

1 **Title**

2 The functional repertoire encoded within the native microbiome of the model nematode
3 *Caenorhabditis elegans*

4 **Running title**

5 *C. elegans* microbiome functions

6 **Authors**

7 Johannes Zimmermann^{1*}, Nancy Obeng^{2*}, Wentao Yang², Barbara Pees³, Carola
8 Petersen^{2,3}, Silvio Waschina¹, Kohar Annie Kissoyan², Jack Aidley², Marc P. Hoepfner⁴,
9 Boyke Bunk⁵, Cathrin Spröer⁵, Matthias Leippe³, Katja Dierking², Christoph Kaleta^{1#},
10 Hinrich Schulenburg^{2,6#}

11 **Affiliations**

12 1 Research Group Medical Systems Biology, Institute of Experimental Medicine,
13 Christian-Albrechts University, Kiel, Germany

14 2 Research Group of Evolutionary Ecology and Genetics, Zoological Institute, Christian-
15 Albrechts University, Kiel, Germany

16 3 Research Group of Comparative Immunobiology, Zoological Institute, Christian-
17 Albrechts University, Kiel, Germany

18 4 Institute of Clinical Molecular Biology, Christian-Albrechts University, Kiel, Germany

19 5 Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures,
20 Braunschweig, Germany

21 6 Max-Planck Institute for Evolutionary Biology, Ploen, Germany

22 * These authors contributed equally to this work: Shared first authorship

23 # These authors contributed equally to this work: Shared senior authorship

24 **Correspondence**

25 Christoph Kaleta, Research Group Medical Systems Biology, Institute of Experimental
26 Medicine, Christian-Albrechts University, Michaelisstraße 5, 24105 Kiel, Germany; Tel:
27 +49-431-50030340; Fax: +49-431-50030344; Email: c.kaleta@iem.uni-kiel.de

28 Hinrich Schulenburg, Research Group of Evolutionary Ecology and Genetics, Zoological
29 Institute, Christian-Albrechts University, Am Botanischen Garten 9, 24118 Kiel,
30 Germany; Tel.: +49-431-8804141; Fax: +49-431-8802403; Email:
31 hschulenburg@zoologie.uni-kiel.de

32 **Conflict of interest**

33 All authors declare no competing financial interests in relation to the work described.

34 **Keywords**

35 Microbiome, Microbiota, *Caenorhabditis elegans*, *Ochrobactrum*, *Pseudomonas*,
36 Metabolic networks

37 Financial support: German Science Foundation Collaborative Research Center CRC
38 1182 on Origin and Function of Metaorganisms, projects A1 (KD, ML, HS), A4 (HS), and
39 INF (MPH, CK). Excellence Cluster Precision Medicine in Chronic Inflammation (PMI;
40 CK, HS); the Competence Center for Genome Analysis Kiel (CCGA Kiel; HS); the Max-
41 Planck Society (Fellowship to HS); and the International Max-Planck Research School
42 for Evolutionary Biology (NO).

43 **Abstract**

44

45 The microbiome is generally assumed to have a substantial influence on the biology of
46 multicellular organisms. The exact functional contributions of the microbes are often
47 unclear and cannot be inferred easily from 16S rRNA genotyping, which is commonly
48 used for taxonomic characterization of the bacterial associates. In order to bridge this
49 knowledge gap, we here analyzed the metabolic competences of the native microbiome
50 of the model nematode *Caenorhabditis elegans*. We integrated whole genome
51 sequences of 77 bacterial microbiome members with metabolic modelling and
52 experimental characterization of bacterial physiology. We found that, as a community,
53 the microbiome can synthesize all essential nutrients for *C. elegans*. Both metabolic
54 models and experimental analyses further revealed that nutrient context can influence
55 how bacteria interact within the microbiome. We identified key bacterial traits that are
56 likely to influence the microbe's ability to colonize *C. elegans* (e.g., pyruvate
57 fermentation to acetoin) and the resulting effects on nematode fitness (e.g.,
58 hydroxyproline degradation). Considering that the microbiome is usually neglected in the
59 comprehensive research on this nematode, the resource presented here will help our
60 understanding of *C. elegans* biology in a more natural context. Our integrative approach
61 moreover provides a novel, general framework to dissect microbiome-mediated
62 functions.

63 **Introduction**

64 Multicellular organisms are continuously associated with microbial communities. The
65 ongoing interactions are likely to have influenced evolution of the involved microbes and
66 hosts, affecting bacterial growth characteristics or host development, metabolism,
67 immunity, and even behavior (1). Host organisms and their associated microorganisms
68 (i.e., the microbiome) are thus widely assumed to form a functional unit, the
69 metaorganism, where microbial traits expand host biology (2). To date, most microbiome
70 studies focus on describing the taxonomic composition of associated communities, using
71 16S rRNA amplicon sequencing (3). These studies revealed that specific taxa reliably
72 associate with certain hosts, for example Bacteroidetes and Firmicutes with humans,
73 *Snodgrassella* and *Gilliamella* with honeybees, or *Lactobacillus* and *Acetobacter* with
74 *Drosophila* (4–6). 16S profiling, however, is insufficient to identify bacterial functions of
75 importance for the interaction (7). More detailed information can be obtained from
76 bacterial genome sequences. For example, genomic analysis of the dominant members
77 of the bee microbiome revealed complementary functions in carbohydrate metabolism,
78 suggesting syntrophic interactions among coexisting bacteria (8). Further, the systems
79 biology approach of constraint-based modeling permits inference of genome-scale
80 metabolic models as a basis for predicting microbial phenotypes (9), as previously
81 demonstrated for the interaction between whiteflies and their endosymbionts (10,11) and
82 also hosts with more complex microbiomes (12,13).

83 The nematode *Caenorhabditis elegans* is one of the main model organisms in
84 biomedical research. Yet, almost all research with this nematode has been performed in
85 the absence of its native microbiome. In fact, its microbiome was only characterized

86 recently, consisting mostly of Gammaproteobacteria (*Enterobacteriaceae*,
87 *Pseudomonaceae*, and *Xanthomonodaceae*) and Bacteroidetes (*Sphingobacteriaceae*,
88 *Weeksellaceae*, *Flavobacteriaceae*) (14–17). The little currently available data on
89 microbiome functions highlights an influence on *C. elegans* fitness, stress resistance,
90 and protection against pathogens (15). Previous studies also combined *C. elegans* with
91 various soil bacteria, revealing that these can provide specific nutrients (18–22) or affect
92 the response to drugs against cancer and diabetes (23–26). To date, the functions of the
93 native microbiome have not yet been systematically explored.

94 The aim of this study was to establish the natural *C. elegans* microbiome as a model for
95 studying microbiome functions. We extended previous 16S rRNA data (15) by
96 sequencing whole genomes for 77 bacteria, which are associated with *C. elegans* in
97 nature, and also *Escherichia coli* OP50, the nematode’s standard laboratory food. We
98 reconstructed metabolic networks from the genome data to explore the metabolic
99 competences and resulting interaction potential of the microbiome. We additionally
100 characterized bacterial physiology and assessed which bacterial traits shape
101 colonization ability and influence *C. elegans* fitness.

102 **Material and Methods**

103 *Material*

104 Microbiome strains were previously isolated from natural *C. elegans* isolates or
105 corresponding substrates in Northern Germany (15; Supplementary Table S1). A
106 representative set of 77 strains was chosen for genome sequencing. For physiological
107 analysis, bacteria were cultured in tryptic soy broth (TSB) at 28 °C. For experiments with
108 *C. elegans* N2, bacterial TSB cultures (500 µl at OD₆₀₀ = 10) were spread onto peptone-
109 free medium (PFM) agar plates. Maintenance and bleaching, to obtain gnotobiotic, age-
110 synchronized worms, followed standard methods (27).

111 *Genome sequencing*

112 Total DNA was isolated from bacterial cultures using a cetyl-trimethyl-ammonium-bromid
113 (CTAB) approach (28). Sequencing was based on Illumina HiSeq and in a subset of nine
114 strains additionally the PacBio platform (Supplementary Table S1). For PacBio long read
115 genome sequencing, SMRTbell™ template library was prepared according to the
116 manufacturer's instructions (Pacific Biosciences, US; Protocol for Greater Than 10 kb
117 Template Preparation). SMRT sequencing was carried out on the PacBio RSII (Pacific
118 Biosciences, US) on one to three SMRT Cells, applying a movie length of 240-minutes.
119 SMRT Cell data was assembled using the RS_HGAP_Assembly.3 protocol (SMRT
120 Portal version 2.3.0). Chromosomes and chromids were circularized, unusual
121 redundancies at the ends of the contigs and artificial contigs were removed after a
122 comparison against all other replicons. Error correction was performed by Illumina reads
123 mapping onto finished genomes using BWA (29) with subsequent variant and
124 consensus calling using VarScan (30). QV60 consensus concordances were confirmed

125 for all genomes. Annotations were obtained with the NCBI Prokaryotic Genome
126 Annotation Pipeline (PGAP). For samples with only Illumina data, low quality reads
127 and/or adaptors were trimmed with Trimmomatic v0.36 (31). *De novo* genomes were
128 assembled using SPAdes v3.8.0 (32). Genomes (contigs greater than 1000 bp) were
129 annotated with PGAP and Prokka v1.11 (33). Genomes were compared with BRIG (34).
130 All sequences are available from NCBI Genbank, Bioproject PRJNA400855.

