



Predictive genomic traits for bacterial growth in culture versus actual growth in soil

Junhui Li¹ · Rebecca L. Mau¹ · Paul Dijkstra^{1,2} · Benjamin J. Koch^{1,2} · Egbert Schwartz^{1,2} · Xiao-Jun Allen Liu^{1,2} · Ember M. Morrissey³ · Steven J. Blazewicz⁴ · Jennifer Pett-Ridge⁴ · Bram W. Stone¹ · Michaela Hayer¹ · Bruce A. Hungate^{1,2}

Received: 15 December 2018 / Revised: 1 April 2019 / Accepted: 3 April 2019
© International Society for Microbial Ecology 2019

Abstract

Relationships between microbial genes and performance are often evaluated in the laboratory in pure cultures, with little validation in nature. Here, we show that genomic traits related to laboratory measurements of maximum growth potential failed to predict the growth rates of bacteria in unamended soil, but successfully predicted growth responses to resource pulses: growth increased with 16S rRNA gene copy number and declined with genome size after substrate addition to soils, responses that were repeated in four different ecosystems. Genome size best predicted growth rate in response to addition of glucose alone; adding ammonium with glucose weakened the relationship, and the relationship was absent in nutrient-replete pure cultures, consistent with the idea that reduced genome size is a mechanism of nutrient conservation. Our findings demonstrate that genomic traits of soil bacteria can map to their ecological performance in nature, but the mapping is poor under native soil conditions, where genomic traits related to stress tolerance may prove more predictive. These results remind that phenotype depends on environmental context, underscoring the importance of verifying proposed schemes of trait-based strategies through direct measurement of performance in nature, an important and currently missing foundation for translating microbial processes from genes to ecosystems.

Introduction

Maximum potential growth rate is a defining feature of organisms, variable among species of plants [1], animals [2], and prokaryotes [3]. It influences the ability of an organism to grow rapidly in response to a pulse of substrate

availability, a warming or cooling toward more optimal temperatures, or a release from competition. Maximum potential growth rate is thought to be an emergent feature of genomic traits, characteristics of species encoded in the genome. For example, maximum growth rate is expected to increase with an organism's capacity to produce proteins [4], a capacity indicated by the number of copies of the rRNA gene [5] because ribosomes are the sites of protein production in the cell. As a genome accumulates copies of the genes for producing ribosomes, the more rapidly ribosome biogenesis can occur, and the more rapidly protein synthesis can keep pace with opportunities in a fluctuating environment. Thus, the number of rRNA gene copies in the genome is thought to be an indicator of an organism's maximum potential growth rate, an indicator of the ability of that organism to ramp up production of the thousands of proteins necessary for metabolic functioning and cellular biosynthesis [5]. There is support for this idea in bacteria: maximum potential growth rate, measured in pure culture under laboratory conditions, is related to rRNA gene copy number across numerous prokaryotic taxa [3].

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-019-0422-z>) contains supplementary material, which is available to authorized users.

✉ Bruce A. Hungate
Bruce.Hungate@nau.edu

¹ Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011, USA

² Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011, USA

³ Department of Biology, West Virginia University, Morgantown, WV 26506, USA

⁴ Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Genome size is another trait thought to influence growth rate: small genomes are expected to be associated with rapid growth [6], because smaller genomes reduce nutrient demand for genome replication [7] and reduce mutational load [6]. Experimental genome reductions can enhance growth [8], and synthetic genomes of minimum size have the capacity for very rapid replication [9]. Decreasing the nitrogen and phosphorus requirements of cell division is postulated to be one advantage of a small genome [7]; nutrient-replete conditions may obscure this advantage during growth in laboratory cultures [3].

In nature, bacterial growth may be less constrained by maximum potential growth rate than by environmental limitations, such as low resource availability [10], stress from conditions that are rarely optimal [11], and interactions with other organisms, many of which reduce growth [12]. These and other factors influencing actual rates of growth in nature are undoubtedly shaped by other genomic traits, such that actual rates of growth in the environment may be unrelated to genomic traits governing growth potential. On the other hand, fluctuations in nature can create conditions of temporary resource abundance where the actual growth rates of organisms may approach their maximum potentials, and thus where genetic indicators of those maxima could be predictive. Such resource fluctuations in soil can be caused by roots' release of carbon compounds through exudation [13], freeze-thaw and wet-up events that create resource pulses [11], or excretion of waste products into the environment by animals [14]. In response to such resource fluctuations, organisms with greater genetic capacity for rapid growth may reveal themselves phenotypically by growing rapidly.

The idea that measurable genomic traits relate to growth provides an attractive foundation for functional interpretation of genomic data from the environment. Indeed, interpretations of microbiomes collected from the environment often rely on inferences about growth rates based on rRNA gene copy numbers of bacterial community members [15–20]. But the relationship between copy number and

maximum potential growth rate is based on laboratory measurements of growth, so the underlying hypothesis that the relationship actually applies in nature remains untested.

Here, we used quantitative stable isotope probing (qSIP) with ^{18}O -labeled water to measure the rate of DNA synthesis of individual bacterial taxa in soil, using the 16S rRNA gene as a taxonomic marker [21].

Materials and methods

Summary

Measurements occurred in natural soil assemblages collected from a climatic gradient in northern Arizona, USA (Table 1). We measured growth rates in unamended soils and in response to substrate addition simulating pulses of resource availability, glucose alone and glucose plus ammonium, in order to test the hypotheses that genome size and the number of copies of the ribosomal gene are genomic functional traits that predict taxon-specific growth rates in soil bacteria across a range of resource conditions. 16S rRNA gene copy numbers were estimated by comparing observed sequences with BLAST against a database we constructed to contain only complete genomes, using four different thresholds for sequence identity at the 16S locus to test for sensitivity to any particular cutoff: 94.5% (genus), 98.7% (species), 99.5%, and 100% [22]. Patterns observed were insensitive to the selected cutoff, and we used the 98.7% for the analyses presented in the main text (results for other cutoff values are provided in the Online Supplementary Materials). We observed that estimated 16S rRNA gene copy numbers corresponded well with strains with known copy number ($r \geq 0.99$, Fig. S1). We calculated the sequence-weighted mean 16S rRNA gene copy number of each operational taxonomic unit (OTU). Similarly, by matching the sequences to genome assemblies with known genome sizes, we estimated the sequence-weighted mean genome size of each operational taxonomic unit.

