Deconstructing *taxa x taxa x environment* interactions in the microbiota: A theoretical examination

Yitbarek, Senay[1], Guittar, John[2,3], Knutie, Sarah A.[4,5]., Ogbunugafor, C. Brandon[6,7]

1. *Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA*
2. *Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824 USA*
3. *Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, USA*
4. *Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA*
5. *Institute for Systems Genomics, University of Connecticut, Storrs, CT, 06269, USA*
6. *Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520*
7. *Vermont Complex Systems Center, University of Vermont, Burlington, VT 05405*

Abstract

1. **A major objective of microbial ecology is to identify how the composition of gut microbial taxa shapes host phenotypes. However, most studies focus solely on community-level patterns and pairwise interactions and ignore the potentially significant effects of higher-order interactions involving three or more component taxa.**
2. **Studies on higher-order interactions among microbial taxa are scarce for many reasons, including experimental intractability, daunting diversity and complexity of many microbial systems, and the potential confounding role of the environment. Moreover, we still lack the empirical and statistical tools to isolate and understand the role of higher-order interactions on the host.**
3. **Here, we apply a mathematical approach to quantifying the effects of higher-order interactions among taxa on host infection risk. To do so, we adapt the Hadamard-Walsh method recently used in evolutionary genetics to quantify the nonlinear effects of mutations on fitness. We apply our approach to an *in silico* dataset built to resemble a population of insect hosts with gut-associated microbial communities at risk of infection from an intestinal parasite. Critically, we examine these interactions across a breadth of environmental contexts, using nutrient content of the insect diet as a model for context.**
4. **We find that the effect of higher-order interactions is considerable and can change appreciably across environmental contexts. Strikingly, the relative eminence of different orders (pairwise vs. third order, fourth order, and fifth order) changes as a function of environmental context. Furthermore, we show– in our theoretical microcosm– that higher-order interactions can stabilize community structure thereby reducing host susceptibility to parasite invasion.**
5. **Our approach illustrates how incorporating the effects of higher-order interactions among gut microbiota across environments can be essential for understanding their effects on host phenotypes. We conclude that higher-order interactions among taxa can profoundly shape important organismal phenotypes, and they deserve greater attention in host-associated microbiome studies.**

Keywords: Higher-order interactions, insect microbiota, microenvironments, network theory

Introduction

Animal guts contain complex microbial communities whose structure and function depend upon the interactions among microbes and the host. Gut microbiota serves as key actors in host health, impacting development, metabolism, and the immune system (Brugman et al., 2018; McFall-Ngai et al., 2013). The development of axenic and gnotobiotic model hosts has made it possible to experimentally study how the microbiota influences host traits of interest (Douglas, 2018). However, most studies rely on correlations between the relative abundances of individual microbial taxa and host traits (e.g. immune function), and also community-level patterns at family level taxonomic resolutions, ignoring the potential influence of higher-order interactions among taxa within the community (Hooper et al., 2012; Knutie et al., 2017; Macpherson & Harris, 2004; Round & Mazmanian, 2009).

The field of complex systems is increasingly interested in understanding the emergent properties of higher-order interactions (Battiston et al., 2020). Higher order interactions have been the object of relatively rigorous inquiry in the realm of genetics, where they are discussed in terms of epistasis, or non-linear interactions between genes and mutations (Mackay & Moore, 2014; Weinreich et al., 2013, 2018a). A useful non-technical definition of epistasis is "surprise at the phenotype when mutations are combined, given the constituent mutations' individual effects" (Weinreich et al., 2013). In particular, higher-order epistasis is of interest, as these interactions comprise all of the complexity and challenges of understanding and studying higher-order interactions in other systems, and even in microbes (Gould et al., 2018). Not unlike genomes, communities or neural circuits, insect gut microbiomes are complex systems defined by the interaction between individual entities or parcels of information (in this case, component

72 taxa in the microbiota). Consequently, we might predict that higher-order interactions between taxa in the
73 microbiota might underlie microbiota-associated organismal phenotypes.
74
75 A long-standing goal of ecology is to capture the vast diversity of multispecies interactions—the
76 unpredictable effects that arise when multiple species are present in an ecosystem (Barabás et al., 2016;
77 Chesson, 2000; Hutchinson, 1961; Mayfield & Stouffer, 2017; Vandermeer, 1969). For example, animals
78 harbor diverse microbial communities that are variable in their composition, governed by stochastic
79 processes, which influences the overall behavior of the system (Douglas, 2018). This problem has more
80 recently become the object of inquiry in communities of microbes (Enke et al., 2019; Guittar et al., 2019;
81 Mickalide & Kuehn, 2019; Sanchez-Gorostiaga et al., 2018). Many ecological studies involving complex
82 network structures typically focus on pairwise interactions (Kareiva, 1994; Levine et al., 2017; Mayfield
83 & Stouffer, 2017). Only very recently has the literature demonstrated that higher-order interactions are at
84 play in these systems, an important area for further inquiry, given how they may potentially complicate
85 (or even undermine) simple models of microbial community function (Sanchez-Gorostiaga et al., 2018).
86
87 Higher-order interactions in the gut microbiota of *Drosophila specie*s impact lifespan, fecundity,
88 development time, and community composition (Gould et al., 2018). With a gut community comprising
89 five core taxa, Gould et al. found that three-way, four-way, and five-way interactions accounted for 13-
90 44% of all possible cases depending on the host trait. Yet, lower-order interactions (2-pairs) still
91 accounted for at least half of all the observed phenotypes in the system. Work by Sanchez-Gorostiage et
92 al. (2018), examined the contributions of multispecies interactions to determining community function
93 (i.e. amylase expression). In the presence of higher-ordered interactions, the predictive power of the
94 additive null model (absence of interactions) in predicting community function decreases. However, by
95 accounting for both behavioral and population dynamics effects into their null model, higher-order
96 interactions did provide good predictions for community function. Higher-order interactions can have
97 important implications for the predictive power of bottom-up approaches to designing complex

