Holobiont Evolution:
Population Genetic Theory for the Hologenome

Joan Roughgarden[1]

Manuscript as of June 22, 2020

Abstract

This paper develops a mathematical theory for holobiont evolution that parallels the population-genetic theory of classical evolutionary biology. It presents theory for hologenomes having two haploid microbial strains and two diploid host alleles. The theory shows how selection on holobionts causes the joint evolution of microbial and host components of the hologenome. The theory also reveals the distribution of microbiome configurations across hosts as well as stable strategies for microbiome-host coadaptation.

1 Introduction

This paper develops a mathematical theory for holobiont evolution that parallels the population-genetic theory of classical evolutionary biology. All organisms host a variety of microbes, their "microbiome". The combination of a host with its microbiome has been termed a "holobiont" (Margulis 1991). Studies reveal astonishing coadaptation between microbiomes and their hosts extending to development, metabolism, physiology and behavior (Gilbert, Sapp and Tauber 2012, McFall-Ngai, Hadfield, Bosch, Carey, Domazet-Loo, Douglas, Dubilier et al. 2013, Bordenstein and Theis 2015). Some evolutionary process must have produced this coadaptation.

Because the phenotype of a holobiont results from the expression of the genes in the host's nucleus plus the genes in the microbes within the host, the union of the host genes with all the microbial genes has been termed a "hologenome" (Zilber-Rosenberg and Rosenberg 2008). Selection on holobionts with their hologenomes is envisioned as analogous to classical natural selection on individuals with their nuclear genomes. This holobiont selection might be the evolutionary process that accounts for microbiome-host coadaptation.

Although the idea of an extended genome consisting of host genes plus microbial genes may seem appealing, conceptual difficulties remain. Nuclear genes and microbial genes are inherited differently—nuclear genes vertically and microbial genes mostly horizontally. This difference in modes of inheritance between different parts of the hologenome might invalidate any useful analogy between holobiont selection and natural selection. So this paper develops mathematical theory for holobiont evolution that allows for the differing inheritance mechanisms of the microbial and host components of the hologenome.

[1]Hawaii Institute of Marine Biology, University of Hawaii and Department of Biology, Stanford University.

This paper first reviews previous research about how to define a theory of holobiont evolution, how to define a "hologenotype", and what the significance is of horizontal vs. vertical microbiome transmission. The paper then presents new research in stages: holobionts with two haploid microbial strains, holobionts with two diploid host alleles, holobionts having both two haploid microbial strains and two diploid host alleles combined, and finally, strategies of coadaptation. The mathematical theory here shows how selection on holobionts causes the joint evolution of the microbial and host components of the hologenome. It also reveals the distribution of microbiome configurations across hosts and stable strategies for microbiome-host coadaptation.

2 Previous Research

2.1 Definition of a Theory of Holobiont Evolution

The classical population-genetic theory for evolution supplies an equation (or equations) that comprise a dynamical system, that is, equations that map gene or genotype frequencies at time t to the gene or genotype frequencies at time $t+1$, or symbolically $p_t \rightarrow p_{t+1}$ where p denotes a vector of gene or genotype frequencies. Therefore, a theory of holobiont evolution should do the same, that is, the theory should supply equations that map the current state of a population of holobionts at time t to the state at time $t+1$, or symbolically, $H_t \rightarrow H_{t+1}$ where H denotes the state of the holobiont population.

Holobiont selection is distinct from coevolutionary selection. Holobiont selection operates on a single hologenome that itself is conceptually merged from the genomes of the host and its resident microbes. A theory of holobiont selection maps the state of a single "hologene pool" from t to $t+1$. In holobiont selection, the evolutionary interests of the host and the microbiome are partially aligned. Also, for a holobiont the selection events on host and microbiome may be synchronized. That is, the death or reproduction of a holobiont may imply the death or reproduction simultaneously of both host and microbiome.

With coevolutionary selection, distinct gene pools exist for the host and for the microbes within it. For example, in a typical coevolutionary setup for two species, the fitnesses influencing a gene frequency in say, species-1 are a function of the abundances and gene frequencies in both species-1 and species-2, and conversely for the fitnesses in species-2 (Dieckmann and Law 1996; Carmona, Fitzpatrick and Johnson 2015). This setup leads to two coupled equations for changes in the gene pools of the two species, or symbolically, simultaneous mappings of $\{p(t) \rightarrow p(t+1), q(t) \rightarrow q(t+1)\}$ where p and q are allele frequencies in species-1 and species-2, respectively. The gene pools of the microbes and host are not merged conceptually and remain separate. In coevolutionary selection, the evolutionary interests of the host and the microbiome are not aligned. Also, the selection events on the species are generally not synchronized–that is, for

example, the death or birth of a pollinator is generally not synchronized with the death or seed production of a plant that it pollinates.

The dynamic equilibrium attained with holobiont selection is different from that with coevolutionary selection. Two species subject to coevolutionary selection play an evolutionary game against each other, and the coevolutionary equilibrium is a Nash competitive equilibrium (Brown and Vincent 1987). In contrast, holobiont selection acting on the combination of a host with a microbe living in it leads to a different evolutionary equilibrium as will be detailed below.

A coevolutionary perspective on the host-microbiome relationship has led to the claim that "the host and each individual microbial strain are distinct entities with potentially divergent selective pressures" and further, that host control is needed to constrain the microbiome, to keep the microbiome on a "leash" (Foster, Schluter, Coyte, and Rakoff-Nahoum 2017, 43). On the other hand, although the evolution of reduced virulence for myxomatosis in Australian rabbits (Fenner 1953) has conventionally been assumed to result from coevolutionary selection on host and pathogen, this coadaptation might actually result from holobiont selection.

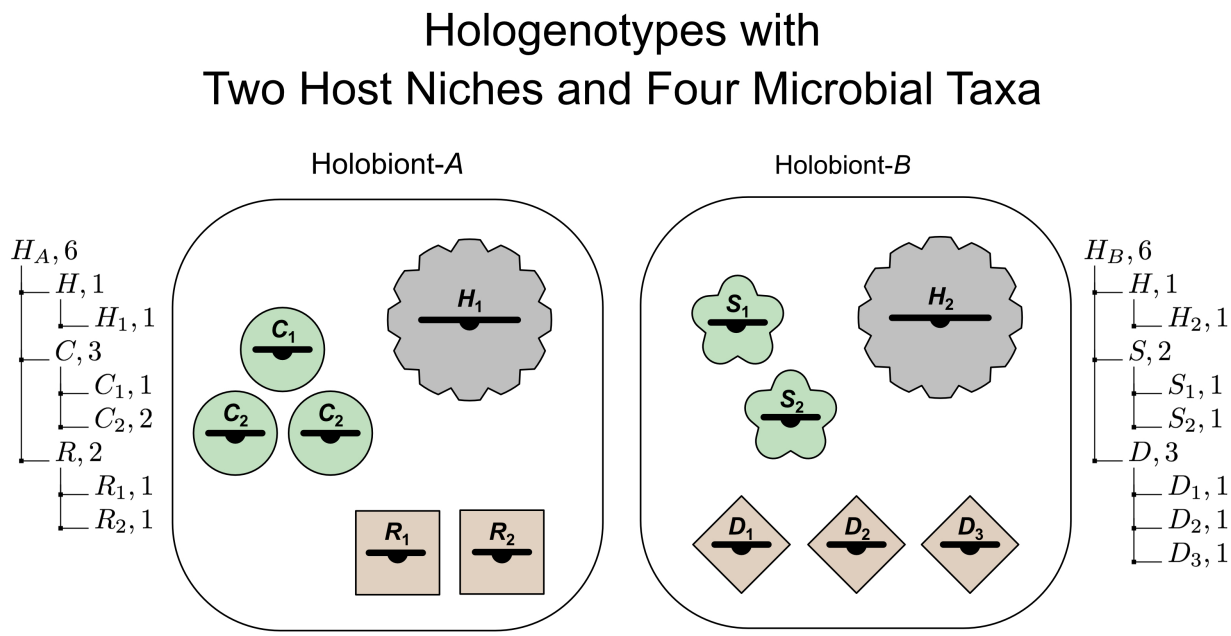
2.2 The Hologenotype

Corresponding to a genotype in classical population genetics, an instance of a hologenome has been called a "hologenotype" (Roughgarden 2020). The hologenotype of a particular holobiont may be represented as a hierarchical data structure as depicted in Figure 1. The outer list indicates the number of genomes in the holobiont. (The host nucleus is counted as a taxon containing one genome.) Nested within this are lists for the abundance and identity of each genome. And nested within these are lists for the number and identity of each allele.

