Supplementary Information for: Life history strategies and niches of soil bacteria emerge from interacting thermodynamic, biophysical, and metabolic traits

Gianna L. Marschmann[1], Jinyun Tang[1], Kateryna Zhalnina[1,2], Ulas Karaoz[1], Heejung Cho[1,3], Beatrice Le[3], Jennifer Pett-Ridge[4,5], Eoin L. Brodie[1,6*]

[1]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, 94720, CA, USA.
[2]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, 94720, CA, USA.
[3]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, 94720, CA, USA.
[4]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, 94550, CA, USA.
[5]Life and Environmental Sciences Department, University of California Merced, Merced, 95343, CA, USA.
[6]Department of Environmental Science, Policy and Management, University of California Berkeley, Berkeley, 94720, CA, USA.

*Corresponding author(s). E-mail(s): elbrodie@lbl.gov;
Contributing authors: glmarschmann@lbl.gov; jinyuntang@lbl.gov; kzhalnina@lbl.gov; ukaraoz@lbl.gov; beeologai@berkeley.edu; beatricele@berkeley.edu; pettridge2@llnl.gov;

Table 1: Overview of *DEBmicroTrait* trait integration.

Trait	Source	Formula	Unit	Source	Formula	Unit
Life-history	Genome-derived			Modelled		
Maximum specific growth rate	microTrait[a]	r_{max}	h^{-1}			
Genome size	microTrait/phylogeny	L_{DNA}	bp			
16S rRNA operon copy number	microTrait/rrnDB[b]	rrn	-			
DNA volume				Scaling law[c]	$V_{DNA} = v_N L_{DNA}$	m^3
Translation power				Biophysical model[c]	k_E	h^{-1}
Translation efficiency				Scaling law[d]	$9.5 - 1.22\log_2(\text{rrn})$	-
Biophysical	Literature			Modelled		
Cell volume				Scaling law[c]	$V_c = (\frac{V_{DNA}}{D_0})^{1/\beta_D}$	m^3
Protein volume				Scaling law[c]	$V_P = P_0 V_c^{\beta_P}$	m^3
Average protein length	Bionumber[e] 108986	$\bar{l}_P$	bp			
Average ribosome length	Bionumber 101439	$\bar{l}_R$	bp			
Bacterial cell density				Scaling law[f]	λ_B	-
Cell stoichiometry	Biophysical model[g]	$CH_uN_vO_w$	-			
Metabolic	Literature			Modelled		
Ribosome volume				Biophysical model[c]	V_R	m^3
Specific ribosome degradation rate				Biophysical model[c]	$\eta = \phi$	s^{-1}
Maximum ribosome processing rate	Bionumber 100059	$\bar{r}_R$	bp s^{-1}			
Specific protein degradation rate				Biophysical model[c]	$\phi = 6.2e^{-7}$	s^{-1}
Basal maintenance rate				Scaling law[h]	$0.39V_c^{0.88}$	h^{-1}
Average transport protein cost	Estimated[i]	$1.09e^{-19}$	W			
Stress tolerance				Modelled		
Intrinsic mortality rate				Scaling law[j]	$\gamma_{V,0} = 0.23e^{0.88 r_{max}}$	h^{-1}
Mortality half-saturation constant				Estimated	$\gamma_{V,1} = 98.8$	μM

Overview of *DEBmicroTrait* trait integration - continued.

