**Genomic traits associated with transcription and *in situ* growth rates of soil bacteria upon rewetting**

**Is this too forward:**

**Genomic traits predict growth and transcription in soil bacteria after rewetting**

**OR**

**Growth of soil bacteria during rewetting is related to genomic traits**

Peter F. Chuckran, Katerina Estera-Molina, Jeff Kimbrell, Alexa M. Nicolas, Ella T. Sieradzki, Alex Greenlon, Paul Dijkstra, Mary K. Firestone, Jennifer Pett-Ridge, Steve J. Blazewicz

ABSTRACT

Genomic traits are informative metrics for assessing life-strategies of soil microorganisms; however, their relationship to growth remains unclear. In soil bacteria, traits such as codon usage bias and genome size have been linked to growth, yet often through indirect measurements or modeling approaches. Here, we utilize metagenome-assembled genomes in combination with metatranscriptomics and 18O-water stable isotope probing to track traits associated with the activity and growth of soil microorganisms over the course of one week following rewetting after a prolonged dry period. Following rewetting, we observed higher growth rates in bacteria with smaller genomes, indicating that reduced genome size may play an important role in bacterial growth in response to sudden change in nutrient availability. However, we found that the predominant trait of fast-growing bacterial taxa was high levels of codon bias in ribosomal protein genes and lower GC content. Similarly, we found that high levels of codon bias corresponded to higher upregulation of ribosomal protein genes. Nucleotides that require less energy to produce were more common at synonymous sites in fast-transcribing taxa, and weakly related to faster growth. Together, these results provide *in situ* evidence for the influence of genomic characteristics on growth rate and transcription, and provide insight into the fitness advantage of these traits in soil bacteria.

INTRODUCTION

Rewetting represents an ecologically important phenomenon where microbial traits associated with growth and activity may play an important role for stabilizing soil carbon (C). In ecosystems characterized by dry and wet seasons, the first rain event following the dry season results in the rapid mineralization of soil carbon (known as the Birch Effect) (1), which can account for a large portion of the annual soil C loss from respiration (2). The rapid stimulation of microbial activity is driven by both the release of water stress as well as an influx in the bioavailability of carbon compounds sourced from osmolytes (3–5), microbial necromass (6, 7), slaking of microaggregates (8), and increased connectivity in the soil matrix (9, 10). The microbial response to rewetting therefore has important consequences for the fate of soil organic carbon; however, we lack an understanding of which traits control the growth and activity of this response.

Trait-based frameworks have long-been used in ecology to help understand complex community structures (11, 12), and this approach has naturally has gained considerable attention in microbial ecology (13, 14), where advancements in sequencing technology have greatly expanded the ability to probe previously unexplored microbial community structures. Traits help uncover the evolutionary forces which dictate when and how traits emerge (15), as well as serve as valuable metrics for assessing community dynamics across scales (16, 17) and with ecosystem function (14). Traits associated with growth are of particular importance, as growth is often central in life-strategy frameworks (14, 18, 19), functional metrics such as carbon use efficiency (20), and understanding changing community dynamics.

Trait-based analyses using genomic data can generally be separated intro two broad categories: functional traits (those which describe specific function) and genomic traits (those which do not describe a specific function; e.g., GC-%, genome size, codon usage, etc.). Functional traits are often observed through functional gene composition and the expression of these genes under a specific set of environmental conditions. During growth, expression of specific genes varies considerably depending on function (21) and environmental constraints (22). Ribosomal genes are highly upregulated during growth, and the number of rRNA gene copies in a genome increases with maximum growth rate. Although the link between rRNA levels and both growth and activity (23) are inconsistent, the tendency for fast-growing bacteria to have a greater number of rRNA gene copies is a useful trait dimension used in several trait-based frameworks in microbial ecology (18, 19, 24). Although specific functional genes have been shown to be associated with growth in soil microbial communities (25), this is often under a specific set of conditions that could favor specific nutrient acquisition strategies or metabolic pathways, and therefore might not be predictive outside of that environment.

However, genomic traits (such as genome size, nucleotide frequency, and codon usage), may provide more generalized metrics that could be compared across systems. For example, codon usage has also been indicated as a strong determinant of both transcription and growth rate. A high alignment of gene codon frequencies to the anticodons of the tRNA pool increases the rate of both transcription and translation (26)—commonly referred to as codon optimization. The level of codon optimization of a transcript has important consequences for translation; affecting the rate of elongation, protein folding, initiation, and termination (27–29). Codon usage is also a strong determinant of mRNA abundance, as a high abundance of optimized codons generally increases mRNA stability (30, 31) and has also been shown to be independently related to higher levels of transcription (32, 33). Codon bias describes codon redundancy in the genetic code for a particular gene or genome, and generally correlates to the degree of codon optimization (34). High levels of codon bias in ribosomal proteins is associated with rapid growth in bacteria (35, 36) and has been increasingly used as a predictor of maximum growth rate (37, 38).

