ocean. *Appl Environ Microbiol* 70: 4129–4135. <https://doi.org/10.1128/AEM.70.7.4129-4135.2004>.
81. Malmstrom RR, Cottrell MT, Elifantz H, Kirchman DL. 2005. Biomass production and assimilation of dissolved organic matter by SAR11 bacteria in the Northwest Atlantic Ocean. *Appl Environ Microbiol* 71: 2979–2986. <https://doi.org/10.1128/AEM.71.6.2979-2986.2005>.
82. Piwosz K, Salcher MM, Zeder M, Ameryk A, Pernthaler J. 2013. Seasonal dynamics and activity of typical freshwater bacteria in brackish waters of the Gulf of Gdańsk. *Limnol Oceanogr* 58:817–826. <https://doi.org/10.4319/lno.2013.58.3.0817>.
83. Laghdass M, Catala P, Caparros J, Oriol L, Lebaron P, Obernosterer I. 2012. High contribution of SAR11 to microbial activity in the north west Mediterranean Sea. *Microb Ecol* 63:324–333. <https://doi.org/10.1007/s00248-011-9915-7>.
84. Salter I, Galand PE, Fagervold SK, Lebaron P, Obernosterer I, Oliver MJ, Suzuki MT, Tricoire C. 2015. Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest Mediterranean Sea. *ISME J* 9:347–360. <https://doi.org/10.1038/ismej.2014.129>.
85. Leizeaga A, Estrany M, Forn I, Sebastián M. 2017. Using click-chemistry for visualizing in situ changes of translational activity in planktonic marine bacteria. *Front Microbiol* 8:2360. <https://doi.org/10.3389/fmicb.2017.02360>.
86. Kirchman DL. 2016. Growth rates of microbes in the oceans. *Annu Rev Mar Sci* 8:285–309. <https://doi.org/10.1146/annurev-marine-122414-033938>.
87. Samo TJ, Smriga S, Malfatti F, Sherwood BP, Azam F. 2014. Broad distribution and high proportion of protein synthesis active marine bacteria revealed by click chemistry at the single cell level. *Front Mar Sci* 1:48. <https://doi.org/10.3389/fmars.2014.00060>.
88. Smriga S, Samo TJ, Malfatti F, Villareal J, Azam F. 2014. Individual cell DNA synthesis within natural marine bacterial assemblages as detected by “click” chemistry. *Aquat Microb Ecol* 72:269–280. <https://doi.org/10.3354/ame01698>.
89. Bradley JA, Amend JP, LaRowe DE. 2019. Survival of the fewest: microbial dormancy and maintenance in marine sediments through deep time. *Geobiology* 17:43–59. <https://doi.org/10.1111/gbi.12313>.
90. Zhao Y, Temperton B, Thrash JC, Schwalbach MS, Vergin KL, Landry ZC, Ellisman M, Deerinck T, Sullivan MB, Giovannoni SJ. 2013. Abundant SAR11 viruses in the ocean. *Nature* 494:357–360. <https://doi.org/10.1038/nature11921>.
91. Martínez-Hernández F, Fornas O, Lluésma Gomez M, García-Heredia I, Maestre-Carballea L, López-Pérez M, Haro-Moreno JM, Rodríguez-Valera F, Martínez-García M. 2019. Single-cell genomics uncover Pelagibacter as the putative host of the extremely abundant uncultured 37-F6 viral population in the ocean. *ISME J* 13:232–236. <https://doi.org/10.1038/s41396-018-0278-7>.
92. Chen L-X, Zhao Y, McMahon KD, Mori JF, Jessen GL, Nelson TC, Warren LA, Banfield JF. 2019. Wide distribution of phage that infect freshwater SAR11 bacteria. *mSystems* 4:e00410-19. <https://doi.org/10.1128/mSystems.00410-19>.
93. Thingstad TF, Lignell R. 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat Microb Ecol* 13:19–27. <https://doi.org/10.3354/ame013019>.
94. Våge S, Storesund JE, Thingstad TF. 2013. SAR11 viruses and defensive host strains. *Nature* 499:E3–E4. <https://doi.org/10.1038/nature12387>.
95. Thingstad TF, Våge S, Storesund JE, Sandaa RA, Giske J. 2014. A theoretical analysis of how strain-specific viruses can control microbial species diversity. *Proc Natl Acad Sci U S A* 111:7813–7818. <https://doi.org/10.1073/pnas.1400909111>.
96. Giovannoni S, Temperton B, Zhao Y. 2013. Giovannoni et al. reply. *Nature* 499:E4–E5. <https://doi.org/10.1038/nature12388>.
97. Alonso-Sáez L, Morán XAG, Clokie MR. 2018. Low activity of lytic pelagiphages in coastal marine waters. *ISME J* 12:2100–2102. <https://doi.org/10.1038/s41396-018-0185-y>.