98 communities and determining their functional traits (Sanchez-Gorostiaga et al., 2018). The
99 aforementioned studies provide examples of how higher-order interactions can be measured and suggest
100 that they are relevant for understanding how microbial taxa influence certain phenotypes. While the
101 importance of diversity and host interactions is clear, to our knowledge no studies have attempted to
102 specifically disentangle effects of higher-order interactions across environmental contexts.
103
104 One major barrier to more of these studies is the paucity (or non-existence) of the datasets structured like
105 those in an evolutionary genetics framework, such that existing statistical methods might be used to
106 resolve interactions (Tekin et al., 2017; Wood et al., 2012). For example, the problem of constructing a
107 set of insects that each carry a different combination of constituent taxa of interest grows exponentially
108 with the number of taxa. And unlike some genetic systems, constructing a different insect with a different
109 set of bacterial taxa (corresponding with the possible combinations of taxa) is currently a non-trivial
110 technical challenge. Nonetheless, the use of combinatorial complete datasets—insects containing all
111 combinations of taxa (even few in number)— to explore higher-order interactions (beyond a single taxon
112 or pairwise interactions) could help to inform how taxa interact in framing organismal phenotypes.
113 Higher-order interactions could, in principle, be used to examine how our predictions for taxa-taxa
114 interactions will be contingent on the host context in which a certain distribution of taxa exists.
115
116 In this study, we reframe how we consider higher-order interactions in an insect gut using theoretical
117 approaches. We apply a relatively simple mathematical method called the Walsh-Hadamard transform
118 (WHT), which has been used to demonstrate how higher-order interactions between mutations influence
119 fitness or other organismal traits (Poelwijk et al., 2016; Weinreich et al., 2013, 2018b). We use this
120 method to explore how higher-order interactions among gut taxa can influence host fitness, across micro-
121 environments. In this study, we use it to quantify higher-order interactions in an *in silico* dataset
122 resembling the type of data that can be collected, presently in genetic systems, and plausibly in the future
123 in microbiota experimental systems.

We have chosen to consider the nutritional environment of the host, as resources can vary due to spatial and temporal differences, and in terms of the quantity and quality of required resources. A key component of resource availability is nutrition, which is likely to influence host resistance to natural enemies. In microbial systems, increased resource availability resulted in greater host resistance to parasites (Gómez et al., 2015; Lopez-Pascua & Buckling, 2008). Lower resource levels have been found to be costly for resistance to parasites in Drosophila melanogaster (McKean et al., 2008). Nutritional content (quality and quantity) is a well-known stressor for insect microbes in many settings, including the gut microbiota (Engel & Moran, 2013; Gurung et al., 2019; Mereghetti et al., 2017; Skidmore & Hansen, 2017). However, experimental studies involving model systems rely on high nutritional diets to understand factors affecting susceptibility to infectious diseases (Roberts et al., 2019). In this work, we consider how varying nutritional environments influence host susceptibility to disease risk.

Using this framework, we are able to examine underappreciated aspects of the microbiota: questions surrounding the notion that the microenvironment of the insect gut may shape higher-order interactions between taxa, with important consequences for host health and fitness. Our study examines the consequences of higher-order microbial interactions for host susceptibility (i.e. phenotype of interest) to disease risk. We hypothesize that higher-order interactions underlie host microbiome robustness to intestinal parasite invasion, reducing host susceptibility to disease risk, and that these interactions are highly dependent on environmental context. While this study is designed to address standing questions about interactions within the microbiota, it also offers future directions. We introduce this approach with the hope that it, or a related method may eventually be applied to a tractable experimental system for real-world validation and believe that insect systems are among the most promising candidates for these examinations.

Methods

Data source

The data used in this study arise from raw data used to generate theoretical fitness landscapes, composed of five-bit strings that were generated from an *in silico* data set introduced in a prior study (Meszaros et al., 2019). The data set was originally generated in order to provide large empirical data sets that could be used to study advanced topics in population genetics, including higher-order epistasis. The datasets are constructed such that they can serve as an exploratory space for any theoretical set of interactors, and therefore, is well-structured for the study of interacting microbial taxa. That is, there is nothing about the datasets that renders it a better fit for any one biological problem than another: these data could just as well be used to study interacting genes as taxa or any parcel of information. The data are defined as strings of information (e.g. 01011 or 11001), each with a corresponding "phenotype" value. Therefore, this data is equipped for the analyses as proposed in this study. Here, we use it to generate theoretical microbiota in an insect gut. For more information on the data set and its origin, see Meszaros et. al 2019 and the Supplementary Information.

For the purpose of this study, it is important that we are transparent with regards to the data source, the notation, and the method for transforming the data into a microcosm for taxa in an *in silico* insect microbiota. In this study, our hypothetical insect guts are encoded as strings of bits. Bits can either be 0 or 1. 1 indicates the presence (+) of a taxa. 0 indicates absence (-) of a taxon. For example, we can write a string of 0 and 1 corresponding to an insect gut with five interacting taxa (A-D) as demonstrated in Table 1.

Taxa	A	B	C	D	E
Presence (+) or absence (-)	+	-	+	+	-
Binary representation	1	0	1	1	0

Table 1. Data structure of a hypothetical insect gut.

As we can see, the "10110" string corresponds to an insect gut where taxa A is present, B is absent, C is present, D is present, and E is absent. The data set that we are mining, originally derived for studying combinatorial data sets that are common in the study of fitness landscapes, offers tens of thousands of combinatorial sets that correspond to a hypothetical insect gut with interacting taxa (Meszaros et al., 2019). We have randomly chosen one such set, containing five individual bits, to explore the central biological concepts of interest in this study: the measurement of higher-order interactions between taxa, and how these interactions might be influenced by the environmental context.

The data set that we will use is a five-bit string (a string of five numbers), combinations of presence and absence (+ or -) of five taxa (A-D). The combinatorial possibility corresponds to $2^5 = 32$ theoretical combinations of taxa across four different insect environments. In Table 1, we show the "fitness" (host infection risk in our study) values for all 32 *in silico* combinations of taxa across four different insects.

