Supplementary Material

Exploring the Determinants of Organic Matter Bioavailability through Substrate-Explicit Thermodynamic Modeling

Firnaaz Ahamed[1], Yaqi You[2], Amy Burgin[3], James C. Stegen[4,5], Timothy D. Scheibe[4], Hyun-Seob Song[1,6,*]

[1]Department of Biological Systems Engineering, University of Nebraska–Lincoln, Lincoln, NE, United States

[2]Department of Environmental Resources Engineering, SUNY College of Environmental Science and Forestry, Syracuse, NY, United States

[3]Departments of Ecology & Evolutionary Biology, and Kansas Biological Survey-Center for Ecological Research, University of Kansas, Lawrence, KS, United States

[4]Pacific Northwest National Laboratory, Richland, WA, United States

[5] School of the Environment, Washington State University, Pullman, WA, United States

[6]Department of Food Science and Technology, Nebraska Food for Health Center, University of Nebraska–Lincoln, Lincoln, NE, United States

*Correspondence: Hyun-Seob Song, hsong5@unl.edu

Supplementary Texts

1 Derivation of balanced C/N requirement of microbiome for growth and maintenance

Threshold element ratio (TER) is the balanced requirement of C/N ratio for growth and maintenance of microbiome evaluated based on the carbon and nitrogen use efficiencies (CUE and NUE), as well as the assumed biomass composition. In this work, we assume the biomass elemental compositions as $CH_{1.8}N_{0.2}O_{0.5}$ (Stephanopoulos et al., 1998; Kleerebezem and Van Loosdrecht, 2010). Accordingly, TER, CUE and NUE for the oxidative respiration of organic matter (OM) are given as:

$$\mathrm{TER}_i = \left(\frac{C_{\mathrm{biom}}}{N_{\mathrm{biom}}}\right) \times \frac{\mathrm{NUE}_i}{\mathrm{CUE}_i}; \quad \mathrm{CUE}_i = \frac{y_{\mathrm{biom},i} \times C_{\mathrm{biom}}}{\left|y_{\mathrm{OC},i}\right| \times C_{\mathrm{OC},i}}; \quad \mathrm{NUE}_i = \frac{y_{\mathrm{biom},i} \times N_{\mathrm{biom}}}{\left|y_{\mathrm{OC},i}\right| \times N_{\mathrm{OC},i} + \left|y_{\mathrm{NH}_4,i}\right| \times 1} \qquad \text{(S1)}$$

Here, C_{biom} and N_{biom} are the number of C and N in the biomass, $C_{\mathrm{OC},i}$ and $N_{\mathrm{OC},i}$ are the number of C and N in the ith organic compound, $y_{\mathrm{biom},i}$, $y_{\mathrm{OC},i}$ and $y_{\mathrm{NH}_4,i}$ are the stoichiometric coefficients of biomass, organic carbon, and NH_4, in the metabolic reaction.

REFERENCES

Kleerebezem, R., and Van Loosdrecht, M. C. M. (2010). A generalized method for thermodynamic state analysis of environmental systems. *Crit. Rev. Environ. Sci. Technol.* 40, 1–54. doi:10.1080/10643380802000974.

Stephanopoulos, G. N., Aristidou, A. A., and Nielsen, J. (1998). *Metabolic Engineering: Principles and Methadologies*. Amsterdam: Elsevier doi:10.1016/B978-0-12-666260-3.50015-7.