Figure 1 diagrams two instances of a conceptual holobiont, A and B. A representative host is depicted as a rectangle with rounded corners. A host's genome contains nuclear genes depicted in the gray gear box. Holobiont-A has allele H_1 and Holobiont-B has H_2 at a locus of interest in its host's nucleus. Each holobiont has two niches for microbes depicted as green or brown that can represent different physical locations within the host or different functional roles. Two taxa of greens, circle and star, and two taxa of browns, rectangle and diamond, are available for each niche. The abundance of these taxa in each niche can vary among holobionts—Holobiont-A has three circles and Holobiont-B two stars. The allelic composition can vary among microbes—the circles have one C_1 allele and two C_2 alleles in Holobiont-A.

In the hologenotype description the number of microbes is normalized to a standard host cell. For example, if Holobiont-A as a whole is comprised of 1000 host cells, 3000 green circle microbes and 2000 brown rectangle microbes, then its hologenotype description, $\mathbf{H}_A$, records 1 host nucleus, 3 green circles, and 2 brown rectangles as shown in Figure 1.

Gene duplication is covered by the hologenotype description of Figure 1. A duplicated nu-



$$\mathbf{H}_A = \{H_A, 6, \{\{H, 1, \{\{H_1, 1\}\}\}, \{C, 3, \{\{C_1, 1\}, \{C_2, 2\}\}\}, \{R, 2, \{\{R_1, 1\}, \{R_2, 1\}\}\}\}\}$$

$$\mathbf{H}_B = \{H_B, 6, \{\{H, 1, \{\{H_2, 1\}\}\}, \{S, 2, \{\{S_1, 1\}, \{S_2, 1\}\}\}, \{D, 3, \{\{D_1, 1\}, \{D_2, 1\}, \{D_3, 1\}\}\}\}\}$$

Figure 1: *Examples of hologenotypes for two holobionts.*

clear gene with say, two copies of the H_1 allele might be shown in the gray gear box of one hologenotype in comparison with a single copy of the H_1 allele in the gray gear box of another hologenotype.

Moreover, microbe abundance is the equivalent of gene duplication—duplication of microbe genes rather than host nuclear genes. A hologenotype with say, two C_1 circle microbes could be compared with a hologenotype having say, three C_1 circle microbes.

The role of the microbes can be looked at from either of two sides, as though an Escher painting. From the host's standpoint, the microbes are no more than encapsulated genes whose diversity represents polymorphism and whose abundance represents a gene copy number. From the microbiome's standpoint, the microbes are tiny living entities whose pattern of species abundance represents community structure. From the host's standpoint, holobiont selection controls gene identity and copy number, whereas from the microbiome standpoint, holobiont selection molds microbial community structure.

Thus, in parallel with classical population genetics, a theory of holobiont evolution predicts the trajectory through time of the frequencies of these hologenotypes in a holobiont population.

Although Figure 1 points to how a complex hologenotype can be represented, for theoretical development much simpler cases of hologenotypes may be examined.

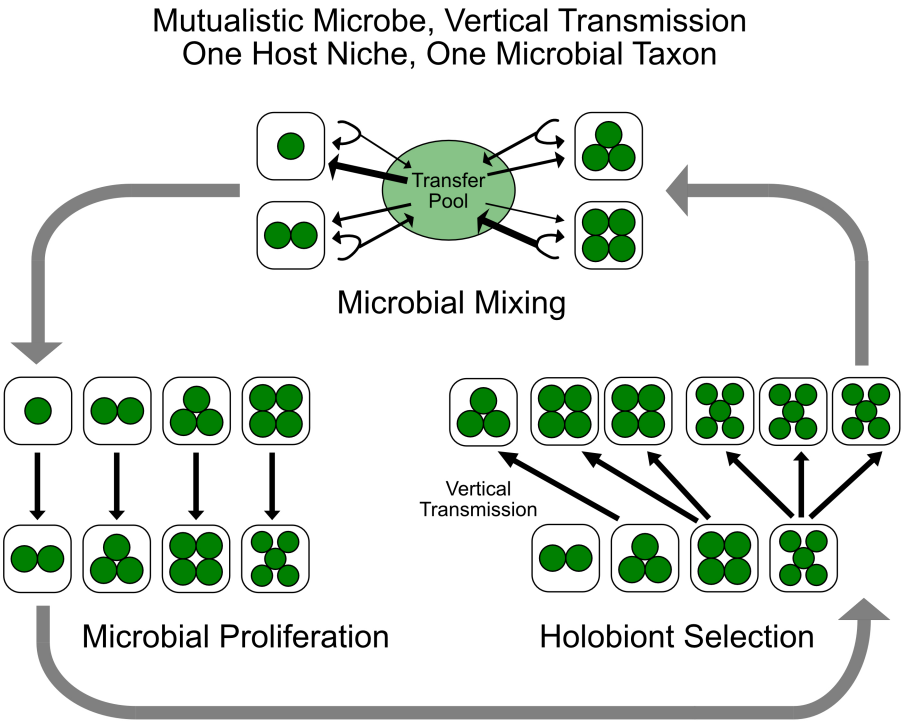


Figure 2: *Life-cycle diagram for vertical transmission assuming a mutualistic microbe. Three processes occur as sequential stages within each host generation. The host's generation time is the macro time step. The macro time step begins in the upper right of the diagram,* i.e. *after the holobiont selection stage and before the microbial mixing stage. The macro time step begins with microbial mixing, as shown at the top of the diagram. Each microbe has a small probability,* m*, of exiting its holobiont and entering a temporary "transfer pool." Then the total transfer pool is divided equally across the holobionts. Each holobiont receives the same number of transfers. After microbial mixing, the microbes proliferate within the holobionts, as shown on the left side of the diagram. The microbes within each holobiont increase according to a density-dependent ecological population model such as the logistic equation. Next the holobionts incur holobiont selection, as shown on the right side of the diagram. Each holobiont reproduces as a whole. The number of progeny a holobiont produces depends on the number of microbes in it. Parental holobionts transmit their microbes to their juveniles such that the number of microbes in each juvenile equals the number of microbes in its parent. With microbial mutualists (shown), the holobionts with the highest number of microbes leave the largest number of progeny and conversely for microbial parasites (not shown). Holobiont reproduction and survival are independent of the density of holobionts. The macro time step ends after holobiont selection, whereupon another macro time step begins at the microbial mixing stage. Iteration of the macro time step with these three stages generates a trajectory of holobiont abundance and hologenotype frequencies through time.*

2.3 Vertical *vs* Horizontal Microbe Transmission

An important issue to settle is whether holobiont evolution is even possible. Most of the reservations about the idea of holobiont selection revolve around the rarity of vertical microbiome transmission. For example, in a few species of corals the zooxanthellae are transmitted vertically in the coral's gametes (Hirose and Hidaka 2006). But the vast majority of zooxanthellae are acquired from the open sea water surrounding the coral (Babcock, Bull, Harrison, Heyward, Oliver, Wallace, and Willis 1986, Trench 1993). Even terrestrial groups may offer less vertical transmission than might be expected. In a live-bearing cockroach, researchers found only one component of the microbiome to be transmitted vertically (Jennings 2019). Indeed, juvenile hosts usually obtain their microbiomes from the surrounding environment—this is horizontal transmission.

A horizontally transmitted microbiome in a holobiont is an acquired trait not directly inherited from parent to offspring. As such, holobiont selection presumably cannot cause much evolution because most microbiomes are acquired rather than inherited traits. Nonetheless, host-microbiome coadaptation has indeed evolved in many groups even though the transmission is primarily horizontal. How could this be?

To investigate holobiont evolution even with horizontal microbiome transmission, a dynamical model was developed for both vertical and horizontal microbiome transmission to allow a comparison between these modes of transmission (Roughgarden 2020)[2]. The hologenotypic variation was solely in microbe number and assumed one host niche, one microbial taxon, and no allelic variation in host genomes. The microbe was either a pathogen or a mutualist.