Outside of an intrinsic relationship between GC content and codon usage (26), there appears to be no direct relationship between doubling time and GC content in bacteria (39), and the link between GC content and growth in soil bacteria specifically remains unclear. However, the cost of nucleotide synthesis has been shown to influence transcription, where rapidly transcribed genes often use cheaper nucleotides at synonymous sites to reduce the cost of transcription (40). Since guanine requires more energy to produce than cytosine, and adenine more energy than uracil, genes with cytosine and thymine at synonymous sites tend to be transcribed faster than genes with guanine and adenine (40). Conversely, there is an inverse relationship between nucleotide cost at nonsynonymous sites and the cost of their encoded amino acid sites (40)—such that more expensive nucleotides at non-synonymous sites are associated with higher levels of expression.

Genome size is also thought to influence growth; however, the relationship is less clear than codon usage or rRNA copy number. In oligotrophic marine environments, extremely small genomes are thought to arise in response to nutrient limitation to curb the cost of reproduction (termed genomic streamlining; 30), and often have slower growth rates than copiotrophs with larger genomes (18). Although streamlined genomes have been documented in soils (42), soil bacterial genomes generally tend to be large relative to other systems (43)—a potential result of the increased metabolic diversity required to utilize a complex substrate (44). It’s been hypothesized that these large genomes might come at the expense of growth rate (45), as the increased energy required for reproduction in large genomes may slow growth. However, the evidence regarding the relationship between genome size and growth rate in soil bacteria is inconclusive (46, 47), necessitating further investigation.

In soils, a major hurdle in assessing how traits relate to growth rate has been our inability to effectively measure these traits ­*in situ*. Much of the work linking genomic traits to growth in soils has been conducted using isolate data from databases where the growth rate is often assessed from culture. It has been shown that traits associated with growth of microbes on media might not translate to free-living microbes in their natural environments (47), making it difficult to rely on these trends when assessing the functional potential of soil bacteria. Further, it is unclear how these traits translate to growth during real-world phenomena. Recent methodological and technological advancements have created new opportunities to track the growth of microbes *in situ*. Quantitative stable isotope probing (qSIP) uses the incorporation of added stables isotopes into DNA to track the growth rate of microbes based on sequence (48). This approach has provided valuable insights into metagenomic features and metagenome-assembled genomes (MAGs) associated with growth (7, 25, 49), and shows promise for linking measured microbial traits and growth. In this study, we pair metagenomic, metatranscriptomic, and qSIP data to assess how genomic traits are associated with activity and growth during the rewetting of dry soil at the end of the dry season in a Mediterranean grassland. Specifically, we investigate how traits such as genome size, nucleotide selection, codon usage, and ribosomal protein gene frequency relate to growth and transcription.

METHODS

The field study was located at Hopland Extension and Research Center (HERC), Hopland, California, USA (39.00092, -123.07962), which resides on the ancestral home of the Shóqowa and Hopland people. The region has a Mediterranean climate, with warm dry summers and cool wet winters. The field site was in a field dominated by *Avena barbata* (wild oat grass). This study consisted of 16 plots 3.24 m2 in a randomized block design across X. Rain out shelters were constructed around each plot, either allowing full or 50% precipitation. A full description of the field experimental set-up can be found in *[Rina’s paper]*

Soils were collected in August 2018 before the first rain event at the end of the dry season. At the time of collection the soil gravimetric water content was approximately 3%. Topsoil samples (0-15 cm, roughly 0.5 m3) were taken from 8 randomly selected plots (4 full and 4 reduced precipitation). Samples were transported to Lawrence Livermore National Laboratory where soil from each field plot was individually homogenized and sieved (2 mm) to remove large rocks and roots.

*Wet-up experiment*

Details of the H218O labeling can be found in Nicolas et al. 2023 and Sieradzki et al. 2022 (49, 50). In brief, soils from each of the 3 plots were separated into 12 chambers containing 5 g of soil each, for 48 samples in total. Soils were brought to 22% gravimetric water content either with natural abundance water or with 98-at% H218O. Three replicates each of labeled and unlabeled samples from each plot were destructively harvested at 0, 3, 24, 48, 72, and 168 h post wet-up. Samples were immediately frozen in liquid-N2 and stored in a -80°C freezer.