Table 1 Median growth rates, copy number of the 16S rRNA gene, number of bacterial taxa mean annual temperature (MAT), mean annual precipitation (MAP), aboveground net primary productivity (ANPP), and net ecosystem carbon exchange (NEE) across the four soils for the unamended (control) condition

	Growth rate median \pm S.E. (h^{-1})	Copy number median \pm S.E. (copies cell^{-1})	Genome size median \pm S.E. (Megabyte cell^{-1})	Count (# of taxa)	MAT $^{\circ}\text{C}$	MAP mm	ANPP $\text{g m}^{-2} \text{y}^{-1}$	NEE $\mu\text{mol m}^{-2} \text{s}^{-1}$
Grassland	0.00057 \pm 0.00007	3.00 \pm 0.23	7.01 \pm 0.25	82	13.0	190.8	94.4 \pm 5.6	-0.453 \pm 0.144
Pinyon-Juniper	0.00077 \pm 0.00005	3.00 \pm 0.32	6.40 \pm 0.28	59	10.5	282.1	125.7 \pm 5.0	-0.551 \pm 0.076
Ponderosa	0.00099 \pm 0.00007	3.00 \pm 0.23	6.11 \pm 0.19	115	9.1	520.6	151.3 \pm 5.1	-0.660 \pm 0.098
Mixed Conifer	0.0012 \pm 0.00005	3.00 \pm 0.20	6.31 \pm 0.19	105	6.6	661.2	159.4 \pm 8.1	-0.766 \pm 0.106

Experimental design

Soils (0–10 cm) were collected from four ecosystems: high desert grassland, piñon-juniper woodland, ponderosa pine forest, and mixed conifer forest, representing a range of climate and ecosystem variation (Supplementary Table 1) along the C. Hart Merriam elevation gradient [23]. Soils were air-dried at room temperature for 24 h, passed through a 2-mm sieve, and stored at 4 °C prior to experiments.

Two grams of dry weight soil from each ecosystem were placed into 15 mL falcon tubes. Soils received water (adjusted to 70% water-holding capacity) or water spiked with C alone or C plus N at concentrations of 1000 $\mu\text{g C g}^{-1}$ soil (as glucose) soil and 100 $\mu\text{g N g}^{-1}$ soil (as $(\text{NH}_4)_2\text{SO}_4$) in the following isotope and nutrient treatments ($n = 3$ per treatment): (1) ^{18}O -enriched water (97 atom %); (2) glucose at natural abundance $\delta^{13}\text{C}$ and ^{18}O -enriched water (97 atom %); (3) glucose and $(\text{NH}_4)_2\text{SO}_4$ at natural abundance $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$, and ^{18}O -enriched water (97 atom %); (4) ^{13}C -enriched glucose (99 atom %) and water at natural abundance $\delta^{18}\text{O}$; and (5) ^{13}C -enriched glucose (99 atom %), and $(\text{NH}_4)_2\text{SO}_4$ and water at natural abundance $\delta^{15}\text{N}$ and $\delta^{18}\text{O}$. All the treatment combinations above had an associated natural abundance isotopic control that received the same amount of water and nutrients, but all at natural abundance $\delta^{18}\text{O}$, $\delta^{13}\text{C}$, and $\delta^{15}\text{N}$ values. Soils were incubated for 1 week at room temperature (~ 23 °C).

Quantitative stable isotope probing

After the incubation, samples were frozen at -80 °C. DNA was extracted using a PowerSoil DNA extraction kit following manufacturer's instructions (Mobio laboratories, Carlsbad, CA, USA). Ultracentrifugation, fractionation, quantitative PCR, and 16S rRNA gene amplicon sequencing were performed as previously described [21] with minor modifications. For density centrifugation, 1 μg of DNA was added to 2.6 mL of saturated CsCl solution and gradient buffer (200 mM Tris, 200 mM KCl, 2 mM EDTA) in a 3.3 mL OptiSeal ultracentrifuge tube (Beckman Coulter, Fullerton, CA, USA), which had a final density of 1.71 g cm^{-3} . Samples were centrifuged in an Optima Max bench top ultracentrifuge (Beckman Coulter) with a Beckman TLN-100 rotor ($127,000 \times g$ for 72 h) at 18 °C. The resulting density gradient was immediately fractionated into ~ 20 fractions (150 μL) per sample using a modified fraction recovery system (Beckman Coulter). Fraction density was determined using a Reichert AR200 digital refractometer (Reichert Analytical Instruments, Depew, NY, USA). DNA in each fraction was purified using a standard isopropanol precipitation method and

16S rRNA gene copies were quantified via quantitative PCR as described previously [24]. Fractions with densities between 1.650 and 1.735 g cm^{-3} (~ 15 per sample) were amplified using 515F and 806R primers [25] and sequenced on an Illumina MiSeq instrument using a v2 300 cycle reagent kit (Illumina, Inc., San Diego, CA, USA).

Sequence data processing and analysis

Sequences were subject to quality filtering following the protocol previously described [26]. Samples with fewer than 4000 sequences were eliminated from downstream analyses using the custom shell script `filter_samples.sh` (<https://bitbucket.org/junhuilinau/manuscript-supplementary/src/master/sequence/>). The filtered sequences were clustered using open-reference OTU (operational taxonomic unit) picking with 97% sequence identity against the SILVA 132 database (https://www.arb-silva.de/download/archive/qiime/Silva_132_release.zip) in QIIME 1.9.1 [27]. OTUs with less than 0.001% relative abundance were discarded. After filtering, 89.7% of total sequences were retained, representing 7665 OTUs. Samples were rarefied to 4000 reads 100 times using the QIIME script `multiple_rarefactions_even_depth.py`. Using the custom shell script `rarefaction_average.sh` (<https://bitbucket.org/junhuilinau/manuscript-supplementary/src/master/sequence/>), we calculated the average counts of each OTU of all 100 rarefied OTU tables for downstream excess atom fraction (EAF) estimation.

Estimating 16S rRNA gene copy number and genome size

In order to estimate 16S rRNA copy number for each individual sequence, and best match that estimate to available information, we developed an algorithm that was performed separately from OTU assignment. Specifically, we evaluated each unclustered, quality-filtered sequence for genome size and copy number of the 16S rRNA gene against a locally constructed complete genome database, as described below. After assigning an estimate of copy number and genome size to each sequence from our 16S rRNA amplicon dataset, we compiled these into an estimate for each OTU, based on the assignment of sequences to OTUs using QIIME (described above). We used this procedure in lieu of publicly available software often used for assigning traits to taxa, in order to capitalize on the most up-to-date available information on whole genome sequences, because software updates are not always able to keep pace with the growing complement of available data [28].

