

# 1 Deconstructing higher-order interactions in the microbiota: A theoretical 2 examination 3

4 Yitbarek, Senay<sup>1</sup>, Guittar, John<sup>2,3</sup>, Knutie, Sarah A.<sup>4</sup>, Ogbunugafor, C. Brandon<sup>5</sup>  
5

- 6 1. Department of Integrative Biology, University of California Berkeley, Berkeley, CA 94720 USA  
7 2. Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824  
8 USA  
9 3. Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, USA  
10 4. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269 USA  
11 5. Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912 USA  
12

## 13 Abstract

- 14 1. The animal gut is a complex ecosystem containing many interacting species. A major  
15 objective of microbiota research is to identify the scale at which gut taxa shape hosts.  
16 However, most studies focus solely on pairwise interactions and ignore higher-order  
17 interactions involving three or more component taxa. Higher-order interactions  
18 represent non-additive effects that cannot be predicted from first-order or pairwise  
19 interactions.  
20 2. Possible reasons as to why studies of higher higher-order interactions have been  
21 scarce is that many host-associated systems are experimentally intractable, gut  
22 microbiota are prohibitively species rich, and the influence of any given taxon on  
23 hosts is often context-dependent. Furthermore, quantifying emergent effects that  
24 represent higher-order interactions that are not simply the result of lower-order  
25 interactions, present a combinatorial challenge for which there are few well-  
26 developed statistical approaches in host-microbiota studies.  
27 3. In this perspective, our goal is to quantify the existence of emerging higher-order  
28 effects and characterize their prevalence in the microbiota. To do so, we adapt a  
29 method from evolutionary genetics used to quantify epistatic effects between  
30 mutations and use it to quantify the effects of higher-order microbial interactions on  
31 host infection risk.  
32 4. We illustrate this approach by applying it to an *in silico* dataset generated to resemble  
33 a population of hosts with gut-associated microbial communities. We assign each host  
34 a pathogen load, and then determine how emergent interactions between gut taxa  
35 influence this host trait.  
36 5. We find that the effect of higher-order interactions generally increases in magnitude  
37 with the number of species in the gut community. Based on the average magnitude of  
38 interaction for each order, we find that 9<sup>th</sup> order interactions have the largest non-  
39 linear effect on determining host infection risk.  
40 6. Our approach illustrates how incorporating the effects of higher-order interactions  
41 among gut microbiota can be essential for understanding their effects on host  
42 infection risk. We conclude that insofar as higher-order interactions between taxa may  
43 profoundly shape important organismal phenotypes (such as susceptibility to  
44 infection), that they deserve greater attention in microbiome studies.

45

46

## 47 **Introduction**

48 Animal guts contain complex microbial communities whose structure and function  
49 depend upon the interactions among microbes and the host. Gut microbiota serve as  
50 key actors in host health, impacting development, metabolism, and pathogen  
51 susceptibility (Brugman et al., 2018). The development of microbe-free (also known as  
52 germ-free) model hosts has made it possible to experimentally study how the  
53 microbiota influences host susceptibility to infection (Goodman et al., 2011; Ridaura et  
54 al., 2013). However, most studies rely on correlations between the relative abundances  
55 of individual bacterial taxa and host infection risk (e.g. pathogen load), ignoring the  
56 potential influence of higher-order interactions between taxa within the community.  
57 The field of complex systems is increasingly interested in understanding the emergent  
58 properties of higher-order interactions between objects (Lambiotte, Rosvall, & Scholtes,  
59 2019a). Relatedly, a long-standing issue in ecology is to capture the vast diversity of  
60 multispecies species interactions—the unpredictable effects that arise when multiple  
61 species are present in an ecosystem (Hutchinson 1962). For example, the order of  
62 arrival of species into an ecosystem, and other factors (deterministic or stochastic in  
63 nature) can dictate species composition and the overall behavior of the system  
64 (Saavedra et al., 2017; Uricchio, Daws, Spear, & Mordecai, 2019). This problem has more  
65 recently become the object of inquiry in communities of microbes (Enke et al., 2019;  
66 Mickalide & Kuehn, 2019; Sanchez-Gorostiaga, Baji •, Osborne, Poyatos, & Sanchez,

67 2018). Many ecological studies involving complex network structures typically focus on  
68 pair-wise interactions and tend to ignore higher-order effects among three or more  
69 components (Kareiva, 1994; Levine, Bascompte, Adler, & Allesina, 2017; Mayfield &  
70 Stouffer, 2017). For example, in a system with two interacting microbes—*A* and *B*—the  
71 addition of a third microbe *C* may alter the pairwise interaction between *A* and *B* in a  
72 non-linear or non-intuitive fashion. This would constitute an emergent higher-order  
73 interaction between *A*, *B* and *C*. This is in contrast to a scenario where the microbe *C*  
74 interacts with either *A* or *B* in isolation, which constitute pairwise interactions with their  
75 own interaction effects. Therefore, quantifying emergent higher-order effects between  
76 microbial taxa is necessary to fully capture the structure and dynamics of biological  
77 systems.

78  
79 Higher-order interactions have recently been the object of study in the realm of genetics,  
80 where they are discussed in light of epistasis, or non-linear interactions between genes  
81 and mutations (Mackay & Moore, 2014; Weinreich, Lan, Jaffe, & Heckendorn, 2018a;  
82 Weinreich, Lan, Wylie, & Heckendorn, 2013). A useful non-technical definition of  
83 epistasis is the “surprise at the phenotype when mutations are combined, given the  
84 constituent mutations’ individual effects (Weinreich, Lan, Jaffe, & Heckendorn, 2018b).  
85 This effectively captures what makes epistasis a provocative concept: the notion that  
86 interacting objects or parcels can have effects that are non-additive. In particular,  
87 higher-order epistasis is of interest, as it comprises all of the complexity and challenges

88 of understanding and studying higher-order interactions in other systems (Lambiotte et  
89 al., 2019a).