*Metagenomic qSIP*

A full description of the DNA extraction, sequencing, assembly, binning, and qSIP calculations can be found in Sieradzki et al., 2022. DNA from 3 plots per treatment were extracted in triplicate using a phenol chloroform extraction and then pooled. Samples were then spun on an ultracentrifuge in a cesium chloride solution to create a density gradient. The contents of the ultracentrifuge tube were then separated into 36 fractions, each of which was assessed for DNA density. Fractions were combined into 5 groups along the density gradient, purified and concentrated, and then sequenced on an Illumina Novaseq platform.

Adapters were trimmed and reads were QC filtered using bbduk (51) and Sickle (52). QC filtered reads were assembled into contigs using MEGAHIT (v1.2.9; (53), binned with MaxBin 2.0 (54) and MetaBAT2 (55), and refined with MetaWrap (56). We combined this set of MAGs with another set binned from a study at a nearby site at HERC (25) and dereplicated MAGs using dRep (57). Reads from each fraction were then mapped against the set of MAGs using BBMap (51). The mapping of reads from each fraction was used to calculate atom fraction excess (AFE) for each MAG—which can be used as a metric of growth. DRAM (58) was used to identify open reading frames (ORFs) using Prodigal (59) and annotate functional genes against the KEGG database (60).

*RNA Extraction and sequencing*

Since no isotopic work was conducted with the RNA samples, we selected 4 samples from each precipitation treatment to track RNA transcription at all timepoints. RNA was extracted from soils using a RNeasy PowerSoil Total RNA kit (Qiagen) according to manufacturer instructions. Extracted RNA was treated with RNase-free DNase (Qiagen) and stored at -80°C. RNA concentration was determined using a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA) and quality was assessed using a Nanodrop One Spectrophotometer (Thermo Scientific Invitrogen). Samples were sent to the Joint Genome Institute (JGI; Berkeley, California, USA), for sequencing. Paired-end 2 x 151 bp libraries were sequenced on an Illumina NovaSeq platform. Metatranscriptomes were run through the JGI Integrated Microbial Genomes (IMG) pipeline v.5.1.5 (61) and can be found under the GOLD project ID Gp0612223. IMG assemblies are not included in this analysis, and a more detailed description can be found in the data release (DOI: and citation).

*Metatranscriptome analysis*

Raw reads were downloaded from the JGI genome portal in September 2022. Reads were QC filtered using bbduk (51) and Sickle (52). BBmap was then used to map QC filtered reads to a concatenated reference mag file (minid=0.95). FeatureCounts (62) was then used to extract the number of reads mapped to genes identified by DRAM. Expression and normalization values of mapped transcripts for each annotated gene (excluding rRNA genes) were generated using DESeq2 (63). Nonannotated genes were not included in the normalization, as some of these may represent rRNA genes—the abundance of which would be representative of the efficiency of rRNA depletion, as opposed to any biologically meaningful mechanism.

*Genomic traits*

Genome size was calculated for medium to high quality mags by multiplying the total base pairs assembled by the completeness of the genome. For each gene in each MAG we calculated the nucleotide frequencies and effective number of codons considering the genomic background nucleotide frequency (ENC’), as described by *Novembre 2002* (64). We also calculated the ENC’ of the ribosomal proteins, which was used to calculate ΔENC’ (36):

Where ENC’all is the ENC’ for the whole genome, and ENC’ribo is the ENC’ of the ribosomal protein genes. Higher ΔENC’ values represent greater codon bias in ribosomal protein genes.

Nucleotide frequencies and skews were determined at synonymous and non-synonymous substitution sites on ribosomal protein genes according to nucleotide degeneracy detailed in Chen et al. 2016. Synonymous site nucleotide frequencies were calculated from fourfold degenerative sites, and nonsynonymous frequencies from sites where any substitution would change the encoded amino acid. GC and AT skew were calculated as:

And:

The cost of amino acid synthesis—measured at the total number of phosphate bonds used for synthesis—was derived from Akashi & Gojobori 2002 (65).