98. Chapman-McQuiston E, Wu XL. 2008. Stochastic receptor expression allows sensitive bacteria to evade phage attack. I. Experiments. *Biophys J* 94:4525–4536. <https://doi.org/10.1529/biophysj.107.120212>.
99. Bertozzi Silva J, Storms Z, Sauvageau D. 2016. Host receptors for



- bacteriophage adsorption. *FEMS Microbiol Lett* 363:fnw002. <https://doi.org/10.1093/femsle/fnw002>.
100. Williamson SJ, Houchin LA, McDaniel L, Paul JH. 2002. Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl Environ Microbiol* 68:4307–4314. <https://doi.org/10.1128/aem.68.9.4307-4314.2002>.
  101. Stough JMA, Tang X, Krausfeldt LE, Steffen MM, Gao G, Boyer GL, Wilhelm SW. 2017. Molecular prediction of lytic vs lysogenic states for *Microcystis* phage: metatranscriptomic evidence of lysogeny during large bloom events. *PLoS One* 12:e0184146. <https://doi.org/10.1371/journal.pone.0184146>.
  102. Zhao Y, Qin F, Zhang R, Giovannoni SJ, Zhang Z, Sun J, Du S, Rensing C. 2019. Pelagiphages in the Podoviridae family integrate into host genomes. *Environ Microbiol* 21:1989–2001. <https://doi.org/10.1111/1462-2920.14487>.
  103. Morris RM, Cain KR, Hvorecny KL, Kollman JM. 2020. Lysogenic host–virus interactions in SAR11 marine bacteria. *Nat Microbiol* <https://doi.org/10.1038/s41564-020-0725-x>.
  104. Warwick-Dugdale J, Buchholz HH, Allen MJ, Temperton B. 2019. Host-hijacking and planktonic piracy: how phages command the microbial high seas. *Virol J* 16:15. <https://doi.org/10.1186/s12985-019-1120-1>.
  105. Yu Z-C, Chen X-L, Shen Q-T, Zhao D-L, Tang B-L, Su H-N, Wu Z-Y, Qin Q-L, Xie B-B, Zhang X-Y, Yu Y, Zhou B-C, Chen B, Zhang Y-Z. 2015. Filamentous phages prevalent in *Pseudoalteromonas* spp. confer properties advantageous to host survival in Arctic sea ice. *ISME J* 9:871–881. <https://doi.org/10.1038/ismej.2014.185>.
  106. Shkoporov AN, Khokhlova EV, Fitzgerald CB, Stockdale SR, Draper LA, Ross RP, Hill C. 2018. ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects *Bacteroides* intestinalis. *Nat Commun* 9:4781. <https://doi.org/10.1038/s41467-018-07225-7>.
  107. Teira E, Martínez-García S, Lønborg C, Alvarez-Salgado XA. 2009. Growth rates of different phylogenetic bacterioplankton groups in a coastal upwelling system. *Environ Microbiol Rep* 1:545–554. <https://doi.org/10.1111/j.1758-2229.2009.00079.x>.
  108. Lankiewicz TS, Cottrell MT, Kirchman DL. 2016. Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. *ISME J* 10:823–832. <https://doi.org/10.1038/ismej.2015.156>.
  109. Ferrera I, Gasol JM, Sebastián M, Hojerová E, Koblížek M. 2011. Comparison of growth rates of aerobic anoxygenic phototrophic bacteria and other bacterioplankton groups in coastal Mediterranean waters. *Appl Environ Microbiol* 77:7451–7458. <https://doi.org/10.1128/AEM.00208-11>.
  110. Carini P, Campbell EO, Morré J, Sañudo-Wilhelmy SA, Thrash JC, Bennett SE, Temperton B, Begley T, Giovannoni SJ. 2014. Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea. *ISME J* 8:1727–1738. <https://doi.org/10.1038/ismej.2014.61>.
  111. Grant SR, Church MJ, Ferrón S, Laws EA, Rappé MS. 2019. Elemental composition, phosphorous uptake, and characteristics of growth of a SAR11 strain in batch and continuous culture *mSystems* 4:e00218-18. <https://doi.org/10.1128/mSystems.00218-18>.
  112. Simon M, Azam F. 1989. Protein content and protein synthesis rates of planktonic marine bacteria. *Mar Ecol Prog Ser* 51:201–213. <https://doi.org/10.3354/meps051201>.
  113. White AE, Giovannoni SJ, Zhao Y, Vergin K, Carlson CA. 2019. Elemental content and stoichiometry of SAR11 chemoheterotrophic marine bacteria. *Limnol Oceanogr* 4:44–51. <https://doi.org/10.1002/lol2.10103>.