90

91 Higher-order epistasis can have powerful effects on organismal phenotypes, which has  
92 complicated the genotype-phenotype mapping problem in genetics (Sackton & Hartl,  
93 2016). To study higher-order epistasis in model organisms, molecular biologists  
94 engineer genes and mutations of interest in all possible permutations, a method labeled  
95 the “combinatorial approach.” (Weinreich et al., 2018b, 2013). Other studies resolve  
96 higher-order epistasis through more advanced statistical methods (Guerrero, Scarpino,  
97 Rodrigues, Hartl, & Ogbunugafor, 2019; Otwinowski, McCandlish, & Plotkin, 2018;  
98 Poelwijk, Krishna, & Ranganathan, 2016; Sailer & Harms, 2017).

99

100 Insect gut microbiota have been used as model systems to study the formation and  
101 assembly of microbial communities. Insect guts harbor relatively fewer microbial  
102 species, as compared to higher eukaryote hosts, with restricted core-members that can  
103 be grown axenically and manipulated genetically (Zheng, Steele, Leonard, Motta, &  
104 Moran, 2018). The protective function of microbes against invading pathogens have  
105 been studied across a range of insect hosts. For example, previous studies with bees  
106 found that core gut species were associated with increased host health, while non-core  
107 taxa were associated with decreased host health and increased pathogen infection  
108 (Cariveau, Elijah Powell, Koch, Winfree, & Moran, 2014; Koch & Schmid-Hempel, 2011;  
109 Raymann & Moran, 2018). However, other studies have also shown that pathogens alter

110 the gut microbiota and facilitate gut infections (Abraham et al., 2017; Wei et al., 2017).

111 Although many studies have shown correlations between core species and host traits,

112 the extent to which individual versus species interactions facilitate or resist gut

113 infections remains understudied.

114

115 Not unlike genomes, societies or neural circuits, insect gut microbiomes are complex

116 systems defined by the interaction between individual parcels (component taxa in the

117 microbiota). Consequently, we might predict that higher-order interactions between

118 taxa in the microbiota might underlie microbiota-associated organismal phenotypes,

119 such as susceptibility to infection. Recent work by Gould et al. 2018 found that higher-

120 order interactions in the gut microbiota impact lifespan, fecundity, development time,

121 and bacterial composition of *Drosophila* sp. With a gut community composed of 5 core

122 taxa, they found that three-way, four-way, and five-way interactions accounted for 13-

123 44% of all possible cases depending on the host trait. Yet, lower-order interactions (2-

124 pairs) still accounted for at least half of all the observed phenotypes in the system

125 (Gould et al., 2018).

126

127 Studies like Gould et al. 2018 provide an example of how higher-order interactions can

128 be measured and suggest that they might be relevant for understanding how taxa

129 influence certain phenotypes. But while the importance of diversity and host

130 interactions is clear, no studies have attempted to specifically disentangle effects

131 beyond four or five-way interactions. One major barrier to more of these studies is the

132 paucity (or non-existence) of the datasets structured like those in an evolutionary  
133 genetics framework, such that existing statistical methods might be used to resolve  
134 interactions. (Tekin, Savage, & Yeh, 2017; Wood, Nishida, Sontag, & Cluzel, 2012). For  
135 example, the problem of constructing a set of insects that each carry a different  
136 combination of constituent taxa of interest grows exponentially with the number of  
137 taxa. And (perhaps) unlike genetics, constructing a different insect with a different set  
138 of bacterial taxa (corresponding with the possible combinations of taxa) is a non-trivial  
139 technical challenge. Nonetheless, the use of combinatorial complete datasets—insects  
140 containing all combinations of taxa— to explore higher-order interactions (beyond a  
141 single taxon or pairwise interactions) could help to inform how taxa interact in framing  
142 organismal phenotypes.

143

144 In this commentary, we propose a theoretical examination of higher order interactions  
145 in the gut microbiome. Specifically, we employ the Walsh-Hadamard transform  
146 (WHT), a mathematical regime that has been used to demonstrate how higher-order  
147 interactions between mutations influence fitness or other organismal traits (Poelwijk et  
148 al., 2016; Weinreich et al., 2013), to explore how higher-order interactions among gut  
149 taxa can influence host infection risk. We use it to quantify higher-order interactions in  
150 an *in silico* dataset resembling the type of data that can be empirically—that can be  
151 developed in the future—collected from insect guts. We introduce this approach with  
152 the hope that it may eventually be applied to a tractable experimental system for real-

153 world validation, and believe that insect systems are among the most promising  
154 empirical systems.

155

## 156 **Methods**

157 The Walsh-Hadamard Transform allows one to quantify the eminence of interaction  
158 effects of different order in a system of potentially-interacting objects or parcels. It  
159 yields a Walsh coefficient, which communicates the magnitude and sign of how a  
160 particular order interaction influences an output of interest. It implements phenotypic  
161 values in the form of a vector, before reformatting it into a Hadamard matrix (and is  
162 then scaled by a diagonal matrix). The output is a collection of coefficients which  
163 measure the degree to which the map is linear, or second order, third, and so forth. We  
164 provide a brief primer on the method, and refer readers to two published  
165 manuscripts—Poelwijk et al. (2016) and Weinreich et. al. (2013)—that outline and apply  
166 the method in good detail. Also see the Supplementary Information for a brief primer.

167

168 The Walsh-Hadamard Transform relies on the existence of combinatorial data sets,  
169 where the objects for which we are interested in understanding the interactions between  
170 (taxa in this study) are constructed in all possible combinations. Another limitation of  
171 the WHT is that it can only accommodate two variants per site, that is, two states per  
172 actor. In the case of taxa, we can think of this in terms of the presence/absence of a  
173 certain taxon, and we can encode this in terms of 0 (absence) or 1 (presence). For each  
174 hypothetical insect with a different presence/absence combination, we have a

175 corresponding phenotypic measurement (e.g. parasite load). For example, if we wanted  
176 to measure the higher-order interactions between 4 taxa within an insect with regards to  
177 their role in parasite load (as a model phenotype), we would need  $2^L = 16$  individual  
178 measurements (insects in this case), with  $L$  corresponding to the number of different  
179 taxa whose effects we were interested in disentangling. We can encode this  
180 combination of 4 taxa in bit string notation (see Figure 1).