*Analysis and model selection*

MAGs were grouped by transcriptional response according to when each MAG was most transcriptionally active, as measured by the total number of upregulated genes. We divided these responses into 4 groups: early responders (most active 3 h post wet-up), middle responders (most active 24, 48, 72 h), late responders (168 h), and sensitive (down-regulated post wet-up). Differences in genomic traits between transcriptional response groups were determined using an analysis of variance (ANOVA). Tukey’s HSD was used for pairwise comparison between groups. Relationships between traits and growth rates were determined using multiple linear regression. Model comparison was conducted using Akaike information criteria (AIC)—with a threshold of -Δ 4 indicating an improvement of a model with the addition of a parameter. All statistical analyses were conducted in R version 4.2.1 (66) and visualized with the ggplot2 package (67).

RESULTS AND DISCUSSION

*Codon bias, GC content, and genome size*

A diagram of ribosome protein

Description automatically generatedA graph of a graph showing a number of red dots

Description automatically generated with medium confidenceWe initially conducted model comparisons of traits most often associated with growth, specifically, codon bias, genome size, and GC content. Atom fraction excess (AFE) values were used as the response variable, with higher levels of isotopic enrichment being used as index of growth. R2 and AIC values indicated that the codon bias of the ribosomal proteins (ENC’ribo) better predicted AFE values than the codon bias of the whole genome (ENC’) or the codon bias of ribosomal proteins relative to the whole genome (ΔENC’). ENC’ribo also predicted growth better than estimated genome size, GC content, or ribosomal GC content (Supplemental Table 1). This demonstrates that codon bias can effectively predict growth not only in pure cultures, as previously established, but also in complex soil microbial communities responding to a natural pulse phenomenon. Although no broad-scale analysis of codon usage bias has been conducted in soils, previous work has found that in soils with lower C, higher pH, and less rainfall, there tends to be a higher redundancy in the codon usage in the bacterial population (68). High codon bias could be a fundamental life-strategy trait for soil communities experiencing transient water and nutrient availability, enabling quicker responses to sudden pulse events. This highlights the importance of codon usage as a critical trait-dimension when considering the adaptation of soil bacteria to changing precipitation patterns in the context of climate change.

**a**

**b**

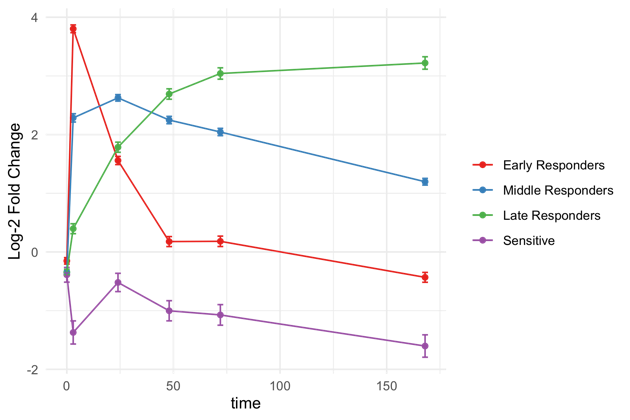
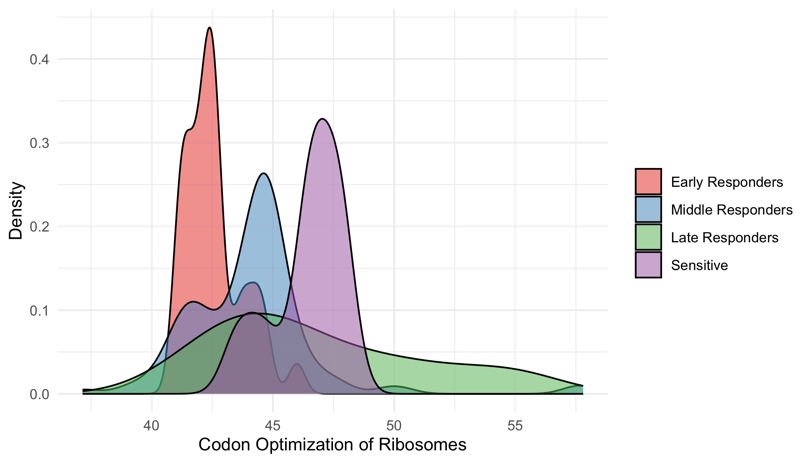
*Figure 1: Atom fraction excess (AFE) values as they related to effective number of codons in ribosomal protein genes (ENC`) and estimated genome size (Mbp) at each time point* ***(A)*** *as well as the GC content of ribosomal protein genes at 72 h post wet-up, with color indicating ENC values* ***(B).*** *Lower ENC values indicate a higher level of codon bias.*

Although ENC’ribo explained the most variance of any single metric, a multiple regression model incorporating ENC’ribo, genome size, and ribosomal GC content performed better than ENC’ribo alone (Supplemental Table 1, R2 = 0.49). Ribosomal GC content was negatively correlated with growth rate when codon bias was high (i.e., lower ENC`; Figure 1b). We believe this could be attributed to the higher metabolic cost of GC base pairs vs AT base pairs. We found that early growth at 24 h was predominantly driven by codon bias and GC content, and that the influence of genome size increased over time (Figure 1a).