131 *Reconstruction of metabolic networks*

132 Metabolic networks were reconstructed as a basis for all subsequent computational
133 metabolic analyses and followed a two-step pipeline (Fig. 1a). First, the sequenced
134 genomes were used to create draft metabolic models, using ModelSEED version 2.0
135 (35) and associated SEED reaction database. Second, we corrected errors and
136 extended drafts by (i) finding futile cycles, (ii) allowing growth with the isolation medium
137 (TSB), (iii) improving biosynthesis of biomass components, (iv) extending capacities to
138 use different carbon sources, and (v) checking for additional fermentation products. This
139 curation was based on combining topological- and sequenced-based gap filling using
140 gapseq (version 0.9 “darwinian turtle”; <https://github.com/jotech/gapseq>), pathway
141 definitions of the MetaCyc database release 22 (36), and sequence data from UniProt
142 (37). The presence of enzymatic reactions was inferred by BLAST with bitscore of at
143 least 50 (≥ 150 for a more conservative estimation), and a 75% minimum query
144 coverage. Moreover, reactions were assumed to be present if overall pathway
145 completeness was higher than 75% or if it was higher than 66% and key enzymes of the
146 pathway were present (36). We also searched for genes possibly relevant in host-
147 microbe interactions using the virulence factor database (38). The resulting curated

148 models (Supplementary data S1) were used for further metabolic network analysis.
149 Computations were done with GNU parallel (39).

150 *Phylogenetic correlation and clustering of metabolic pathways*

151 We assessed whether similarity of metabolic reactions correlated with phylogenetic
152 relationship, using pairwise comparisons of bacteria. For each pair, the overlap of
153 present and absent pathways (predicted by gapseq) was calculated. The corresponding
154 16S rRNA similarity was scored as percent identity of the global alignment using
155 biostrings (40). 16S data was obtained from the SILVA database (41) based on the best
156 hit of the extracted genomic 16S rRNA using RNAmmer (42). To determine overall
157 metabolic distances between isolates, metabolic networks were treated as vectors,
158 clustered horizontally, and metabolic distances computed as Euclidean distances
159 between vectors. Cluster similarity was estimated by average linkage and assessed via
160 multi-scale bootstrapping (10,000 replications) using pvclust (43).

161 *BIOLOG experiments*

162 We used BIOLOG GN2 plates to assess the metabolic competence of selected bacterial
163 strains, including MYb10, MYb11, MYb71, MYb237, and OP50. Bacterial cultures were
164 washed three times using phosphate buffered saline (PBS) and density adjusted to
165 $OD_{600} = 1$. 150 μ l bacterial suspension per each well of BIOLOG plate were incubated at
166 28 °C for 46 h. Tetrazolium dye absorption (OD_{595}) was measured every 30 min (three
167 replicates per strain). We defined the magnitude of substrate reduction as the fold-
168 change in tetrazolium absorbance:

$$169 \quad \text{fold change} = \frac{OD_{t46} - OD_{t0}}{OD_{t0}} - OD_{control}$$

170 Fold-changes in water were subtracted as background. Hierarchical clustering of strains
171 was based on average fold-change profiles (Ward's clustering; Euclidean distance) and
172 bootstrapping ($n = 100$). To analyze metabolic specialization, k-means clustering of
173 substrates ($k = 7$, $n = 10^3$; (44)) was performed (Supplementary Fig. S1). Statistical
174 analyses were performed in R version 3.3.1 (45) and ggplot2 (46).

175 *Bacterial growth experiments*

176 To validate BIOLOG results, we assessed growth of MYb11, MYb71, and a co-culture of
177 both in defined media with either alpha-D-glucose or D-(+)-sucrose as carbon sources.
178 Our defined medium is related to S medium (27), and contains 0.3% NaCl, 1 mM
179 $MgSO_4$, 1 mM $CaCl_2$, 25 mM KPO_4 , 0.1% NH_4NO_3 , 0.05 mM EDTA, 0.025 mM $FeSO_4$,
180 0.01 mM $MnCl_2$, 0.01 mM $ZnSO_4$, 0.01 mM $CuSO_4$, and 1% carbon source. Defined
181 medium without carbon source served as negative and TSB as positive control.
182 Overnight cultures were washed and adjusted to 3.94×10^7 CFUs for growth
183 experiments. Microtiter plates were incubated as BIOLOG plates above. OD_{600} was
184 measured every 30 min, and cultures plated after 48 h. Selective plating of MYb71 using
185 kanamycin (10 $\mu g/ml$) allowed to quantify MYb11/MYb71 proportions in co-culture.
186 Three independent runs with technical replicates were performed and assessed with
187 Mann-Whitney U-tests and P-value adjustment by false discovery rate (fdr).

188 *Simulation of bacterial in silico growth*

189 We used the curated models to simulate the growth of MYb11 and MYb71 with sucrose
190 as carbon source. We searched for sucrose invertases using gapseq
191 (<https://github.com/jotech/gapseq>) and secreted peptides with SignalP 4.1 (47). *In silico*
192 growth was simulated with BacArena (48). The MYb71 extracellular sucrose invertase

193 was modeled as independent species with a single sucrose invertase reaction and
194 exchange reactions for sucrose, glucose, and fructose. Carbon source utilization and
195 metabolic by-products were predicted using flux balance analysis and flux variability
196 analysis in R with sybil (49). A carbon source was assumed to be utilizable if the minimal
197 solution of the corresponding exchange was negative (i.e., uptake) and a byproduct
198 producible if the maximal solution of exchange positive (i.e., production).

199 *Simulation of ecological interactions*

200 We assessed possible interactions among bacteria based on joined models, assuming a
201 common compartment for metabolite exchange between microbes. Activity of individual
202 reactions (i.e., fluxes) was linearly coupled to biomass production to prevent unrealistic
203 exchange fluxes, such as those that solely benefit the partner but not the producer (50).
204 The objective function was set to maximize the sum of fluxes through both biomass
205 reactions. Two growth media were used for simulations, including TSB and a glucose
206 minimal medium with thiamine and traces (0.001 mM) of sucrose and methionine to
207 allow initial bacterial growth (Supplementary Table S3). Joined growth rates (j_1 , j_2) were
208 compared to single growth rates (s_1 , s_2). Mutualism was defined as $j_1 > s_1$ and $j_2 > s_2$,
209 competition as $j_1 < s_1$ and $j_2 < s_2$, parasitism as $j_1 < s_1$ and $j_2 > s_2$ (or *vice versa*), and
210 commensalism as $j_1 = s_1$ and $j_2 > s_2$ (or the reverse).

211 *Experimental analysis of bacterial colonization and bacterial effects on C. elegans* 212 *population growth*

213 We examined bacterial colonization by quantifying CFUs extracted from young adult
214 worms exposed to bacteria for 24 h. In detail, L4 larvae were placed on bacterial isolates
215 (500 μ l, $OD_{600} = 10$) for 24 h, washed in a series of buffers (2x M9 buffer with 25 mM

216 tetramisole, 2x M9 with 25 mM tetramisole and 100 µg/ml gentamicin, 1x PBS with
217 0.025% Triton-X) to remove bacteria from the surface of nematodes, and homogenized
218 in the GenoGrinder 2000 using 1 mm zirconia beads (1200 strokes/min, 3 min). Worm
219 homogenate and supernatant control were plated onto TSA for quantification.

220 We further measured worm population growth as a proxy for worm fitness. We counted
221 worms in the population initiated with three L4 after five days at 20 °C on bacterial
222 lawns.

223 *Regression models*

224 We analyzed the association between phenotypic measurements (i.e., bacterial
225 colonization and worm fitness) and metabolic as well as virulence characteristics using
226 Spearman rank correlation and random forest regression analysis. Significance for the
227 correlation analysis was assessed with permutation tests using 100 randomly generated
228 features and FDR-adjusted P-values. For random forest regression, the R package
229 VSURF was used to select features based on permutation-based score of importance
230 (51) and otherwise default settings (ntree = 2000, ntry = p/3).

231 *Adaptive strategies*

232 According to the universal adaptive strategy theory (UAST) (52,53), heterotrophic
233 bacteria follow one of three strategies: i) rapid growth and thus good competitor, ii) high
234 resistance and thus stress-tolerator, or iii) fast niche occupation and thus ruderal. We
235 categorized bacterial isolates using published UAST criteria (53), based on three scores,
236 inferred from the genomes and metabolic models. In detail, the components of a
237 competitive strategy were a large genome size, antibiotics production (presence of
238 pathways belonging to 'Antibiotic-Biosynthesis' category in MetaCyc), high catabolic

239 diversity (Metacyc: 'Energy-Metabolism'), and siderophore biosynthesis (Metacyc:
240 'Siderophores-Biosynthesis'). The criteria for stress-tolerators were auxotrophies, slow
241 growth rates in TSB, few rRNA copies, and exopolysaccharides production (MetaCyc
242 pathways: PWY-6773, PWY-6655, PWY-6658, PWY-1001, PWY-6068, PWY-6082,
243 PWY-6073). The hallmarks of a ruderal strategy were fast growth in TSB, multiple rRNA
244 copies, and low catabolic diversity (Metacyc: 'Energy-Metabolism'). The characteristics
245 of each isolate were related to those of the other microbiome members, yielding a
246 relative score, thereby assuming that different strategies are present in the microbial
247 community as a whole. For each isolate, we assessed whether the inferred value
248 belonged to the lower or upper quantile of this criterium (in case of growth rates we used
249 the mean instead). The total adaptive score per strategy was scaled by the number of
250 features considered for a particular strategy. An isolate was assumed to follow the
251 strategy, for which it produced the highest score. If two strategies had the same score,
252 then isolates were considered to follow a mixed strategy.