Calculating the strength of interactions

As mentioned in the introduction, there are myriad methods for resolving higher-order interactions, and many such methods have been explored in genomic studies (Crona, 2020; Domingo et al., 2018; Guerrero et al., 2019; Otwinowski et al., 2018; Poelwijk et al., 2016; Sailer & Harms, 2017). A full treatment of the strengths and weaknesses of every method would require a review that is beyond the scope of our study, but some existing work has interrogated multiple methods in the study of epistasis (Poelwijk et al., 2016).

In describing the methods as applied in this study, we have erred on the side of redundancy in our explanations. We believe that this is appropriate, given that our method of choice – the Walsh Hadamard Transform – has never been applied to the study of the microbiota and so could benefit from further explanations.

The Walsh-Hadamard Transform

The Walsh-Hadamard Transform allows one to quantify the eminence of interaction effects among potentially interacting objects or parcels. Its main output is a Walsh coefficient, which communicates the magnitude (how large the interaction is) and sign (positive interaction or negative direction) for a given interaction. The method implements phenotypic (host infection risk in our study) values in the form of a vector, before reformatting it into a Hadamard matrix (and is then scaled by a diagonal matrix). The output is a collection of coefficients which correspond to the strength of interaction between taxa.

For example, we can define the Walsh Hadamard coefficient for the following:

*B*DE

The asterisks (*) correspond to taxa that could either be present or absent. This can reencoded in binary as:

01011

This Walsh Hadamard coefficient for this string would correspond to the magnitude of the interaction between the B, D and E taxa. Importantly, we would label the interaction between B, D and E as a "third order" interaction, as the calculation provides the average strength of interaction between three different taxa: B, D and E. Understanding the different orders of interaction is the key to gaining a perspective on "higher-order" interactions. In a gut containing five taxa that we are interested in understanding the interaction between, there are five different "orders" of potential interaction.

For example:

0^{th} (zeroth) order interaction would be the insect containing none of the taxa of interest (A-E) present. First order interactions correspond to the influence of individual taxa on the infection risk. There are five such first order terms in this theoretical insect microbiota:

A****

*B***

C

***D*

****E

Similarly, there are ten second order coefficients, ten third order, five fourth order, and one fifth order (corresponding to the interaction between all five taxa; ABCD or 11111). These Walsh Hadamard coefficients can be summed within an order. Consequently, a whole theoretical "insect gut" can be described in terms of the overall magnitude of its $0 - 5^{th}$ order interactions. For example, we can examine the strength of third-order interactions (in sum) and compare them to the strength of fourth order interactions.

The Walsh Hadamard coefficient describes the magnitude to which an interaction map is linear, or second order, third, and so forth. We refer interested readers to two published manuscripts—Poelwijk et al. (2016) and Weinreich et. al. (2013)—that outline and apply the method in good detail. Also, see the Supplementary Information for a brief primer.

The Walsh-Hadamard Transform relies on the existence of combinatorial data sets, where the objects for which we are interested in understanding the interactions between (taxa in this study) are constructed in all possible combinations. Another limitation of the Walsh-Hadamard Transform is that it can only accommodate two variants per site, that is, two states per actor. In the case of taxa, we can think of this in

terms of the presence/absence of a certain taxon, and we can encode this in terms of 0 (absence) or 1 (presence). For each hypothetical insect with a different presence/absence combination, we have a corresponding phenotypic measurement (e.g. host infection risk). For example, if we wanted to measure the higher-order interactions between 4 taxa within an insect with regards to their role in parasite load (as a model phenotype), we would need $2^L = 16$ individual measurements (insects in this case), with L corresponding to the number of different taxa whose effects we were interested in disentangling. We can encode this combination of 4 taxa in bit string notation (see Figure 1).

As described above (Methods section), each site (0 or 1) in the string corresponds to the presence or absence of a given taxa in a given insect. This notation allows us to keep a mental picture of which taxa are in which insect for which we have a phenotypic measurement and can be used to construct a vector of values. Again, the string 01011 corresponds to an insect with the pattern of absent (0), present (1), absent (0), present (1), present (1). The full data set includes a vector of phenotypic values for all possible combinations of taxa—(see Table 1). Note, again, that these can be divided into different classes based on the "order" of the interaction. This vector of phenotypic values for the 32 will be multiplied by a (32 x 32) square matrix, which is the product of a diagonal matrix V and a Hadamard matrix H. These matrices are defined recursively by:

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

$$H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}, H_0 = 1 \quad \textbf{[2]}$$

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

$$\gamma = VIIx$$

Where V and H are the matrices described in [1] and [2] above, and γ is the Walsh coefficient, the measure of the interaction between parcels of information in a string. Using this, we compute γ values for every possible interaction between bits in a given string. These methods measure every one of these interactions (e.g. all ten 2nd order interactions) between taxa. While our use of a five-bit string structure (as opposed to an three or fifteen bit string) is arbitrary, it communicates the nature of the higher-interaction problem: Even if we suspect that only five taxa in an insect microbiota are meaningfully influencing a phenotype of interest (Cagnolo et al., 2011; Ferrari & Vavre, 2011; McLean et al., 2016), the possible ways that these species are interacting, and the number of measurable coefficients between them can be meaningful.

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

The theoretical environment of the insect gut microbiota

Here, we explore how varying nutrient diets influence host susceptibility to parasites in the gut microbiota. We chose to focus on the nutrient diet content in our study design because the resource environment is highly relevant to the insect gut microbiota. In insects, nutrition content of the host's food can be controlled by the addition of methyl cellulose (an indigestible bulk agent) in the standard food medium (Boots & Begon, 1994). Resource-levels varying from high-quality diets (containing no methyl cellulose in the food medium) to lower-quality diets (replacing 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of the food medium with methyl cellulose) have been utilized to empirically study the role of varying nutrition environments to parasite resistance in lepidopteran pest species (Boots et al., 2011). In

our theoretical study, we define "nutrient content" as a diet compromising a range of nutrients in a standard insect diet. A diet of 0 % would correspond to an extremely low nutrition diet, and 100% to a high-quality diet composed of the standard food amount for insects. Consequently, the nutrient gradient 0 – 100% represents varying degrees of resource availability.