We built a database using a total of 10,684 prokaryotic assemblies deposited in Genbank that were annotated as a "Complete Genome" downloaded from NCBI in 16 June

2018 (ftp://ftp.ncbi.nlm.nih.gov/genomes/GENOME_REPORTS/prokaryotes.txt). This yielded a single, local database with a total of 20,652 individual chromosomes and plasmids, because many bacterial genomes contain more than a single chromosome. We used BLAST (with the parameters bit score ≥ 400 and sequence alignment length ≥ 100) to identify which genome in the constructed database had the best match (defined by highest bit score) at the 16S rRNA locus to each environmental amplicon sequence. In total, 21% of all the sequences assigned to an OTU were matched to a chromosome or plasmid at 98.7% identity. Ten percent of sequences were matched at 100% identity, 14 percent at 99.5% identity, and 44 percent at 94.5% identity. Copies of the 16S rRNA gene were enumerated for each genome match as the number of iterations in the chromosome or plasmid that contained the match to our sequence at four thresholds (94.5, 98.7, 99.5, and 100%). The total number of 16S rRNA gene copies for each OTU was determined as the abundance-weighted average across sequences assigned to that OTU. When assemblies in the NCBI database indicated the presence of multiple chromosomes or plasmids in a genome, copy number estimates were summed prior to averaging for the sample. The 16S rRNA gene copy number estimation was done using the R script `best.hit.assembly.R` (<https://bitbucket.org/junhuilinau/manuscript-supplementary/src/master/16s/>).

The procedure and criteria for best BLAST hit described above were also used for genome size estimates: when our environmental sequences matched a 16S rRNA sequence in the database, the size of the chromosome or plasmid in which the match occurred (at a given identity threshold) was recorded. Chromosomes and plasmids assigned to the same assembly in the database were summed to arrive at a single genome size estimate for that assembly. More specifically, we manually created a mapping file (`assembly-genome.size.txt`) for genome assemblies in NCBI. We compared the sequence-assembly file generated during the estimation of 16S rRNA gene copy number to the mapping file to identify matches, and then we compiled chromosome and plasmid size estimates for each OTU (`silva_otu_seq.txt`). Genome size was estimated as the average size weighted by relative abundance in the database. The estimation of genome size can be reproduced using the R script `genome_size.R` (<https://bitbucket.org/junhuilinau/manuscript-supplementary/src/master/size/>).

EAF estimation

The EAF ^{18}O and ^{13}C of each taxon were estimated following the procedures described previously [21, 29]. Relative growth rate (RGR) was estimated as a function of the rate of ^{18}O assimilation into DNA (as measured by EAF at the end of the incubation), assuming that 60% of the

oxygen in DNA is derived from water [30], and that populations are at steady state: $\text{RGR} = \text{EAF}/(0.6 * 168 \text{ h})$. This estimate further assumes that oxygen is a conservative tracer of DNA replication, and that DNA replication primarily occurs during cellular division (i.e., growth). The soils were incubated for one week, and it is possible that bacterial necromass was re-utilized by other bacteria, cross-feeding that would introduce error into our estimates of RGRs. As of yet, we are unable to quantify cross-feeding, so this is a potential source of error in our growth estimates. Still, soils were exposed to $^{18}\text{O}\text{-H}_2\text{O}$ during the entire incubation period, so even growing cross-feeders would become labeled with ^{18}O from water, and this is likely to be a larger signal than the re-cycled ^{18}O -labeled organic matter.

Sensitivity analyses

To assess the robustness of our findings, we performed sensitivity analyses on the estimation of 16S rRNA gene copy number for OTUs across the four thresholds for percent match (94.5% (genus), 98.7% (species), 99.5, and 100%) at the 16S rRNA locus (hereinafter referred to as ‘percent identity’) when comparing sequences against the constructed complete genome database described above. To do so, we calculated the sequence-weighted mean 16S rRNA gene copy number of each OTU (Fig. S2), then we evaluated the results in Figs. 1 and 2, and Tables 1–3 by repeating the analyses for different thresholds of percent identity: 94.5% (genus), 99.5, and 100%. With few minor statistical differences, patterns were identical for growth versus 16S rRNA gene copy number and genome size (Figs. S3 and S4, and Tables S1–S3), indicating that the results were robust to variation in thresholds used to assess sequence identity. We also compared our 16S rRNA gene copy number estimates with those derived from the `rrnDB` [31], and correspondence was excellent (slopes all indistinguishable from the 1:1 line, with correlation coefficients very near 1; Fig. S1).

Data on growth rate and 16S rRNA gene copy number in Fig. 1 were obtained from the literature [3, 32–36]. Data were extracted from tables or digitized from figures. In 12 out of 199 cases, more than one report for a given bacterial genus and species were encountered, where separate laboratories measured maximum potential growth rate and 16S rRNA copy number for lab strains identified by the same genus and species. Likely, these represent estimates for genetically distinct strains that were related enough to be identified by the same name, similar to strain-level variation within organisms identified by the same OTU in our soil studies. In order to have comparable approaches to statistical independence

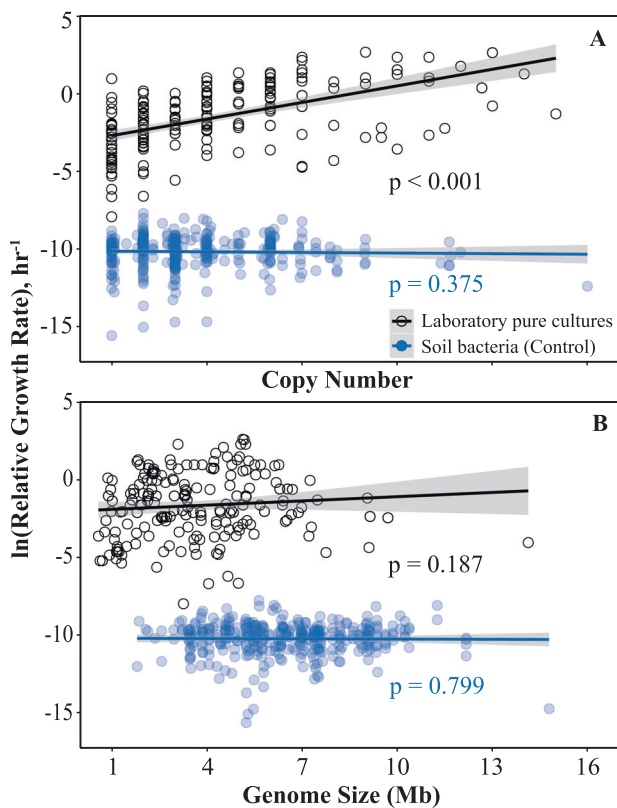


Fig. 1 Relationship between growth rates of bacteria and genomic traits (copy number, panel **a**; and genome size, panel **b**) for bacteria grown in the laboratory in pure culture (gray circles; data from the literature [3, 32–36] or bacteria naturally occurring in unamended soils (blue filled circles). Solid lines show results from simple linear regression (lm function in R), with 95% confidence intervals shown in the shaded areas. To facilitate comparison with the laboratory data, data from four soils are combined, here; statistical analyses of each soil independently are summarized in Table 2

between the soil studies and lab synthesis, for the laboratory data we used the averages of reported growth and 16S rRNA copy numbers across studies that evaluated organisms identified by the same name. Data on bacterial genome size were obtained by matching bacterial name to the complete assembly in NCBI with known genome size, and the average genome size was used if matching to multiple assemblies.