  114. Smith DP, Cameron Thrash J, Nicora CD, Lipton MS, Burnum-Johnson KE, Carini P, Smith RD, Giovannoni SJ. 2013. Proteomic and transcriptomic analyses of “*Candidatus Pelagibacter ubique*” describe the first PII-independent response to nitrogen limitation in a free-living alphaproteobacterium. *mBio* 4:e00133-12. <https://doi.org/10.1128/mBio.00133-12>.
  115. van Vliet S, Dal Co A, Winkler AR, Spriewald S, Stecher B, Ackermann M. 2018. Spatially correlated gene expression in bacterial groups: the role of lineage history, spatial gradients, and cell-cell interactions. *Cell Syst* 6:496–507. <https://doi.org/10.1016/j.cels.2018.03.009>.
  116. Thrash JC, Weckhorst JL, Pitre DM. 2017. Cultivating fastidious microbes, p 57–78. In McGinty TJ, Timmis KN, Nogales B (ed), *Hydrocarbon and lipid microbiology protocols*. Springer, Berlin, Germany.
  117. Vila-Costa M, Simó R, Harada H, Gasol JM, Slezak D, Kiene RP. 2006. Dimethylsulfoniopropionate uptake by marine phytoplankton. *Science* 314:652–654. <https://doi.org/10.1126/science.1131043>.
  118. Mou X, Hodson RE, Moran MA. 2007. Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. *Environ Microbiol* 9:2025–2037. <https://doi.org/10.1111/j.1462-2920.2007.01318.x>.
  119. Stein LY. 2015. Microbiology: cyanate fuels the nitrogen cycle. *Nature* 524:43–44. <https://doi.org/10.1038/nature14639>.
  120. Repeta DJ, Ferrón S, Sosa OA, Johnson CG, Repeta LD, Acker M, Delong EF, Karl DM. 2016. Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat Geosci* 9:884–887. <https://doi.org/10.1038/ngeo2837>.
  121. Curson ARJ, Liu J, Bermejo Martínez A, Green RT, Chan Y, Carrión O, Williams BT, Zhang S-H, Yang G-P, Bulman Page PC, Zhang X-H, Todd JD. 2017. Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat Microbiol* 2:17009. <https://doi.org/10.1038/nmicrobiol.2017.9>.
  122. Widner B, Mulholland MR. 2017. Cyanate distribution and uptake in North Atlantic coastal waters. *Limnol Oceanogr* 62:2538–2549. <https://doi.org/10.1002/lno.10588>.
  123. Widner B, Mulholland MR, Mopper K. 2016. Distribution, sources, and sinks of cyanate in the coastal North Atlantic Ocean. *Environ Sci Technol Lett* 3:297–302. <https://doi.org/10.1021/acs.estlett.6b00165>.
  124. Henson MW, Hanssen J, Spooner G, Fleming P, Pukonen M, Stahr F, Thrash JC. 2018. Nutrient dynamics and stream order influence microbial community patterns along a 2914 kilometer transect of the Mississippi River. *Limnol Oceanogr* 63:1837–1855. <https://doi.org/10.1002/lno.10811>.
  125. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541. <https://doi.org/10.1128/AEM.01541-09>.
  126. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO, 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>.
  127. 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* 41:590–596. <https://doi.org/10.1093/nar/gks1219>.
  128. Fujimoto M, Cavaletto J, Liebig JR, McCarthy A, Vanderploeg HA, Denef VJ. 2016. Spatiotemporal distribution of bacterioplankton functional groups along a freshwater estuary to pelagic gradient in Lake Michigan. *J Great Lakes Res* 42:1036–1048. <https://doi.org/10.1016/j.jglr.2016.07.029>.
  129. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42:D633–D642. <https://doi.org/10.1093/nar/gkt1244>.
  130. R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
  131. McMurdie PJ, Holmes S. 2013. phyloseq: an R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
  132. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
  133. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
  134. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
  135. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
  136. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>.
  137. Junier T, Zdobnov EM. 2010. The Newick utilities: high-throughput

- phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26: 1669–1670. <https://doi.org/10.1093/bioinformatics/btq243>.
138. Han MV, Zmasek CM. 2009. phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* 10:356. <https://doi.org/10.1186/1471-2105-10-356>.
139. Stingl U, Cho JC, Foo W, Vergin KL, Lanoil B, Giovannoni SJ. 2008. Dilution-to-extinction culturing of psychrotolerant planktonic bacteria from permanently ice-covered lakes in the McMurdo Dry Valleys, Antarctica. *Microb Ecol* 55:395–405. <https://doi.org/10.1007/s00248-007-9284-4>.
140. Page KA, Cannon SA, Giovannoni SJ. 2004. Representative freshwater bacterioplankton isolated from Crater Lake, Oregon. *Appl Environ Microbiol* 70:6542–6550. <https://doi.org/10.1128/AEM.70.11.6542-6550.2004>.