Results

Norm of reaction. The norm of reaction demonstrates that two insect guts, corresponding to 00000 (no taxa) and 11111 (the presence of taxa of every kind) have the largest parasite loads relative to other insect microbiota combinations. The high parasite load pattern is consistent across the nutrient content that insects consume (Fig. 2). In contrast, we find that parasite load is drastically reduced for all other insect microbiota combinations (examples include combinations 001100; 11011; 11101).

Comparison of the orders of interactions among taxa across microenvironments. Figure 3 demonstrates the sum of the absolute values of the interaction coefficients. Here, we can observe the raw magnitude (whether positive or negative in sign) of higher-order interactions as a function of interaction order. Note how the eminence of the higher-order effects changes as a function of nutrient content. At low nutrient contents, fourth order effects are the most impactful on the overall parasite load. At approximately 20%, the fifth order effects (corresponding to the five-way interaction of taxa in the *in silico* insect gut represented by 11111). The change in order of eminence also applies to the second order (pairwise) and third order interactions. At low nutrient contents, the pairwise interactions exert a more meaningful influence on the parasite load than the three-way interactions. At approximately 20% nutrient content--not far from that nutrient percentage where a switch between fourth and fifth order effects manifests--the three-way interactions supplant the pairwise effects in their overall influence on parasite load. Note that all of these values—the *in silico* parasite load data, the interaction coefficients for all individual interactions, and the scaled, absolute value coefficients—can be found in the Supplementary Material.

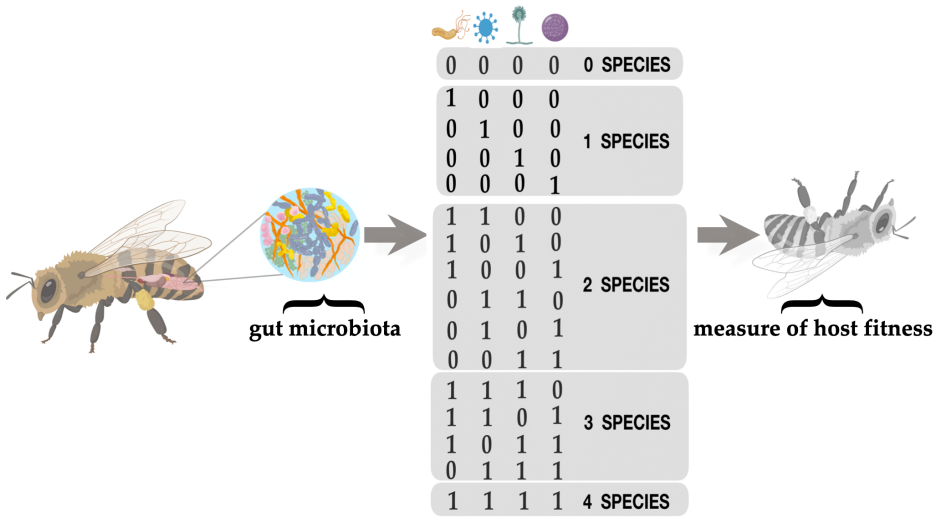

Figure 1. Schematic representation of higher-order interactions in the insect gut microbiota. We represent the presence of microbial species in the gut similarly to the presence of a genetic locus. Species composition of gut microbiota is represented in binary strings. In this configuration, the combination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

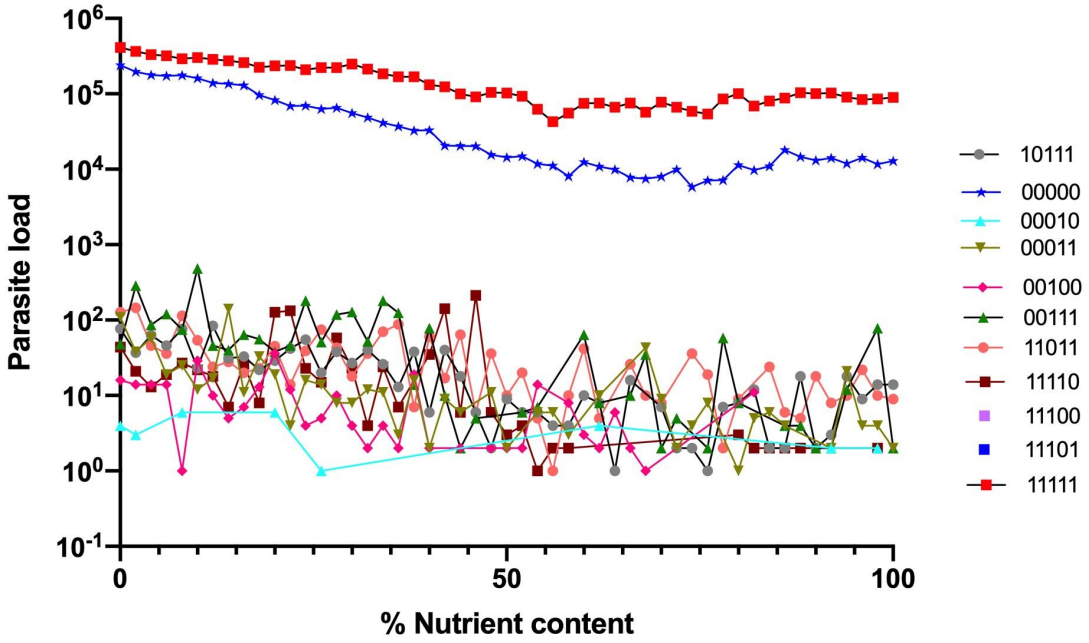

Figure 2. *A Norm of reaction representing the parasite load of in silico insect guts as a function of nutrient microenvironment*. The *x-axis* represents the nutrient content that insects consume, ranging from 0% (deprived) to 100% (a full, standard nutrient content. Individual data points correspond to insect cuts containing different combinations of taxa. The *y-axis* represents the parasite load, a proxy for the susceptibility of a given insect to infection by parasites. Note that only a subset of the 32 taxa are represented in this, as many of the *in silico* insect guts have parasite loads that are very low. The data for all 32 can be found in the supplementary material.