Code and data availability

More details of the pipeline and all scripts used for computational analyses in this study are available at <https://bitbucket.org/junhuilinau/manuscript-supplementary/src/master/>. All sequence data and sample metadata have been deposited in MG-RAST under the project ID mgp88472. All other data that support the findings of this study are available from the corresponding author upon reasonable request.

Statistical analyses

All statistical analyses were performed in R version 3.4.1 [37]. The linear regression analyses between copy number and RGR, and genome size and RGR were performed using the lm function in R, and in general, the residuals were normally distributed. All figures were created in ggplot2. To tease apart the relative importance of genomic traits on the RGR, we used multiple regression models including 16S rRNA copy number and genome size, and the best model was selected with the smallest AIC value (Table 2).

Results

In unamended soils, the number of copies of the 16S rRNA gene was unrelated to bacterial growth rate ($r^2 = 0.002$, Fig. 1a). Genome size was also a poor predictor of growth ($r^2 < 0.001$, Fig. 1b). When soils were considered independently, the patterns were similar, with low explanatory power of 16S rRNA gene copy number for growth (Table 2); the highest r^2 value of 0.07 actually applied to a negative relationship observed for the ponderosa pine soil. Genome size was also a poor predictor of growth rate for soils considered independently, with at most 1.4% of the variation explained for the desert grassland (Table 2). By contrast, growth rates of bacteria in culture under nutrient replete conditions in the laboratory were strongly correlated with the number of copies of the 16S rRNA gene (Fig. 1a), where this single trait explained nearly a third of the observed variation ($r^2 = 0.317$).

Growth rates measured in the natural soil assemblages were nearly three orders of magnitude lower compared to potential growth rates in pure cultures under laboratory conditions (Fig. 1). Even though growth rates in soils were substantially lower than laboratory growth potentials, there was still substantial variation in bacterial growth rates in unamended soil, from 0.0000088 to 0.0047 h⁻¹ (Fig. 1). Median growth rates increased along the gradient, from the xeric and warm grassland site to the mesic and cool mixed conifer site, and from low to high net primary production and net ecosystem exchange, which are ecosystem-scale measures of carbon input to soil (Table 1). Variation in the growth rates of these bacteria spanned more than two orders of magnitude. Yet nearly all (>99%) observed variation was unexplained by rRNA gene copy number and genome size, indicating that the variation in the growth rates of bacteria in unamended soil reflects constraints other than the maximum potential.

With added glucose, with or without supplemented nitrogen, the number of rRNA copies in the genome

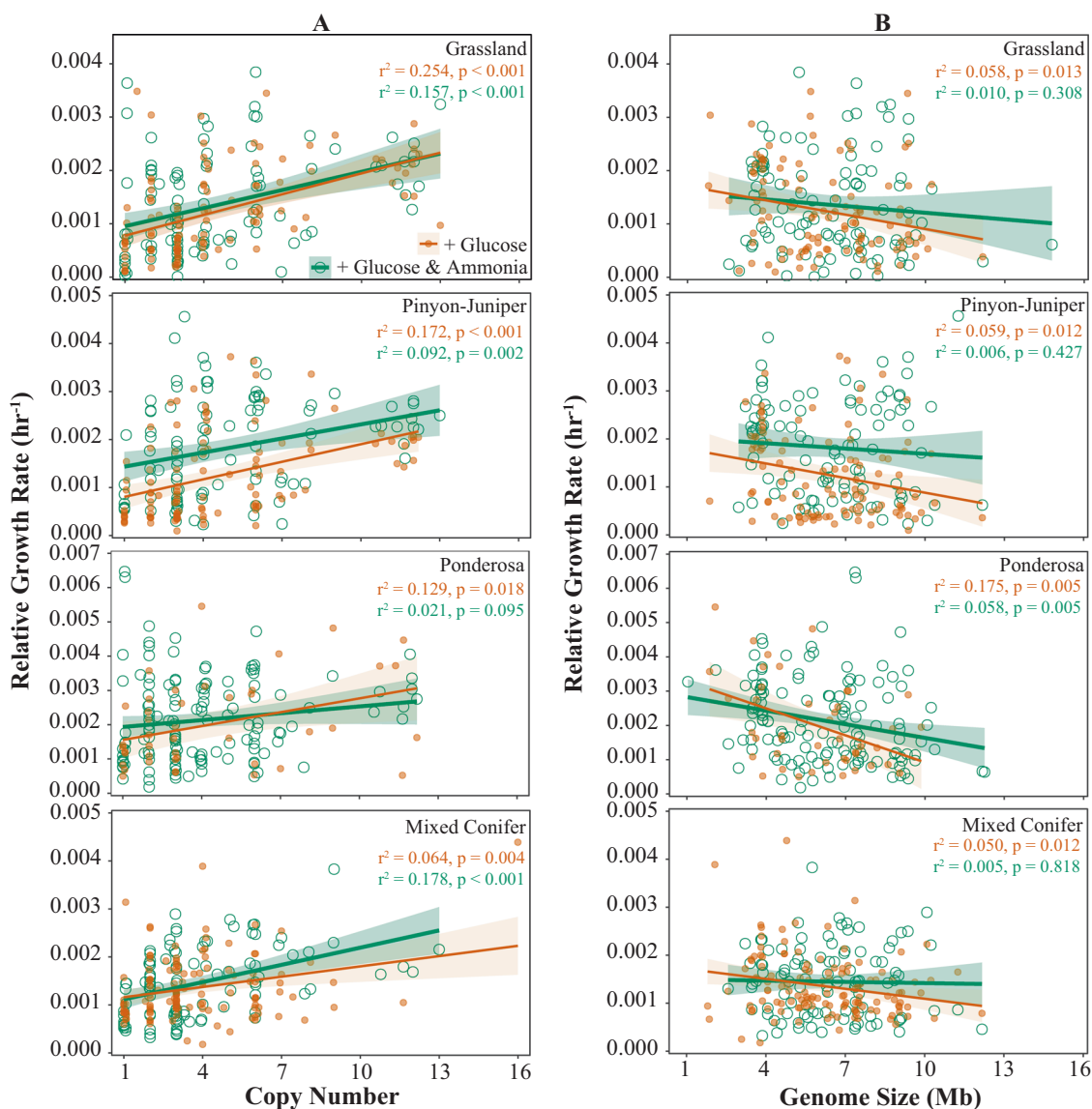


Fig. 2 Relationships between observed growth rates of bacteria in soil in response to resource amendment and copy number of the 16S rRNA gene (a) or genome size (b) for four different ecosystems along an elevation gradient in northern Arizona. Resource additions were either