The model variants with vertical and horizontal transmission both feature three processes that occur as sequential stages: microbial mixing $\rightarrow$ within-host microbial proliferation $\rightarrow$ holobiont selection. One go-around of the life cycle is a macro time step corresponding to the host generation time. Microbial proliferation within a host occurs in micro time steps corresponding to the microbial generation time. The two variants of the model differ in how the microbial mixing process is visualized and how the microbial inheritance occurs.

The model variant for the life cycle with vertical transmission is shown in Figure 2. This variant of the model allows for limited horizontal microbial exchange using the parameter m that indicates the fraction of the microbiome that is exchanged with other hosts. This number is assumed to be small.

To model horizontal transmission one possible modification of the vertical-transmission version is to allow the mixing probability, m, to equal 1, implying that every microbe leaves its

[2]Short programs to iterate both the vertical and horizontal transmission lifecycles written in *Mathematica* are available for download at: https://github.com/JoanKauai/Holobiont-Dynamics-Evolution. The programs are supplied as both *Mathematica* Notebooks with an .nb extension and as Computable Document Format (CDF) files with an .cdf extension. A five-minute video demonstrating the software is available at https://youtu.be/19w7ZDkRe18.

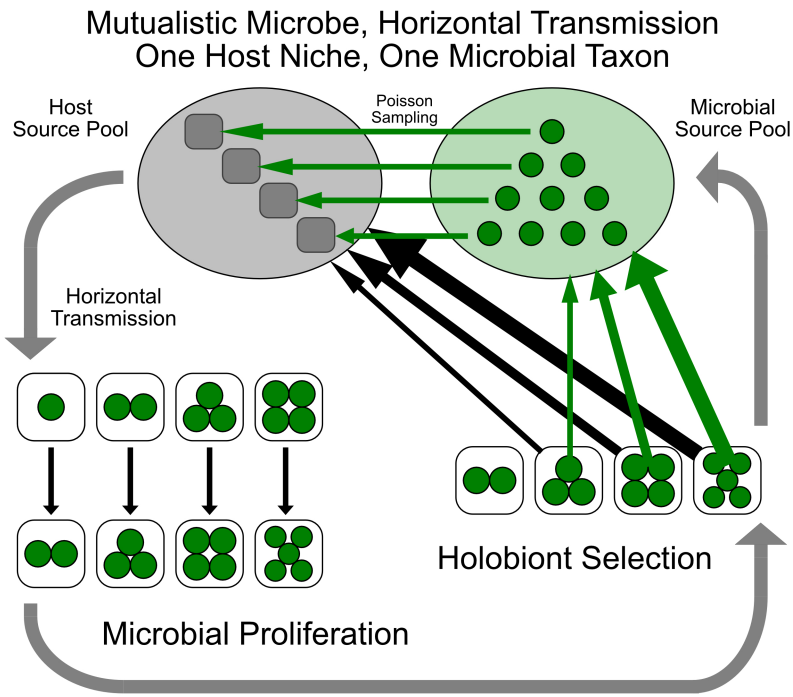


Figure 3: *Life-cycle diagram for horizontal transmission assuming a mutualistic microbe. Each generation begins with a normalized Poisson distribution of microbes across holobionts based on a density parameter equal to the number of microbes in its source pool divided by the number of juvenile hosts in its source pool. Next the microbes within each holobiont increase according to density-dependent ecological population models, as before. Finally, each holobiont reproduces as a whole. The number of progeny a holobiont produces depends on the number of microbes in it. Upon reproduction all the microbes produced from the holobionts enter a common microbe source pool and all the juvenile hosts join a common host source pool. For microbial mutualists (shown), holobionts with the highest number of microbes leave the largest number of progeny and conversely for the microbial parasites (not shown).*

host to enter the transfer pool only to be redistributed in equal measure back into now empty hosts. However, that would instantaneously stock all the hosts with identical microbiomes, denying holobiont selection any hologenotypic variation to operate on. Instead, a sampling process during microbe transmission was introduced. This sampling produces hologenotypic variation across holobionts and leads to the life cycle diagram of Figure 3.

The model has shown that directional holobiont selection causes the evolution of hosts with a reduced number of parasites and an increased number of mutualists, both for vertical and for horizontal microbiome transmission. Therefore, horizontal transmission does not prevent holobiont selection from being effective.

Horizontal transmission involves a kind of inheritance of its own, "collective inheritance"

as distinct from the "lineal inheritance" of a vertically transmitted microbiome. In collective inheritance each generation's hologenome is assembled from the collective parental hologenomes that have passed through the filter of holobiont selection.

The contrast between holobiont evolution with vertical *vs* horizontal transmission somewhat resembles the distinction between multi-level selection 1 (MLS1) and multi-level selection 2 (MLS2) (Mayo and Gilinsky 1987, Damuth and Heisler 1988, Okasha 2006). According to Okasha (2006, 58), for MLS2 "it is essential that the collectives reproduce in the ordinary sense, that is, they 'make more' collectives; otherwise, determining parent-offspring relations at the collective level will not exist." In contrast, "the role of the collectives in MLS1 is to generate a population structure for the particles, which affects their fitnesses. For MLS1 to produce sustained evolutionary consequences, collectives must 'reappear' regularly down the generations, but there is no reason why the collectives themselves must stand in parent-offspring relations." Thus, a "collective" in MLS1 is a recurring feature in the environment, possibly inanimate, that confines organisms into a social group. Indeed, Maynard Smith (1964) developed a model for MLS1 that referenced the recurring haystacks in farmers' fields as the collectives within which mice (particles) can interact.

A holobiont can, to some extent, also be identified as a "collective". If their hologenomes are vertically inherited, then parent-offspring relations exist among the holobionts of successive generations as required by the definition of ML2. In contrast, holobiont evolution with horizontal transmission is only weakly similar to MLS1 because the holobiont as a"collective" is not merely a recurring feature of the environment. Both microbes *and* hosts evolve. Although Maynard Smith's mice may evolve within haystacks, the haystacks themselves do not.

This model has demonstrated theoretically that holobiont selection can produce directional evolutionary change in holobionts with both horizontally and vertically transmitted microbiomes. In particular, extensive horizontal microbiome transmission poses no challenge to considering holobiont selection as a force in holobiont evolution. Therefore, the stage is now set to extend holobiont theory to include multiple microbial strains within the microbiome and multiple alleles within the host nucleus.

3 New Research

3.1 New Assumptions

Two new assumptions are now introduced to permit the theoretical development to move from computer iteration to mathematical analysis.

First, the generation times of the microbes are assumed to be fast relative to the generation time of the hosts, fast enough that the microbiomes within the hosts come to population dynamic

equilibrium within each macro time step. Because the generation times of microbes are usually on the order of minutes to days, whereas the generation times of hosts are on the order of years to decades, this assumption seems generally reasonable.

Second, for horizontal microbiome transmission, the microbial source pool is assumed to be dilute. To see why, suppose the source pool contains two strains of microbes. If the source pool is dense, then juvenile hosts will acquire microbes from both strains. Because of the microbe sampling process, the juvenile hosts will generally have differing initial numbers of each strain, but the point is, all the juvenile hosts will acquire at least some propagules from every strain in the source pool. In contrast, if the source pool is dilute, then some juveniles will acquire propagules from say, only strain-1, others from only strain-2, and others from both strains.

With a dense source pool, even though the initial numbers of each microbial strain might differ across juvenile hosts, this difference becomes erased as the microbes in each host approach the same population dynamic equilibrium. In contrast, with a dilute source pool, the microbiomes within each host come to different population dynamic equilibria reflecting the differing strain composition that each host begins with. Thus, microbe colonization from a dense source pool leads to the loss of hologenotypic variation among hosts whereas microbe colonization from a dilute source pool preserves the hologenotypic variation. Because holobiont selection relies on hologenotypic variation among hosts, the source pool must be dilute enough to guarantee that juvenile hosts acquire different sets of species from the microbe source pool.

3.2 Two Horizontally Transmitted Haploid Microbial Strains

3.2.1 Model

The life cycle of a holobiont with a horizontally transmitted microbiome consisting of two microbial strains in diagrammed in Figure 4.

A macro time step begins with the state variables $H(t)$, $G_1(t)$ and $G_2(t)$ where $H(t)$ is number of empty hosts in the host pool at time t, and $G_1(t)$ and $G_2(t)$) are the number of microbes of strain-1 and strain-2 in the microbial pool at time t. The total number of microbes is $G(t) \equiv G_1(t) + G_2(t)$. The frequencies of strain-1 and strain-2 in the microbial pool at time t are $p_1(t) \equiv G_1(t)/G(t)$ and $p_2(t) \equiv G_2(t)/G(t)$ where $p_1(t) + p_2(t) \equiv 1$.