Trait	Source	Formula	Unit	Source	Formula	Unit
Thermodynamic	Literature					
Reserve chemical potential	Estimated[k]	$\mu_E = 33$	kJ mol^{-1}			
Structure chemical potential	Estimated[k]	$\mu_V = 107$	kJ mol^{-1}			
Resource acquisition	Genome-derived			Modelled		
Membrane transport proteins	microTrait/TransportDB[l]	z_ρ	-			
Carbohydrate active enzymes	microTrait/CAZy[m]	z_X	-			
Membrane binding site density				Estimated[n]	ρ_{porter}	-
Biomass-specific site density				ECA parameter[f]	$N_{SB} = \frac{\lambda_B N_{porter}}{12.011}$	mol gC^{-1}
Maximum specific uptake rate				ECA parameter[f]	$V_{max} = k_{cat} N_{SB}$	h^{-1}
Half-saturation constant				ECA parameter[f]	K_D	μM
	Literature					
Max. substrate processing rate	Bionumber 114686	k_{cat}	s^{-1}			
Resource use	Modelled					
Reserve maintenance fraction	Default[o]	$y_{EM} = 1.0$	-			
Constitutive exoenzyme rate	Estimated[p]	$1e^{-2} z_X$	-			

References: [a]: [1], [b]: [2], [c]: [3], [d]: [4], [e]: [5], [f]: [6], [g]: [7], [h]: [8], [i]: [9], [j]: [10], [k]: [11], [l]: [12], [m]: [13], [n]: [14], [o]: [15], [p]: [16]

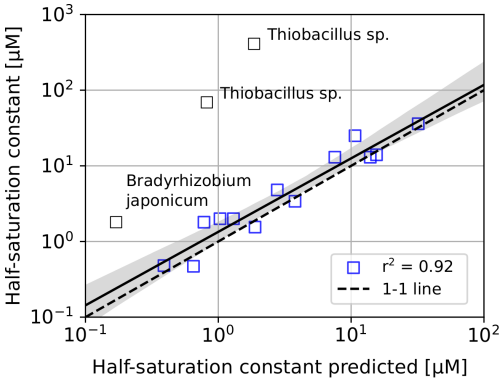

Fig. 1: Comparison of predicted half-saturation constants and measured half-saturation constants for 13 reference genomes as reported in [17]. The solid line indicates the regression line excluding outliers (r^2=0.92), while shaded areas indicate the 95% confidence bands. Data points that were excluded from the regression are labelled by species name. The dashed line corresponds to the 1-1 line for log-scaled observed vs. predicted values.

Table 2: Measured and modelled half-saturation constants for substrate uptake based on existing literature data [17].

Substrate	Taxonomy	Strain	Method, Incubation time	g_{max} [h^{-1}]	V_{max} [h^{-1}]	K_{meas} [μM]	K_{model} [μM]	
Glucose	Flavobacterium johnsoniae	C-21	Steady-state in continuous culture	0.2	[-]	1.55	1.89	[18]
Glucose	Escherichia coli	ML 308	Batch culture growth	1.238	[-]	13.0	13.93	[19]
Lactose	Lacticaseibacillus casei	64 H	Radioactivity uptake, 10 min	[-]	2.19	14.0	15.53	[20]
Glucose	Corynebacterium sp.	198	$^{14}CO_2$ from continuous culture samples	0.15	[-]	0.48	0.39	[21]
Tyrosine	Brevibacterium linens	47	Initial uptake, 5 min	[-]	1.05	3.40	3.78	[22]
Phenylalanine	Brevibacterium linens	47	Initial uptake, 5 min	[-]	1.75	25.0	10.7	[22]
Tryptophan	Brevibacterium linens	[-]	Initial uptake, 5 min	[-]	0.35	1.8	0.78	[22]
Isoleucine	Streptococcus thermophylus	302	Radioactivity uptake, 1 min	[-]	0.95	36	32	[23]
Valine	Streptococcus thermophylus	302	Radioactivity uptake, 1 min	[-]	1.20	2.0	1.3	[23]
Glycerol-3-phosphate	Escherichia coli	[-]	Radioactivity uptake, 1 min	[-]	0.006	2.0	1.03	[24]
Succinate	Rhizobium leguminosarum	[-]	Radioactivity uptake, 2 min	[-]	0.06	2.0	1.3	[25]
Toluene	Pseudomonas sp	[-]	Total ^{14}C-labelled product production, 4 hrs	[-]	0.111	0.47	0.65	[26]
Methanol	Pseudomonas sp	MA	Radioactivity uptake, 10 min	[-]	0.05	4.8	2.78	[27]
Fructose	Thiobacillus sp.	A2	Dyalisis rate of ^{14}C labelled substrate, 5 min	0.39	0.68	410	1.86	[28]
Ribose	Thiobacillus sp.	A3	Dyalisis rate of ^{14}C labelled substrate, 5 min	0.03	0.02	69	0.86	[28]
Succinate	Bradyrhizobium japonicum	[-]	Radioactivity uptake, 10 min	[-]	0.001	1.8	0.17	[29]