181  
182 Each site (0 or 1) in the string corresponds to the presence or absence of a given taxa in a  
183 given insect. This notation allows us to keep a mental picture of which taxa are in which  
184 insect for which we have a phenotypic measurement and can be used to construct a  
185 vector of values. For example, the string 1010 corresponds to an insect with the pattern  
186 of present (1), absent (0), present (1), absent (0). The full data set includes a vector of  
187 phenotypic values for all possible combinations of taxa—0000, 0001, 0010, 0100, 1000,  
188 0011, 0101, 0110, 1001, 1010, 1100, 0111, 1101, 1011, 1110, 1111. Note that these can be  
189 divided into different classes based on the “order” of the interaction. Order corresponds  
190 to the number of interacting actors. “Zeroth order” would correspond to the 0000  
191 variant. This would translate to an insect that has none of the insect taxa present. There  
192 are 4, 1<sup>st</sup> order interactions (0001, 0010, 0100, 1000), 6, 2<sup>nd</sup> order (or pairwise) interactions  
193 (0011, 0101, 0110, 1001, 1010, 1100), 4, third-order interactions (0111, 1101, 1011, 1110),  
194 and 1 fourth order interaction (1111). The WHT will quantify



195 This vector of phenotypic values for the 16 will be multiplied by a (16 x 16) square  
196 matrix, which is the product of a diagonal matrix  $V$  and a Hadamard matrix  $H$ . These  
197 matrices are defined recursively by:

$$V_{n+1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -V_n \end{pmatrix}, V_0 = 1 \quad [1]$$

198  
199

$$H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}, V_0 = 1 \quad [2]$$

200  
201

202  $n$  is the number of loci ( $n = 4$  in this hypothetical example). This matrix multiplication  
203 gives an output:

$$\gamma = VIIx$$

204  
205

206 Where  $V$  and  $H$  are the matrices described in [1] and [2] above, and  $y$  is the Walsh  
207 coefficient, the measure of the interaction between parcels of information in a string.  
208 Using this, we compute  $y$  values for every possible interaction between bits in a given  
209 string. The *in silico* generated data discussed in this commentary are composed of 10-bit  
210 strings, each corresponding to the presence/absence of a different microbial taxa. Such  
211 a case would have  $2^{10} = 1024$  total combinations of taxa, and corresponding phenotypic  
212 measurements (parasite load).

213

214 Similar to the 4-bit string example used to explain the method, note that each order has  
215 a different number of possible combinations. That is, the number of insects that can  
216 carry a combination of interacting taxa of a certain order. These are as follows: 0<sup>th</sup> = 1; 1<sup>st</sup>  
217 = 10, 2<sup>nd</sup> = 45; 3<sup>rd</sup> = 120; 4<sup>th</sup> = 210; 5<sup>th</sup> = 252; 6<sup>th</sup> = 210; 7<sup>th</sup> = 120; 8<sup>th</sup> = 45; 9<sup>th</sup> = 10; 10<sup>th</sup> = 1. The  
218 methods offered here measure every one of these interactions (e.g. all 210 of the possible  
219 6<sup>th</sup> order interactions) between taxa. While our use of a 10-bit string (as opposed to an 8  
220 or 15 bit string) is rather arbitrary, it is meant to highlight the vastness of the higher-  
221 interaction problem: Even if we suspect that only 10 taxa are meaningfully influencing a  
222 phenotype of interest (many studies contain more), the possible ways that these species  
223 are interacting, and the number of measurable coefficients between them can be  
224 astronomical in number.

225  
226 Having outlined the method used to quantify higher-order interactions above, it is  
227 important to directly explain the presumptive biological interpretation of the values.  
228 The WHT returns a Walsh coefficient for each “order” of interaction. This corresponds  
229 to the relative strength or importance of that “order” in the phenotype being measured.  
230 Therefore, the Walsh-Hadamard Transform can help to interpret the overall presence  
231 and eminence of higher-order interactions between taxa in a microbiota.

232

## 233 **Results**

234

235 Figure 2 depicts an *in silico* generated collection of 1024 insects, each containing one of  
236 the combinations of 10 taxa ( $2^{10} = 1024$ ), organized into a fitness graph (see

237 Supplemental Information for details on the in silico code and dataset). Each individual  
238 also has a parasite load. While other statistical methods may not require all possible  
239 combinations of taxa in order to extract meaningful information on the magnitude of  
240 higher-order interactions, creating the combinatorial set demonstrates the size and  
241 shape of the problem, all of the possible ways that taxa could interact.

242

243 Figure 3 depicts the raw calculations of the Walsh coefficients for all of the higher-order  
244 interactions (orders 2 – 9). Here we observe that the magnitude and direction of the  
245 interaction effect (Walsh coefficient) varies across different combinations of taxa. That  
246 the Walsh Hadamard Transform can disentangle these types of effects is a feature of the  
247 calculation and reveals the possibilities that exist in complex systems—like the  
248 microbiota—where many different objects are interacting. It is especially important to  
249 note that the specific identity of the taxa present is very important to understand in  
250 determining their interaction. We cannot assume that, for example, all third-order  
251 interactions (interactions between three taxa) will have the same magnitude or direction  
252 of interaction (e.g. positive or negative).