Smaller genomes were associated with higher growth rates at later timepoints, particularly 72 and 168 h post wet-up. Smaller genomes may reduce the cost of reproduction, and thus might be key in responding to sudden pulses of nutrients. This idea compliments the hypothesis that large genomes with metabolic diversity come at the expense of growth rate (45), and poses a trade-off where smaller genomes may therefore be better adapted to pulse-driven systems where carbon is limited. This would also explain why smaller genomes are highly prevalent in arid and carbon poor soils (68–70), since the ability to quickly respond to the momentary availability of water and nutrients may be fundamental to survival and persistence.

This contrasts with bacteria in marine systems, where small streamlined genomes tend to have slower growth rates (41). This inconsistency between growth and genome size is potentially a reflection of fundamentally different life-strategies which both leverage genome reduction as a cost-saving measure. In marine systems, reduced cell and genome complexity are a response to a more static nutrient-limited environment which favors high-surface area to volume ratios with less need for rapid growth (41). In this soil system, the lower cost of smaller genomes might allow for faster growth in response to pulses of nutrients entering the environment. In this way, the reduced cost of smaller genomes might serve both strategies well, and the relationship of genome size on growth right might be heavily reliant on life-strategy. Assessing the impact of genome size on growth therefore likely requires putting genome size in the context of other traits associated with growth, such as number of rRNA gene copies or codon bias.

The influence of codon bias was also observed in transcription. We separated transcriptional responses into 4 categories according to their peak transcriptional responses (see methods). Early responding taxa had a higher level of codon bias (Fig 2a), and a greater upregulation of ribosomal protein genes early on (Fig 2b). The peak time of transcription followed a logical pattern with growth response, where early transcribing taxa were represented more during early growth and late transcribing taxa were more often growing later (Supplemental Figure Xa); however, we did not find a strong relationship between the expression of ribosomal protein genes and growth rate (Supplemental Fig Xb). These results demonstrate that although the transcription of ribosomal proteins may be necessary for growth, it cannot be reliably used as a quantitative proxy for growth rate. This aligns with prior research indicating rRNA as a poor metric for activity (23) and highlights important limitations in the use of metatranscriptomics for quantitatively assessing certain metrics between organisms in microbial communities.



a

b

*Figure 2:* ***(A)*** *Codon usage bias of ribosomal protein genes colored by transcriptional response type.* ***(B)*** *Ribosomal protein gene expression over time, averaged by response type*

*Nucleotide selection*

Nucleotide cost was linked to transcriptional response, with early responding organisms displaying a higher AT (A-T/A+T) and GC skew (G-C/G+C) at synonymous sites for ribosomal protein genes (Fig 3 A&B). Our findings are in line with results from Chen et al. (2016) (40), who demonstrated that the lower cost U and C at nonsynonymous sites were associated with increased gene transcription. However, this relationship did not directly correlate with growth rate. GC skew of synonymous substitutions on ribosomal protein genes was weakly negatively correlated to growth (R2 = 0.12, p < 0.01; Fig 3C) and AT skew was not significantly negatively correlated with growth (R2 = 0.01, p = 0.06; Fig 3D). These results underscore the disparity between traits affecting transcription and those having a more pronounced influence on growth.

There was no relationship between nonsynonymous nucleotide usage and transcription (Fig 3 A&B). We found a slight positive relationship between GC skew and growth (Fig. 3E), and a surprisingly positive relationship between AT skew and growth (Fig. 3F, R2 = 0.22, p < 0.01). This could potentially be reflecting the inverse relationship between nucleotide cost and amino acid cost. More expensive nucleotides tend to encode for less expensive amino acids (40), and this relationship could reflect this cost conservation. However, we did not find that amino acid cost was associated with growth (Supplemental Fig X). When we examined the amino acid composition as it relates to these nonsynonymous substitutions, we found that AT skew was most closely correlated with the abundance of lysine (encoded for by AAG and AAA)—yet it is unclear what specifically about lysine would be causing this relationship. Lysine is not notably cheaper to make and has elemental ratios similar to other amino acids. This relationship might therefore be due to features of the ribosomal protein structure, the analysis of which is beyond the scope of this study. Alternatively, the correlation could also be caused by correlations between AT skew and other genomic features, especially GC content (linear regression R2= 0.46). Further, the addition of nonsynonymous AT skew to the model outlined in Supplemental Table 1 did not explain much more variation as opposed to the base model.