Supplementary Figures

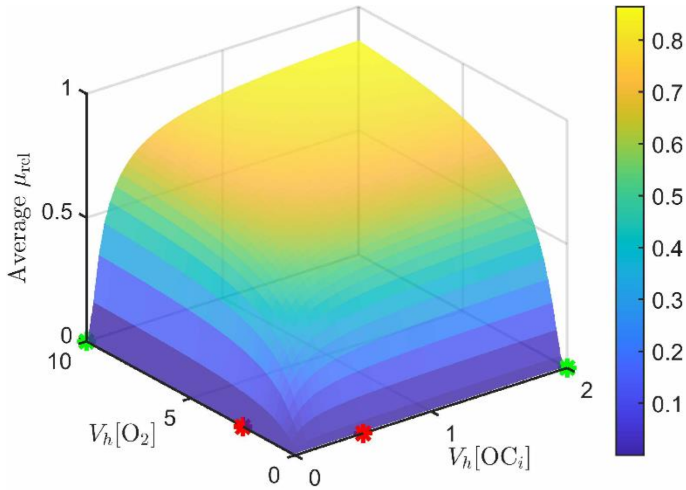

Figure S1. Three-dimensional surface plot illustrating the distribution of the average microbiome relative specific growth rates (averaged across all detected compounds) as functions of model parameters $V_{\mathrm{h}}[\mathrm{OC}_i]$ (assuming that all compounds are equally abundant) and $V_{\mathrm{h}}[\mathrm{O_2}]$. We chose two values for each parameter to represent low (green markers) and severe levels (red markers) of substrate limitations. The parameter values for severe substrate limitations are chosen such that the average $\mu_{\mathrm{rel}} \approx 0.5$ when the other substrate is non-limiting, and the parameter values that offer maximum possible μ_{rel} are selected to represent low substrate limitations. For the data used in this work, we estimate $\left[V_{\mathrm{h}}[\mathrm{OC}_i], V_{\mathrm{h}}[\mathrm{O_2}]\right] = [2,10]$ and $\left[V_{\mathrm{h}}[\mathrm{OC}_i], V_{\mathrm{h}}[\mathrm{O_2}]\right] = [0.5,2.5]$ to represent low and severe substrate limitations, respectively. Accordingly, different combinations of these values can be used concomitantly to realize different environmental conditions, e.g., $\left[V_{\mathrm{h}}[\mathrm{OC}_i], V_{\mathrm{h}}[\mathrm{O_2}]\right] = [0.5,10]$ corresponds to C-limited condition.

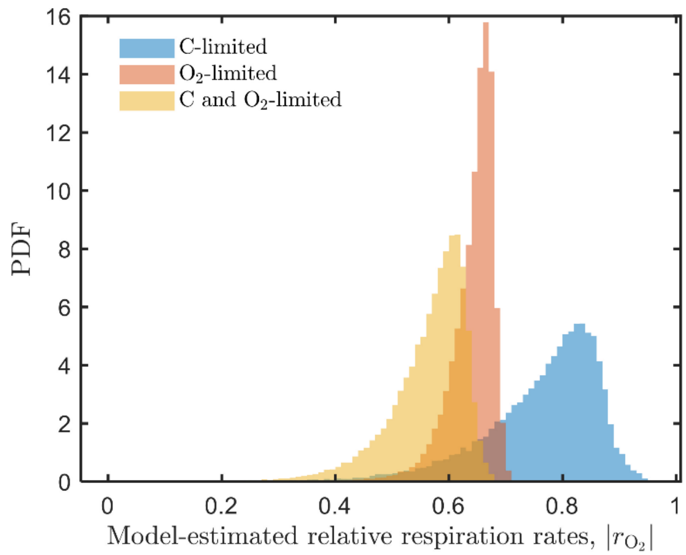

Figure S2. Three separate continuums of model-estimated respiration rates of organic compounds collectively detected in all samples ($n = 31014$) under varying environmental conditions – C-limited, O_2-limited and both C- and O_2-limited conditions.