Given that the microbe pool in the water column surrounding the empty hosts is dilute, one may posit that two microbe propagules colonize each host at random according to a binomial distribution. Thus, after the microbiome colonization phase,

$$\begin{aligned} H'(t,2,0) &= p_1(t)^2 H(t) \\ H'(t,1,1) &= 2\, p_1(t)\, p_2(t) H(t) \\ H'(t,0,2) &= p_2(t)^2 H(t) \end{aligned}$$

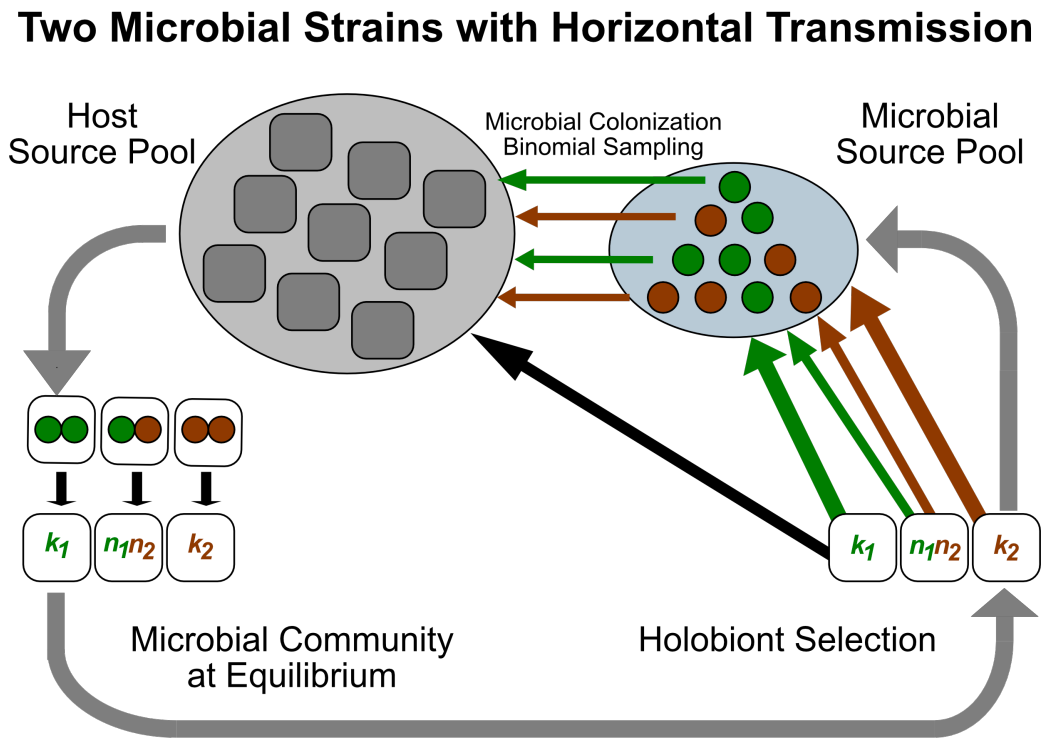


Figure 4: *Holobiont life cycle where microbiome consists of one or of two microbial strains (green and/or brown circle). In each generation the microbiomes for the juvenile hosts are assembled by sampling from the microbial source pool. The microbial source pool is dilute—hologenotypes are assembled via binomial sampling of microbial strains. Generation time of microbes is short relative to host—microbiome comes to community equilibrium within each host generation.*

where $H(t, n_1, n_2)$ is the number of holobionts with a hologenotype consisting of n_1 microbes of strain-1 and n_2 microbes of strain-2 at time t.

Once the juvenile hosts have been initially populated, the microbes proliferate within their hosts, coming to an equilibrium microbiome community structure. The equilibrium abundances in a dual-strain microbiome consisting of strain-1 and strain-2 are $\widehat{n}_1$ and $\widehat{n}_2$. The equilibrium abundances of pure strain microbiomes consisting of only strain-1 or strain-2 are k_1 and k_2. Therefore, after the microbiome proliferation phase

$$\begin{aligned} H''(t, k_1, 0) &= p_1(t)^2 H(t) \\ H''(t, \widehat{n}_1, \widehat{n}_2) &= 2\, p_1(t)\, p_2(t) H(t) \\ H''(t, 0, k_2) &= p_2(t)^2 H(t) \end{aligned}$$

Next, the holobiont fitness depend on the hologenotype, $W(n_1, n_2)$. Accordingly, after holo-

biont selection, the hologenotype numbers are

$$\begin{aligned} H'''(t,k_1,0) &= W(k_1,0)\, p_1(t)^2 H(t) \\ H'''(t,\widehat{n_1},\widehat{n_2}) &= W(\widehat{n_1},\widehat{n_2})\, 2\, p_1(t)\, p_2(t) H(t) \\ H'''(t,0,k_2) &= W(0,k_2)\, p_2(t)^2 H(t) \end{aligned}$$

To simplify notation put $(W(k_1,0),\ W(\widehat{n_1},\widehat{n_2}),\ W(0,k_2))$ as (W_1, W_{12}, W_2), so the the equations above are rewritten as

$$\begin{aligned} H'''(t,k_1,0) &= W_1\, p_1(t)^2 H(t) \\ H'''(t,\widehat{n_1},\widehat{n_2}) &= W_{12}\, 2\, p_1(t)\, p_2(t) H(t) \\ H'''(t,0,k_2) &= W_2\, p_2(t)^2 H(t) \end{aligned}$$

These holobionts then release their microbiomes into the microbe pool and also produce empty juvenile hosts ready for the newly released microbes to colonize.

So, to obtain the state variables at time $t+1$ given the state at time t, collect the previous terms. The total number of empty juvenile hosts at $t+1$ is

$$\begin{aligned} H(t+1) &= H'''(t,k_1,0) + H'''(t,\widehat{n_1},\widehat{n_2}) + H'''(t,0,k_2) \\ &= W_1\, p_1(t)^2 H(t) + W_{12}\, 2\, p_1(t)\, p_2(t) H(t) + W_2\, p_2(t)^2 H(t) \\ &= p_1(t)\, \overline{W_1}(t) + p_2(t)\, \overline{W_2}(t) \\ &= \overline{\overline{W}}(t)\, H(t) \end{aligned}$$

where the holobiont mean fitness, $\overline{\overline{W}}(t)$, and the marginal host fitnesses, $\overline{W_1}(t)$ and $\overline{W_2}(t)$, at time t are

$$\begin{aligned} \overline{W_1}(t) &\equiv p_1(t)\, W_1 + p_2(t) W_{12} \\ \overline{W_2}(t) &\equiv p_1(t) W_{12} + p_2(t)\, W_2 \\ \overline{\overline{W}}(t) &\equiv p_1(t)^2\, W_1 + 2\, p_1(t)\, p_2(t)\, W_{12} + p_2(t)^2\, W_2 \end{aligned}$$

An overline over the W indicates an average taken over the microbiome compositions, ie, with respect to p_1 and p_2. Two overlines indicate two successive averages with respect to p_1 and p_2.