A collage of diagrams

Description automatically generated

*Figure 3: Nucleotide skew as it relates to transcriptional responses and growth rate. The synonymous and non-synonymous GC (****A****) and AT* ***(B)*** *skew by transcriptional response, with color indicating site-type. Growth rate as a function of synonymous nucleotide skew for GC-skew* ***(C)*** *and AT skew* ***(D)****, and nonsynonymous GC* ***(E)*** *and AT skew* ***(F)****.*

*Limitations and future directions*

Codon bias, GC content, and genome size may be particularly relevant in pulse-driven systems, but it is still unknown how they may relate to growth in soils where water and nutrients are either less variable or more available. For example, the frequency of cycles of drying and rewetting are known to influence microbial community responses to rewetting (71). Although this may be tied to shifts in community composition, it does not preclude plasticity in the physiological response of individual taxa. More or less frequent rain events could potentially influence the response of these traits to rewetting events.

How genomic traits relate to growth in environments where change is gradual is also not well understood. A complete picture of the relationship between genomic traits and activity of soil bacteria requires assessing these relationships across environments and stimuli. Although the growth rates of soil bacteria under stable conditions is often difficult to measure—recent advancements in isotopic vapor labeling allow for the study the growth rates of bacteria without the addition of water or nutrients (72, 73) and provides a promising avenue to study growth in undisturbed environments.

It is essential to underscore the potential of these traits for predicting growth and transcription. Even though the most predictive traits could account for approximately 50% of the variance in growth, it's important to acknowledge the limitations in terms of sample size and the potential for bias resulting from metagenomic binning. Consequently, the extent of predictive power these traits possess for forecasting growth in response to real-world phenomena remains uncertain; additional studies and data will offer more clarity on these relationships. In light of this, it's crucial to recognize that these traits may not consistently correlate with growth in non-pulse-driven systems. These considerations should guide us when making conclusions about the growth potential of microbial communities.

*Conclusion*

These results highlight trait dimensions that correspond with growth rate and bacteria in soil microbial communities. Specifically, we found that fast growth was related to high codon usage bias, lower GC content, and smaller genome sizes. We also found that transcription was associated with nucleotide synthesis cost as well as codon usage bias. These relationships offer important perspective for examining the broad scale ecological distribution of these traits, as the importance of short-term responses may play a vital role in systems characterized by pulses of nutrient availability.

REFERENCES

1. H. F. Birch, The effect of soil drying on humus decomposition and nitrogen availability. *Plant and Soil* **10**, 9–31 (1958).

2. P. Jarvis, *et al.*, Drying and wetting of Mediterranean soils stimulates decomposition and carbon dioxide emission: the “Birch effect.” *Tree Physiology* **27**, 929–940 (2007).

3. N. Fierer, J. P. Schimel, A Proposed Mechanism for the Pulse in Carbon Dioxide Production Commonly Observed Following the Rapid Rewetting of a Dry Soil. *Soil Science Society of America Journal* **67**, 798 (2003).

4. E. W. Slessarev, *et al.*, Cellular and extracellular C contributions to respiration after wetting dry soil. *Biogeochemistry* **147**, 307–324 (2020).

5. C. R. Warren, Response of osmolytes in soil to drying and rewetting. *Soil Biology and Biochemistry* **70**, 22–32 (2014).

6. S. J. Blazewicz, E. Schwartz, M. K. Firestone, Growth and death of bacteria and fungi underlie rainfall-induced carbon dioxide pulses from seasonally dried soil. *Ecology* **95**, 1162–1172 (2014).

7. A. M. Nicolas, *et al.*, Isotope-enrichment reveals active viruses follow microbial host dynamics during rewetting of a California grassland soil. 2022.09.30.510406 (2022).

8. P. M. Homyak, *et al.*, Effects of altered dry season length and plant inputs on soluble soil carbon. *Ecology* **99**, 2348–2362 (2018).

9. F. E. Moyano, S. Manzoni, C. Chenu, Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. *Soil Biology and Biochemistry* **59**, 72–85 (2013).