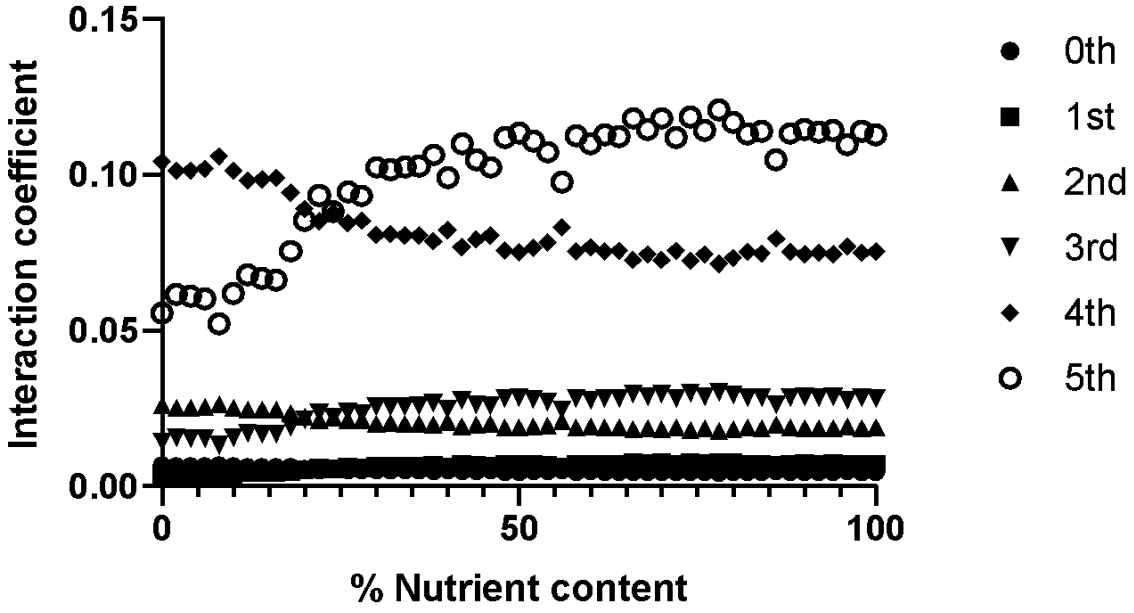

Figure 3. *The magnitude of higher-order interaction between taxa as a function of environment (nutrient content).* Absolute value, averaged magnitude interactions across interaction orders. The purpose of this depiction is to illustrate how the magnitude (not sign) of the interactions change with interaction order. Of special note is how importance of the order of interactions changes as a function of nutrient content (x-axis). In this scenario, there is a nutrient content threshold (~20%) where the patterns of the interactions change. In our theoretical insect microbiota system, we define "nutrient content" as a diet compromising a range of nutrients in a standard insect diet. In insect populations, the nutrient content of the host's food can be controlled by the addition of methyl cellulose (an indigestible bulk agent) in the standard food medium (Boots & Begon, 1994). In our model, a diet of 0 % would correspond to an extremely low nutrition diet, and 100% to a high-quality diet composed of the standard food amount for insects. Consequently, the nutrient gradient 0 – 100% represents varying degrees of nutrient content.

Discussion

In this study, we explore the possibility of higher-order interactions between taxa that compose an insect gut microbiota. Using *in silico* and mathematical approaches, we demonstrate how higher-order interactions can be measured in a complex system of interacting microbial taxa. In our theoretical scenario, higher-order interactions are present and generally increase in relevance with the order of interaction. Notably, the environment (nutrient content in this case) has a meaningful influence on how higher-order interactions among taxa manifest. This result highlights an aspect of higher-order interactions that is so far largely under-appreciated: that the environment and context in which taxa exist can have a meaningful impact on how taxa interact. Consequently, simply noting that non-linear and higher order interactions between taxa may exist is no longer sufficient in how the insect microbiota is discussed: we must consider, and measure, how environments may influence how interactions manifest. Though our results arise from a theoretical examination of *in silico* insect microbiota, they are results nonetheless (Goldstein, 2018) and highlight the potentially vast scope of the higher-order interaction problem that could define the true dynamics of gut microbiota. Specifically, the outlining of a method that can be used to deconstruct higher-order interactions in biological systems, across environmental contexts, represents a potentially useful contribution to the study of the microbiota.

Though empirical data of the size and scope used in this study are currently challenging to generate, this intractability may be temporary, and future methods may permit the generation of data similar in structure to those explored in our theoretical examination. Note that the calculations of higher-order interactions, and their dynamic nature, can be considered without knowing the specific mechanism that underlies the nature of these interactions, determining the magnitude of coefficients provides relevant information on the eminence of a given order in the microbiota. One additional benefit of these results is that they can identify those settings (combinations between microbiota and a given microenvironment) that should be the focus on mechanistic study. For example, by identifying the taxa involved in large pairwise

389 interactions, one can then examine the mechanistic basis underlying this pairwise interaction through
390 manipulative experiments.
391
392 Our results are consistent from recent findings, where diverse communities are more effective at resisting
393 invasions including E. coli invasion of soil communities (Elsas et al., 2012), plant root bacterial
394 communities (Wei et al., 2015), and experimental invasions in bacterial communities (Lu et al., 2018).
395 Collectively these studies show that outcome of invasions are determined by available resources in the
396 microbiota. Our main result showing that higher-order microbial interactions limits the invasion of
397 parasites across nutrient environments is in agreement with studies that interactions are mediated by
398 underlying resource dynamics. The nutritional status of the gut microbiome plays an important role in the
399 health of hosts. Simple gut microbiotas have been engineered to provide hosts with novel functions, such
400 as the ability to use more complex nutrient sources and to fight against pathogens. Recent work by Sun et
401 al. 2020 shows that in Caenorhabditis elegans, the colonization of cellulolytic bacteria enables *C. elegans*
402 to utilize cellulose, an otherwise indigestible carbon substrate. At the community level, cellulolytic
403 bacteria can also support resident bacterial species with additional functional roles, such as the protection
404 by Lactobacillus in the gut against *Salmonella* infection (Sun et al., n.d.). To test our model, insect gut
405 microbiota could be engineered to explore how higher-order microbial endosymbiont interactions protect
406 against pathogen infection by enhancing the nutritional status of the host.
407
408 The mathematical approach used in this study—the Walsh-Hadamard Transform—has been previously
409 used by theoretical population geneticists to measure non-linear interactions between mutations
410 (Weinreich et al., 2013). Several empirical data sets in genetics and genomics have demonstrated that the
411 sign of interaction effects can change readily with the identity of the interacting parcels (Guerrero et al.,
412 2018; Weinreich et al., 2013, 2018a). Given this, we predict that the taxa that compose the gut microbiota
413 might be similarly defined by higher-order interactions, and that these interactions will change
414 appreciably with insect microenvironment. The capacity for measuring the effects of higher-order