To simplify notation further, put $(\widehat{n_1},\ \widehat{n_2})$ as $(n_1,\ n_2)$. Then the number of strain-1 microbes in the microbe pool at time $t+1$ is

$$\begin{aligned} G_1(t+1) &= k_1\, H'''(t,k_1,0) + \widehat{n_1}\, H'''(t,\widehat{n_1},\widehat{n_2}) \\ &= k_1\, W_1\, p_1(t)^2 H(t) + n_1\, W_{12}\, 2\, p_1(t)\, p_2(t) H(t) \end{aligned}$$

The products of microbe population sizes and holobiont fitnesses can be combined into measures of multilevel microbe fitness per holobiont that include both success within the holobiont and holobiont success itself. These multilevel microbe fitnesses per holobiont are

$$\begin{aligned} w_{1,1} &\equiv k_1\, W_1 \\ w_{1,12} &\equiv n_1\, W_{12} \\ w_{2,12} &\equiv n_2\, W_{12} \\ w_{2,2} &\equiv k_2\, W_2 \end{aligned}$$

These new coefficients refer to the multilevel success of a specific microbe strain within a specific microbiome. Thus, $w_{1,1}$ is the multilevel fitness per holobiont of a strain-1 microbe in a single-strain microbiome consisting only of strain-1. $w_{1,12}$ is the multilevel fitness per holobiont of a strain-1 microbe in a dual strain microbiome, and so forth. The notation convention is that upper case W's refer to holobiont fitnesses and lower case w's refer to microbe fitnesses. With these definitions, define strain-1's marginal multilevel fitness per holobiont as

$$\overline{w_1}(t) \equiv p_1(t)\, w_{1,1} + p_2(t)\, 2\, w_{1,12}$$

For dual-strain microbiomes, a factor of 2 is included because a strain-1 microbe can be the first or second of the two microbes that colonize the holobiont as a juvenile—this effectively doubles the number of holobionts that carry strain-1. The number of strain-1 microbes placed into the microbe source pool at time $t+1$ from both single and dual strain microbiomes then is

$$G_1(t+1) = \overline{w_1}(t)\, p_1(t)\, H(t)$$

Similarly, the number of strain-2 microbes in the microbe pool at time $t+1$ is

$$\begin{aligned} G_2(t+1) &= \widehat{n_2}\, H'''(t, \widehat{n_1}, \widehat{n_2}) + k_2\, H'''(t, 0, k_2) \\ &= n_2\, W_{12}\, 2\, p_1(t)\, p_2(t) H(t) + k_2\, W(0, k_2)\, p_2(t)^2 H(t) \\ &= \overline{w_2}(t)\, p_2(t)\, H(t) \end{aligned}$$

where strain-2's marginal multilevel fitness per holobiont is

$$\overline{w_2}(t) \equiv p_1(t)\, 2\, w_{2,12} + p_2(t)\, w_{2,2}$$

The total number of microbes in the microbe pool at time $t+1$ then is

$$\begin{aligned} G(t+1) &= \left(p_1(t)\, \overline{w_1}(t) + p_2(t)\, \overline{w_2}(t)\right)\, H(t) \\ &= \overline{\overline{w}}(t)\, H(t) \end{aligned}$$

where the microbiome multilevel mean fitness per holobiont is

$$\overline{\overline{w}}(t) \equiv p_1(t)\,\overline{w_1}(t) + p_2(t)\,\overline{w_2}(t)$$

Therefore the microbe frequencies in the microbial pool at time $t+1$ are

$$p_1(t+1) = \frac{G_1(t+1)}{G(t+1)} = \frac{\overline{w_1}(t)}{\overline{\overline{w}}(t)}\,p_1(t)$$

and

$$p_2(t+1) = \frac{G_2(t+1)}{G(t+1)} = \frac{\overline{w_2}(t)}{\overline{\overline{w}}(t)}\,p_2(t)$$

To summarize, the equations above for $H(t+1)$, $G(t+1)$, $p_1(t+1)$ and $p_2(t+1)$ together constitute a dynamical system for holobiont evolution with two microbe strains assuming (a) horizontal microbe transmission, (b) dilute microbe source pools and (c) macro time steps much longer than micro time steps. These equations correspond to those from classical population genetics for evolution at one locus with two alleles. Those equations predict the trajectory of alleles in the gene pool whereas these equations predict the trajectory of microbial strains in the hologene pool.

The expressions for the multilevel microbe fitnesses suggest that microbes "view" their holobiont as a vehicle to increase their productivity. Because the multilevel microbe fitness is a product of microbe abundance within the holobiont times the fitness of the holobiont as a whole, the microbes increase in the hologene pool through a combination of K-selection within the holobionts and r-selection between holobionts.

3.2.2 Boundary Analysis—Conditions for Increase When Rare

The condition for strain-1 to increase when rare

$$p_1(t+1) > p_1(t)\big|_{p_1\approx 0, p_2\approx 1}$$

works out to be

$$E_1 \equiv 2\,w_{1,12} - w_{2,2} > 0$$

Similarly, the condition for strain-2 to increase when rare works out to be

$$E_2 \equiv 2\,w_{2,12} - w_{1,1} > 0$$

E_1 is the excess by which a rare strain-1 microbe out-produces a common strain-1 microbe. E_2 is the excess by which a rare strain-2 microbe out-produces a common strain-1 microbe. Strain-1 increases when it is rare if it out-produces strain-2 when strain-2 is common. Similarly, strain-2 increases when it is rare if it out-produces strain-1 when strain-1 is common. If both strain-1

and strain-2 can increase when rare, then the microbiome community consists of two strains. Otherwise, if say, strain-1 can increase when rare and strain-2 cannot, then strain-1 excludes strain-2 from the microbiome. And conversely if strain-1 cannot increase when rare while strain-2 can, then strain-2 excludes strain-1 from the microbiome.

3.2.3 Illustration—Altruistic and Selfish Microbes

To illustrate, consider two microbes, one altruistic and one selfish. Suppose an initially rare strain-1 (green) is altruistic in the holobiome because it transfers some of its resources to the host thereby lowering its own carrying capacity while increasing the holobiont fitness. Meanwhile the selfish common strain-2 (brown) keeps its resources to itself. Thus, $W_{12} > W_2$ while $k_1 < k_2$.

Suppose further that the two microbes compete according the the Lotka-Volterra competition equations and that the reciprocal competition coefficients are equal, $a_{12} = a_{21} = a$. The equilibrium abundances of the strains then are

$$n_1 = \frac{k_1 - a\,k_2}{1 - a^2},\ n_2 = \frac{k_2 - a\,k_1}{1 - a^2}$$

Now, as a concrete example, let, $a \to 1/2$, $k_1 \to 300$, $k_2 \to 400$, $W_1 \to 3$, $W_{12} \to 2$, $W_2 \to 1$. Then the condition for strain-1 to increase when rare works out to be

$$E_1 \equiv 2\,w_{1,12} - w_{2,2} = 400/3 > 0$$

Thus, strain-1 which has lowed its carrying capacity, k, from 400 to 300 but increased the fitness of a holobiont it resides in from 1 to 2, increases when rare against the selfish strain.

But after increasing when rare does this altruistic strain-1 proceed to exclude the selfish strain-2? To find out check the condition for increase when rare for strain-2, which works out to be

$$E_2 \equiv 2\,w_{2,12} - w_{1,1} = 1300/3 > 0$$

This condition is positive too, showing that altruistic strain-1 does not exclude the selfish strain-2. Therefore, both strain-1 and strain-2 coexist in this numerical example.

3.2.4 Microbiome Polymorphism

Seeing that strain-1 and strain-2 coexist in the numerical example above, the next question is how do they coexist? Do they come to come an equilibrium polymorphism, or do they oscillate in a limit cycle? Solving for the equilibrium polymorphism between the two strains, and analyzing the stability of that equilibrium will answer this question.

Putting $p_1(t+1) = p_1(t) \equiv \hat{p}$ and solving for the equilibrium frequency of strain-1 yields

$$\hat{p} = \frac{E_1}{E_1 + E_2}$$

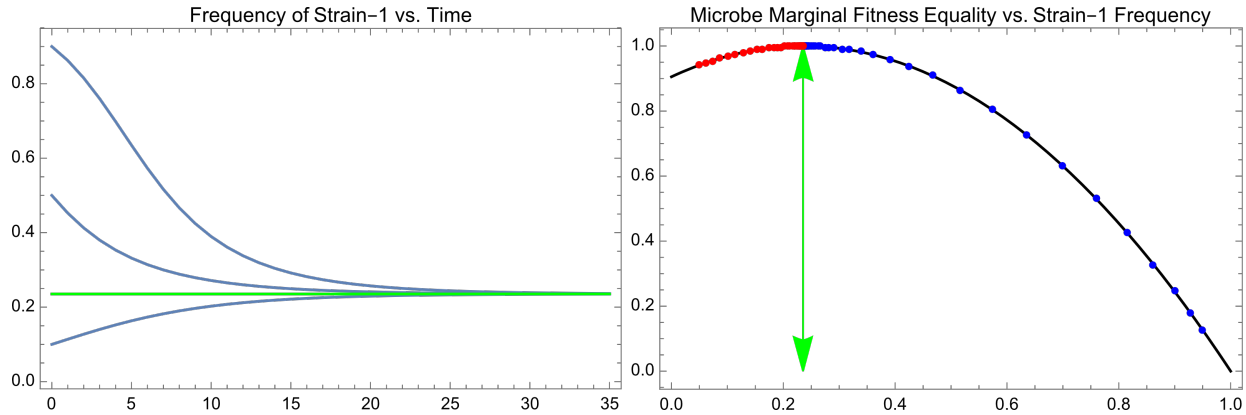


Figure 5: *Left: Frequency of the strain-1 microbe in the microbial pool vs. time, assuming strain-1 is altruistic and strain-2 is selfish. Equilibrium strain-1 frequency is horizontal line in green. Figure shows that different initial conditions all converge to the equilibrium. Right: Degree of marginal multilevel fitness equality in the microbiome vs. strain-1's frequency, assuming strain-1 is altruistic and strain-2 is selfish. Curve for degree of marginal multilevel microbial fitness equality is in black. Equilibrium strain-1 frequency is vertical arrow in green. Figure shows two different initial conditions, one (red) at* $p = 0.05$ *and another (blue) at* $p = 0.95$ *that both converge to the equilibrium at the top of the curve of fitness equality.* ($a \to 1/2$, $k_1 \to 300$, $k_2 \to 400$, $n_1 \to 400/3$, $n_2 \to 1000/3$, $W_1 \to 3$, $W_{12} \to 2$, $W_2 \to 1$)

