The macroecological dynamics of sojourn trajectories in the human gut microbiome

William R. Shoemaker[1,*] and Jacopo Grilli[1]

1 Quantitative Life Sciences, The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, 34151, Italy.

*** Contact:** `williamrshoemaker@gmail.com`

Abstract

The human gut microbiome is a dynamic ecosystem. Host behaviors (e.g., diet) provide a regular source of environmental variation that induces fluctuations in the abundances of resident microbiota. Despite these displacements, microbial community members remain highly resilient. Population abundances tend to fluctuate around characteristic steady-state abundances over long timescales in healthy human hosts. These temporary excursions from steady-state abundances, known as *sojourn trajectories*, have the potential to inform our understanding of the fundamental dynamics of the microbiome. However, to our knowledge, the macroecology of sojourn trajectories has yet to be systematically characterized. In this study, we leverage theoretical tools from the study of random walks to characterize the duration of sojourn trajectories, their shape, and the degree that diverse community members exhibit similar qualitative and quantitative dynamics. We apply the Stochastic Logistic Model as a theoretical lens for interpreting our empirical observations. We find that the typical timescale of a sojourn trajectory does not depend on the mean abundance of a community member (i.e., carrying capacity), although it is strongly related to its coefficient of variation (i.e., environmental noise). This work provides fundamental insight into the dynamics, timescales, and fluctuations exhibited by diverse microbial communities.

Importance

Microorganisms in the human gut often fluctuate around a characteristic abundance in healthy hosts over extended periods of time. These typical abundances can be viewed as steady-states, meaning that fluctuating abundances do not continue towards extinction or dominance, but rather return to a specific value over a typical timescale. Here we empirically characterize the 1) length (i.e., number of days), 2) relationship between length and area, and 3) typical deviation of a sojourn trajectory. These three patterns can be explained and unified through an established minimal model of ecological dynamics, the Stochastic Logistic Model of growth.

Main

Microbial communities in the human gut often maintain steady-state abundances over extended timescales. This tendency has been the subject of observational macroecological investigations [1] as well as experimental efforts imposing external

perturbations [2,3]. However, we lack a quantitative understanding of the dynamics of microorganisms fluctuating around their steady-state. Excursions away from, and subsequent return to, said steady-state values, known as *sojourn trajectories*, have been systematically studied for various stochastic models [4,5] and quantitatively examined in neurological [6] and geophysical systems [7]. Such an approach may also aid our understanding of microbial community dynamics.

Here we characterize the sojourn trajectories of microbial communities in the human gut, investigating the duration of sojourn trajectories ($\mathcal{T}$) and how this quantity relates to the typical deviation from steady-state abundance. We interpret our findings and generate quantitative predictions using the Stochastic Logistic Model of growth (SLM), a minimal ecological model where species dynamics are governed by self-limiting growth and environmental noise. This choice of this model is reasonable as it has been successfully applied to investigate the macroecology of microbial communities within and across human hosts [8,9], community-level measures of ecological distances and dissimilarities [10], the temporal dynamics of alternative stable-states [1], experimental communities [11], dynamics within and patterns across human hosts at the level of ecological strains [12,13], the macroecology of microbial communities across taxonomic and phylogenetic scales [14], and the dependency between correlations in abundance between community members and phylogenetic distance [15].

We identified and re-processed time-resolved human gut microbiome 16S rRNA datasets from healthy hosts with daily or near-daily sampling [16–19] (Supplemental Material; Table S1). We inferred the mean and CV of each ASV using sampling-aware maximum likelihood (Supplemental Material; Fig. S1). As typical abundances can vary by orders of magnitude in microbial communities, abundances were rescaled by their time-averaged mean and log-transformed, $y_i(t) \equiv \ln\left(\frac{x_i(t)}{\bar{x}_i}\right)$ (Fig. S2).