253 **Results**

254 *Genomes of bacterial isolates, reconstruction and validation of metabolic networks*

255 We obtained whole genome sequences for 77 bacterial isolates of the *C. elegans*
 256 microbiome (Table 1). Of these, nine were sequenced with PacBio technology, allowing
 257 their full assembly, yielding either a single circular chromosome (four strains) or three
 258 circular chromosomes/chromids in case of the five isolates of the genus *Ochrobactrum*,
 259 which is known to have more than one chromosome (54)(Supplementary Table S1,
 260 underlined). The remaining isolates were sequenced with Illumina only, resulting in
 261 assemblies with 11 up to 243 contigs. For four genera (*Ochrobactrum*, *Pseudomonas*,
 262 *Arthrobacter*, *Microbacterium*), we included more than five strains and identified
 263 substantial intra-generic genome variation (Supplementary Fig. S2).

264

265 **Table 1: Overview of bacterial isolates from the natural microbiome of *C. elegans***
 266 **included in this study**

Phylum	Order	Genus	Isolate
Proteobacteria	Xanthomonadales	<i>Stenotrophomona</i>	MYb238, MYb57
Proteobacteria	Pseudomonadales	<i>Pseudomonas</i>	MYb1, MYb114, MYb115, MYb117, MYb12, MYb13, MYb16, MYb17, MYb184, MYb185, MYb2, MYb22, MYb3, MYb60, MYb75, MYb11 , MYb187, MYb193
Proteobacteria	Pseudomonadales	<i>Acinetobacter</i>	MYb10
Proteobacteria	Enterobacterales	<i>Erwinia</i>	MYb121
Proteobacteria	Enterobacterales	<i>Escherichia</i>	MYb137, MYb5, OP50
Terrabacteria group	Actinobacteria	<i>Micrococcaceae</i>	MYb211, MYb213, MYb214, MYb216, MYb221, MYb222, MYb224, MYb227, MYb229, MYb23, MYb51
Terrabacteria group	Actinobacteria	<i>Microbacteriaceae</i>	MYb24, MYb32, MYb40, MYb43, MYb45, MYb50, MYb54, MYb62, MYb64, MYb66, MYb72

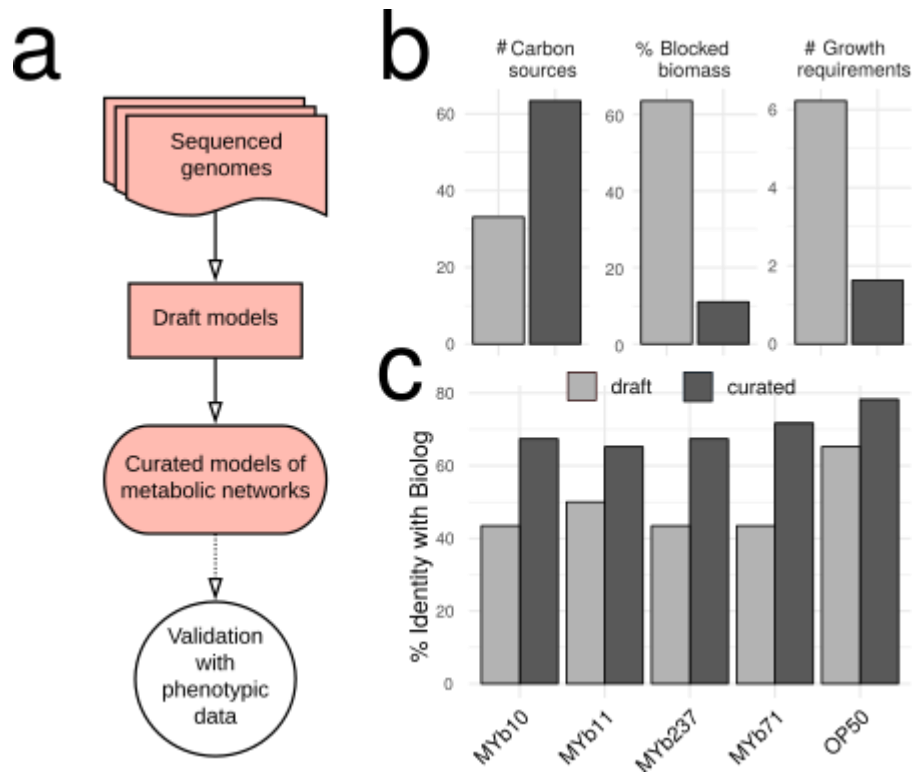
FCB group	Bacteroidetes	<i>Flavobacteriales</i>	MYb25, MYb44, MYb7
Proteobacteria	Caulobacterales	<i>Brevundimonas</i>	MYb31, MYb33, MYb46, MYb52
Terrabacteria group	Bacilli	<i>Paenibacillaceae</i>	MYb63
Proteobacteria	Rhizobiales	<i>Ochrobactrum</i>	MYb6 , MYb14, MYb15 , MYb18, MYb19, MYb29, MYb49 , MYb58 , MYb68, MYb71 , MYb237
Proteobacteria	Burkholderiales	<i>Achromobacter</i>	MYb9, MYb73
Terrabacteria group	Bacilli	<i>Bacillaceae</i>	MYb48, MYb56, MYb67, MYb78, MYb209, MYb212, MYb220
Bacteroidetes	Sphingobacteriales	<i>Sphingobacterium</i>	MYb181
Actinobacteria	Actinomycetales	<i>Rhodococcus</i>	MYb53

267 Strains with PacBio sequencing data are given in bold.

268

269 To study the functional repertoire of the microbiome, we reconstructed genome-scale
270 metabolic models (Fig. 1a, Supplementary data S1). The initial metabolic models were
271 curated by screening for transporter proteins and filling of missing reactions (gap-filling).
272 Curation increased model quality, including doubling of the number of utilized carbon
273 sources, reduction in the absence of essential biosynthesis pathways (e.g., for
274 nucleotides or amino acids) from 60% to below 10%, and reduction in the required
275 additional compounds for growth on defined media from on average six to one (Fig. 1b).
276 In order to validate our metabolic models, we experimentally quantified the ability of five
277 selected bacterial isolates to utilize 46 carbon sources using the BIOLOG approach. The
278 BIOLOG results produced a 49.6% overlap with the initial draft models and an increase
279 to 70% overlap with the curated models (Fig. 1c and Supplementary Fig. S9). These
280 curated models were subsequently used to explore bacterial metabolic competences.

281



282

283 **Fig. 1. Genomes of bacterial isolates, reconstruction and validation of metabolic**

284 **networks.** (a) Pipeline for metabolic network reconstruction. Sequenced genomes were

285 used to create draft metabolic models. Draft models were curated using topological- and

286 sequenced-based gap filling. The resulting models were validated with physiological

287 data (BIOLOG GN2; see Fig. 3); these models represent the metabolic networks of

288 microbiome isolates and were used for functional inference. (b) Model improvements by

289 curation, leading to an increase in accurate prediction of uptake of carbon sources, and

290 decreases in the prediction of non-producible biomass components and the number of

291 components needed for growth. (c) Model curation improved agreement with

292 experimental data, as for example the BIOLOG results.