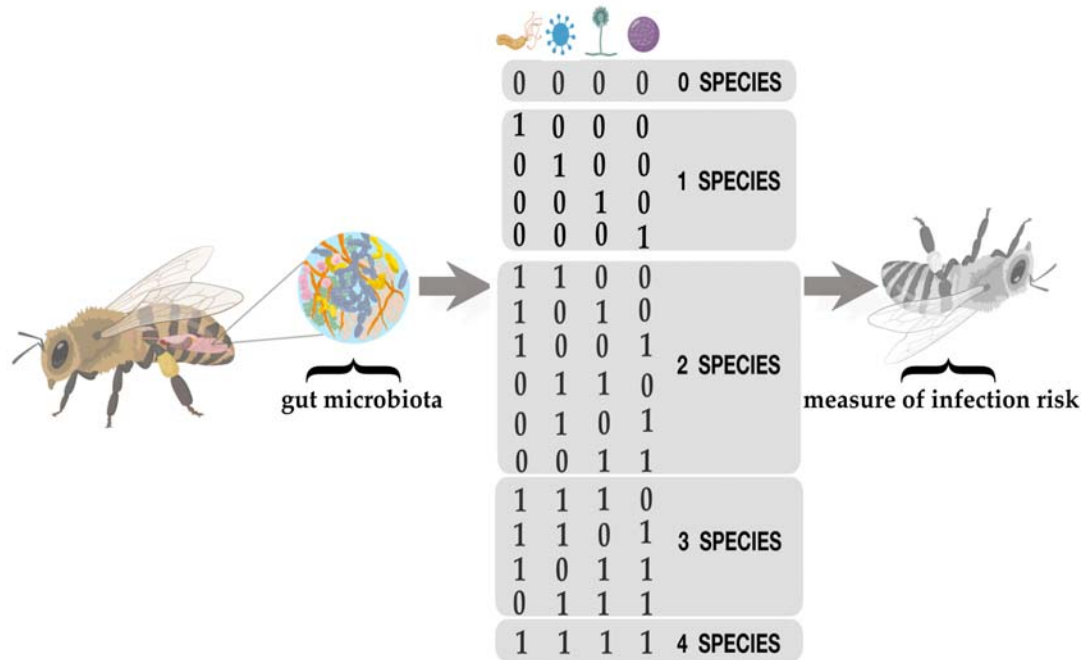
253

254 Figure 4 demonstrates the sum of the absolute values of the interaction coefficients  
255 highlighted in Figure 3. Here, we can observe the raw magnitude (leaving the sign—  
256 positive or negative—of it aside) of higher-order interactions as a function of interaction  
257 order. Between 1<sup>st</sup> and 9<sup>th</sup> order, higher-order effects increase, suggesting that they  
258 become more *meaningful* with the number of interacting microbes. Without knowing

259 the specific mechanism at work, determining the mean magnitude of coefficients  
260 provides relevant information on the eminence of a given order in the microbiota. For  
261 example, in our *in silico* microbiota the 9th order taxa represents the highest magnitude  
262 of interaction relative to other taxa orders (Figure 4). As this is a theoretical, *in silico*  
263 generated microbiota, we can interpret this finding as meaning that 9<sup>th</sup> order  
264 interactions contain the largest average deviation from additivity. That is, knowledge of  
265 how any given 9 taxa will interact requires very specific information on the identity of  
266 which 9 taxa are interacting. This is a characteristic of a highly non-linear, complex  
267 systems.

268

269 Note that all of these values—the raw *in silico* parasite load data, the interaction  
270 coefficients for all individual interactions, and the scaled, absolute value coefficients—  
271 can be found in the Supplementary Material.

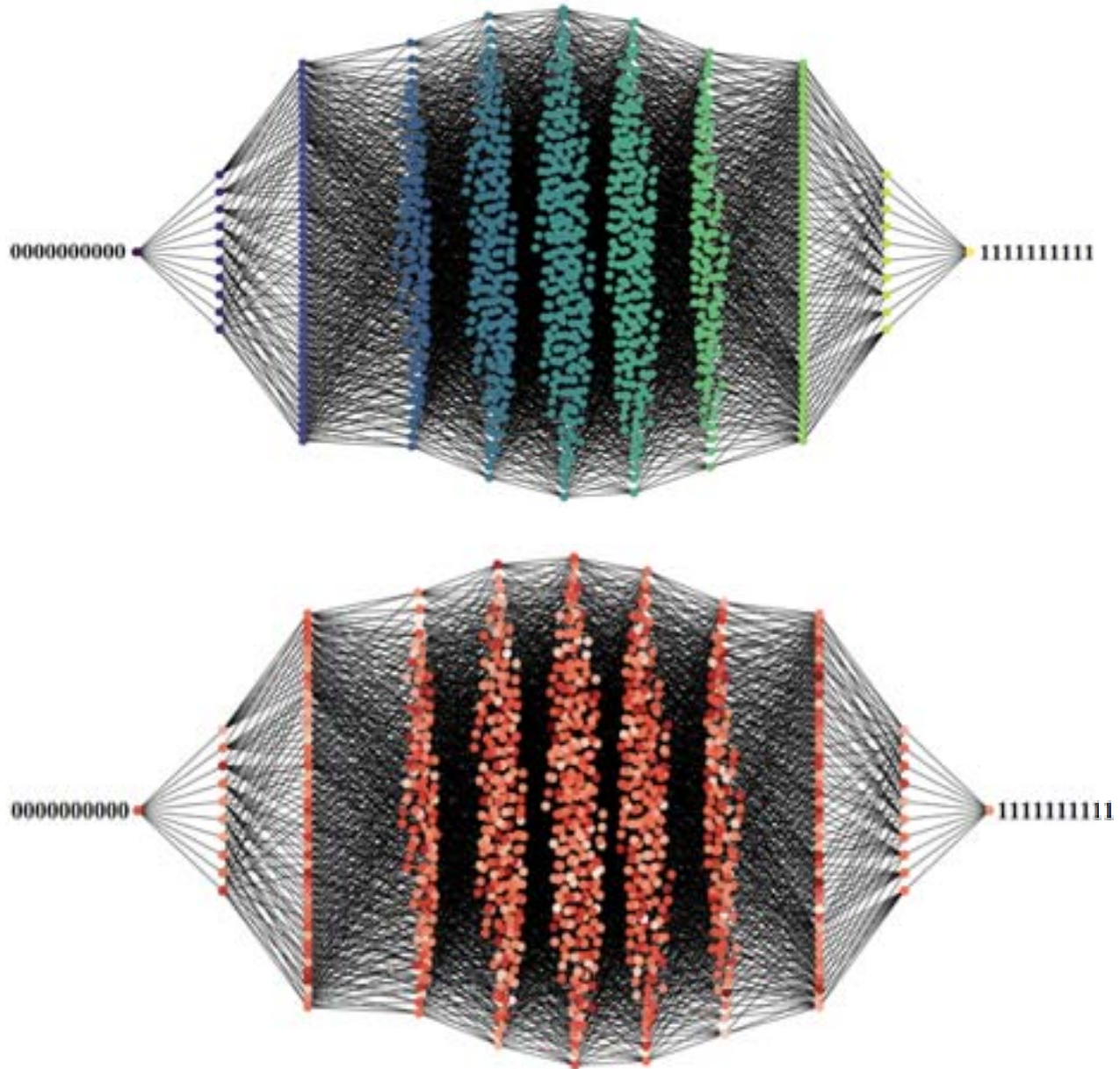


272

273 Figure 1. Schematic representation of higher-order interactions in the insect gut  
 274 microbiota. We represent the presence of microbial species in the gut similarly to the  
 275 presence of a genetic locus. Species composition are represented in binary strings. In  
 276 this configuration, the combination 0011 represents both the presence and absence of  
 277 two species. For each string combination, we associated a phenotypic measurement,  
 278 such as infection risk. We quantify “epistatic” interactions between microbes in  $n$   
 279 dimensional space, where  $n$  represents the number of species interacting.