glucose alone (orange, filled circles) or glucose plus ammonium additions (green, open circles). Statistical analyses of the relationships are presented in Table 2

was a significant predictor of actual growth rates across taxa in the assemblage (Fig. 2a). This response is most simply explained by glucose utilization for growth: the observed ^{18}O increase in DNA in response to added glucose was strongly correlated with increased utilization of glucose for growth, as evidenced by ^{13}C assimilation from ^{13}C -labeled glucose (Fig. 3). Bacterial taxa with more copies of the rRNA gene grew faster in response to glucose addition, 0.00011 h^{-1} per additional gene copy, a finding that was statistically significant for all ecosystems (Table 2). Similarly, in the resource pulse treatments, growth rates increased as genome size declined (Fig. 2b). The relationship was strongest with added glucose alone,

where growth rate decreased by $0.00013 \pm 0.000044 \text{ h}^{-1} \text{ Mb}^{-1}$ genome.

When ammonium was added with glucose, the slope declined by more than a factor of two, to $0.000055 \pm 0.000027 \text{ h}^{-1} \text{ Mb}^{-1}$ (Table 2). The decline in the slope means that ammonium addition weakened the growth advantage of a smaller genome. Similarly, for the glucose-only treatment, models predicting growth and selected based on AIC consistently included both copy number and genome size as predictors of growth rate (Table 2), whereas for the treatment with added ammonium, copy number was always included in the best model, but genome size was only statistically significant for one soil,

Table 2 Model selection for predicting ¹⁸O content (excess atom fraction ¹⁸O) in prokaryotic growth rates based on copy number (C) and genome size (G)

	Model	Slope C	Slope G	Intercept	<i>P</i> -value C	<i>P</i> -value G	<i>P</i> -value C & G	R ²	AIC
No amendment									
Desert Grassland	C	0.0041		0.06393	0.2271		0.2271	0.0182	-214.4
	G		0.0034	0.05499		0.2824	0.2824	0.0144	-214.1
	C & G	0.0037	-0.0030	0.04518	0.2827	0.3554	0.3151	0.0288	-213.3
Pinyon-Juniper	C	0.0013		0.07344	0.5346		0.5346	0.0068	-208.5
	G		-0.0007	0.08254		0.7839	0.7839	0.0013	-208.2
	C & G	0.0013	-0.0006	0.07743	0.5450	0.8042	0.8008	0.0079	-206.6
Ponderosa-Pine	C	-0.0073		0.14523	<i>0.0079</i>		<i>0.0079</i>	0.0609	-275.1
	G		-0.0022	0.13127		0.5200	0.5200	0.0037	-268.3
	C & G	-0.0073	-0.0018	0.15604	<i>0.0087</i>	0.5911	<i>0.0257</i>	0.0633	-273.4
Mixed Conifer	C	0.0016		0.12470	0.5254		0.5254	0.0039	-316.1
	G		0.0012	0.12246		0.6334	0.6334	0.0022	-315.9
	C & G	0.0015	-0.0010	0.11872	0.5680	0.6970	0.7580	0.0054	-314.3
Glucose									
Desert Grassland	C	0.0131		0.06528	<i><0.0001</i>		<i><0.0001</i>	0.2540	-259.3
	G		-0.0090	0.18107		<i>0.0128</i>	<i>0.0128</i>	0.0576	-234.3
	C & G	0.0125	-0.0068	0.11032	<i><0.0001</i>	<i>0.0323</i>	<i><0.0001</i>	0.2863	-262.0
Pinyon-Juniper	C	0.0123		0.06881	<i><0.0001</i>		<i><0.0001</i>	0.1724	-227.8
	G		-0.0102	0.19058		<i>0.0120</i>	<i>0.0120</i>	0.0586	-214.0
	C & G	0.0114	-0.0076	0.12109	<i><0.0001</i>	<i>0.0433</i>	<i><0.0001</i>	0.2044	-230.0
Ponderosa-Pine	C	0.0136		0.14325	<i>0.0182</i>		<i>0.0182</i>	0.1286	-55.0
	G		-0.0264	0.35728		<i>0.0052</i>	<i>0.0052</i>	0.1754	-57.4
	C & G	0.0097	-0.0216	0.28604	0.0839	<i>0.0230</i>	<i>0.0047</i>	0.2354	-58.6
Mixed Conifer	C	0.0072		0.10965	<i>0.0040</i>		<i>0.0040</i>	0.0643	-334.3
	G		-0.0070	0.17980		<i>0.0117</i>	<i>0.0117</i>	0.0498	-332.4
	C & G	0.0074	-0.0072	0.15376	<i>0.0026</i>	<i>0.0073</i>	<i>0.0004</i>	0.1173	-339.7
Glucose+Ammonium									
Desert Grassland	C	0.0115		0.08456	<i><0.0001</i>		<i><0.0001</i>	0.1570	-211.7
	G		-0.0042	0.16368		0.3084	0.3084	0.0102	-195.0
	C & G	0.0113	-0.0016	0.09548	<i><0.0001</i>	0.6808	<i>0.0002</i>	0.1584	-209.9
Pinyon-Juniper	C	0.0099		0.13478	<i>0.0021</i>		<i>0.0021</i>	0.0919	-179.3
	G		-0.0037	0.20714		0.4269	0.4269	0.0064	-170.2
	C & G	0.0097	-0.0018	0.14759	<i>0.0028</i>	0.6817	<i>0.0082</i>	0.0935	-177.5
Ponderosa-Pine	C	0.0067		0.18824	0.0947		0.0947	0.0209	-179.1
	G		-0.0133	0.29832		<i>0.0050</i>	<i>0.0050</i>	0.0577	-184.3
	C & G	0.0071	-0.0136	0.27164	0.0710	<i>0.0040</i>	<i>0.0039</i>	0.0807	-185.6
Mixed Conifer	C	0.0121		0.10045	<i><0.0001</i>		<i><0.0001</i>	0.1780	-258.5
	G		-0.0009	0.15169		0.8181	0.8181	<i>0.0005</i>	-238.6
	C & G	0.0122	-0.0021	0.11338	<i><0.0001</i>	0.5420	<i><0.0001</i>	0.1811	-256.9

P-values in italics indicate statistical significance at *P* < 0.05

All models have the generic form: excess atom fraction ¹⁸O = copy number * slope C + genome size * slope G + intercept. Blank cells indicate cases where a term was excluded from the model (e.g., a model based on C, copy number, will not have a slope or *P*-value estimate for G, genome size). For cases where at least one model was statistically significant, the best model is indicated in bold, underline

ponderosa pine (Table 2). In pure cultures assessing maximum potential growth rates, where bacteria are grown with ample nutrients [3], there was no relationship between growth rate and genome size (Fig. 1b).