If both strains increase when rare, then p_1 is in $(0, 1)$ as it should be. The value of the equilibrium hologenotype frequency reflects the relative speeds at which the strains can increase when rare. If strain-1 can charge ahead at gang-buster speed whereas strain-2 can barely squeeze in, then the equilibrium consists mostly of strain-1 with just a bit of strain-2.

Returning to the numerical example of altruistic and selfish microbes mentioned above, the equilibrium frequency of strain-1 works out to be

$$\hat{p} = \frac{400/3}{400/3 + 1300/3} = \frac{4}{17}$$

Therefore the full picture of the holobiont population along with its microbiomes at equilibrium is: $(4/17) = 0.235$ of the microbial pool consists of the altruistic strain-1 microbes and $(13/17) = 0.764$ of the microbial pool consists of the selfish strain-2 microbes; after microbiome assembly, $(4/17)^2 = 0.055$ of the holobionts contain solely strain-1 microbes, $2 \times (4/17) \times (13/17) = 0.359$ of the holobionts contain both microbes and $(13/17)^2 = 0.584$ of the holobionts contain only strain-2 microbes; the holobionts with only strain-1 contain 300 microbes, holobionts with both strains each contain $(400/3) = 133$ strain-1 microbes and $(1000/3) = 333$ strain-2 microbes, and holobionts with only strain-2 contain 400 microbes in their microbiomes.

Figure 5 (Left) presents a plot of trajectories for the frequency of strain-1 through time from

several initial conditions obtained by numerically iterating equation the equation for $p_1(t+1)$. The red line is the predicted equilibrium frequency. The trajectories all converge to the equilibrium.

3.2.5 Stability and The Direction of Microbiome Evolution

The observation that trajectories of $p_1(t)$ converge numerically to an equilibrium as shown in Figure 5 (Left) suggests that the polymorphism frequency for strain-1, p_1 is a stable polymorphism.

To analyze the stability of this equilibrium, write the dynamics as difference equation for $\Delta p_1(t) \equiv p_1(t+1) - p_1(t)$ yielding, after rearrangement,

$$\Delta p = \frac{p\,(1-p)\,(\overline{w_1} - \overline{w_2})}{\overline{\overline{w}}}$$

where to simplify notation, the subscript and time index are dropped leaving $p \equiv p_1(t)$.

Next, consider the expression,

$$-(\overline{w_1} - \overline{w_2})^2$$

which measures the degree of equality between the marginal multilevel mean fitnesses per holobiont of the two microbes. This is a quadratic function of p, a $\cap$-shaped curve whose top is at the equilibrium value of p. By taking the derivative of this function with respect to p, rearranging and substituting into the equation for Δp one obtains,

$$\Delta p = \frac{1}{2\,(E_1 + E_2)}\;\frac{p\,(1-p)}{\overline{\overline{w}}}\;\left(\frac{d\;\left(-(\overline{w_1} - \overline{w_2})^2\right)}{d\,p}\right)$$

This equation for Δp shows that holobiont selection causes p to ascend the $\cap$-shaped curve to its top where the equilibrium occurs. The expression in large parentheses is the curve's slope at p. This expression shows that if p is at a spot where the curve's slope is positive, i.e., p is to the left of the top, then holobiont selection increases p, moving p to the right and closer to the top. Conversely, if p is at a spot where the curve's slope is negative, i.e., p is to the right of the top, then holobiont selection decreases p, moving p to the left and closer to the top. If p is exactly at the top, then Δp is zero, and equilibrium has been attained. Thus, $-(\overline{w_1} - \overline{w_2})^2$ constitutes a curve of the degree of equality between the marginal multilevel fitness per holobiont of the two strains. The result here that p ascends to the top of the curve of multilevel marginal fitness equality is the counterpart for holobiont selection to the result in classical population genetics where p climbs an "adaptive surface". Unlike classical population genetics however, the peak of the curve of marginal fitness equality per holobiont between the two microbe strains does not coincide with the peak of the overall mean multilevel microbe fitness per holobiont, $\overline{\overline{w}}(t)$, (nor with the overall mean holobiont fitness $\overline{\overline{W}}(t)$).

Although $-(\overline{w_1} - \overline{w_2})^2$ represents a curve that p ascends during holobiont evolution, this expression can be scaled and a constant added to it to produce a similar curve that is conveniently graphed between 0 and 1. So, put

$$L_G(p) \equiv \frac{-(\overline{w_1} - \overline{w_2})^2}{\max(E_1^2, E_2^2)} + 1$$

Figure 5 (Right) illustrates both $L_G(p)$ and some trajectories climbing toward its peak.

Even though the equation for Δp above shows that trajectories climb the surface of marginal multilevel mean fitness equality for the two microbes, this equation alone is not sufficient to show that the equilibrium at the top is stable. It remains theoretically conceivable that as trajectories approach the top they overshoot, and even overshoot enough to diverge away from the equilibrium. This would be realized as trajectories that constitute a multi-point limit cycle, first to the left of the equilibrium, then overshooting to the right of the equilibrium, then overshooting back to the left and so forth. This theoretical possibility can be ruled out by considering the eigenvalue of the mapping of $p(t) \to p(t+1)$ evaluated at the equilibrium p. For overshoots to occur, this quantity must be negative at the equilibrium p. And for divergent overshoots, this quantity must be strongly negative, < -1. However, this quantity works out to be positive for any p, including the equilibrium p. Therefore, overshoots do not occur, either convergent or divergent. Thus, p the top of the surface of mean fitness equality per holobiont between the two microbe strains is a globally stable equilibrium that trajectories approach smoothly.

3.3 Two Vertically Transmitted Diploid Host Alleles

The life cycle of a holobiont with two possible alleles in the host nucleus, A_1 and A_2, and with a single-strain horizontally transmitted microbiome is diagrammed in Figure 6.

A macro time step begins with the gamete pool. The frequencies of A_1 and A_2 in the gamete pool at time t are $q_1(t)$ and $q_2(t)$ where $q_1(t) + q_2(t) \equiv 1$. Then gametes unite at random leading to a (Hardy-Weinberg) binomial distribution such that initial hologenotypes occur in the ratios: $(q_1(t)^2, 2\,q_1(t)\,q_2(t), q_2(t)^2)$. These juvenile hosts are then colonized by the single strain of microbes.

Once the juvenile hosts have been colonized, the microbes proliferate within their hosts, coming to an equilibrium microbe population size that depends on the host genotype: (k_{11}, k_{12}, k_{22}) for host nuclear genotypes (A_1A_1, A_1A_2, A_2A_2) respectively.

Next, the holobiont fitness depend on the hologenotype after microbiome proliferation, ($W(k_{11})$, $W(k_{12})$, $W(k_{22})$) which, to simplify notation, are rewritten as (W_{11}, W_{12}, W_{22}). Based on these fitnesses, the holobionts release their gametes into the gamete pool whereupon they unite at random to yield the empty juvenile hosts ready for microbe colonization. Meanwhile, the holobionts also release microbes that replenish the microbial pool.