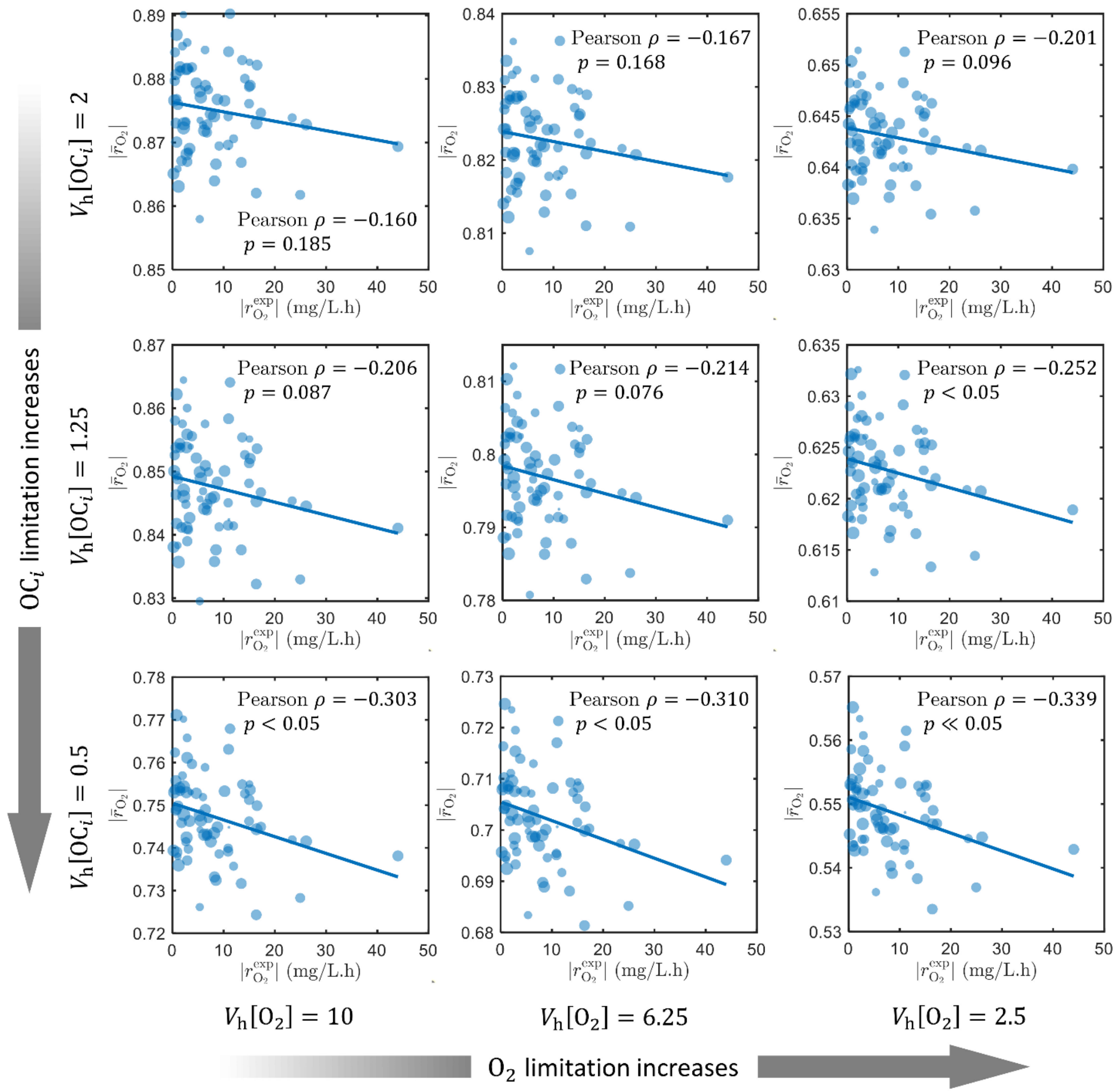

Figure S3. Correlations of model oxidative respiration rates against experimental respiration rates under various levels of organic carbon and O_2 limitations as defined by the model parameters $V_h[OC_i]$ and $V_h[O_2]$. The correlations are shown when assuming that all detected compounds in respective samples are respired. The size of markers denotes the number of compounds detected in respective samples.