We consider the SLM as a lens to investigate empirical sojourn patterns, defined as:

$$\frac{dx_i}{dt} = \frac{x_i}{\tau_i}\left(1 - \frac{x_i}{K_i}\right) + \sqrt{\frac{\sigma_i}{\tau_i}} x_i \cdot \eta_i(t) \ . \tag{1}$$

The SLM contains three parameters: 1) the timescale τ_i (inverse of the maximum growth rate), 2) the carrying capacity K_i, and 3) the strength of environmental noise σ_i. The term $\eta_i(t)$ introduces stochasticity into the equation as white noise: the expected value of is $\langle \eta_i(t) \rangle = 0$ and the time correlation is a delta function $\langle \eta_i(t)\eta_i(t') \rangle = \delta(t - t')$ (i.e., noise at time t is uncorrelated with the noise at time t').

We first investigated sojourn trajectories as consecutive observations where a community member deviated from and proceeded to return to its steady-state abundance (Fig. 1a). By calculating sojourn times for all ASVs in all datasets, we were able to obtain an empirical probability distribution of $\mathcal{T}$ (Fig. 1b). We first compared the empirical distribution to those obtained under two null models that both reflect the absence of temporal dynamics contributing to $\mathcal{T}$. Neither the null distribution obtained by 1) independently drawing observations from the gamma distribution (stationary distribution of the SLM) nor 2) permuting time labels in the data reproduced the empirical distribution (Figs. 1b; S3). In contrast, the time-dependent solution of the SLM (function of τ) captured the bulk of the empirical distribution, though we note the existence of a few high values of $\mathcal{T}$ that were not captured by the SLM. To assess whether $\mathcal{T}$ contains greater temporal information than previously investigated macroecological patterns, we examined the length of time that ASVs were consistently observed of unobserved, quantities known as the *residence* and *return* times that have been the subject of prior macroecological investigations [20–22]. For each quantity we calculated the divergence between the empirical and time-permuted null distributions, finding that $\mathcal{T}$ displayed greater divergence than the other two quantities in all hosts (Supplemental Material; Figs. S4-S7). These results suggest that sojourn times have

greater sensitivity to underlying temporal dynamics. We note that ASVs that are periodically absent (i.e., those used to calculate residence and return times) may reflect dynamics where abundances occasionally and rapidly become highly abundant (i.e., Conditionally Rare Taxa) [23]. For ASVs consistently observed throughout a given timeseries (i.e., those used to calculate sojourn times) we note that periods of atypically large abundances would be interpreted as sojourn trajectories.

We then investigated how $\mathcal{T}$ as a measure of length relates to the area under a sojourn trajectory ($\mathcal{A}(\mathcal{T})$; Fig. 1a), a dependency that has been previously characterized as a scaling relationship [4, 5] (Fig. 1a; Supplemental Material). This relationship determines whether and how different trajectories must be rescaled to have the same shape. In empirical data, we found that the area of a sojourn trajectory was effectively independent of its sojourn time ($\mathcal{A}(\mathcal{T}) \to \mathcal{A}$; Fig. 1c). The absence of a relationship between the two quantities was predicted by the SLM when the timescale of growth is much smaller than the typical sojourn time ($\tau \ll \mathcal{T}$; Supplemental Material). Motivated by this result, we tested the prediction that under this parameter regime $\mathcal{A}$ would scale with the CV of an ASV (a function of σ_i), finding that the exponent predicted by the SLM sufficiently captured the empirical trend (inset of Fig. 1c).

The fact that $\mathcal{A}$ is independent of $\mathcal{T}$ informs our understanding of the typical deviation from the steady-state *within* a sojourn trajectory, $\langle y_i(t) - \bar{y}_i \rangle_{\mathcal{T}}$. Prior theoretical efforts have found that a wide range of stochastic dynamics can be captured by a scaling relationship that depends on $\mathcal{T}^{\alpha}$ for a given exponent α (Supplemental Material) [4, 5]. The absence of a relationship in Fig. 1c implies that $\alpha = 0$, meaning that sojourn deviations should exhibit similar trajectories for different values of $\mathcal{T}$ *without* additional rescaling. This prediction held, as empirical trajectories $\langle y_i(t) - \bar{y}_i \rangle$ did not systematically change with increasing $\mathcal{T}$ (Figs. 1d; S8).