293

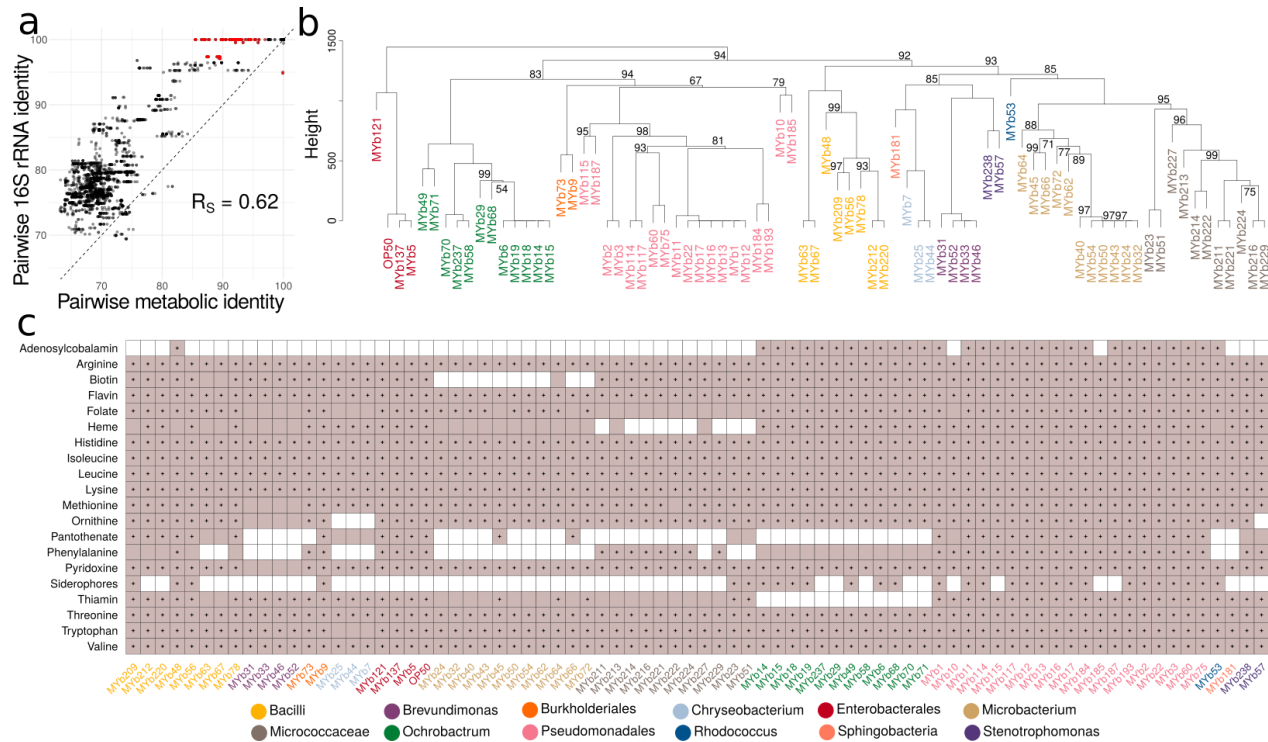
294

295

296 *Metabolic diversity within the microbiome of C. elegans*

297 Using the metabolic networks, we assessed a possible relationship between metabolic
298 and phylogenetic similarities and explored the metabolic potential of the isolates. We
299 found that the information contained in pairwise 16S rRNA phylogenetic relationships is
300 generally indicative of the corresponding similarities in metabolic networks (Fig. 2a;
301 Spearman rank correlation, $R_s = 0.6199$, $P < 0.0001$). Metabolic similarities appeared to
302 be higher than phylogenetic relationships, suggesting a considerable overlap in
303 metabolic competences across the included isolates. Nevertheless, some variation was
304 identified, even among isolates from the same genus. Such variation within taxonomic
305 groups was confirmed through hierarchical clustering of the inferred metabolic networks
306 (Fig. 2b), as for example seen for the *Pseudomonas* isolates, which contain three clearly
307 separated clusters. Similar patterns are also observed for other genera, for example
308 *Enterobacter*, *Ochrobactrum*, or *Microbacterium*. We conclude that variation in metabolic
309 competences is generally related to the bacterial phylogeny albeit some variation being
310 present within genera.

311



312

313 **Fig. 2. Metabolic network clustering and distribution of important pathways.** (a)

314 Correlation between pairwise similarities in 16S rRNA sequences and metabolic

315 networks is shown. Red indicates pairs with a 16S rRNA identity above 97% and

316 metabolic identity below 97% and vice versa. (b) Hierarchical clustering of metabolic

317 networks based on pathway prediction. P-values were calculated via multiscale

318 bootstrap resampling. In case of full support (i.e., $P = 100$), P-values are not shown (For

319 a complete list of different unbiased P-values and bootstrap values see Supplementary

320 Figure S11). (c) Prediction of bacterial capacity to produce metabolites favoring *C.*

321 *elegans* growth. Filled squares in light purple indicate that the metabolic networks

322 predict presence of the biosynthetic pathway required to produce essential amino acids

323 and co-factors. Black dots within the filled squares indicate that pathway presence is

324 supported by more conservative parameters (BLAST bit score ≥ 150). Different

325 bacterial genera in (b) and (c) are indicated by different colors of the strain names (Table
326 1).

327

328 We next assessed the metabolic competences of the microbiome isolates
329 (Supplementary Table S4). In general, the inferred metabolic competences are
330 consistent with the aerobic and heterotrophic lifestyle of the *C. elegans* host. The
331 glycolysis, at least the partial pentose phosphate pathway, the tricarboxylic acid cycle,
332 and enzymes enabling oxidative phosphorylation (cytochrome oxidases) were present in
333 all genomes. Almost all isolates possessed enzymes enabling tolerance to microaerobic
334 conditions (e.g., cytochrome bd oxidase). Some isolates from Bacilli, *Pseudomonas*, and
335 *Ochrobactrum* showed sequence-evidence for chemolithotrophic life style (nitrite and
336 formate oxidation) and anaerobic respiration (nitrate, arsenate reduction). Pathways
337 related to CO₂ fixation (reductive TCA or anaplerosis) were found in a few
338 *Pseudomonas*, Bacilli, or *Microbacterium* isolates. Two Bacillales strains further showed
339 capacity to degrade polysaccharides, such as starch, cellulose, mannan,
340 rhamnogalacturonan (e.g., *Paenibacillus* MYb63, *Bacillus* MYb67). The microbiome
341 members are able to produce all essential substances required for *C. elegans* growth,
342 which the nematode cannot synthesize on its own (i.e., all essential amino acids and
343 vitamins; Fig. 2c). Most variation among isolates was observed in the biosynthetic
344 pathways of B12, pantothenate, phenylalanine and siderophores (Fig. 2c). Simulation of
345 *in silico* growth (Supplementary Fig. S9) suggests that simple sugars, such as glucose,
346 ribose or arabinose, can be used by all organisms while the ability to degrade lactose,
347 maltodextrin, or sucrose varies among strains. Short chain fatty acids were among the

348 compounds that can be generated by all organisms (Supplementary Fig. S9), while there
349 was variation in the ability to produce succinate, cysteine, and valine. Moreover, several
350 microbiome members possessed potential virulence genes, especially the
351 *Pseudomonas* and *Escherichia* isolates (Supplementary Table S5).

352 We subsequently focused our analysis on *Ochrobactrum* and *Pseudomonas* isolates.
353 These two genera are enriched in the native microbiome of *C. elegans*, comprising 10–
354 20 % of the associated bacteria, they are also particularly well able to colonize the
355 nematode gut (15), and some isolates can protect *C. elegans* from pathogen infection
356 (15,55). Most *Pseudomonas* isolates can provide all required substances for nematode
357 growth. *Ochrobactrum* isolates are able to produce vitamin B12, like *Pseudomonas*
358 isolates, but unlike almost any of the other microbiome members (Fig. 2c). Moreover,
359 the *Ochrobactrum* isolates vary from other microbiome members in degradation
360 pathways, energy metabolism, vitamin biosynthesis, and presence of potential virulence
361 factors (Supplementary Table S6). These isolates appear to lack some for *C. elegans*
362 relevant vitamin biosynthetic pathways, such as those leading to thiamine and
363 panthothenate. They possess a unique Brucella-like putatively immune-modulating LPS
364 (Supplementary Table S5).

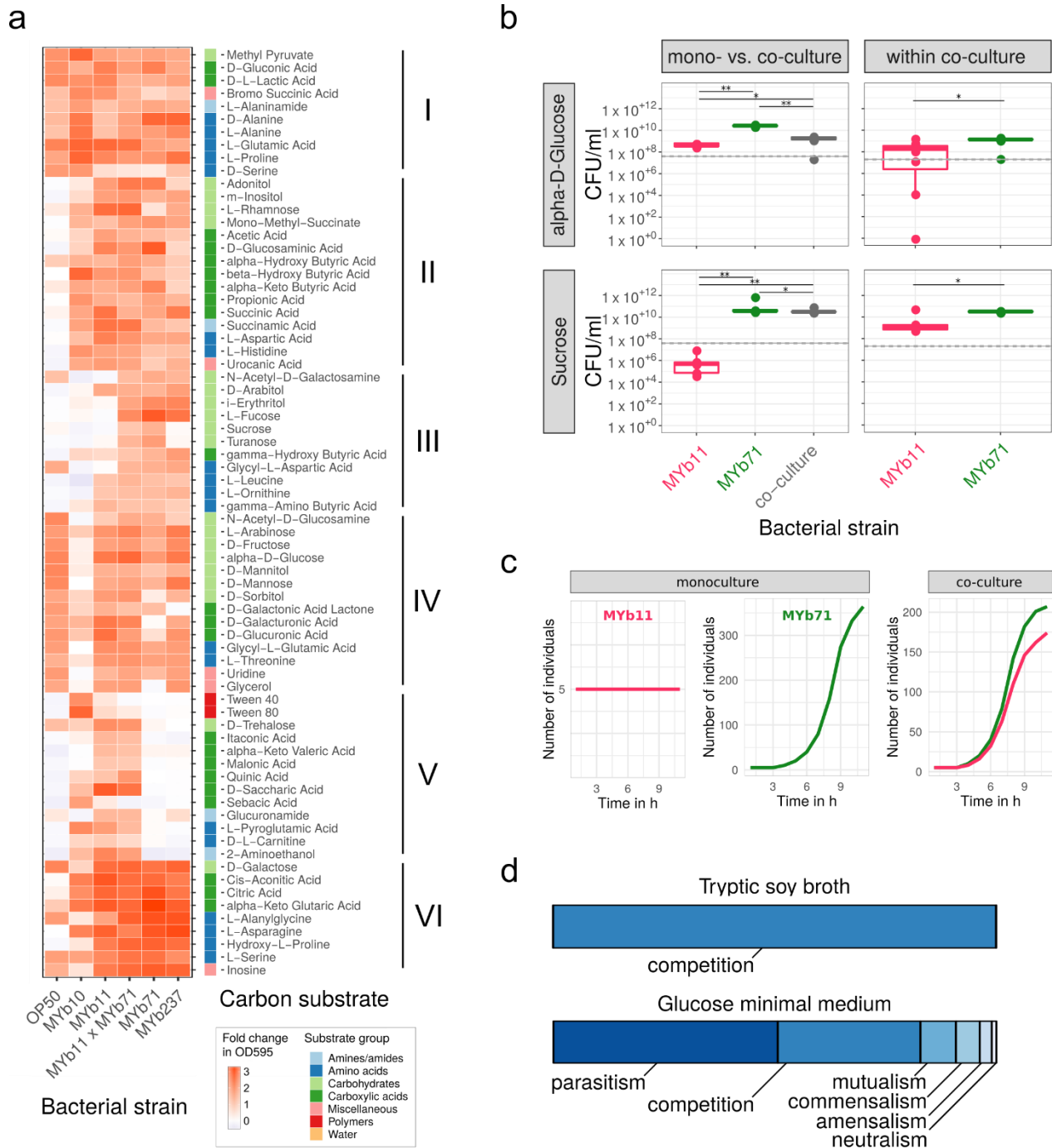
365 In summary, we found that *C. elegans* harbors a microbial community with diverse
366 metabolic competences, which can supply all essential nutrients for *C. elegans* and
367 which includes several *Ochrobactrum* and *Pseudomonas* isolates capable of producing
368 important vitamins such as vitamin B12.