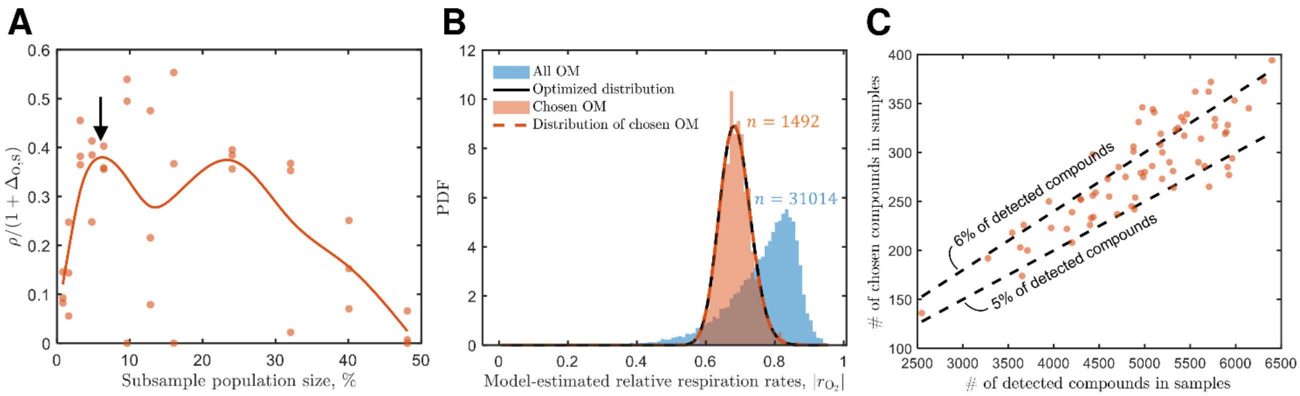

Figure S4. Outcome of optimization routine for identifying bioavailable compounds under C-limited condition. **(A)** The optimal population size to subsample bioavailable compounds from optimized distribution is 5% (of compounds collectively detected in all samples, $n = 31014$), where the scaled Pearson coefficients for correlations of model respiration rates against experimental respiration rates is maximized (for a minimal subsample size). The line shows a non-linear spline fit. **(B)** The optimized and the chosen bioavailable OM (5% subsample population size or 1492 compounds) distributions match adequately and make up a narrow range of respiration rates with a lower mean than the full OM. **(C)** In individual samples, the chosen bioavailable OM mostly make up between 5 – 6% (100 – 400 compounds) of detected compounds.

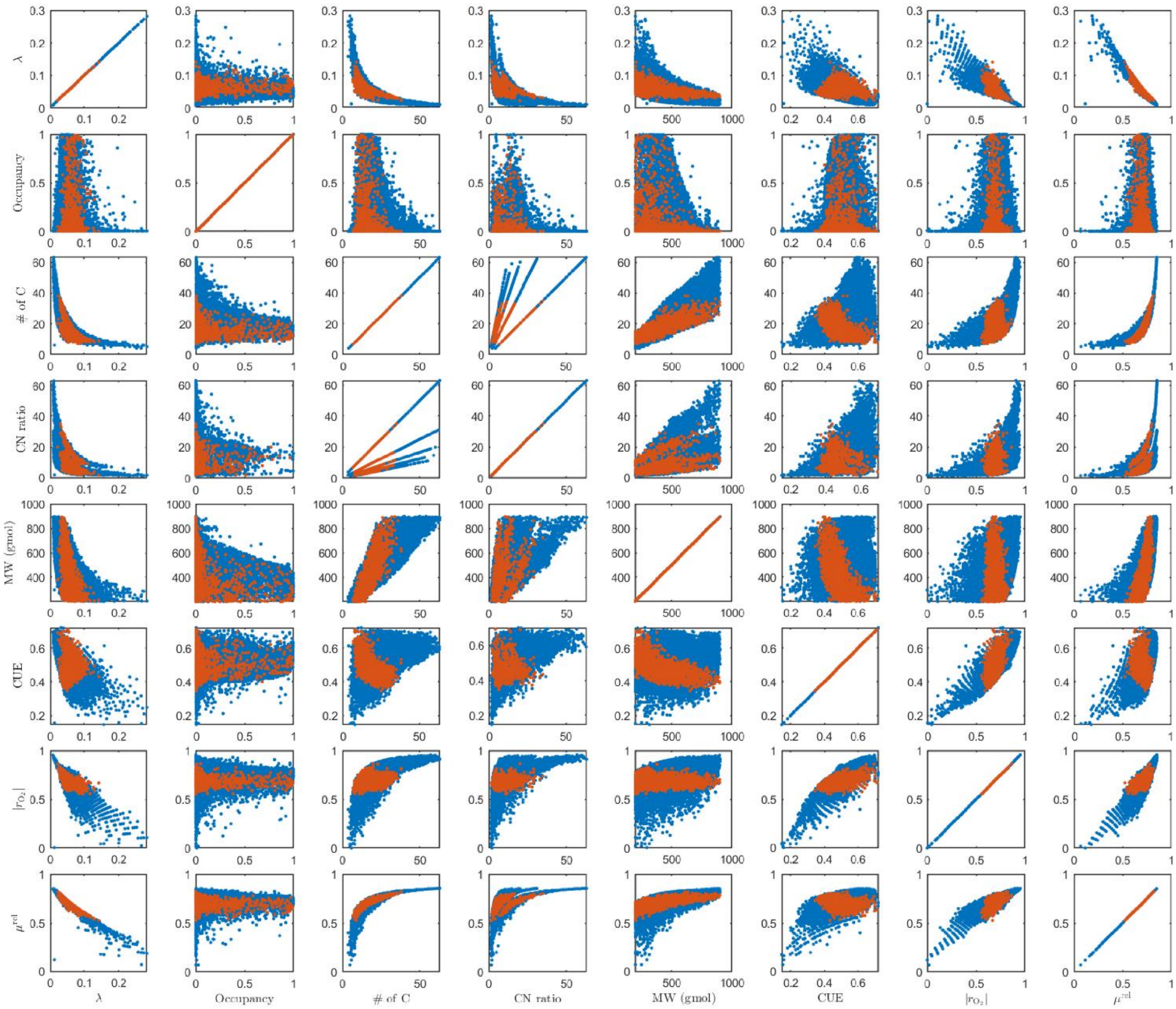

Figure S5. Matrix plot displaying pair-wise interdependencies between various properties of chosen bioavailable compounds (orange) compared against detected compounds from all samples (blue) under C-limited condition. MW, molecular weight; CUE, carbon use efficiency; TER, threshold element ratio.