Up to this point, we have investigated the similarities in the sojourn dynamics of microbial community members but have neglected how variation in fluctuations can shape typical sojourn times. Therefore, we examined the relationship between the mean sojourn time (i.e., $\langle \mathcal{T} \rangle$) and both the mean abundance (function of K_i and σ_i) and the coefficient of variation of abundance (solely a function of σ_i) (Fig. 2a). We focused on the dataset with the highest temporal resolution in order to estimate sojourn times $\mathcal{T}$ of the lowest possible values (Table S1; [17]). The SLM predicted the absence of a relationship between $\langle \mathcal{T} \rangle$ and mean abundance, which held in the data (Supplemental Material; Figs. 2b, S9). The SLM also predicted the existence of a relationship between $\langle \mathcal{T} \rangle$ and the CV, which was also recapitulated in the data (Fig. 2c). These results indicate that the typical sojourn time in the gut is highly dependent on the CV, which is interpreted under the SLM as the strength of environmental noise that a community member experiences.

Building on prior investigations into the temporal dynamics of microbial communities [9, 20, 21, 24, 25], this study focused on the macroecology of excursions from steady-state abundances by leveraging a theoretical framework from statistical physics [4, 5]. We found that while the timescale of growth was necessary to recapitulate the distribution of sojourn times, sojourn trajectories were effectively scale-free with respect to sojourn time. Similarly, the mean sojourn time of an ASV could be predicted by the CV. These results can be interpreted under the SLM as excursions from the steady-state reflecting the strength of environmental noise.

Due to the available data, this study focused on healthy hosts, though it would be of interest to extend our analyses to the scenario where microbiomes are effectively out of steady-state as such perturbations have the potential to impact stability [26]. We note that our predictions were obtained from the SLM in a parameter limit where the model reduced to a form amenable to analytic solutions (Supplemental Material). This detail implies that while the SLM remains a viable minimal model for interpreting empirical

microbial ecological dynamics, it may not be the sole model capable of explaining our observations. We emphasize that here the length of time between observations (δt) is similar to the intrinsic timescale of the system (i.e., the timescale of growth) due to the inherent limitations of fecal sample collection, in contrast to other systems where sojourn times have been investigated [6, 7]. Alternative sojourn dynamics are likely to be found in community sampling scenarios where the timescale of growth exceeds that of sampling ($\tau \gg \delta t$), though models such as the SLM may continue to provide useful interpretations.

Data and code availability

Public raw sequence data used in this study were obtained from previous studies (Table S1). Data processed for this study are available on Zenodo DOI: `10.5281/zenodo.17809697`. All code is available on GitHub under a GNU General Public License: `sojourn_macroeco`

Acknowledgments

We thank O. Mazzarisi for their feedback on this manuscript. This work was supported by Fondo Italiano per la Scienza - FIS (CUP J53C23002290001; J.G.). Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Conceptual diagrams created in BioRender under a CC BY 4.0 license.

Author contributions

W.R.S. and J.G. conceptualized the project, performed the derivations, and wrote the manuscript. W.R.S. performed all analyses.

References

1. Zaoli S, Grilli J. A macroecological description of alternative stable states reproduces intra- and inter-host variability of gut microbiome. Science Advances. 2021 Oct;7(43):eabj2882. Available from: `https://www.science.org/doi/10.1126/sciadv.abj2882`.

2. Ng KM, Aranda-Díaz A, Tropini C, Frankel MR, Van Treuren W, O'Loughlin CT, et al. Recovery of the Gut Microbiota after Antibiotics Depends on Host Diet, Community Context, and Environmental Reservoirs. Cell Host & Microbe. 2019 Nov;26(5):650–665.e4.

3. Xue KS, Walton SJ, Goldman DA, Morrison ML, Verster AJ, Parrott AB, et al. Prolonged delays in human microbiota transmission after a controlled antibiotic perturbation. bioRxiv: The Preprint Server for Biology. 2023 Oct:2023.09.26.559480.

4. Baldassarri A, Colaiori F, Castellano C. Average Shape of a Fluctuation: Universality in Excursions of Stochastic Processes. Physical Review Letters. 2003 Feb;90(6):060601. Available from: `https://link.aps.org/doi/10.1103/PhysRevLett.90.060601`.