369

370 *Nutrient context influences ecological interactions within the microbiome*

371 To study how metabolic repertoires affect bacterial growth and interactions within the
372 microbiome, we characterized carbon source utilization of selected isolates and tested
373 growth in different nutrient environments *in vitro* and *in silico*. Using the BIOLOG
374 approach, we focused on prominent *C. elegans* microbiome members that colonize
375 worms and affect host fitness, including *Ochrobactrum* sp. MYb71, *Ochrobactrum* sp.
376 MYb237, *Acinetobacter* sp. MYb10, *Pseudomonas lurida* MYb11, and as a contrast the
377 laboratory food strain *E. coli* OP50 (Supplementary Fig. S3; (15)). For a first insight into
378 bacterial interactions, we additionally included a MYb11-MYb71 mixture (two strains that
379 can co-exist in *C. elegans* (15)). We found that the metabolic repertoires of the strains
380 differ and that the four microbiome isolates can be distinguished from OP50 based on
381 the metabolism of carboxylic and amino acids (Fig. 3a, cluster II; Supplementary Fig.
382 S4). Within the microbiome, MYb10 was least versatile at using carboxylic acids and
383 sugar alcohols (Fig. 3a, cluster IV), while MYb11 and the two *Ochrobactrum* strains
384 could additionally metabolize unique sets of carboxylic acids and sugar alcohols,
385 respectively (Fig. 3a, cluster V and III). Notably, the disaccharides sucrose and turanose
386 were only metabolized by MYb71 (alone and in co-culture) (Fig. 3a, cluster III), although
387 sucrose invertases were present in the genomes of both MYb71 and MYb11 (cf.
388 pathway: sucrose degradation I, Supplementary Table S4). In co-culture, the metabolic
389 repertoires of MYb11 and MYb71 appeared additive.

390



391

392 **Fig. 3. Realized carbon metabolism and growth.** (a) Profiles of carbon substrate use

393 of *Acinetobacter* sp. (MYb10), *Pseudomonas lurida* (MYb11), *Ochrobactrum* sp.

394 (MYb71), *Ochrobactrum* sp. (MYb237), and *E. coli* OP50 in BIOLOG GN2 plates over 46

395 h. The fold-change in indicator dye absorption from 0 to 46 h indicates that the particular

396 compound is metabolized. K-means clustering ($k = 7$) of substrates by fold-change
397 highlights metabolic differences between strains. See Supplementary Fig. S5 for cluster
398 VII with substrates used poorly across most strains. (b) Colony forming units per ml
399 (CFU/ml) of MYb11 and MYb71 in mono- and co-culture at 48 h in alpha-D-glucose and
400 sucrose-containing minimal media. The horizontal and dashed lines indicate mean and
401 SD of CFU/ml at inoculation. Statistical differences were determined using Mann-
402 Whitney U-tests and corrected for multiple testing using *fdr*, where appropriate.
403 Significant differences are indicated by stars (** for $p < 0.01$; * for $p < 0.05$). Data from
404 three independent experiments is shown. (c) *In silico* growth of MYb11 and MYb71 in
405 mono- and co-culture in sucrose-thiamine medium using BacArena with an arena of
406 20x20 and five initial cells per species. (d) Bacterial interaction types observed during *in*
407 *silico* co-cultures of all combinations of the 77 microbiome isolates and OP50.

408
409
410 We next assessed whether the differences in MYb11 and MYb71 metabolic
411 competences shape bacterial interactions in growth media with only a single carbon
412 source. We did not observe any growth in a control medium without a carbon source,
413 and thus conclude that the tested bacteria are not chemoautotrophic (Supplementary
414 Fig. S6). In minimal medium with alpha-D-glucose, both MYb11 and MYb71 grew, yet
415 exhibited distinct growth dynamics (Fig. 3b; Supplementary Fig. S6). MYb71 produced
416 more CFUs than MYb11 in co-culture (Fig. 3b), suggesting that MYb71 has a growth
417 advantage over MYb11 and/or interferes with MYb11 in some other way. In agreement
418 with the BIOLOG results, a medium including sucrose as the sole carbon source

419 supported only growth of MYb71 but not MYb11 in monoculture (Fig. 3b, Supplementary
420 Fig. S6). Surprisingly, MYb11 increased in CFUs in co-culture, while in monoculture
421 MYb11 CFUs declined over time, indicating parasitic growth (Fig. 3b). Thus, the
422 presence of different carbon sources can change the interaction type between two
423 isolates.

424 We subsequently analysed in more detail the basis for co-growth of MYb11 and MYb71
425 in sucrose medium, using genome sequence information and *in silico* growth
426 simulations. Interestingly, we found a secreted sucrose invertase in the genome of
427 MYb71 but not MYb11 (Supplementary Fig. S10). *In silico* growth simulations
428 demonstrated that MYb71 can grow in sucrose medium, MYb11 alone does not, while in
429 co-culture both increased in numbers (Fig. 3c). The simulations thereby re-captured the
430 *in vitro* findings of distinct growth patterns in sucrose medium. Genome sequence
431 information strongly suggests that growth of both in co-culture is mediated by a secreted
432 enzyme from MYb71.

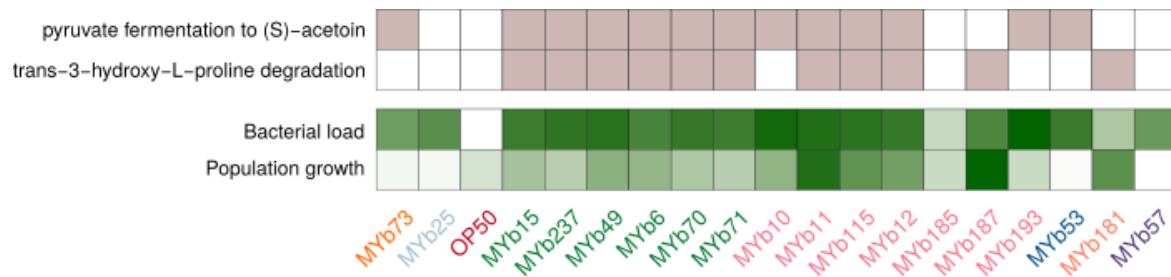
433 Taking a more global perspective, we next investigated *in silico* the potential ecological
434 interactions among the 77 microbiome isolates and *E. coli* OP50. We compared the
435 growth characteristics of single bacteria with co-growth rates of all 3003 possible
436 bacterial pairs in different nutrient environments. In a rich medium (TSB), the exclusive
437 interaction type was competition, indicated by lower growth rates in co- vs. mono-culture
438 (Fig. 3d). This changed completely when the nutrient environment was simplified to a
439 glucose minimal medium: 50% of the interactions were parasitic (i.e., the growth rate for
440 one isolate was higher in co-culture than in monoculture, while this pattern was opposite
441 for the other isolate of a pair), one out of three interactions were competitive, and 8%

442 mutualistic (i.e., growth rates for both isolates higher in co-culture than the
443 monocultures; Fig. 3d). Under these minimal medium conditions, the most frequently
444 exchanged metabolites across bacteria were glyceraldehyde, acetate, and ethanol
445 (Supplementary Fig. S7). We conclude that the nutrient context modulates bacterial
446 growth, consistently identified both *in silico* and *in vitro*, and thereby shapes bacteria-
447 bacteria interactions within the microbiome.

448
449 *Specific metabolic competences predict bacterial colonization ability and bacterial*
450 *effects on nematode fitness*

451 To characterize traits involved in the interaction between bacteria and *C. elegans*, we
452 identified genomic and metabolic features that are associated with the bacteria's
453 colonization ability and their effects on worm fitness. We focused on 18 microbiome
454 isolates based on (i) their abundance in the *C. elegans* microbiome, (ii) enrichment in
455 worms, and (iii) effects on worm population growth (15,56). OP50 was included as
456 control. Our phenotypic analysis revealed substantial variation among bacterial isolates
457 in both their ability to colonize *C. elegans* and their effects on nematode fitness (Fig. 4;
458 Supplementary Fig. S3). Importantly, these two microbiome characteristics were
459 significantly related with certain metabolic competences of the bacteria. Pyruvate
460 fermentation to (S)-acetoin was significantly associated with bacterial load and the
461 degradation of trans-3-hydroxyproline with nematode population growth (Fig. 4,
462 Supplementary Table S8).