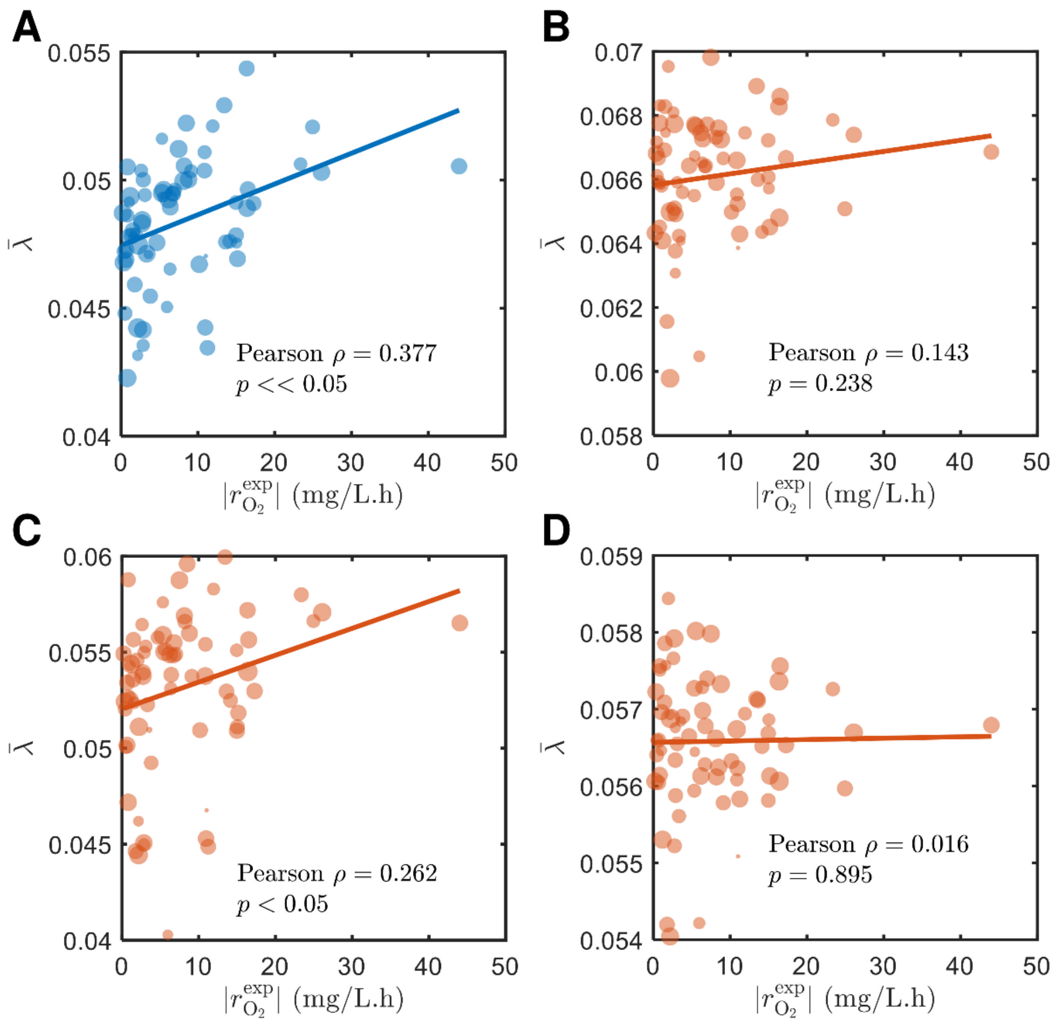

Figure S6. Correlations of thermodynamic favorability ($\bar{\lambda}$) against experimental respiration rates by including **(A)** all detected compounds in respective samples, as well as only the chosen bioavailable compounds detected in respective samples under **(B)** C-limited, **(C)** O_2-limited, and **(D)** both C- and O_2-limited conditions. The size of markers denotes the number of compounds detected (blue) and the number of chosen bioavailable compounds detected (orange) in respective samples. The directions of all correlations here are opposed to the typical expectation based on thermodynamic theory.