Discussion

Resource pulses elicited growth responses in soil bacteria, and the patterns in growth after resource additions were

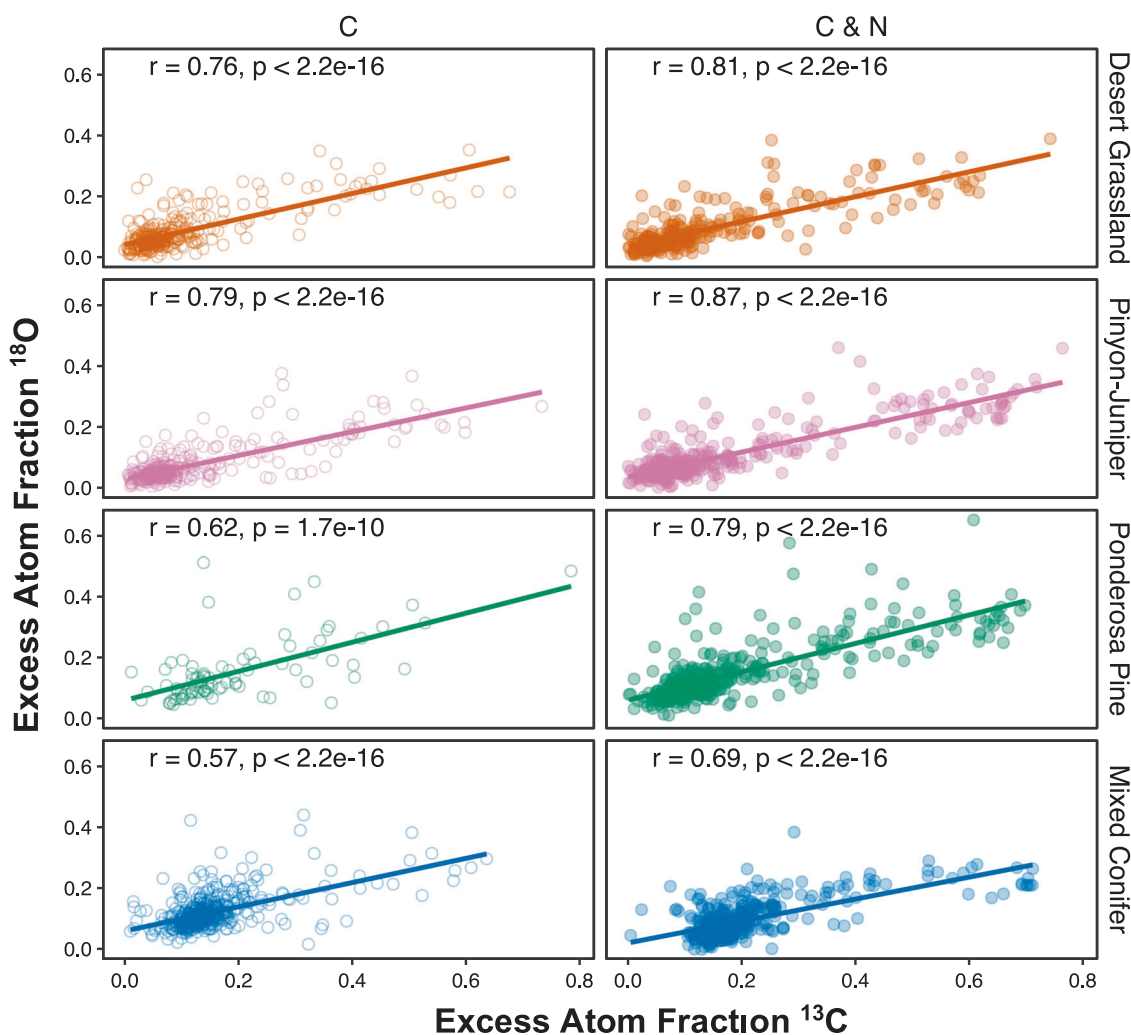


Fig. 3 Positive correlation (Pearson) between ^{13}C and ^{18}O atom fraction excess in treatments with added glucose, indicating that increased

growth rates in response to glucose addition were associated with utilization of glucose as a growth substrate

positively related to 16S rRNA gene copy number and negatively to genome size (Fig. 2). These patterns, recapitulated across four different ecosystems, show that genomic traits can predict growth response to resource pulses in intact soil bacterial assemblages. In other words, under conditions where resource pulses bring natural soils closer to laboratory conditions, the phenotypes of bacteria in nature are predictable from genomic traits. This finding supports the translation of general principles developed from pure culture studies to the ecological performance of organisms in nature: resource pulses that occur in biodiverse soil microbial assemblages elicit growth responses analogous to those observed in pure cultures under nutrient-replete conditions.

However, the success of this translation was qualified. The increase in growth we measured in natural soils was small: growth increased by $0.00011 \text{ h}^{-1} \text{ copy}^{-1}$, whereas maximum potential growth rate in the laboratory cultures

increased far more rapidly with each accumulated 16S rRNA gene copy, $0.25 \text{ h}^{-1} \text{ copy}^{-1}$ (Fig. 1a). The proportion of the variance explained in the soils (r^2 averaged 0.02 for unamended and 0.13 for glucose and glucose+ammonium treatments across ecosystems) was less than that explained in the lab ($r^2 = 0.317$). This pattern is similar to that observed for the RNA-to-DNA ratio, which has sometimes been found to increase with growth rate and metabolic activity in laboratory cultures [38, 39], but fails to reflect variation among taxa in either growth or metabolic activity in soil bacterial communities [40, 41]. Soils are far more complex than laboratory cultures, and the large unexplained variation in the soils indicates the importance of other limitations to growth not related to maximum potential growth rate.

The simplest explanation for increased growth in response to glucose addition is direct utilization of glucose as a growth substrate. This is consistent with the idea that

Table 3 Variance partitioning for regression analyses for growth versus 16S rRNA gene copy number

	Positive	Negative
Desert Grassland	93.9%	5.4%
Pinyon-Juniper	94.6%	5.3%
Ponderosa Pine	97.5%	2.4%
Mixed Conifer	74.5%	24.9%

We predicted ('predict' function in R) the excess atom fraction (EAF) values of OTUs in the glucose-amended treatment of each ecosystem using the regression model between EAF and 16S rRNA gene copy number of the control treatment, and we calculated the residuals using the observed EAF value in the glucose-amended treatment. We then subsetted the positive and negative responders: an OTU was considered to be a positive responder if the smaller residual of the OTU at 95% confidence was greater than 0; an OTU was considered to be a negative responder if the larger residual of the OTU at 95% confidence was less than 0. We finally calculated the partitioning variance by dividing the residual sum of squares for positive or negative responders by the total residual sum of squares

microbial growth in soil is limited by carbon [42]. However, some taxa responded to glucose addition by reducing growth, a response most simply explained by competition [12]: taxa whose growth rates were stimulated by glucose addition may have produced compounds inhibitory to other bacteria, or they may have appropriated resources limiting to the growth of other bacteria in the community, whose growth rates declined in response. Glucose addition may cause toxicity to microbes, but the amounts added are likely too low to elicit toxic responses caused by osmotic stress [43] or acidification [44]. We have no direct evidence for either mechanism, but we find competition the more likely.