280

281

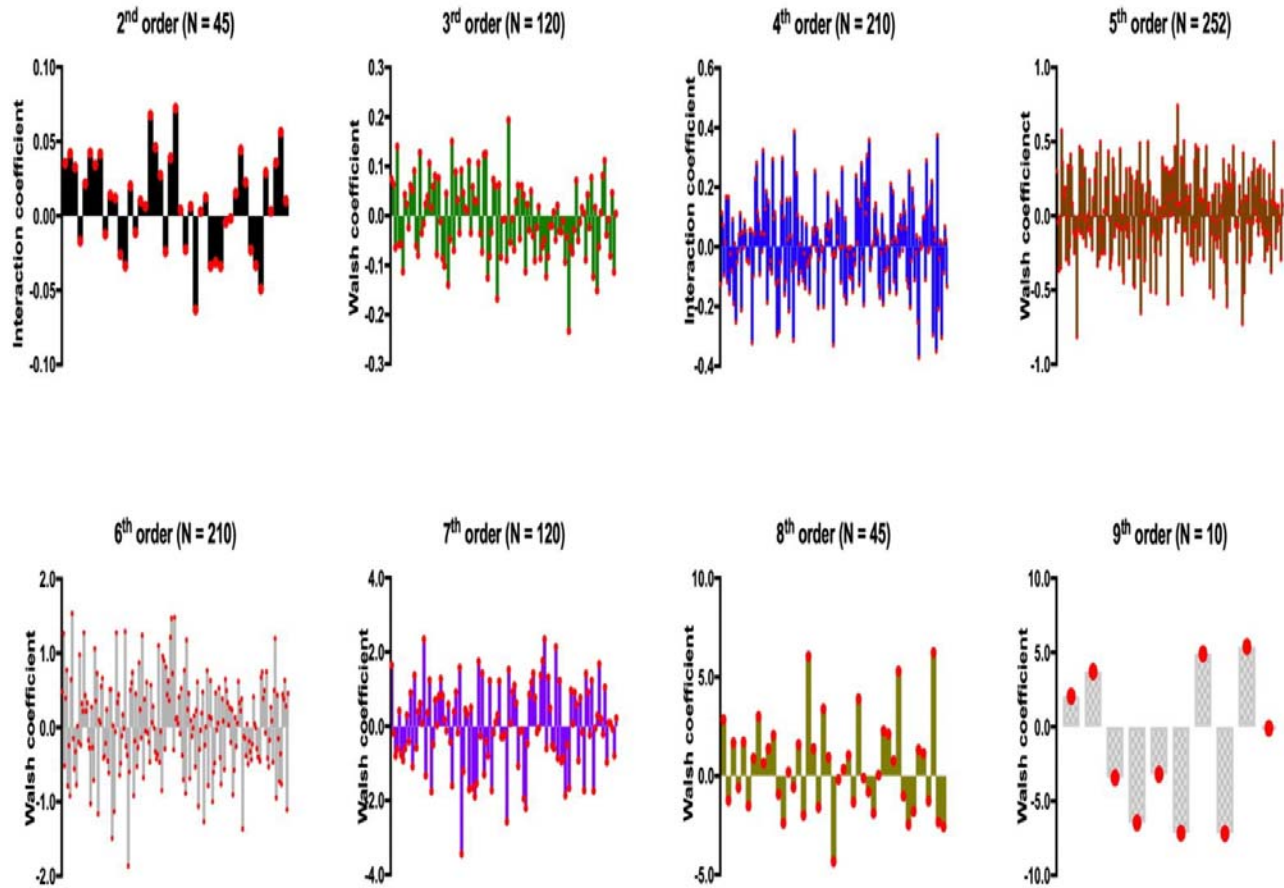


282

283

284 **Figure 2. An *in silico* generated microbiota with a combinatorial complete set of 10 taxa,**  
285 **composing 1024 total nodes ( $2^{10}$ ).** Edges link neighboring strings. In genetics, these edges  
286 represent genotypes that are mutational neighbors; in this data set they just denote  
287 insect microbiomes that differ by the presence/absence of a single taxa). (A) Color  
288 gradient represents the mutational class (order): lighter colors correspond to higher  
289 orders (5 taxa, 6 taxa, etc). (B) Color gradient represents actual parasite load  
290 measurements. Darker colors represent higher parasite loads.