Abbreviations: maximum specific growth rate: g_{max} [h^{-1}], maximum specific uptake rate: V_{max} [h^{-1}], measured half-saturation constant: K_{meas} [μM], estimated half-saturation constant: K_{model} [μM]. Literature maximum uptake rates reported in units of nmol/min/mg were converted to the given unit by scaling with estimates of cellular dry mass [3].

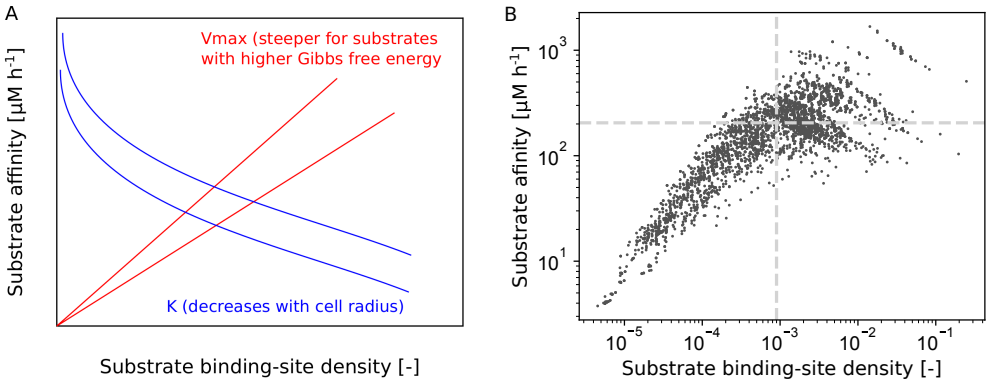

Fig. 2: **A** Relationship between the substrate uptake affinity at low external substrate concentration ($J_{S\to 0} = V_{max}/K$) and substrate binding-site density in the equilibrium chemistry approximation (ECA [6]) for substrate uptake. **B** Estimated substrate affinity as a function of the substrate binding-site density of rhizosphere isolates. The dashed lines indicate the substrate binding-site density corresponding to the locally weighted maximum substrate affinity. Substrates: n=82, consumers: n=39.

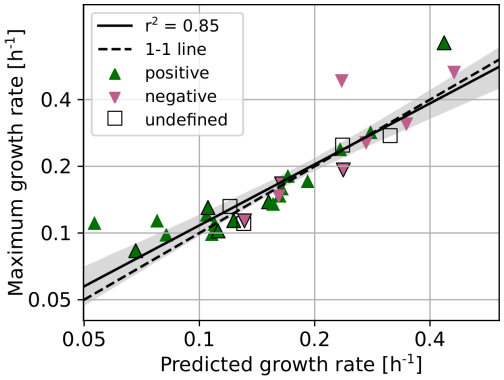

Fig. 3: Comparison of predicted maximum specific growth rates of rhizosphere isolates (n=39) and confirmed genome-predicted maximum specific growth rates through laboratory growth rate experiments [30]. Isolates are colored by their response to plant growth (green: positive, magenta: negative, white: undefined). Following guidelines developed for genome-scale models [31], the in-silico growth medium was designed to match the original R2 1/10 medium, on which the isolates were originally cultured at 28 °C. The solid line indicates the regression line (r^2=0.85), while shaded areas indicate the 95% confidence bands. The dashed line corresponds to the 1-1 line for observed vs. predicted values.

Table 3: Measured, genome-predicted, and modelled minimum generation times (minGT) of Hopland isolates.