463



464

465 **Fig. 4. Relationship of bacterial metabolic competences with their colonization**

466 **ability and their effects on nematode fitness.** Presence of metabolic traits (light

467 purple color), which were found to be associated with the bacteria's ability to colonize *C.*

468 *elegans* or affect nematode population growth as a proxy for worm fitness (green color).

469 Regression models suggested that the pathway of pyruvate fermentation to acetoin

470 influences bacterial load while the presence of hydroxyproline degradation is associated

471 with *C. elegans* population growth. Colonization and population growth data was

472 normalized; darker colors indicate increased capacities. Different bacterial genera are

473 indicated by the different colors of the strain names (Table 1).

474

475

476 To further explore the potential behavior of the microbiome isolates in an ecological

477 context, we interpreted their genomic and metabolic traits in light of the universal

478 adaptive strategy theory (52,53). We found that 26 isolates were associated with a

479 competitive, 9 with a stress-tolerating, and 37 with a ruderal (fast niche occupiers)

480 strategy (Fig. 5a). The remaining 6 isolates showed a mixed strategy (same score for

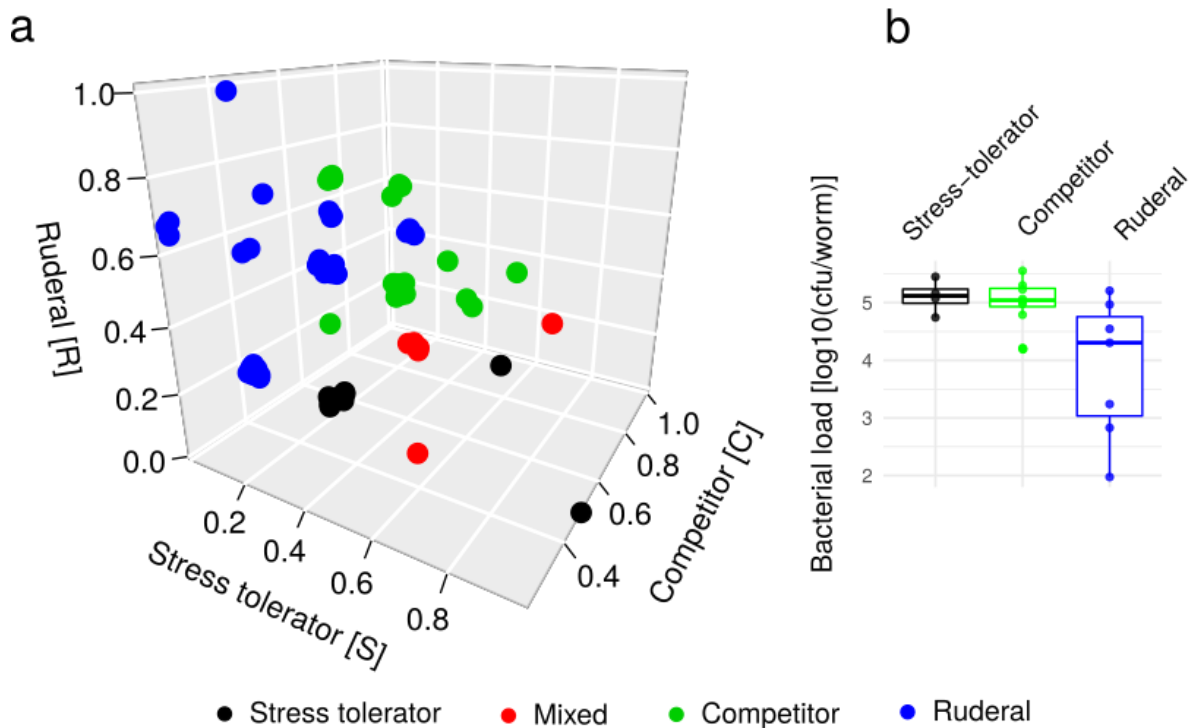
481 competition and stress-tolerance). Interestingly, bacterial isolates with different adaptive

482 strategies also varied in their colonization ability (Fig. 5b): Bacterial isolates with

483 competitive or stress-tolerance strategies showed higher bacterial load in *C. elegans*

484 than those with ruderal strategy (Wilcoxon rank sum test, $P = 0.01$). Moreover, for the
485 competitive and stress-tolerance isolates, we identified a positive correlation between
486 bacterial load and the inferred score (Spearman, $R_s = 0.37$, $P = 0.1$; Supplementary Fig.
487 S8). Taken together, the competitive and stress-tolerating strategies are most prevalent
488 within the microbiome of *C. elegans* and relate to bacterial colonization capacity.

489



490 ● Stress tolerator ● Mixed ● Competitor ● Ruderal

491 **Fig. 5. Different adaptive strategies within the microbiome and their relationship to**
492 **worm colonization.** We applied the universal adaptive strategy theory proposed for soil
493 bacteria (48) to categorize the bacterial isolates. (a) Based on genomic and metabolic
494 features, each isolate obtained a score for the competitive (C), stress tolerating (S), and
495 ruderal (R) strategy, which is represented in the 3D-coordinate system. (b) Bacterial
496 colonization behavior in comparison to adaptive strategies. Isolates which were
497 categorized as mixed strategists (i.e. same score for competitive and stress-tolerance)

498 produced the lowest bacterial load, whereas stress-tolerator and competitors had the
499 highest values. The difference in bacterial load between ruderal and other strategies
500 was significant (Wilcoxon rank sum test, $P = 0.01$).

501 **Discussion**

502 We here present the first overview of the functional repertoire encoded within the native
503 microbiome of the model organism *C. elegans* and provide a metabolic framework for
504 the functional analysis of host-associated microbial communities, which combines
505 genome sequence data, metabolic network modeling, and physiological
506 characterizations. For our analysis, we obtained whole genome sequences and
507 reconstructed the metabolic network of 77 microbiome members. We identified variation
508 in the metabolic competences within the microbiome and found that the community as a
509 whole is able to produce nutrients essential for *C. elegans* growth. For selected bacteria,
510 we were able to validate the model predictions with the help of physiological analyses.
511 Moreover, we used both *in vitro* and *in silico* approaches to demonstrate that the nutrient
512 environment can lead to a shift in the interaction between bacteria, for example from
513 competition to mutualism. We further identified specific metabolic modules that appear
514 to shape the interaction with the nematode host, including pyruvate fermentation to (S)-
515 acetoin and the degradation of trans-3-hydroxyproline. Finally, we considered a
516 combination of genomic, metabolic and cellular traits to infer bacterial life history
517 strategies according to the universal adaptive strategy theory (52,53), finding that
518 bacterial colonization ability is associated with a competitive or stress-tolerant strategy.
519 In the following, we will discuss in more detail (i) the diversity of metabolic competences
520 in the microbiome and possible implications for *C. elegans* biology, (ii) how the
521 metabolic networks shape bacteria-bacteria interactions, and (iii) in which ways bacterial
522 traits can affect colonization and *C. elegans* fitness.

523 Our analysis revealed that the microbiome members are jointly able to synthesize all
524 essential nutrients required by *C. elegans*. Individual bacterial isolates are not able to do
525 so. The considered isolates varied in their capacity to produce vitamins essential to *C.*
526 *elegans*, such as folate, thiamine, and vitamin B12, which are known to affect nematode
527 physiology and life history (20,21,23,57–59). For example, vitamin B12 influences
528 propionate breakdown, it can accelerate development, and reduce fertility (58,59).
529 Within the microbiome isolates studied here, only *Pseudomonas* and *Ochrobactrum*
530 strains had the pathways to produce vitamin B12. An enrichment of these genera in the
531 microbiome should therefore affect both the metabolic state and fitness of *C. elegans*.

532 Our study demonstrated that bacterial interactions can change depending on the
533 nutrient environment. In our simulations, competitive interactions dominated in rich
534 medium (TSB), while parasitic and mutualistic interactions in minimal medium.
535 Furthermore, interactions between *Pseudomonas lurida* (MYb11) and *Ochrobactrum* sp.
536 (MYb71) shifted from parasitic to competitive in a sucrose- vs. glucose-supplemented
537 medium. In line with this, we have detected a secreted sucrose invertase in the genome
538 of MYb71, which otherwise lacks any known sucrose transporter. Thus, we propose that
539 MYb71 breaks down sucrose extracellularly, and the monosaccharides glucose and
540 fructose become exploitable by MYb11. While a similar phenomenon has been
541 described in cultures of yeast with engineered auxotrophies (60,61), we here observed
542 this capacity in strains from a natural community of host-associated microbes. This
543 emphasizes the relevance of nutrient context in host-microbiome interactions.
544 Importantly, no single growth medium might reliably predict all possible interaction types
545 among bacteria. It is therefore essential to consider the nutrient context in order to fully
546 understand the range of possible interactions within the microbiome (e.g. (62)).

547 Our analysis further identified two bacterial traits that appear to influence colonization
548 ability and impact *C. elegans* fitness. Colonization ability was associated with pyruvate
549 fermentation to (S)-acetoin. This fermentation pathway includes the ketone diacetyl as
550 an intermediate, whose buttery odor attracts *C. elegans* and promotes feeding behaviour
551 (63). In detail, diacetyl binds the transmembrane odor receptor *odr-10* and induces a
552 shift in odortaxis (63–65). As a result, worms are more attracted to and less likely to
553 leave bacterial lawns with this particular smell (63). Indeed, lactic acid bacteria in rotting
554 citrus fruits were more attractive to worms when releasing diacetyl (66). Similarly,
555 entomopathogenic nematodes of the genus *Steinernema* were more attracted to insect
556 cadavers, when they were infected with the diacetyl-producing bacterial symbionts of the
557 nematode (67). Thus, if worms are attracted to diacetyl-producing bacteria, they should
558 spend more time in their presence. This alone could increase uptake of bacteria and
559 subsequently bacterial colonization.