291

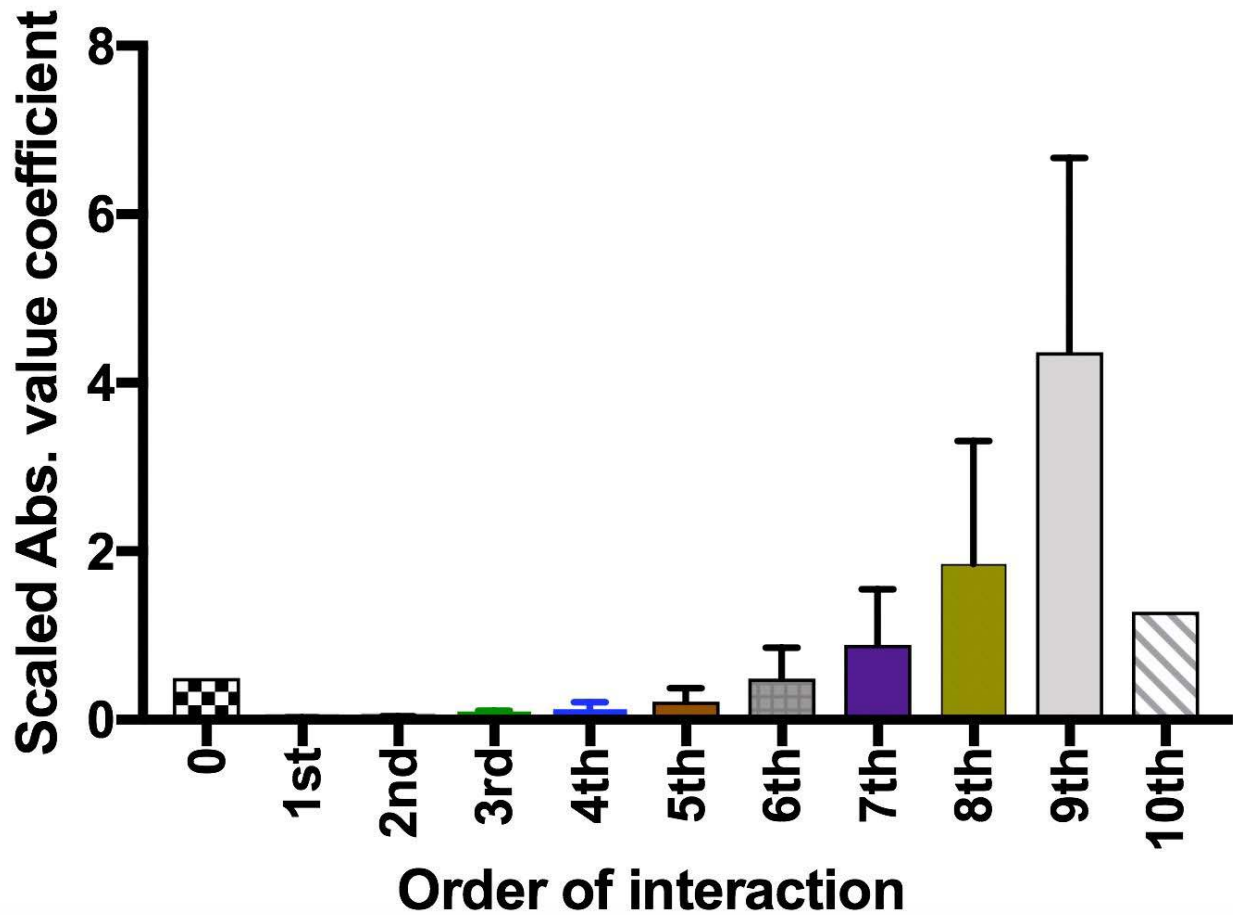


292

293 *Figure 3. The entire space of higher-order interactions between taxa in all possible combinations,*  
294 *organized by higher-order class. Each line represents an individual combination of the 10*  
295 *taxa, and the y-axis is the epistatic coefficient corresponding to that taxa. That is, for*  
296 *1001000001, there is a value corresponding to the magnitude of the 3<sup>rd</sup> order epistatic*  
297 *interaction between the presence of the taxa corresponding to first site, fourth site, and*  
298 *tenth site. 0<sup>th</sup> and 10<sup>th</sup> orders are missing from this figure because they are composed of a*  
299 *single interaction. 1<sup>st</sup> order is missing because it doesn't constitute "higher-order"*  
300 *interactions, but rather, the "main effects" of each taxon acting in isolation.*  
301 *Alternatively, higher order interactions (2<sup>nd</sup> through 9<sup>th</sup>) are all composed of multiple*  
302 *parcels interacting. Please note that the y-axes are different across figures.*

303

304



305

306 Figure 4. *The magnitude of interaction between taxa.* Absolute value, averaged magnitude  
307 interactions across interaction orders. These are the averaged, scaled, absolute values of  
308 the data show in Figure 3. The purpose of this depiction is to illustrate how the  
309 magnitude (not sign) of the interactions change with interaction order. In this *in silico*  
310 microbiota data set, this translates to the 9<sup>th</sup> order interactions (interactions between 9-  
311 different taxa) containing the most non-linearity, overall.

312

313

314

315

316

317



318

## 319 **Discussion**

320 In this commentary, we explore the possibility of higher-order interactions between  
321 taxa composing an insect gut microbiota. Using *in silico* and applied mathematical  
322 approaches, we demonstrate how higher-order interactions can be measured in a  
323 complex system of interacting microbial taxa. In our theoretical scenario, higher-order  
324 interactions are present and generally increase in relevance with the order of  
325 interaction. Though our results are theoretical, they are results nonetheless (Goldstein,  
326 2018), highlighting the vast scope of the higher-order interaction problem, and outline  
327 one method that can be used to deconstruct them in biological systems. Though  
328 empirical data of the size and scope used in this study are currently challenging to  
329 generate, this intractability may be temporary, and future methods may permit the  
330 generation of data similar in structure to those explored in our theoretical examination.

331

332 The approach used in this study—the Walsh-Hadamard Transform—has been  
333 previously used by theoretical population geneticists to measure non-linear interactions  
334 between mutations (Weinreich et al., 2013). Several empirical data sets in genetics and  
335 genomics have demonstrated that the sign of interaction effects can change readily with  
336 the identity of the interacting parcels (Guerrero et al., 2019; Weinreich et al., 2018a, 2013).  
337 Given this, we predict that the taxa that compose the gut microbiota might be similarly  
338 defined by higher-order interactions. The capacity for measuring the effects of higher-  
339 order interactions on host fitness is an important step towards understanding the effects

340 of microbiota on their host. Indeed, considering higher-order interactions can enable  
341 more robust information on non-linear interactions in microbiome communities.