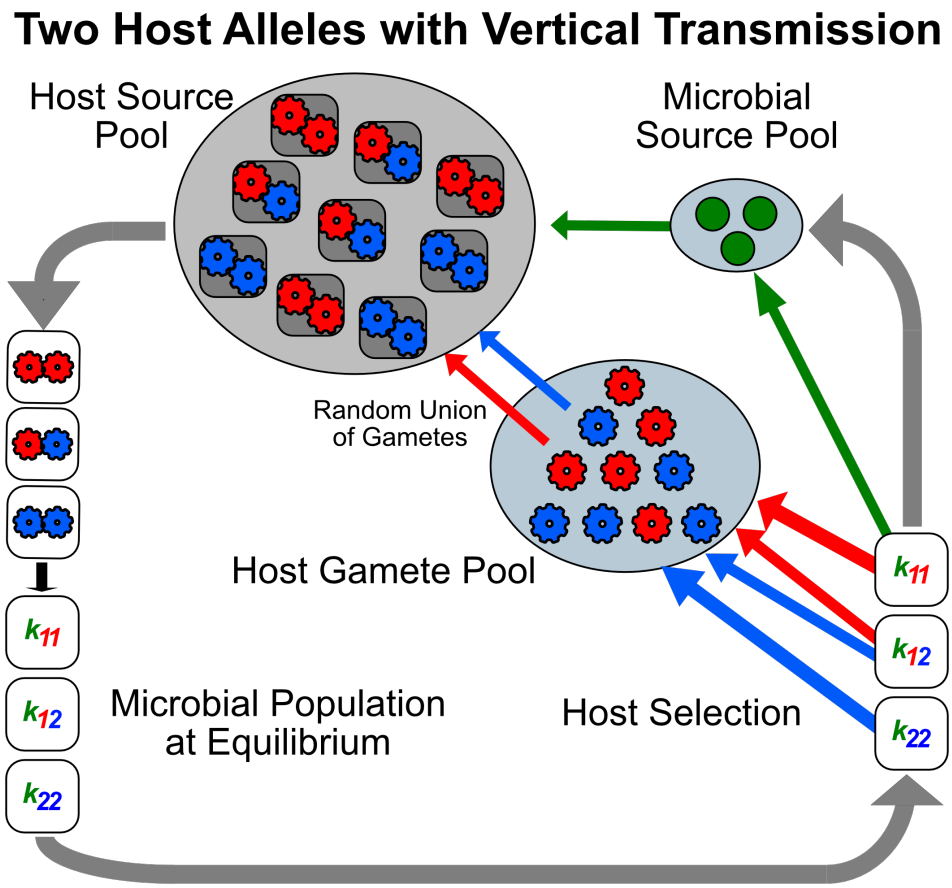


Figure 6: *Holobiont life cycle where the host nuclei contain two diploid alleles (red and/or blue cogs). In each generation the host's nuclear genotype is assembled by by random union of gamete from the preceding generation. The microbial source pool contains one microbial strain (green circles). Generation time of microbes is short relative to host—microbe population comes to equilibrium within each host generation.*

The equation to map $q_{1,t} \to q_{1,t+1}$ is the standard equation for selection at one locus with two alleles from classical population genetics. The only new wrinkle here is that the fitness of a host depends on the population size of the microbes in it, as though the microbe population size were itself a host trait, which it effectively is because k, the host's microbe carrying capacity, is a host trait. Thus, from classical population genetics,

$$\begin{aligned} q_1(t+1) &= \frac{q_1(t)W_{11} + q_2(t)W_{12}}{q_1(t)^2 W_{11} + 2\,q_1(t)\,q_2(t)W_{12} + q_2(t)^2 W_{22}}\; q_1(t) \\ &= \frac{\widetilde{W}_1(t)}{\widetilde{\widetilde{W}}(t)}\; q_1(t) \end{aligned}$$

where

$$\begin{aligned}
\widetilde{W_1}(t) &\equiv q_1(t)W_{11} + q_2(t)W_{12} \\
\widetilde{W_2}(t) &\equiv q_2(t)W_{22} + q_1(t)W_{12} \\
\widetilde{\widetilde{W}}(t) &\equiv q_1(t)^2 W_{11} + 2\,q_1(t)\,q_2(t)W_{12} + q_2(t)^2 W_{22} \\
&\equiv q_1(t)\widetilde{W_1}(t) + q_2(t)\widetilde{W_2}(t)
\end{aligned}$$

Here, $\widetilde{W_1}(t)$ is the marginal fitness of the A_1 allele, $\widetilde{W_2}(t)$ is the marginal fitness of the A_2 allele, and $\widetilde{\widetilde{W}}(t)$ is the overall mean fitness. (The *wide tilde* above any fitness coefficient indicates an average taken over the host allele frequencies, ie, with respect to q_1 and q_2. Contrast in the previous section where an *overline* above a fitness coefficient indicates an average taken with respect to microbiome strain frequencies.)

The boundary analysis for this equation is well known. q_1 can increase when rare if

$$q_1(t+1) > q_1(t)\big|_{q_1 \approx 0, q_2 \approx 1}$$

which works out to be

$$F_1 \equiv W_{12} - W_{22} > 0$$

And similarly, q_2 can increase when rare if

$$F_2 \equiv W_{12} - W_{11} > 0$$

Polymorphism occurs if both alleles can increase when rare, $F_1 > 0$ and $F_2 > 0$. The polymorphism allele frequency where $q(t+1) = q(t)$ then works out to be

$$\hat{q} = \frac{F_1}{F_1 + F_2}$$

Next, develop the equation for $\Delta q \equiv q_1(t+1) - q_1(t)$ as

$$\Delta q = \frac{q(1-q)(\widetilde{W}_1 - \widetilde{W}_2)}{\widetilde{\widetilde{W}}}$$

Equilibrium results when the marginal fitness of the two alleles are equal. In classical population genetics, the change in allele frequencies is customarily associated with climbing the mean fitness function, $\widetilde{\widetilde{W}}$. Although the polymorphism frequency for the host alleles does lie at the peak of $\widetilde{\widetilde{W}}$, selection maximizes other functions as well.

To parallel the formulation in the last section, consider the expression

$$-(\widetilde{W}_1 - \widetilde{W}_2)^2$$

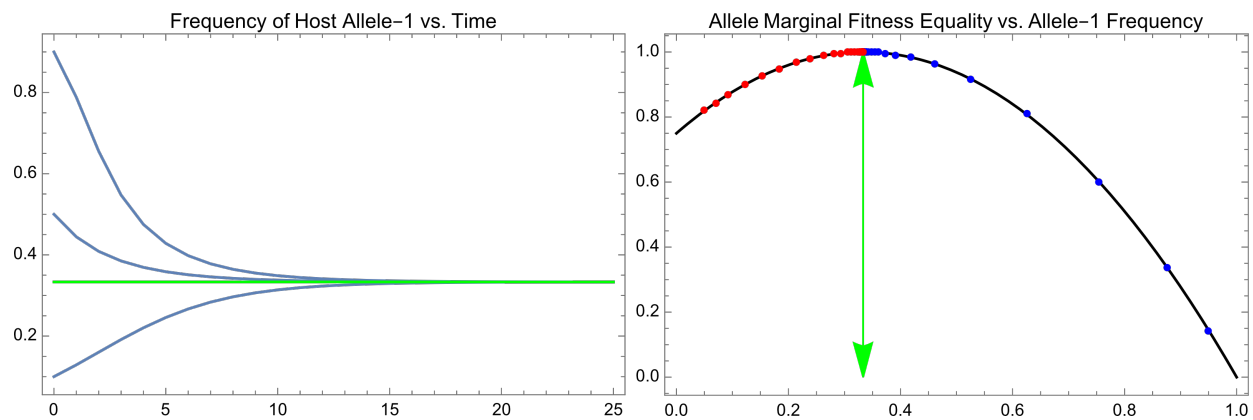


Figure 7: *Left: Frequency of the allele-1 in the host gamete pool vs. time, assuming heterozygote superiority. Equilibrium allele-1 frequency is horizontal line in green. Figure shows that different initial conditions all converge to the equilibrium. Right: Degree of allele marginal fitness equality in host vs. allele-1's frequency in the gamete pool, assuming heterozygote superiority. Curve for degree of allele marginal equality is in black. Equilibrium allele-1 frequency is vertical arrow in green. Figure shows two different initial conditions, one (red) at* $p = 0.05$ *and another (blue) at* $p = 0.95$*, whose trajectories both converge to the equilibrium at the top of the curve.* ($W_{11} \to 1$, $W_{12} \to 3$, $W_{22} \to 2$)

This expression represents the degree of equality in marginal fitness of the two alleles. It is a $\cap$-shaped quadratic function of q whose peak is at $\hat{q}$. Rewriting the equation for Δq in terms of the derivative of this expression yields

$$\Delta q = \frac{1}{2\,(F_1 + F_2)}\,\frac{q\,(1-q)}{\widetilde{\widetilde{W}}}\,\left(\frac{d\,\left(-(\widetilde{W_1} - \widetilde{W_2})^2\right)}{d\,q}\right)$$

This equation indicates that $q(t)$ climbs the curve for the degree of marginal fitness equality between the two alleles.