Isolate	minGT measured [h]	minGT predicted [h]	minGT modelled [h]
HE68	-	7.00	6.43
HB09	-	1.31	1.50
HD36	1.55	2.78	2.93
HA02	6.05	6.09	8.94
HA13	5.9	5.73	6.63
HD69	5.93	5.32	6.58
HA54	2.31	2.43	2.48
HA33	-	2.90	2.97
HD24	3.57	2.52	2.20
HE23	2.34	3.57	2.90
HA28	3.1	6.24	13.03
HE60	5.96	6.09	5.64
HB58	3.59	3.33	3.08
HA31	3.28	2.72	2.54
HB48	4.67	3.60	2.92
HB62	3.62	2.24	2.00
HB13	2.55	2.25	2.11
HA19	1.87	5.26	5.77
HB36	3.99	3.88	3.67
HA56	4.25	4.70	4.28
HB07	5.58	5.01	4.57
HC08	5.07	5.12	4.45
HA36	2.45	1.82	1.57
HE70	-	4.16	4.23
HB44	4.34	4.36	4.22
HD25	3.87	3.83	4.07
HA20	5.41	6.28	5.31
HA32	4.44	4.82	3.87
HA14	-	1.43	2.95
HB15	2.52	0.96	1.59
HD17	-	7.05	8.46
HD88	3.69	7.49	7.94
HD82	4.95	6.11	5.29
HD59	6.29	4.74	4.30
HD07	6.22	6.30	6.33
HD57	-	6.78	6.20
HA57	-	0.75	0.81
HA41	-	4.04	3.62
HB20	7.13	8.33	10.17

Measured and genome-predicted minimum generation times as originally reported in [30].

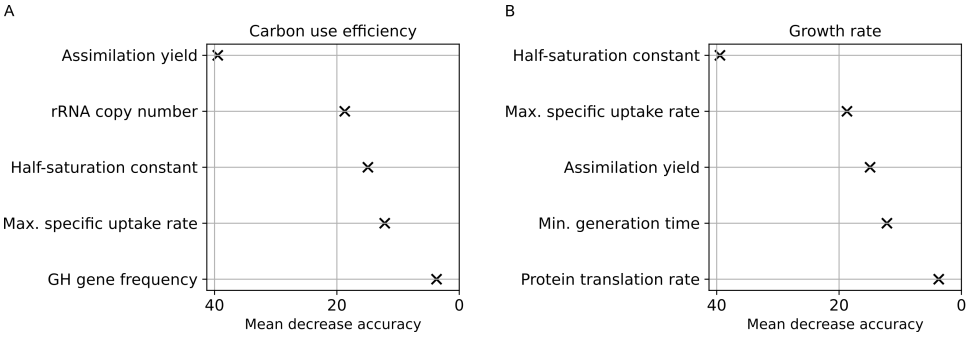

Fig. 4: Representation of the five most influential predictors of carbon use efficiency (**A**) and growth rate (**B**) in batch simulations as determined by mean decrease in accuracy. The mean decrease in accuracy is a measurement of the change in the accuracy of the random forest's predictions when the variable in question is randomly permuted. Labels on the y axis indicate the feature names. Feature contributions for all case studies were computed on predictions for the *undefined* rhizosphere isolate response group as out-of-bag samples.

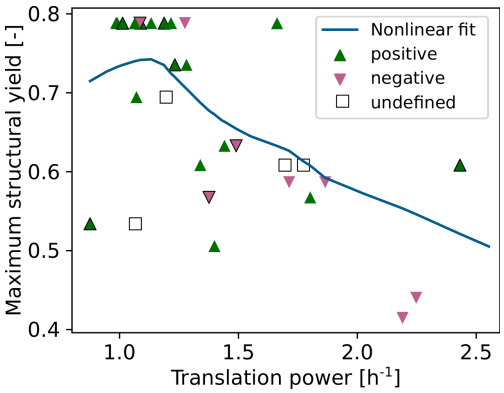

Fig. 5: Relationship between maximum structural biomass yield and translation power across rhizosphere isolates (n=39). Isolates are colored by their response to plant growth (green: positive, magenta: negative, white: undefined). The blue line corresponds to a locally weighted linear regression model.