5. Colaiori F, Baldassarri A, Castellano C. Average trajectory of returning walks. Physical Review E. 2004 Apr;69(4):041105. Available from: https://link.aps.org/doi/10.1103/PhysRevE.69.041105.

6. Dutta S, Iyer KK, Vanhatalo S, Breakspear M, Roberts JA. Mechanisms underlying pathological cortical bursts during metabolic depletion. Nature Communications. 2023 Aug;14(1):4792. Available from: https://www.nature.com/articles/s41467-023-40437-0.

7. Consolini G, De Michelis P, Tozzi R. On the Earth's magnetospheric dynamics: Nonequilibrium evolution and the fluctuation theorem. Journal of Geophysical Research: Space Physics. 2008;113(A8). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1029/2008JA013074.

8. Grilli J. Macroecological laws describe variation and diversity in microbial communities. Nature Communications. 2020 Sep;11(1):4743. Available from: https://www.nature.com/articles/s41467-020-18529-y.

9. Descheemaeker L, de Buyl S. Stochastic logistic models reproduce experimental time series of microbial communities. eLife. 2020 Jul;9:e55650. Available from: https://doi.org/10.7554/eLife.55650.

10. Zaoli S, Grilli J. The stochastic logistic model with correlated carrying capacities reproduces beta-diversity metrics of microbial communities. PLOS Computational Biology. 2022 Apr;18(4):e1010043. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010043.

11. Shoemaker WR, Sánchez Grilli J. Macroecological patterns in experimental microbial communities. PLOS Computational Biology. 2025 May;21(5):e1013044. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1013044.

12. Shoemaker WR. A macroecological perspective on genetic diversity in the human gut microbiome. PLOS ONE. 2023 Jul;18(7):e0288926. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0288926.

13. Wolff R, Shoemaker W, Garud N. Ecological Stability Emerges at the Level of Strains in the Human Gut Microbiome. mBio. 2023 Feb;14(2):e02502–22. Available from: https://journals.asm.org/doi/10.1128/mbio.02502-22.

14. Shoemaker WR, Grilli J. Macroecological patterns in coarse-grained microbial communities. eLife. 2023 Aug;12. Available from: https://elifesciences.org/reviewed-preprints/89650.

15. Sireci M, Muñoz MA, Grilli J. Environmental fluctuations explain the universal decay of species-abundance correlations with phylogenetic distance. Proceedings of the National Academy of Sciences. 2023 Sep;120(37):e2217144120. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2217144120.

16. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, et al. Moving pictures of the human microbiome. Genome Biology. 2011 May;12(5):R50. Available from: https://doi.org/10.1186/gb-2011-12-5-r50.

17. David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, et al. Host lifestyle affects human microbiota on daily timescales. Genome Biology. 2014;15(7):R89.

18. Poyet M, Groussin M, Gibbons SM, Avila-Pacheco J, Jiang X, Kearney SM, et al. A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. Nature Medicine. 2019 Sep;25(9):1442–1452.

19. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods. 2016 Jul;13(7):581–583. Available from: https://www.nature.com/articles/nmeth.3869.

20. Ji BW, Sheth RU, Dixit PD, Tchourine K, Vitkup D. Macroecological dynamics of gut microbiota. Nature Microbiology. 2020 May;5(5):768–775. Available from: https://www.nature.com/articles/s41564-020-0685-1.

21. Ho PY, Good BH, Huang KC. Competition for fluctuating resources reproduces statistics of species abundance over time across wide-ranging microbiotas. eLife. 2022 Apr;11:e75168. Available from: https://doi.org/10.7554/eLife.75168.

22. Wang XW, Liu YY. Characterizing scaling laws in gut microbial dynamics from time series data: caution is warranted. bioRxiv; 2021. Pages: 2021.01.11.426045 Section: Contradictory Results. Available from: https://www.biorxiv.org/content/10.1101/2021.01.11.426045v1.

23. Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, et al. Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. mBio. 2014 Jul;5(4):10.1128/mbio.01371–14. Publisher: American Society for Microbiology. Available from: https://journals.asm.org/doi/10.1128/mbio.01371-14.

24. George AB, O'Dwyer J. Universal abundance fluctuations across microbial communities, tropical forests, and urban populations. Proceedings of the National Academy of Sciences. 2023 Oct;120(44):e2215832120. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.2215832120.

25. Wang XW, Liu YY. Origins of scaling laws in microbial dynamics. Physical Review Research. 2023 Jan;5(1):013004. Available from: https://link.aps.org/doi/10.1103/PhysRevResearch.5.013004.

26. Dean AM, Shnerb NM. Stochasticity-induced stabilization in ecology and evolution: a new synthesis. Ecology. 2020 Sep;101(9):e03098. Available from: https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.3098.

Figures

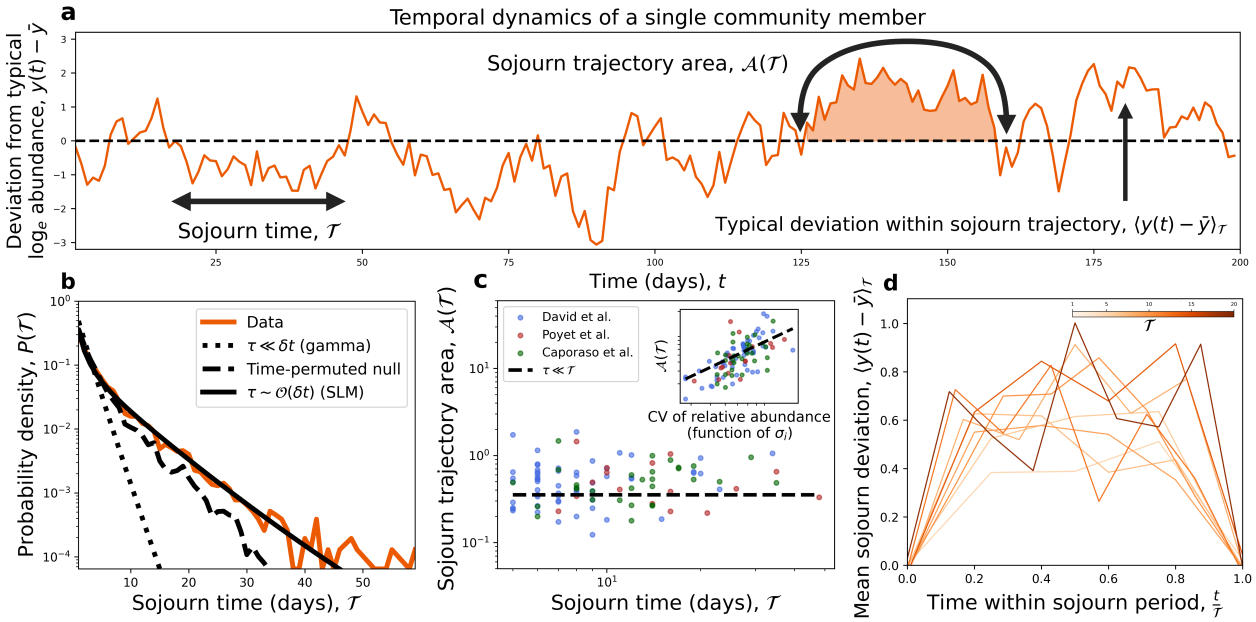

Figure 1. Patterns of sojourn trajectories inform our understanding of microbial community dynamics. a) We investigated the macroecology of sojourn trajectories using log-transformed relative abundances that have been rescaled by their mean ($y_i(t) \equiv \ln\left(\frac{x_i(t)}{\bar{x}_i}\right)$). Consecutive fluctuations of $y_i(t)$ either above or below the time-averaged mean were investigated, focusing on three previously identified patterns [4, 5]: 1) the distribution of sojourn times $P(\mathcal{T})$, 2) the relationship between $\mathcal{T}$ and the area under a sojourn trajectory ($\mathcal{A}$), and 3) how the mean deviation *within* a sojourn trajectory depends on $\mathcal{T}$. **b)** The empirical $P(\mathcal{T})$ could not be captured by null distributions obtained by 1) assuming abundances are stationary (i.e., independent draws from a gamma distribution) nor 2) permuting time labels. However, the distribution predicted by the SLM captured the bulk of the distribution ($\tau = 4$; Eq. S23). **c)** We observed a lack of dependency between $\mathcal{T}$ and $\mathcal{A}$ across datasets, but a consistent relationship between the inverse strength of environmental noise and $\mathcal{A}$, results that are predicted by the SLM under the parameter limit where the growth rate is typically smaller than $\mathcal{T}$. Each dot represents a single sojourn trajectory from a given ASV. **d)** Given that $\mathcal{A}$ is a constant, we predicted that the mean sojourn deviation would *not* vary with $\mathcal{T}$, which was observed in the data. The mean deviation was estimated by pooling over ASVs and sojourn trajectories with the same value $\mathcal{T}$. Predictions were obtained using ASV-specific values of σ_i.

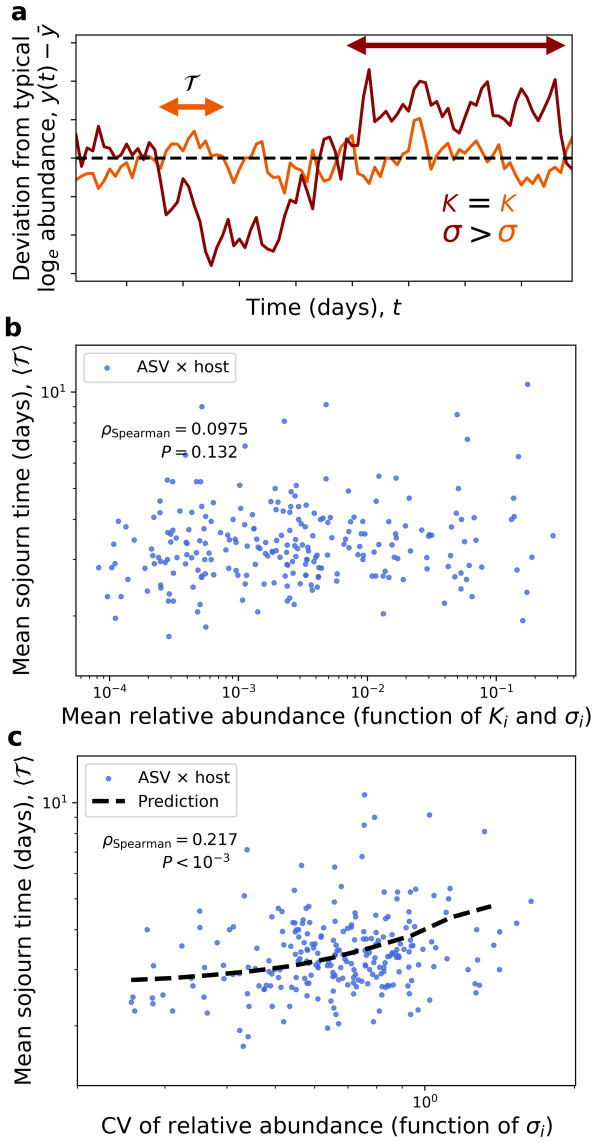

Figure 2. Mean sojourn times are independent of mean abundance but dependent on the CV. a) The SLM predicted that the mean sojourn time (i.e., $\langle \mathcal{T} \rangle$) of a community member would be independent of its mean relative abundance (i.e., K_i) but would increase with its CV (i.e., higher σ_i). **b)** Consistent with our predictions, we found that mean relative abundance $\bar{x}_i$ (function of K_i and σ_i) had no relationship with $\langle \mathcal{T} \rangle$. This absence can be interpreted as a lack of dependency of sojourn trajectories on the carrying capacity. **c)** For the CV (solely a function of σ_i), a relationship existed which could be captured using predictions derived from the SLM. Spearman's rank correlation coefficient was used to assess the statistical significance of each relationship and predictions were obtained using ASV-specific values of σ_i.