560 We also found that trans-3-hydroxyproline degradation in bacteria is associated with
561 increased nematode fitness. In worms, hydroxyproline is present in collagen type IV, a
562 major component of the extracellular matrix in the pharynx, intestine and cuticle (68–70).
563 The breakdown of hydroxyproline can generate reactive oxygen species (71). These
564 may act as signaling molecules, which could affect cellular proliferation (72) and *C.*
565 *elegans* reproduction (73). Whether ROS in the gut increases brood size is unknown.
566 Alternatively, bacteria with the degradation pathway may utilize the amino acid as a
567 carbon source, consistent with the “microbiome on the leash” hypothesis, characterized
568 by host-selection of bacterial traits (74) and where nematodes indirectly benefit, if
569 damage is limited and if it allows worms to maintain a beneficial symbiont.

570 Our analysis of bacterial life history strategies (52,53) additionally identified
571 competitiveness and stress tolerance to associate with *C. elegans* colonization. It
572 appears that the ability to outcompete other microbes and/or tolerate stress
573 environments is an important requirement for a close relationship with the nematode. By
574 linking bacterial genomics, metabolism, and physiology with worm phenotypes, we were
575 thus able to generate testable hypotheses on traits and adaptive strategies important for
576 life in the worm.

577 In conclusion, our study provides a resource of naturally associated bacteria, their whole
578 genome sequences, and reconstructed metabolic competences that can be exploited to
579 study and understand *C. elegans* in an ecologically meaningful context. This resource
580 may help to further establish *C. elegans* as a versatile model for studying the genetics of
581 host-microbe interactions. It may also prove useful to characterize a variety of
582 phenotypes and underlying molecular mechanisms in *C. elegans*, which have thus far
583 only been studied in the complete absence of the worm's microbiome.

584

585 **Acknowledgements**

586 We thank Simone Severitt and Nicole Heyer for technical assistance regarding PacBio
587 genome sequencing, Jolantha Swiderski for long read genome assemblies, Peter
588 Deines for advice on BIOLOG assays, and the BiMo/LMB of Kiel University for access to
589 their core facilities. We are grateful for funding from the Deutsche
590 Forschungsgemeinschaft (DFG, German Research Foundation) within the Collaborative
591 Research Center CRC 1182 on Origin and Function of Metaorganisms, projects A1 (HS,
592 KD, ML), A4 (HS), and INF (MPH, CK). We further acknowledge funding by the DFG

593 under Germany's Excellence Strategy – EXC 2167-390884018 (Excellence Cluster
594 Precision Medicine in Chronic Inflammation; CK, HS), the Competence Center for
595 Genome Analysis Kiel (CCGA Kiel; HS), the Max-Planck Society (Fellowship to HS), and
596 the International Max-Planck Research School for Evolutionary Biology (NO).

597

598 **Conflict of Interest**

599 The authors declare no conflict of interest.

600

601 Supplementary information is available at ISME's website.

602 **References**

- 603 1. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE,
604 et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad*
605 *Sci United States Am.* 2013; **110**(9):3229–36.
- 606 2. Bosch TCG, Miller DJ. *The Holobiont Imperative*. Springer Vienna; 2016 [cited
607 2016]. Available from: <http://link.springer.com/10.1007/978-3-7091-1896-2>
- 608 3. Pascoe EL, Hauffe HC, Marchesi JR, Perkins SE. Network analysis of gut
609 microbiota literature: an overview of the research landscape in non-human animal
610 studies. *ISME J.* 2017;**11**(12):2644–51.
- 611 4. Wong CNA, Ng P, Douglas AE. Low-diversity bacterial community in the gut of
612 the fruitfly *Drosophila melanogaster*. *Environ Microbiol.* 2011;**13**(7):1889–900.
- 613 5. Consortium THMP. Structure, function and diversity of the healthy human
614 microbiome. *Nature.* 2012; 486:207–14.
- 615 6. Moran NA, Hansen AK, Powell JE, Sabree ZL. Distinctive gut microbiota of honey
616 bees assessed using deep sampling from individual worker bees. *PloS one.*
617 2012;**7**(4):e36393.
- 618 7. Louca S, Parfrey LW, Doebeli M. Decoupling function and taxonomy in the global
619 ocean microbiome. *Science.* 2016;**353**(6305):1272–7.
- 620 8. Kwong WK, Engel P, Koch H, Moran NA. Genomics and host specialization of
621 honey bee and bumble bee gut symbionts. *Proc Natl Acad Sci United States Am.*
622 2014;**111**(31):11509–14.

- 623 9. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict
624 metabolic and associated cellular functions. *Nat Rev Genet.* 2014;**15**:107–20.
- 625 10. Luan J-B, Chen W, Hasegawa DK, Simmons AM, Wintermantel WM, Ling K-S, et
626 al. Metabolic Coevolution in the Bacterial Symbiosis of Whiteflies and Related Plant
627 Sap-Feeding Insects. *Genome Biol Evol.* 2015;**7**(9):2635–47.
- 628 11. Ankrah NYD, Luan J, Douglas AE. Cooperative Metabolism in a Three-Partner
629 Insect-Bacterial Symbiosis Revealed by Metabolic Modeling. *J Bacteriol.* 2017;**199**(15).
- 630 12. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al.
631 Generation of genome-scale metabolic reconstructions for 773 members of the human
632 gut microbiota. *Nat Biotechnol.* 2017;**35**(1):81–9.
- 633 13. Bauer E, Laczny CC, Magnúsdóttir S, Wilmes P, Thiele I. Phenotypic
634 differentiation of gastrointestinal microbes is reflected in their encoded metabolic
635 repertoires. *Microbiome.* 2015;**3**:55.
- 636 14. Berg M, Stenuit B, Ho J, Wang A, Parke C, Knight M, et al. Assembly of the
637 *Caenorhabditis elegans* gut microbiota from diverse soil microbial environments. *ISME J*
638 2016;**10**(8):1998–2009.
- 639 15. Dirksen P, Marsh SA, Braker I, Heitland N, Wagner S, Nakad R, et al. The native
640 microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-
641 microbiome model. *BMC Biol.* 2016; 14:38.
- 642 16. Samuel BS, Rowedder H, Braendle C, Félix M-A, Ruvkun G. *Caenorhabditis*
643 *elegans* responses to bacteria from its natural habitats. *Proc Natl Acad Sci United States*
644 *Am.* 2016;**113**(27):E3941–9.

- 645 17. Zhang F, Berg M, Dierking K, Félix M-A, Shapira M, Samuel BS, et al. as a Model
646 for Microbiome Research. *Front Microbiol.* 2017; 8:485.
- 647 18. MacNeil LT, Watson E, Arda HE, Zhu LJ, Walhout AJM. Diet-induced
648 developmental acceleration independent of TOR and insulin in *C. elegans*. *Cell.*
649 2013;**153**(1):240–52.
- 650 19. Watson E, MacNeil LT, Ritter AD, Yilmaz LS, Rosebrock AP, Caudy AA, et al.
651 Interspecies Systems Biology Uncovers Metabolites Affecting *C. elegans* Gene
652 Expression and Life History Traits. *Cell.* 2014;**156**(6):1336–7.
- 653 20. Chaudhari SN, Mukherjee M, Vagasi AS, Bi G, Rahman MM, Nguyen CQ, et al.
654 Bacterial Folates Provide an Exogenous Signal for *C. elegans* Germline Stem Cell
655 Proliferation. *Dev cell.* 2016;**38**(1):33–46.
- 656 21. Virk B, Jia J, Maynard CA, Raimundo A, Lefebvre J, Richards SA, et al. Folate
657 Acts in *E. coli* to Accelerate *C. elegans* Aging Independently of Bacterial Biosynthesis.
658 *Cell reports.* 2016;**14**(7):1611–20.
- 659 22. Shapira M. Host-microbiota interactions in *Caenorhabditis elegans* and their
660 significance. *Curr Opin Microbiol.* 2017;**38**:142–7.
- 661 23. Cabreiro F, Au C, Leung K-Y, Vergara-Irigaray N, Cochemé HM, Noori T, et al.
662 Metformin retards aging in *C. elegans* by altering microbial folate and methionine
663 metabolism. *Cell.* 2013;**153**(1):228–39.
- 664 24. Scott TA, Quintaneiro LM, Norvaisas P, Lui PP, Wilson MP, Leung K-Y, et al.
665 Host-Microbe Co-metabolism Dictates Cancer Drug Efficacy in *C. elegans*. *Cell.*
666 2017;**169**(3):442–456.e18.