A simple and sufficient explanation for the positive slope between 16S rRNA gene copy number and growth is increased growth among taxa with high numbers of 16S rRNA gene copies. However, the reduced growth among low 16S rRNA gene-copy-number taxa that we observed (Table 3) also contributed to the positive slope. This means that the number of copies of the rRNA gene might predict not only growth potential but also competitive ability, a metric of performance accessible only when examining organisms in communities. These results align with recent findings of smaller genome size and reduced genetic potential for competitive antimicrobial compound production in coal-fire heated soils [45].

Our finding that ammonium addition alleviated the cost of a larger genome for growth (Fig. 2, Table 2) is consistent with the nutrient stress hypothesis [7]. Carbon addition to soil often stimulates ammonium immobilization by heterotrophic microorganisms, exacerbating N limitation of growth and of associated genome replication, and therefore conferring an advantage to a smaller genome with lower N cost of replication. When N and C are added together, this advantage is diminished. This also explains the results from

the synthesis of pure culture studies (Fig. 1b), because with no nutrient limitation, there would be no growth advantage to having a smaller genome.

Microbiology has advanced by studying microorganisms in pure culture in the laboratory, and calls for improved culturing strategies promise new insights [46]. Ecology has advanced by studying organisms in nature, accepting environmental heterogeneity and community interactions as essential features of the world organisms inhabit. Microbial ecology should bridge these approaches. Attempts to do so are widespread using molecular tools, inferring function based on gene sequence data collected from the field. However, the translation from genes to function rests on a pure culture foundation, and that foundation can fail, as demonstrated here for unamended soils where a commonly applied prediction from culture failed in nature. The failure is not surprising: performance of an organism under near optimal conditions in pure culture will differ from the performance of the very same organism in nature, subject to competitive interactions, limited by resources, and stressed by suboptimal environmental conditions. Results presented here show how microbial ecology can advance by measuring quantitative trait variation of microorganisms in the habitats where they naturally occur. Techniques like qSIP can evaluate where principles derived from the laboratory apply to microorganisms in nature, and where they fail. It is not surprising that growth responses to resource pulses corresponded with traits derived from studies of resource-rich laboratory cultures. At the same time, the high variation in growth rates observed without resource amendment points to important phenotypic variation in growth (Fig. 1), even under resource-limited conditions, and the need to explore the genomic traits and environmental conditions that drive that variation.

Assigning ecological strategies based on taxonomy is a common approach for interpreting microbiome data, but this effort to date is largely divorced from measurements testing whether the organisms actually utilize their assigned strategies in nature. Growth is a useful metric for evaluating ecological strategies, because it integrates ecological and evolutionary processes, from metabolism [2], to resource uptake and use [47] and thus the imprint of biology on element cycles, to fitness [48], the ultimate result of variation in genomic traits. Assessing growth in natural microbial communities, combined with molecular tools, provides access to a richer suite of ecological mechanisms influencing organismal performance than can be assessed in laboratory cultures. Findings presented here show that it is now possible to pair genomic traits of microorganisms with their growth rates in nature. Such efforts hold promise for a refined and evidence-based foundation for proposed ecological strategies, whether best defined by a single axis of copiotrophic to oligotrophic [49], by a triangle of

competitive, ruderal, and tolerance [50], or by multi-dimensional spectra of traits and tradeoffs [51].

Acknowledgements This work was supported by the US National Science Foundation (Dimensions of Biodiversity, DEB-1321792) and the US Department of Energy, Program in Genomic Sciences (DE-SC0016207). We thank Jingrun Sun, Zacchaeus Compson, Bri Finley, and Jeff Propster for assistance in the laboratory, and Alicia Purcell for intellectual contributions. Work at Lawrence Livermore National Laboratory (LLNL) was funded by the Department of Energy through the Genome Sciences Program under contracts SCW1024 and SCW1590, and performed under the auspices of LLNL under Contract DE-AC52-07NA27344.