342

343 We found that higher-order interactions were present, and that taxa interacted both  
344 positively and negatively. Combined interactions among taxa are augmented compared  
345 to what is expected from individual effects when phenotypic effects are positive. In  
346 contrast, higher-order effects are negative when combined interactions among taxa  
347 show a diminished return and are less fit than would be expected from their individual  
348 effects (fig 3). Such combinatorial complete data-sets can tell us what scale microbial  
349 interactions matter in predicting host infection. Moreover, they reveal patterns of  
350 interactions, particularly those combinations that interact synergistically or  
351 antagonistically (Hartl, 2014). One potential limitation of the outlined approach is the  
352 requirement for combinatorial complete datasets. For high-diversity microbiomes,  
353 including humans and plants, it is not currently feasible to carry out experiments  
354 measuring phenotypes for all the possible microbial interactions.

355

356 Microbe-mediated protection against pathogens depends on subtle differences in gut  
357 community structure. In North American wild bumble bees, lower *Chrithidia* parasite  
358 infection loads are associated with higher microbiota diversity. Using transplants to  
359 naive host, it was shown that the core-gut bacteria were responsible for conferring  
360 resistance to the *Chrithidia* parasite, while non-core gut bacteria were found to be less  
361 effective against the parasite (Mockler, Kwong, Moran, & Koch, 2018). In mosquitos, gut

362 bacterial species can trigger an immune defense against *Plasmodium* parasites, the  
363 causative agent of malaria (Bahia et al., 2014). In sandflies, highly diverse midgut  
364 microbiota's were found to be negatively correlated with the parasite that causes the  
365 vector-borne disease leishmaniasis (Kelly et al., 2017). While these studies did not  
366 investigate the effects of higher-order interactions on host fitness, future experimental  
367 studies manipulating microbial communities should consider combinatorial designs.

368

369 Recent theoretical work suggests that higher-order modeling approaches are able to  
370 capture volumes of rich data arising from complex ecological interactions (Lambiotte,  
371 Rosvall, & Scholtes, 2019b). In this perspective, we have adapted approaches from  
372 population genetics to the study of host-associated microbiota. Applying these methods  
373 to the analysis of real experiments will yield important insight into microbiome  
374 dynamics, towards a richer understanding of just how peculiar the microbiota is, and  
375 the many meaningful interactions that it embodies.

### 376 **Acknowledgements**

377 We wish to acknowledge the organizers and participants of the 2017 RCN-IDEAS  
378 arbovirus workshop held in New Orleans. SY acknowledges funding support from NSF  
379 Postdoctoral Fellowship award number 1612302. CBO acknowledges funding support  
380 from NSF RII Track-2 FEC award number 1736253. The authors would like to thank  
381 Victor Meszaros and Miles Miller-Dickson for their input on the *in silico* data, figures  
382 and Walsh-Hadamard primer. We finally thank Lawrence Uricchio for constructive  
383 feedback on our manuscript.

384

## 385 **Data Availability**

386 The *in silico* data used in this study and code used to generate them can be found on

387 github: <https://github.com/OgPlexus/MicrobeTaxa1>

388

## 389 **Supplemental Information**

390 The authors have prepared a simple mathematical primer on the Walsh-Hadamard

391 Transform: <https://github.com/OgPlexus/MicrobeTaxa1>. For a more rigorous

392 understanding, readers are encouraged to engage the works cited in this manuscript.

393

## 394 **Literature cited**

395

396 Abraham, N. M., Liu, L., Jutras, B. L., Yadav, A. K., Narasimhan, S., Gopalakrishnan, V., ...

397 Fikrig, E. (2017). Pathogen-mediated manipulation of arthropod microbiota to promote

398 infection. *Proceedings of the National Academy of Sciences*, 114(5), E781–E790. doi:

399 10.1073/pnas.1613422114

400 Bahia, A. C., Dong, Y., Blumberg, B. J., Mlambo, G., Tripathi, A., BenMarzouk-Hidalgo, O. J., ...

401 Dimopoulos, G. (2014). Exploring Anopheles gut bacteria for Plasmodium blocking

402 activity. *Environmental Microbiology*, 16(9), 2980–2994. doi: 10.1111/1462-2920.12381

403 Cariveau, D. P., Elijah Powell, J., Koch, H., Winfree, R., & Moran, N. A. (2014). Variation in gut

404 microbial communities and its association with pathogen infection in wild bumble bees

405 (*Bombus*). *The ISME Journal*, 8(12), 2369–2379. doi: 10.1038/ismej.2014.68

- 406 Enke, T. N., Datta, M. S., Schwartzman, J., Cermak, N., Schmitz, D., Barrere, J., ... Cordero, O. X.  
407 (2019). Modular Assembly of Polysaccharide-Degrading Marine Microbial  
408 Communities. *Current Biology*. doi: 10.1016/j.cub.2019.03.047
- 409 Goldstein, R. E. (2018). Are theoretical results ‘Results’? *ELife*, 7, e40018. doi:  
410 10.7554/eLife.40018
- 411 Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., ... Ludington, W.  
412 B. (2018). Microbiome interactions shape host fitness. *Proceedings of the National Academy  
413 of Sciences*, 115(51), E11951–E11960. doi: 10.1073/pnas.1809349115
- 414 Guerrero, R. F., Scarpino, S. V., Rodrigues, J. V., Hartl, D. L., & Ogbunugafor, C. B. (2019).  
415 Proteostasis Environment Shapes Higher-Order Epistasis Operating on Antibiotic  
416 Resistance. *Genetics*, genetics.302138.2019. doi: 10.1534/genetics.119.302138
- 417 Hartl, D. L. (2014). What Can We Learn From Fitness Landscapes? *Current Opinion in  
418 Microbiology*, 0, 51–57. doi: 10.1016/j.mib.2014.08.001
- 419 Kareiva, P. (1994). Special Feature: Higher Order Interactions as a Foil to Reductionist Ecology.  
420 *Ecology*, 75(6). doi: 10.2307/1939613
- 421 Kelly, P. H., Bahr, S. M., Serafim, T. D., Ajami, N. J., Petrosino, J. F., Meneses, C., ... Wilson, M.  
422 E. (2017). The Gut Microbiome of the Vector *Lutzomyia longipalpis* Is Essential for  
423 Survival of *Leishmania infantum*. *MBio*, 8(1). doi: 10.1128/mBio.01121-16
- 424 Koch, H., & Schmid-Hempel, P. (2011). Socially transmitted gut microbiota protect bumble bees  
425 against an intestinal parasite. *Proceedings of the National Academy of Sciences*, 108(48),  
426 19288–19292. doi: 10.1073/pnas.1110474108
- 427 Lambiotte, R., Rosvall, M., & Scholtes, I. (2019a). From networks to optimal higher-order models  
428 of complex systems. *Nature Physics*, 15(4), 313. doi: 10.1038/s41567-019-0459-y