415 interactions on host fitness is an important step towards understanding the effects of microbiota on their
416 host.
417
418 The impact of higher-order interactions in the gut microbiota on host fitness may result from a range of
419 possible interactions, ranging from competitive to mutualistic (Fast et al., 2018; Ludington & Ja, 2020;
420 Newell & Douglas, 2014). To test the full suite of all possible combinatorial interactions and their
421 associated effects on host traits, it is important to experimentally manipulate microbial communities. For
422 example, the fruit fly (*Drosophila melanogaster*) is an attractive model system for designing
423 combinatorial studies due to relative ease of rearing gnotobiotic flies and modularity of its microbiome
424 (Ludington & Ja, 2020). For example, combinatorial designs of microbial communities in *D.*
425 *melanogaster* revealed that emerging higher-order effects composed of 3, 4, and 5-way interactions
426 impacted aspects of host fitness such as life span and fecundity (Gould et al., 2018). While the relative
427 simplicity and tractability of fly microbiomes facilitates the study of host-microbe interactions,
428 underlying mechanisms can provide insights for more complex mammalian gut microbiomes. In *D.*
429 *melanogaster*, stable gut colonizers favor specific regions of the foregut, which like mammals, suggest
430 specific niches for gut colonizers (Pais et al., 2018). Therefore, strategies that invertebrates and their
431 microbes employ to form stable associations might be informative for mammalian gut microbiomes
432 (Ludington & Ja, 2020).

433 **Conclusion**

434 Recent theoretical work suggests that higher-order modeling approaches are able to capture volumes of
435 rich data arising from complex ecological interactions (Battiston et al., 2020). We have adapted
436 approaches from evolutionary genetics to the study of host-associated microbiota. In the future, applying
437 these methods to the analysis of experimental data will yield important insight into microbiome dynamics,
438 towards a richer understanding of just how peculiar the microbiota is, and the many meaningful
439 interactions that it embodies.

Acknowledgements

We wish to acknowledge the support of organizers and participants of the 2017 RCN-IDEAS arbovirus workshop held in New Orleans. SY acknowledges funding support from NSF Postdoctoral Fellowship award number 1612302. CBO acknowledges funding support from NSF RII Track-2 FEC award number 1736253. The authors would like to thank Victor Meszaros and Miles Miller-Dickson for their input on the *in silico* data, figures and Walsh-Hadamard primer, and Daniel Weinreich for helpful discussion on topics relevant to this study. The authors would like to thank the Associate Editor of J*ournal of Animal Ecology* and two anonymous reviewers for thoughtful feedback on the manuscript. Finally, the authors would like to thank Lawrence Uricchio for helpful feedback on the manuscript.

Data Availability

The *in silico* data used in this study and code used to generate them can be found on github: https://github.com/OgPlexus/MicrobeTaxa1

Supplemental Information

The authors can find data, code and other information on: https://github.com/OgPlexus/MicrobeTaxa1. This also includes a short mathematical primer on the Walsh-Hadamard Transform as applied to binary datasets (also available at https://github.com/OgPlexus/MicrobeTaxa1). For a more rigorous understanding, readers are encouraged to engage the works cited in this manuscript.

Literature cited

Barabás, G., J. Michalska-Smith, M., & Allesina, S. (2016). The Effect of Intra- and Interspecific Competition on Coexistence in Multispecies Communities. *The American Naturalist*, *188*(1), E1–E12. https://doi.org/10.1086/686901

Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G., & Petri, G. (2020). Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, *874*, 1–92. https://doi.org/10.1016/j.physrep.2020.05.004

Boots, Michael, & Begon, M. (1994). Resource limitation and the lethal and sublethal effects of a viral pathogen in the Indian meal moth, Plodia interpunctella. *Ecological Entomology*, *19*(4), 319–326. https://doi.org/10.1111/j.1365-2311.1994.tb00248.x

Boots, Mike, Jokela, A. E. J., & Shaw, E. R. G. (2011). The Evolution of Resistance to a Parasite Is Determined by Resources. *The American Naturalist*, *178*(2), 214–220. https://doi.org/10.1086/660833

Brugman, S., Ikeda-Ohtsubo, W., Braber, S., Folkerts, G., Pieterse, C. M. J., & Bakker, P. A. H. M. (2018). A Comparative Review on Microbiota Manipulation: Lessons From Fish, Plants, Livestock, and Human Research. *Frontiers in Nutrition*, *5*. https://doi.org/10.3389/fnut.2018.00080

Cagnolo, L., Salvo, A., & Valladares, G. (2011). Network topology: Patterns and mechanisms in plant-herbivore and host-parasitoid food webs. *Journal of Animal Ecology*, *80*(2), 342–351. https://doi.org/10.1111/j.1365-2656.2010.01778.x

Chesson, P. (2000). Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology and Systematics*, *31*(1), 343–366. https://doi.org/10.1146/annurev.ecolsys.31.1.343

Crona, K. (2020). Rank orders and signed interactions in evolutionary biology. *ELife*, *9*, e51004. https://doi.org/10.7554/eLife.51004

Domingo, J., Diss, G., & Lehner, B. (2018). Pairwise and higher-order genetic interactions

during the evolution of a tRNA. *Nature*, *558*(7708), 117–121. https://doi.org/10.1038/s41586-018-0170-7

Douglas, A. E. (2018). *Fundamentals of Microbiome Science: How Microbes Shape Animal Biology*. Princeton University Press.