10. A. P. Smith, *et al.*, Shifts in pore connectivity from precipitation versus groundwater rewetting increases soil carbon loss after drought. *Nature Communications* **8** (2017).

11. J. P. Grime, Evidence for the Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory. *The American Naturalist* **111**, 1169–1194 (1977).

12. C. Violle, *et al.*, Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).

13. M. Westoby, *et al.*, Trait dimensions in bacteria and archaea compared to vascular plants. *Ecology Letters* **24**, 1487–1504 (2021).

14. A. A. Malik, *et al.*, Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. *ISME Journal* **14**, 1–9 (2020).

15. J. B. H. Martiny, S. E. Jones, J. T. Lennon, A. C. Martiny, Microbiomes in light of traits: A phylogenetic perspective. *Science* **350** (2015).

16. A. Barberán, A. Fernández-Guerra, B. J. M. Bohannan, E. O. Casamayor, Exploration of community traits as ecological markers in microbial metagenomes. *Molecular Ecology* **21**, 1909–1917 (2012).

17. J. Wan, T. W. Crowther, Uniting the scales of microbial biogeochemistry with trait-based modelling. *Functional Ecology* **36**, 1457–1472 (2022).

18. F. M. Lauro, *et al.*, The genomic basis of trophic strategy in marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15527–33 (2009).

19. N. Fierer, Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology* **15**, 579–590 (2017).

20. R. L. Sinsabaugh, S. Manzoni, D. L. Moorhead, A. Richter, Carbon use efficiency of microbial communities: Stoichiometry, methodology and modelling. *Ecology Letters* **16**, 930–939 (2013).

21. S. Klumpp, T. Hwa, Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology* **28**, 96–102 (2014).

22. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* **330**, 1099–1102 (2010).

23. S. J. Blazewicz, R. L. Barnard, R. A. Daly, M. K. Firestone, Evaluating rRNA as an indicator of microbial activity in environmental communities: Limitations and uses. *ISME Journal* **7**, 2061–2068 (2013).

24. S. Krause, *et al.*, Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Frontiers in Microbiology* **5**, 251 (2014).

25. A. Greenlon, *et al.*, Quantitative Stable-Isotope Probing (qSIP) with Metagenomics Links Microbial Physiology and Activity to Soil Moisture in Mediterranean-Climate Grassland Ecosystems. *mSystems* **7**, e00417-22 (2022).

26. J. B. Plotkin, G. Kudla, Synonymous but not the same: The causes and consequences of codon bias. *Nature Reviews Genetics* **12**, 32–42 (2011).

27. Y. Liu, Q. Yang, F. Zhao, Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annual Review of Biochemistry* **90**, 375–401 (2021).

28. C. H. Yu, *et al.*, Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Molecular Cell* **59**, 744–754 (2015).

29. M. Zhou, *et al.*, Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature 2013 495:7439* **495**, 111–115 (2013).

30. C. Dressaire, *et al.*, Role of mRNA Stability during Bacterial Adaptation. *PLOS ONE* **8**, e59059 (2013).

31. V. Presnyak, *et al.*, Codon Optimality Is a Major Determinant of mRNA Stability. *Cell* **160**, 1111–1124 (2015).

32. Z. R. Newman, J. M. Young, N. T. Ingolia, G. M. Barton, Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E1362–E1371 (2016).

33. Z. Zhou, *et al.*, Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E6117–E6125 (2016).

34. S. Bahiri-Elitzur, T. Tuller, Codon-based indices for modeling gene expression and transcript evolution. *Computational and Structural Biotechnology Journal* **19**, 2646–2663 (2021).

35. D. L. Kirchman, Growth rates of microbes in the oceans. *Annual Review of Marine Science* **8**, 285–309 (2016).

36. S. Vieira-Silva, E. P. C. Rocha, The Systemic Imprint of Growth and Its Uses in Ecological (Meta)Genomics. *PLoS Genetics* **6**, e1000808 (2010).

37. A. M. Long, S. Hou, • J Cesar, I. -Espinoza, J. A. Fuhrman, Benchmarking microbial growth rate predictions from metagenomes. *The ISME Journal*, 1–13 (2020).

38. J. L. Weissman, S. Hou, J. A. Fuhrman, Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. *Proceedings of the National Academy of Sciences* **118**, e2016810118 (2021).

39. L. Aliperti, *et al.*, r/K selection of GC content in prokaryotes. *Environmental Microbiology* **n/a** (2023).

40. W.-H. Chen, G. Lu, P. Bork, S. Hu, M. J. Lercher, Energy efficiency trade-offs drive nucleotide usage in transcribed regions. *Nature Communications* **7**, 11334 (2016).