- 429 Lambiotte, R., Rosvall, M., & Scholtes, I. (2019b). From networks to optimal higher-order  
430 models of complex systems. *Nature Physics*, 15(4), 313. doi: 10.1038/s41567-019-0459-y
- 431 Levine, J. M., Bascompte, J., Adler, P. B., & Allesina, S. (2017). Beyond pairwise mechanisms of  
432 species coexistence in complex communities. *Nature*, 546(7656), 56–64. doi:  
433 10.1038/nature22898
- 434 Mackay, T. F., & Moore, J. H. (2014). Why epistasis is important for tackling complex human  
435 disease genetics. *Genome Medicine*, 6(6), 42. doi: 10.1186/gm561
- 436 Mayfield, M. M., & Stouffer, D. B. (2017). Higher-order interactions capture unexplained  
437 complexity in diverse communities. *Nature Ecology & Evolution*, 1(3), 0062. doi:  
438 10.1038/s41559-016-0062
- 439 Mickalide, H., & Kuehn, S. (2019). Higher-order interaction inhibits bacterial invasion of a  
440 phototroph-predator microbial community. *BioRxiv*, 564260. doi: 10.1101/564260
- 441 Mockler, B. K., Kwong, W. K., Moran, N. A., & Koch, H. (2018). Microbiome Structure  
442 Influences Infection by the Parasite *Crithidia bombi* in Bumble Bees. *Appl. Environ.*  
443 *Microbiol.*, 84(7), e02335-17. doi: 10.1128/AEM.02335-17
- 444 Otwinowski, J., McCandlish, D. M., & Plotkin, J. B. (2018). Inferring the shape of global epistasis.  
445 *Proceedings of the National Academy of Sciences*, 115(32), E7550–E7558. doi:  
446 10.1073/pnas.1804015115
- 447 Poelwijk, F. J., Krishna, V., & Ranganathan, R. (2016). The Context-Dependence of Mutations: A  
448 Linkage of Formalisms. *PLOS Computational Biology*, 12(6), e1004771. doi:  
449 10.1371/journal.pcbi.1004771
- 450 Raymann, K., & Moran, N. A. (2018). The role of the gut microbiome in health and disease of  
451 adult honey bee workers. *Current Opinion in Insect Science*, 26, 97–104. doi:  
452 10.1016/j.cois.2018.02.012

- 453 Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B., & Levine, J. M. (2017). A  
454 structural approach for understanding multispecies coexistence. *Ecological Monographs*,  
455 87(3), 470–486. doi: 10.1002/ecm.1263
- 456 Sackton, T. B., & Hartl, D. L. (2016). Genotypic Context and Epistasis in Individuals and  
457 Populations. *Cell*, 166(2), 279–287. doi: 10.1016/j.cell.2016.06.047
- 458 Sailer, Z. R., & Harms, M. J. (2017). Detecting High-Order Epistasis in Nonlinear Genotype-  
459 Phenotype Maps. *Genetics*, 205(3), 1079–1088. doi: 10.1534/genetics.116.195214
- 460 Sanchez-Gorostiaga, A., Baji•, D., Osborne, M. L., Poyatos, J. F., & Sanchez, A. (2018). High-  
461 order interactions dominate the functional landscape of microbial consortia. *BioRxiv*,  
462 333534. doi: 10.1101/333534
- 463 Tekin, E., Savage, V. M., & Yeh, P. J. (2017). Measuring higher-order drug interactions: A review  
464 of recent approaches. *Current Opinion in Systems Biology*, 4, 16–23. doi:  
465 10.1016/j.coisb.2017.05.015
- 466 Uricchio, L. H., Daws, S. C., Spear, E. R., & Mordecai, E. A. (2019). Priority Effects and  
467 Nonhierarchical Competition Shape Species Composition in a Complex Grassland  
468 Community. *The American Naturalist*, 193(2), 213–226. doi: 10.1086/701434
- 469 Wei, G., Lai, Y., Wang, G., Chen, H., Li, F., & Wang, S. (2017). Insect pathogenic fungus interacts  
470 with the gut microbiota to accelerate mosquito mortality. *Proceedings of the National  
471 Academy of Sciences*, 114(23), 5994–5999. doi: 10.1073/pnas.1703546114
- 472 Weinreich, D. M., Lan, Y., Jaffe, J., & Heckendorn, R. B. (2018a). The Influence of Higher-Order  
473 Epistasis on Biological Fitness Landscape Topography. *Journal of Statistical Physics*,  
474 172(1), 208–225. doi: 10.1007/s10955-018-1975-3

- 475 Weinreich, D. M., Lan, Y., Jaffe, J., & Heckendorn, R. B. (2018b). The Influence of Higher-Order  
476 Epistasis on Biological Fitness Landscape Topography. *Journal of Statistical Physics*,  
477 *172*(1), 208–225. doi: 10.1007/s10955-018-1975-3
- 478 Weinreich, D. M., Lan, Y., Wylie, C. S., & Heckendorn, R. B. (2013). Should evolutionary  
479 geneticists worry about higher-order epistasis? *Current Opinion in Genetics &*  
480 *Development*, *23*(6), 700–707. doi: 10.1016/j.gde.2013.10.007
- 481 Wood, K., Nishida, S., Sontag, E. D., & Cluzel, P. (2012). Mechanism-independent method for  
482 predicting response to multidrug combinations in bacteria. *Proceedings of the National*  
483 *Academy of Sciences*, *109*(30), 12254–12259. doi: 10.1073/pnas.1201281109
- 484 Zheng, H., Steele, M. I., Leonard, S. P., Motta, E. V. S., & Moran, N. A. (2018). Honey bees as  
485 models for gut microbiota research. *Lab Animal*, *47*(11), 317–325. doi: 10.1038/s41684-  
486 018-0173-x
- 487
- 488