Elsas, J. D. van, Chiurazzi, M., Mallon, C. A., Elhottovā, D., Krištůfek, V., & Salles, J. F. (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proceedings of the National Academy of Sciences*, *109*(4), 1159–1164. https://doi.org/10.1073/pnas.1109326109

Engel, P., & Moran, N. A. (2013). The gut microbiota of insects – diversity in structure and function. *FEMS Microbiology Reviews*, *37*(5), 699–735. https://doi.org/10.1111/1574-6976.12025

Enke, T. N., Datta, M. S., Schwartzman, J., Cermak, N., Schmitz, D., Barrere, J., Pascual-García, A., & Cordero, O. X. (2019). Modular Assembly of Polysaccharide-Degrading Marine Microbial Communities. *Current Biology*. https://doi.org/10.1016/j.cub.2019.03.047

Fast, D., Kostiuk, B., Foley, E., & Pukatzki, S. (2018). Commensal pathogen competition impacts host viability. *Proceedings of the National Academy of Sciences*, *115*(27), 7099–7104. https://doi.org/10.1073/pnas.1802165115

Ferrari, J., & Vavre, F. (2011). Bacterial symbionts in insects or the story of communities affecting communities. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1569), 1389–1400. https://doi.org/10.1098/rstb.2010.0226

Goldstein, R. E. (2018). Are theoretical results 'Results'? *ELife*, *7*, e40018. https://doi.org/10.7554/eLife.40018

Gómez, P., Bennie, J., Gaston, K. J., & Buckling, A. (2015). The Impact of Resource Availability on Bacterial Resistance to Phages in Soil. *PLOS ONE*, *10*(4), e0123752.

518 https://doi.org/10.1371/journal.pone.0123752

519 Gould, A. L., Zhang, V., Lamberti, L., Jones, E. W., Obadia, B., Korasidis, N., Gavryushkin, A.,
520 Carlson, J. M., Beerenwinkel, N., & Ludington, W. B. (2018). Microbiome interactions
521 shape host fitness. *Proceedings of the National Academy of Sciences*, *115*(51),
522 E11951–E11960. https://doi.org/10.1073/pnas.1809349115

523 Guerrero, R. F., Scarpino, S., Rodrigues, J. V., Hartl, D. L., & Ogbunugafor, C. B. (2018).
524 Proteostasis environment shapes higher-order epistasis operating on antibiotic
525 resistance. *BioRxiv*, 470971. https://doi.org/10.1101/470971

526 Guerrero, R. F., Scarpino, S. V., Rodrigues, J. V., Hartl, D. L., & Ogbunugafor, C. B. (2019).
527 Proteostasis Environment Shapes Higher-Order Epistasis Operating on Antibiotic
528 Resistance. *Genetics*, genetics.302138.2019.
529 https://doi.org/10.1534/genetics.119.302138

530 Guittar, J., Shade, A., & Litchman, E. (2019). Trait-based community assembly and succession
531 of the infant gut microbiome. *Nature Communications*, *10*(1), 512.
532 https://doi.org/10.1038/s41467-019-08377-w

533 Gurung, K., Wertheim, B., & Salles, J. F. (2019). The microbiome of pest insects: It is not just
534 bacteria. *Entomologia Experimentalis et Applicata*, *167*(3), 156–170.
535 https://doi.org/10.1111/eea.12768

536 Hooper, L. V., Littman, D. R., & Macpherson, A. J. (2012). Interactions Between the Microbiota
537 and the Immune System. *Science*, *336*(6086), 1268–1273.
538 https://doi.org/10.1126/science.1223490

539 Hutchinson, G. E. (1961). The Paradox of the Plankton. *The American Naturalist*, *95*(882), 137–
540 145. https://doi.org/10.1086/282171

541 Kareiva, P. (1994). Special Feature: Higher Order Interactions as a Foil to Reductionist Ecology.
542 *Ecology*, *75*(6). JSTOR. https://doi.org/10.2307/1939613

543 Knutie, S. A., Wilkinson, C. L., Kohl, K. D., & Rohr, J. R. (2017). Early-life disruption of

544 amphibian microbiota decreases later-life resistance to parasites. *Nature*
545 *Communications*, *8*(1), 86. https://doi.org/10.1038/s41467-017-00119-0
546 Levine, J. M., Bascompte, J., Adler, P. B., & Allesina, S. (2017). Beyond pairwise mechanisms
547 of species coexistence in complex communities. *Nature*, *546*(7656), 56–64.
548 https://doi.org/10.1038/nature22898
549 Lopez-Pascua, L. d C., & Buckling, A. (2008). Increasing productivity accelerates host-parasite
550 coevolution. *Journal of Evolutionary Biology*, *21*(3), 853–860.
551 https://doi.org/10.1111/j.1420-9101.2008.01501.x
552 Lu, N., Sanchez-Gorostiaga, A., Tikhonov, M., & Sanchez, A. (2018). Cohesiveness in microbial
553 community coalescence. *BioRxiv*, 282723. https://doi.org/10.1101/282723
554 Ludington, W. B., & Ja, W. W. (2020). Drosophila as a model for the gut microbiome. *PLOS*
555 *Pathogens*, *16*(4), e1008398. https://doi.org/10.1371/journal.ppat.1008398
556 Mackay, T. F., & Moore, J. H. (2014). Why epistasis is important for tackling complex human
557 disease genetics. *Genome Medicine*, *6*(6), 42. https://doi.org/10.1186/gm561
558 Macpherson, A. J., & Harris, N. L. (2004). Interactions between commensal intestinal bacteria
559 and the immune system. *Nature Reviews Immunology*, *4*(6), 478–485.
560 https://doi.org/10.1038/nri1373
561 Mayfield, M. M., & Stouffer, D. B. (2017). Higher-order interactions capture unexplained
562 complexity in diverse communities. *Nature Ecology & Evolution*, *1*(3), 0062.
563 https://doi.org/10.1038/s41559-016-0062
564 McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A.
565 E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S. F., Hentschel, U., King, N., Kjelleberg,
566 S., Knoll, A. H., Kremer, N., Mazmanian, S. K., Metcalf, J. L., Nealson, K., Pierce, N. E.,
567 … Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life
568 sciences. *Proceedings of the National Academy of Sciences*, *110*(9), 3229–3236.
569 https://doi.org/10.1073/pnas.1218525110