- 667 25. García-González AP, Ritter AD, Shrestha S, Andersen EC, Yilmaz LS, Walhout
668 AJM. Bacterial Metabolism Affects the *C. elegans* Response to Cancer
669 Chemotherapeutics. *Cell*. 2017;**169**(3):431–441.e8.
- 670 26. Norvaisas P, Cabreiro F. Pharmacology in the age of the holobiont. *Curr Opin*
671 *Syst Biol* [Internet]. 10.
- 672 27. Stiernagle T. Maintenance of *C. elegans*. *Wormb: Online Rev C elegans Biol*.
673 2006;1–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18050451>
- 674 28. von der Schulenburg JH, Hancock JM, Pagnamenta A, Sloggett JJ, Majerus ME,
675 Hurst GD. Extreme length and length variation in the first ribosomal internal transcribed
676 spacer of ladybird beetles (Coleoptera: Coccinellidae). *Mol Biol Evol*. **18**(4):648–60.
- 677 29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler
678 transform. *Bioinformatics*. 2009. **25**:1754–60.
- 679 30. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2:
680 somatic mutation and copy number alteration discovery in cancer by exome sequencing.
681 *Genome Res*. 2012;**22**(3):568–76.
- 682 31. Bolger AM, M Lohse BU. Trimmomatic: a flexible trimmer for Illumina sequence
683 data. *Bioinformatics*. 2014;**30**(15):2114-20.
- 684 32. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al.
685 SPAdes: a new genome assembly algorithm and its applications to single-cell
686 sequencing. *J Comput Biol*. 2012;**19**(5):455–77.
- 687 33. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*.
688 2014;**30**(14):2068–9.

- 689 34. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image
690 Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomic.*
691 2011;**12**:402.
- 692 35. Henry CS, DeJongh M, Best AA, Frybarger PM, Lindsay B, Stevens RL. High-
693 throughput generation, optimization and analysis of genome-scale metabolic models.
694 *Nat Biotechnol.* 2010;**28**(9):977–82.
- 695 36. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al.
696 The MetaCyc database of metabolic pathways and enzymes. *Nucleic acids Res.*
697 2018;**46**(D1):D633–9.
- 698 37. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.*
699 2014;**42**:D191–D198–D191–D198.
- 700 38. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined
701 dataset for big data analysis--10 years on. *Nucleic acids Res.* 2016;**44**(D1):D694–7.
- 702 39. Tange O. GNU Parallel - The Command-Line Power Tool. ;login: *USENIX Mag.*
703 2011;**36**(1):42–7.
- 704 40. Pagès H, Aboyou P, Gentleman R, DebRoy S. Biostrings: Efficient manipulation
705 of biological strings. R package version 2.50.1; 2018.
- 706 41. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA
707 ribosomal RNA gene database project: improved data processing and web-based tools.
708 *Nucleic acids Res.* 2013;**41**(Database issue):D590–6.

- 709 42. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.
710 RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic acids Res.*
711 2007;**35**(9):3100–8.
- 712 43. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in
713 hierarchical clustering. *Bioinformatics.* **22**:1540–2.
- 714 44. Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl*
715 *Stat.* 1979; **28**(1); 100-108.
- 716 45. R Foundation For Statistical Computing Austria V. R Development Core Team, R:
717 a language and environment for statistical computing. 2011;2.4.0 edit:3.
- 718 46. Wickham H. ggplot: Elegant Graphics for Data Analysis [Internet]. Use R.
719 Springer; 2009.
- 720 47. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating
721 signal peptides from transmembrane regions. *Nat methods.* 2011;**8**(10):785–6.
- 722 48. Bauer E, Zimmermann J, Baldini F, Thiele I, Kaleta C. BacArena: Individual-
723 based metabolic modeling of heterogeneous microbes in complex communities. *PLoS*
724 *Comput Biol.* 2017; **13**(5):e1005544.
- 725 49. Gelius-Dietrich G, Desouki AA, Fritzemeier CJ, Lercher MJ. Sybil--efficient
726 constraint-based modelling in R. *BMC Syst Biol.* 2013; **7**:125.
- 727 50. Heinken A, Sahoo S, Fleming RMT, Thiele I. Systems-level characterization of a
728 host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes.* 2013;**4**(1):28–
729 40.

- 730 51. Genuer R, Poggi J-M, Tuleau-Malot C. VSURF: an R package for variable
731 selection using random forests. 7th ed. 2015. p. 19–33.
- 732 52. Grime JP. Evidence for the Existence of Three Primary Strategies in Plants and
733 Its Relevance to Ecological and Evolutionary Theory. 1977. *Am Nat.*
734 **111**(982):1169:1194.
- 735 53. Fierer N. Embracing the unknown: disentangling the complexities of the soil
736 microbiome. *Nat Rev Microbiol.* 2017;**15**(10):579–90.
- 737 54. Chain PSG, Lang DM, Comerci DJ, Malfatti SA, Vergez LM, Shin M, et al.
738 Genome of *Ochrobactrum anthropi* ATCC 49188 T, a versatile opportunistic pathogen
739 and symbiont of several eukaryotic hosts. *J Bacteriol.* 2011;**193**(16):4274–5.
- 740 55. Kissoyan KA, Drechsler M, Stange E, Zimmermann J, Kaleta C, Bode H, et al.
741 Natural *C. elegans* microbiota protects against infection via production of a cyclic
742 lipopeptide of the viscosin group. *Curr Biol.* (in press).
- 743 56. Sieber M, Pita L, Weilan-Bräuer N, Dirksen P, Wang J, Mortzfeld B, et al. The
744 Neutral Metaorganism. bioRxiv. 2018. p. 367243.
- 745 57. Bito T, Matsunaga Y, Yabuta Y, Kawano T, Watanabe F. Vitamin B12 deficiency
746 in *Caenorhabditis elegans* results in loss of fertility, extended life cycle, and reduced
747 lifespan. *FEBS open bio.* 2013 ;**3**:112–7.
- 748 58. Watson E, MacNeil LT, Ritter AD, Yilmaz LS, Rosebrock AP, Caudy AA, et al.
749 Interspecies systems biology uncovers metabolites affecting *C. elegans* gene
750 expression and life history traits. *Cell.* 2014;**156**(4):759–70.

- 751 59. Watson E, Olin-Sandoval V, Hoy MJ, Li C-H, Lousse T, Yao V, et al. Metabolic
752 network rewiring of propionate flux compensates vitamin B12 deficiency in *C. elegans*.
753 *eLife*. 2016;5.
- 754 60. Hoek TA, Axelrod K, Biancalani T, Yurtsev EA, Liu J, Gore J. Resource
755 Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-
756 Feeding Mutualism. *PLoS Biol*. 2016 **14**(8):e1002540.
- 757 61. Zomorodi AR, Segrè D. Genome-driven evolutionary game theory helps
758 understand the rise of metabolic interdependencies in microbial communities. *Nat*
759 *Commun*. 2017;**8**(1):1563.
- 760 62. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: Networks,
761 competition, and stability. *Science*. 2015;**350**(6261):663–6.
- 762 63. Ryan DA, Miller RM, Lee K, Neal SJ, Fagan KA, Sengupta P, et al. Sex, age, and
763 hunger regulate behavioral prioritization through dynamic modulation of chemoreceptor
764 expression. *Curr Biol*. 2014;**24**(21):2509–17.
- 765 64. Zhang Y, Chou JH, Bradley J, Bargmann CI, Zinn K. The *Caenorhabditis elegans*
766 seven-transmembrane protein ODR-10 functions as an odorant receptor in mammalian
767 cells. *Proc Natl Acad Sci United States Am*. 1997;**94**(22):12162–7.
- 768 65. Sengupta P, Chou JH, Bargmann CI. odr-10 encodes a seven transmembrane
769 domain olfactory receptor required for responses to the odorant diacetyl. *Cell*.
770 1996;**84**(6):899–909.

- 771 66. Choi JI, Yoon K-H, Subbammal Kalichamy S, Yoon S-S, Il Lee J. A natural odor
772 attraction between lactic acid bacteria and the nematode *Caenorhabditis elegans*. *ISME*
773 *J.* 2016;**10**(3):558–67.
- 774 67. Baiocchi T, Lee G, Choe D-H, Dillman AR. Host seeking parasitic nematodes use
775 specific odors to assess host resources. *Sci reports.*;7(1):6270.
- 776 68. Page AP, Johnstone IL. The cuticle. *Wormb: Online Rev C elegans Biol.* 2007;1–
777 15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18050497>
- 778 69. Hutter H, Vogel BE, Plenefisch JD, Norris CR, Proenca RB, Spieth J, et al.
779 Conservation and novelty in the evolution of cell adhesion and extracellular matrix
780 genes. *Science.* 2000;**287**(5455):989–94.
- 781 70. Graham PL, Johnson JJ, Wang S, Sibley MH, Gupta MC, Kramer JM. Type IV
782 collagen is detectable in most, but not all, basement membranes of *Caenorhabditis*
783 *elegans* and assembles on tissues that do not express it. *J cell Biol.* 1997;**137**(5):1171–
784 83.
- 785 71. Cooper SK, Pandhare J, Donald SP, Phang JM. A novel function for
786 hydroxyproline oxidase in apoptosis through generation of reactive oxygen species. *J*
787 *Biol Chem.* 2008; **283**(16):10485–92.
- 788 72. Halliwell B. Biochemistry of oxidative stress. *Biochem Soc Trans.* 2007;**35**(Pt
789 5):1147–50.
- 790 73. De Henau S, Tilleman L, Vangheel M, Luyckx E, Trashin S, Pauwels M, et al. A
791 redox signalling globin is essential for reproduction in *Caenorhabditis elegans*. *Nat*
792 *Commun.* 2015;**6**:8782.

- 793 74. Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S. The evolution of the host
794 microbiome as an ecosystem on a leash. *Nature* .2017;**548**(7665):43–51.