41. S. J. Giovannoni, J. Cameron Thrash, B. Temperton, Implications of streamlining theory for microbial ecology. *The ISME Journal* **8**, 1553–1565 (2014).

42. T. E. Brewer, K. M. Handley, P. Carini, J. A. Gilbert, N. Fierer, Genome reduction in an abundant and ubiquitous soil bacterium ‘Candidatus Udaeobacter copiosus.’ *Nature Microbiology* **2**, 16198 (2017).

43. M. Cobo-Simón, J. Tamames, Relating genomic characteristics to environmental preferences and ubiquity in different microbial taxa. *BMC Genomics* **18**, 499 (2017).

44. A. Barberán, *et al.*, Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecology Letters* **17**, 794–802 (2014).

45. K. T. Konstantinidis, J. M. Tiedje, Trends between gene content and genome size in prokaryotic species with larger genomes. **10**, 3160–3165 (2004).

46. M. Westoby, *et al.*, Cell size, genome size, and maximum growth rate are near-independent dimensions of ecological variation across bacteria and archaea. *Ecology and Evolution* **11**, 3956–3976 (2021).

47. J. Li, *et al.*, Predictive genomic traits for bacterial growth in culture versus actual growth in soil. *The ISME Journal*, 1 (2019).

48. B. A. Hungate, *et al.*, Quantitative microbial ecology through stable isotope probing. *Applied and Environmental Microbiology* **81**, 7570–7581 (2015).

49. E. T. Sieradzki, *et al.*, Functional succession of actively growing soil microorganisms during rewetting is shaped by precipitation history. *bioRxiv*, 2022.06.28.498032 (2022).

50. A. M. Nicolas, *et al.*, A subset of viruses thrives following microbial resuscitation during rewetting of a seasonally dry California grassland soil. *Nat Commun* **14**, 5835 (2023).

51. B. Bushnell, BBTools Software Package (2014).

52. Joshi NA, Fass JN., Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) (2023) (May 11, 2023).

53. D. Li, C. M. Liu, R. Luo, K. Sadakane, T. W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

54. Y.-W. Wu, B. A. Simmons, S. W. Singer, MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2016).

55. D. D. Kang, *et al.*, MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).

56. G. V. Uritskiy, J. DiRuggiero, J. Taylor, MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).

57. M. R. Olm, C. T. Brown, B. Brooks, J. F. Banfield, dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* **11**, 2864–2868 (2017).

58. M. Shaffer, *et al.*, DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research* **48**, 8883–8900 (2020).

59. D. Hyatt, *et al.*, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 1–11 (2010).

60. M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45**, D353–D361 (2017).

61. I.-M. A. Chen, *et al.*, IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research* **47**, D666–D677 (2019).

62. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

63. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15** (2014).

64. J. A. Novembre, Accounting for Background Nucleotide Composition When Measuring Codon Usage Bias. *Molecular Biology and Evolution* **19**, 1390–1394 (2002).

65. H. Akashi, T. Gojobori, Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 3695–3700 (2002).

66. R. C. Team, R: A language and environment for statistical computing. *R Foundation for Statistical Computing. Vienna, Austria.* (2018).

67. H. Wickham, “Elegant Graphics for Data Analysis” in *Elegant Graphics for Data Analysis*, (2016), pp. 3–10.

68. P. F. Chuckran, *et al.*, Edaphic controls on genome size and GC content of bacteria in soil microbial communities. *Soil Biology and Biochemistry* **178**, 108935 (2023).

69. H. Liu, *et al.*, Warmer and drier ecosystems select for smaller bacterial genomes in global soils. *iMeta* **2**, e70 (2023).

70. G. Piton, *et al.*, Life history strategies of soil bacterial communities across global terrestrial biomes. *Nat Microbiol*, 1–10 (2023).

71. N. Fierer, J. P. Schimel, P. A. Holden, Influence of Drying-Rewetting Frequency on Soil Bacterial Community Structure. *Microbial Ecology* **45**, 63–71 (2003).

72. A. Canarini, *et al.*, Quantifying microbial growth and carbon use efficiency in dry soil environments via 18O water vapor equilibration. *Global Change Biology* **26**, 5333–5341 (2020).

73. D. Metze, *et al.*, Microbial growth under drought is confined to distinct taxa and modified by potential future climate conditions. *Nat Commun* **14**, 5895 (2023).