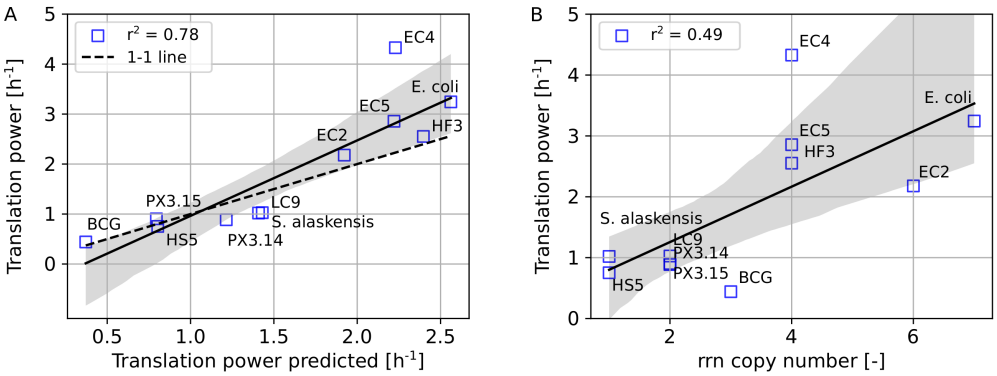

Fig. 6: **A** Regression scatter plot for protein synthesis power of phylogenetically diverse bacteria as reported in [4, 32]. The dashed line indicates the regression line (r^2=0.84), while shaded areas indicate the 95% confidence bands. The solid line corresponds to the 1-1 line for observed vs. predicted values (n=11). **B** Correlation between translation power and $\log_2$ rrn operon copy number (r^2=0.49). Species are the same as in **A**.

References

[1] Karaoz, U. & Brodie, E. L. microtrait: a toolset for a trait-based representation of microbial genomes. *Frontiers in Bioinformatics* **2** (2022).

[2] Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. & Schmidt, T. M. rrn DB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research* **43**, D593–D598 (2015).

[3] Kempes, C. P., Wang, L., Amend, J. P., Doyle, J. & Hoehler, T. Evolutionary tradeoffs in cellular composition across diverse bacteria. *The ISME Journal* **10**, 2145–2157 (2016).

[4] Roller, B. R., Stoddard, S. F. & Schmidt, T. M. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nature Microbiology* **1**, 1–7 (2016).

[5] Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers–the database of key numbers in molecular and cell biology. *Nucleic Acids Research* **38**, D750–D753 (2010).

[6] Tang, J. & Riley, W. J. Competitor and substrate sizes and diffusion together define enzymatic depolymerization and microbial substrate uptake rates. *Soil Biology and Biochemistry* **139**, 107624 (2019).

[7] Vrede, T., Dobberfuhl, D. R., Kooijman, S. & Elser, J. J. Fundamental connections among organism C:N:P stoichiometry, macromolecular composition, and growth. *Ecology* **85**, 1217–1229 (2004).

[8] Lynch, M. & Marinov, G. K. The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences* **112**, 15690–15695 (2015).

[9] Kempes, C. P., Koehl, M. & West, G. B. The scales that limit: the physical boundaries of evolution. *Frontiers in Ecology and Evolution* **7**, 242 (2019).

[10] Biselli, E., Schink, S. J. & Gerland, U. Slower growth of escherichia coli leads to longer survival in carbon starvation due to a decrease in the maintenance rate. *Molecular Systems Biology* **16**, e9478 (2020).

[11] Sousa, T., Mota, R., Domingos, T. & Kooijman, S. M. Thermodynamics of organisms in the context of dynamic energy budget theory. *Physical Review E* **74**, 051901 (2006).

[12] Elbourne, L. D., Tetu, S. G., Hassan, K. A. & Paulsen, I. T. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Research* **45**, D320–D324 (2017).