McKean, K. A., Yourth, C. P., Lazzaro, B. P., & Clark, A. G. (2008). The evolutionary costs of immunological maintenance and deployment. *BMC Evolutionary Biology*, *8*(1), 76. https://doi.org/10.1186/1471-2148-8-76

McLean, A. H. C., Parker, B. J., Hrček, J., Henry, L. M., & Godfray, H. C. J. (2016). Insect symbionts in food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150325. https://doi.org/10.1098/rstb.2015.0325

Mereghetti, V., Chouaia, B., & Montagna, M. (2017). New insights into the microbiota of moth pests. *International Journal of Molecular Sciences*, *18*(11), 2450.

Meszaros, V. A., Miller-Dickson, M. D., & Ogbunugafor, C. B. (2019). Lexical landscapes as large in silico data for examining advanced properties of fitness landscapes. *PloS One*, *14*(8), e0220891.

Mickalide, H., & Kuehn, S. (2019). Higher-order interaction inhibits bacterial invasion of a phototroph-predator microbial community. *BioRxiv*, 564260. https://doi.org/10.1101/564260

Newell, P. D., & Douglas, A. E. (2014). Interspecies Interactions Determine the Impact of the Gut Microbiota on Nutrient Allocation in Drosophila melanogaster. *Applied and Environmental Microbiology*, *80*(2), 788–796. https://doi.org/10.1128/AEM.02742-13

Otwinowski, J., McCandlish, D. M., & Plotkin, J. B. (2018). Inferring the shape of global epistasis. *Proceedings of the National Academy of Sciences*, *115*(32), E7550–E7558. https://doi.org/10.1073/pnas.1804015115

Pais, I. S., Valente, R. S., Sporniak, M., & Teixeira, L. (2018). Drosophila melanogaster establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. *PLoS Biology*, *16*(7). https://doi.org/10.1371/journal.pbio.2005710

Poelwijk, F. J., Krishna, V., & Ranganathan, R. (2016). The Context-Dependence of Mutations: A Linkage of Formalisms. *PLOS Computational Biology*, *12*(6), e1004771. https://doi.org/10.1371/journal.pcbi.1004771

596 Roberts, K., Meaden, S., Sharpe, S., Kay, S., Doyle, T., Wilson, D., Bartlett, L. J., Paterson, S.,
597 & Boots, M. (2019). The genomic basis of evolved virus resistance is dependent on
598 environmental resources. *BioRxiv*, 666404. https://doi.org/10.1101/666404
599 Round, J. L., & Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune
600 responses during health and disease. *Nature Reviews Immunology*, *9*(5), 313–323.
601 https://doi.org/10.1038/nri2515
602 Sailer, Z. R., & Harms, M. J. (2017). Detecting High-Order Epistasis in Nonlinear Genotype-
603 Phenotype Maps. *Genetics*, *205*(3), 1079–1088.
604 https://doi.org/10.1534/genetics.116.195214
605 Sanchez-Gorostiaga, A., Bajić, D., Osborne, M. L., Poyatos, J. F., & Sanchez, A. (2018). High-
606 order interactions dominate the functional landscape of microbial consortia. *BioRxiv*,
607 333534. https://doi.org/10.1101/333534
608 Skidmore, I. H., & Hansen, A. K. (2017). The evolutionary development of plant-feeding insects
609 and their nutritional endosymbionts. *Insect Science*, *24*(6), 910–928.
610 https://doi.org/10.1111/1744-7917.12463
611 Sun, Q., Vega, N. M., Cervantes, B., Mancuso, C. P., Mao, N., Taylor, M., Collins, J. J., Khalil,
612 A. S., & Gore, J. (n.d.). *Colonization with heterologous bacteria reprograms a*
613 *Caenorhabditis elegans nutritional phenotype*. 11.
614 Tekin, E., Savage, V. M., & Yeh, P. J. (2017). Measuring higher-order drug interactions: A
615 review of recent approaches. *Current Opinion in Systems Biology*, *4*, 16–23.
616 https://doi.org/10.1016/j.coisb.2017.05.015
617 Vandermeer, J. H. (1969). The Competitive Structure of Communities: An Experimental
618 Approach with Protozoa. *Ecology*, *50*(3), 362–371. https://doi.org/10.2307/1933884
619 Wei, Z., Yang, T., Friman, V.-P., Xu, Y., Shen, Q., & Jousset, A. (2015). Trophic network
620 architecture of root-associated bacterial communities determines pathogen invasion and
621 plant health. *Nature Communications*, *6*(1), 8413. https://doi.org/10.1038/ncomms9413

622 Weinreich, D. M., Lan, Y., Jaffe, J., & Heckendorn, R. B. (2018a). The Influence of Higher-Order
623 Epistasis on Biological Fitness Landscape Topography. *Journal of Statistical Physics*,
624 *172*(1), 208–225. https://doi.org/10.1007/s10955-018-1975-3
625 Weinreich, D. M., Lan, Y., Jaffe, J., & Heckendorn, R. B. (2018b). The Influence of Higher-Order
626 Epistasis on Biological Fitness Landscape Topography. *Journal of Statistical Physics*,
627 *172*(1), 208–225. https://doi.org/10.1007/s10955-018-1975-3
628 Weinreich, D. M., Lan, Y., Wylie, C. S., & Heckendorn, R. B. (2013). Should evolutionary
629 geneticists worry about higher-order epistasis? *Current Opinion in Genetics &*
630 *Development*, *23*(6), 700–707. https://doi.org/10.1016/j.gde.2013.10.007
631 Wood, K., Nishida, S., Sontag, E. D., & Cluzel, P. (2012). Mechanism-independent method for
632 predicting response to multidrug combinations in bacteria. *Proceedings of the National*
633 *Academy of Sciences*, *109*(30), 12254–12259. https://doi.org/10.1073/pnas.1201281109