Author contributions All authors contributed to the conceptual development of this work. BAH, XJAL, PD, ES, and RLM designed the incubation study. XJAL and RLM performed the incubation study. RLM and MH conducted the molecular analyses. RLM, BJK, ES, BAH, BWS, and EMM analyzed the qSIP data. JL developed and performed the bioinformatics analysis, the meta-analysis of growth estimates of bacteria in the laboratory, and statistical inference tests. EMM, SJB, JPR, and BWS provided additional intellectual input. BAH and JL wrote the manuscript, with editorial contributions from all authors, and particularly substantive contributions from PD, JP, EMM, and SJB.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Reich PB, Walters MB, Ellsworth DS. From tropics to tundra: global convergence in plant functioning. *Proc Natl Acad Sci USA*. 1997;94:13730–4.
2. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. Toward a metabolic theory of ecology. *Ecology*. 2004;85:1771–89.
3. Roller BRK, Stoddard SF, Schmidt TM. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat Microbiol*. 2016;1:16160.
4. Cox RA. Correlation of the rate of protein synthesis and the third power of the RNA:protein ratio in *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology*. 2003;149:729–37.
5. Scott M, Klumpp S, Mateescu EM, Hwa T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol Syst Biol*. 2014;10:747.
6. Lynch M. Streamlining and simplification of microbial genome architecture. *Annu Rev Microbiol*. 2006;60:327–49.
7. Hessen DO, Jeyasingh PD, Neiman M, Weider LJ. Genome streamlining and the elemental costs of growth. *Trends Ecol Evol*. 2010;25:75–80.
8. Lee M-C, Marx CJ. Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet*. 2012;8:e1002651.
9. Wolf YI, Koonin EV. Genome reduction as the dominant mode of evolution. *Bioessays*. 2013;35:829–37.
10. Lever MA, Rogers KL, Lloyd KG, Overmann J, Schink B, Thauer RK, et al. Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiol Rev*. 2015;39:688–728.
11. Schimel J, Balsler TC, Wallenstein M. Microbial stress-response physiology and its implications for ecosystem function. *Ecology*. 2007;88:1386–94.
12. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol*. 2010;8:15–25.
13. Gao J, Sasse J, Lewald KM, Zhahlnina K, Commesser LT, Duncombe TA, et al. Ecosystem Fabrication (EcoFAB) Protocols for The Construction of Laboratory Ecosystems Designed to Study Plant-microbe Interactions. *J Vis Exp*. 2018;134:57170.
14. Williams PH, Haynes RJ. Effect of sheep, deer and cattle dung on herbage production and soil nutrient content. *Grass Forage Sci*. 1995;50:263–71.
15. Kearns PJ, Shade A. Trait-based patterns of microbial dynamics in dormancy potential and heterotrophic strategy: case studies of resource-based and post-press succession. *ISME J*. 2018;12:2575–81.
16. Nemergut DR, Knelman JE, Ferrenberg S, Bilinski T, Melbourne B, Jiang L, et al. Decreases in average bacterial community rRNA operon copy number during succession. *ISME J*. 2016;10:1147–56.
17. Niederdorfer R, Besemer K, Battin TJ, Peter H. Ecological strategies and metabolic trade-offs of complex environmental biofilms. *Npj Biofilms Microbi*. 2017;3:21.
18. Ortiz-Álvarez R, Fierer N, de los Ríos A, Casamayor EO, Barberán A. Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession. *ISME J*. 2018;12:1658–67.
19. Rivett DW, Scheuerl T, Culbert CT, Mombrikotb SB, Johnstone E, Barraclough TG, et al. Resource-dependent attenuation of species interactions during bacterial succession. *ISME J*. 2016;10:2259–68.
20. Wu LW, Yang YF, Chen S, Shi ZJ, Zhao MX, Zhu ZW, et al. Microbial functional trait of rRNA operon copy numbers increases with organic levels in anaerobic digesters. *ISME J*. 2017;11:2874–8.
21. Hungate BA, Mau RL, Schwartz E, Caporaso JG, Dijkstra P, van Gestel N, et al. Quantitative microbial ecology through stable isotope probing. *Appl Environ Microbiol*. 2015;81:7570–81.
22. Yarza P, Yilmaz P, Pruesse E, Glockner FO, Ludwig W, Schleifer KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol*. 2014;12:635–45.
23. Wu ZT, Koch GW, Dijkstra P, Bowker MA, Hungate BA. Responses of ecosystem carbon cycling to climate change treatments along an elevation gradient. *Ecosystems*. 2011;14:1066–80.
24. Hayer M, Schwartz E, Marks JC, Koch BJ, Morrissey EM, Schuettenberg AA, et al. Identification of growing bacteria during litter decomposition in freshwater through H₂¹⁸O quantitative stable isotope probing. *Environ Microbiol Rep*. 2016;8:975–82.
25. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA*. 2011;108:4516–22.
26. Morrissey EM, Mau RL, Schwartz E, Koch BJ, Hayer M, Hungate BA. Taxonomic patterns in the nitrogen assimilation of soil prokaryotes. *Environ Microbiol*. 2018;20:1112–9.
27. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7:335–6.
28. Louca S, Doebeli M, Parfrey LW. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome*. 2018;6:41.

29. Morrissey EM, Mau RL, Schwartz E, McHugh TA, Dijkstra P, Koch BJ, et al. Bacterial carbon use plasticity, phylogenetic diversity and the priming of soil organic matter. *ISME J.* 2017;11:1890–9.
30. Koch BJ, McHugh TA, Hayer M, Schwartz E, Blazewicz SJ, Dijkstra P, et al. Estimating taxon-specific population dynamics in diverse microbial communities. *Ecosphere.* 2018;9:e02090.
31. Stoddard SF, Smith BJ, Hein R, Roller BRK, Schmidt TM. *rrnDB*: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* 2015;43:D593–8.
32. Lankiewicz TS, Cottrell MT, Kirchman DL. Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. *ISME J.* 2016;10:823–32.
33. Valdivia-Anistro JA, Eguiarte-Frutos LE, Delgado-Sapien G, Marquez-Zacarias P, Gasca-Pineda J, Learned J, et al. Variability of rRNA operon copy number and growth rate dynamics of *Bacillus* isolated from an extremely oligotrophic aquatic ecosystem. *Front Microbiol.* 2015;6:1486.
34. Vieira-Silva S, Rocha EPC. The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genet.* 2010;6:e1000808.
35. Zhu M, Dai X. On the intrinsic constraint of bacterial growth rate: *M. tuberculosis*'s view of the protein translation capacity. *Crit Rev Microbiol.* 2018;44:455–64.
36. Dethlefsen L, Schmidt TM. Performance of the translational apparatus varies with the ecological strategies of bacteria. *J Bacteriol.* 2007;189:3237–45.
37. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. <https://www.R-project.org/>.
38. Muttray AF, Mohn WW. Quantitation of the population size and metabolic activity of a resin acid degrading bacterium in activated sludge using slot-blot hybridization to measure the rRNA:rDNA ratio. *Microb Ecol.* 1999;38:348–57.
39. Rosset R, Julien J, Monier R. Ribonucleic acid composition of bacteria as a function of growth rate. *J Mol Biol.* 1966;18:308–20.
40. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* 2013;7:2061–8.
41. Papp K, Hungate BA, Schwartz E. Microbial rRNA synthesis and growth compared through quantitative stable isotope probing with $H_2^{18}O$. *Appl Environ Microbiol.* 2018; 84, pii: e02441-17.
42. Demoling F, Figueroa D, Bååth E. Comparison of factors limiting bacterial growth in different soils. *Soil Biol Biochem.* 2007;39:2485–95.
43. Chirife J, Herszage L, Joseph A, Kohn ES. In vitro study of bacterial growth inhibition in concentrated sugar solutions: microbiological basis for the use of sugar in treating infected wounds. *Antimicrob Agents Ch.* 1983;23:766–73.
44. Dohi M, Mougé A. A coexistence theory in microbial communities. *R Soc Open Sci.* 2018;5:180476.
45. Sorensen JW, Dunivin TK, Tobin TC, Shade A. Ecological selection for small microbial genomes along a temperate-to-thermal soil gradient. *Nat Microbiol.* 2018;4:55–61.
46. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol.* 2017;15:579–90.
47. Morrissey EM, Mau RL, Schwartz E, Caporaso JG, Dijkstra P, van Gestel N, et al. Phylogenetic organization of bacterial activity. *ISME J.* 2016;10:2336–40.
48. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: Networks, competition, and stability. *Science.* 2015;350:663–6.
49. Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. *Ecology.* 2007;88:1354–64.
50. Ho A, Kerckhof FM, Luke C, Reim A, Krause S, Boon N, et al. Conceptualizing functional traits and ecological characteristics of methane-oxidizing bacteria as life strategies. *Environ Microbiol Rep.* 2013;5:335–45.
51. Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, et al. The worldwide leaf economics spectrum. *Nature.* 2004;428:821–7.