[13] Drula, E. *et al.* The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Research* **50**, D571–D577 (2022).

[14] Flynn, K. J., Skibinski, D. O. & Lindemann, C. Effects of growth rate, cell size, motion, and elemental stoichiometry on nutrient transport kinetics. *PLoS Computational Biology* **14**, e1006118 (2018).

[15] Tolla, C., Kooijman, S. A. & Poggiale, J.-C. A kinetic inhibition mechanism for maintenance. *Journal of Theoretical Biology* **244**, 576–587 (2007).

[16] Traving, S. J., Thygesen, U. H., Riemann, L. & Stedmon, C. A. A model of extracellular enzymes in free-living microbes: which strategy pays off? *Applied and Environmental Microbiology* **81**, 7385–7393 (2015).

[17] Button, D. Kinetics of nutrient-limited transport and microbial growth. *Microbiological Reviews* **49**, 270–297 (1985).

[18] Höfle, M. G. Glucose uptake of cytophaga johnsonae studied in batch and chemostat culture. *Archives of Microbiology* **133**, 289–294 (1982).

[19] Koch, A. L. & Houston Wang, C. How close to the theoretical diffusion limit do bacterial uptake systems function? *Archives of Microbiology* **131**, 36–42 (1982).

[20] Chassy, B. & Thompson, J. Regulation of lactose-phosphoenolpyruvate-dependent phosphotransferase system and beta-D-phosphogalactoside galactohydrolase activities in lactobacillus casei. *Journal of Bacteriology* **154**, 1195–1203 (1983).

[21] Law, A. & Button, D. Multiple carbon source-limited growth kinetics of a marine coryneform bacterium. *Journal of Bacteriology* **129**, 115–123 (1977).

[22] Boyaval, P., Moreira, E. & Desmazeaud, M. Transport of aromatic amino acids by Brevibacterium linens. *Journal of Bacteriology* **155**, 1123–1129 (1983).

[23] Akpemado, K. & Bracquart, P. Uptake of branched-chain amino acids by Streptococcus thermophilus. *Applied and Environmental Microbiology* **45**, 136–140 (1983).

[24] Schweizer, H., Argast, M. & Boos, W. Characteristics of a binding protein-dependent transport system for sn-glycerol-3-phosphate in Escherichia coli that is part of the pho regulon. *Journal of Bacteriology* **150**, 1154–1163 (1982).

[25] Finan, T., Wood, J. & Jordan, D. C. Symbiotic properties of C4-dicarboxylic acid transport mutants of Rhizobium leguminosarum. *Journal of Bacteriology* **154**, 1403–1413 (1983).

[26] Robertson, B. & Button, D. Toluene induction and uptake kinetics and their inclusion in the specific-affinity relationship for describing rates of hydrocarbon

metabolism. *Applied and Environmental Microbiology* **53**, 2193–2205 (1987).

[27] Bellion, E., Kent, M. E., Aud, J. C., Alikhan, M. & Bolbot, J. A. Uptake of methylamine and methanol by Pseudomonas sp. strain am1. *Journal of Bacteriology* **154**, 1168–1173 (1983).

[28] Wood, A. P. & Kelly, D. P. Kinetics of sugar transport by Thiobacillus a2. *Archives of Microbiology* **131**, 156–159 (1982).

[29] McAllister, C. F. & Lepo, J. Succinate transport by free-living forms of Rhizobium japonicum. *Journal of Bacteriology* **153**, 1155–1162 (1983).

[30] Zhalnina, K. *et al.* Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nature Microbiology* **3**, 470–480 (2018).

[31] Marinos, G., Kaleta, C. & Waschina, S. Defining the nutritional input for genome-scale metabolic models: A roadmap. *PLoS One* **15**, e0236890 (2020).

[32] Dethlefsen, L. & Schmidt, T. M. Performance of the translational apparatus varies with the ecological strategies of bacteria. *Journal of Bacteriology* **189**, 3237–3245 (2007).