Scaling the expression, $-(\widetilde{W}_1 - \widetilde{W}_2)^2$, to lie between 0 and 1 yields

$$L_H(q) \equiv \frac{-(\widetilde{W}_1 - \widetilde{W}_2)^2}{\max(F_1^2, F_2^2)} + 1$$

Trajectories from holobiont selection on the host alleles climb this function representing the degree of marginal fitness equality for the two host alleles and come to equilibrium when equality is attained.

Figure 7 illustrates a numerical example with heterozygote superiority in the host alleles, $W_{11} \to 1$, $W_{12} \to 3$, $W_{22} \to 2$. The equilibrium, q, works out to be $1/3$. Figure 7 (Left) illustrates trajectories approaching the equilibrium through time from different initial conditions, and Figure 7 (Right) shows $L_h(q)$ and two trajectories climbing toward its peak.

3.4 Two Microbial Strains and Two Host Alleles

The life cycle of a holobiont with both two vertically transmitted alleles in the host nucleus and two horizontally transmitted microbial strains in the microbiome is diagrammed in Figure 8. This figure combines the elements both of Figure 4 for two microbes and of Figure 6 for two host alleles. The task now is to develop the equations that simultaneously map both microbe and allele frequencies at time t to their frequencies at time $t+1$.

3.4.1 Model

A macro time step begins with the basic state variables for the nine hologenotypes. The notation convention is that $H_{ij}(t,k,0)$ refers to the number of holobionts at time t with k strain-1 microbes and no strain-2 microbes in the microbiome together with alleles A_i and A_j in the host nucleus, whereas $H_{ij}(t,n_1,n_2)$ refers to the number of holobionts at time t with n_1 strain-1 and n_2 strain-2 microbes in the microbiome together with alleles A_i and A_j in the host nucleus. The total at time t of the nine hologenotype numbers is $H(t)$.

Let p_1 and p_2 denote the frequencies of strain-1 and strain-2 in the microbe source pool, and let q_1 and q_2 denote the frequencies of the A_1 and A_2 alleles in the host gamete pool. The nuclei of the empty hosts are formed by random union of gametes from the host gamete pool according to the Hardy-Weinberg ratios of $(q_1^2, 2\,q_1q_2, q_2^2)$. (Hardy-Weinberg ratios would also be produced through random mating by the parental hosts.) Then these empty hosts are colonized at random through binomial sampling of the dilute microbe pool according to the ratios of $(p_1^2, 2\,p_1p_2, p_2^2)$. Random union of gametes (or random mating) followed by random colonization of microbes leads to the following nine initial hologenotype numbers:

$$
\begin{aligned}
H'_{11}(t,2,0) &= q_1(t)^2\ p_1(t)^2\ H(t) \\
H'_{11}(t,1,1) &= q_1(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H'_{11}(t,0,2) &= q_1(t)^2\ p_2(t)^2\ H(t) \\
H'_{12}(t,2,0) &= 2q_1(t)q_2(t)\ p_1(t)^2\ H(t) \\
H'_{12}(t,1,1) &= 2q_1(t)q_2(t)\ 2p_1(t)p_2(t)\ H(t) \\
H'_{12}(t,0,2) &= 2q_1(t)q_2(t)\ p_2(t)^2\ H(t) \\
H'_{22}(t,2,0) &= q_2(t)^2\ p_1(t)^2\ H(t) \\
H'_{22}(t,1,1) &= q_2(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H'_{22}(t,0,2) &= q_2(t)^2\ p_2(t)^2\ H(t)
\end{aligned}
$$

The notation reads as follows: $H'_{11}(t,2,0)$ is the number of holobionts after random union of gametes and binomial microbe colonization with solely strain-1 as its microbiome and A_1A_1 as

its nuclear genome—these holobionts have 2 microbes in them of strain-1 and 0 microbes of strain-2, and so forth for the remaining hologenotype numbers.

After the juvenile hosts have been initially populated, the microbes proliferate within their hosts and come to an equilibrium microbiome community structure. The equilibrium microbe abundances in a dual-strain microbiome consisting of strain-1 and strain-2 in a host with an A_iA_j genome are $n_{1,ij}$ and $n_{2,ij}$. The equilibrium microbe abundances of pure strain microbiomes consisting of only strain-1 or only strain-2 in a host with an A_iA_j genome are $k_{1,ij}$ and $k_{2,ij}$ respectively. Therefore, after the microbiome proliferation phase, the hologenotype numbers are:

$$\begin{aligned}
H''_{11}(t, k_{1,11}, 0) &= q_1(t)^2\ p_1(t)^2\ H(t) \\
H''_{11}(t, n_{1,11}, n_{2,11}) &= q_1(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H''_{11}(t, 0, k_{2,11}) &= q_1(t)^2\ p_2(t)^2\ H(t) \\
H''_{12}(t, k_{1,12}, 0) &= 2q_1(t)q_2(t)\ p_1(t)^2\ H(t) \\
H''_{12}(t, n_{1,12}, n_{2,12}) &= 2q_1(t)q_2(t)\ 2p_1(t)p_2(t)\ H(t) \\
H''_{12}(t, 0, k_{2,12}) &= 2q_1(t)q_2(t)\ p_2(t)^2\ H(t) \\
H''_{22}(t, k_{1,22}, 0) &= q_2(t)^2\ p_1(t)^2\ H(t) \\
H''_{22}(t, n_{1,22}, n_{2,22}) &= q_2(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H''_{22}(t, 0, k_{2,22}) &= q_2(t)^2\ p_2(t)^2\ H(t)
\end{aligned}$$

Next, the holobiont fitnesses depend on the full hologenotype which includes both the microbiome and the host nuclear genome, eg., $W_{ij}(k_1)$ is the fitness of a single-strain holobiont with strain-1 in a host with A_iA_j alleles, and $W_{ij}(n_1, n_2)$ is the fitness for a dual-strain holobiont with n_1 strain-1 microbes and n_2 strain-2 microbes with a host genome of A_iA_j. However, to simplify notation, put $W_{ij}(k_1)$ as $W_{1,ij}$ and $W_{ij}(n_1, n_2)$ as $W_{12,ij}$ where the subscript refers first to the microbiome composition followed by the nuclear genome. Accordingly, after holobiont selection, the hologenotype numbers are:

$$\begin{aligned}
H'''_{11}(t, k_{1,11}, 0) &= W_{1,11}\ q_1(t)^2\ p_1(t)^2\ H(t) \\
H'''_{11}(t, n_{1,11}, n_{2,11}) &= W_{12,11}\ q_1(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H'''_{11}(t, 0, k_{2,11}) &= W_{2,11}\ q_1(t)^2\ p_2(t)^2\ H(t) \\
H'''_{12}(t, k_{1,12}, 0) &= W_{1,12}\ 2q_1(t)q_2(t)\ p_1(t)^2\ H(t) \\
H'''_{12}(t, n_{1,12}, n_{2,12}) &= W_{12,12}\ 2q_1(t)q_2(t)\ 2p_1(t)p_2(t)\ H(t) \\
H'''_{12}(t, 0, k_{2,12}) &= W_{2,12}\ 2q_1(t)q_2(t)\ p_2(t)^2\ H(t) \\
H'''_{22}(t, k_{1,22}, 0) &= W_{1,22}\ q_2(t)^2\ p_1(t)^2\ H(t) \\
H'''_{22}(t, n_{1,22}, n_{2,22}) &= W_{12,22}\ q_2(t)^2\ 2p_1(t)p_2(t)\ H(t) \\
H'''_{22}(t, 0, k_{2,22}) &= W_{2,22}\ q_2(t)^2\ p_2(t)^2\ H(t)
\end{aligned}$$

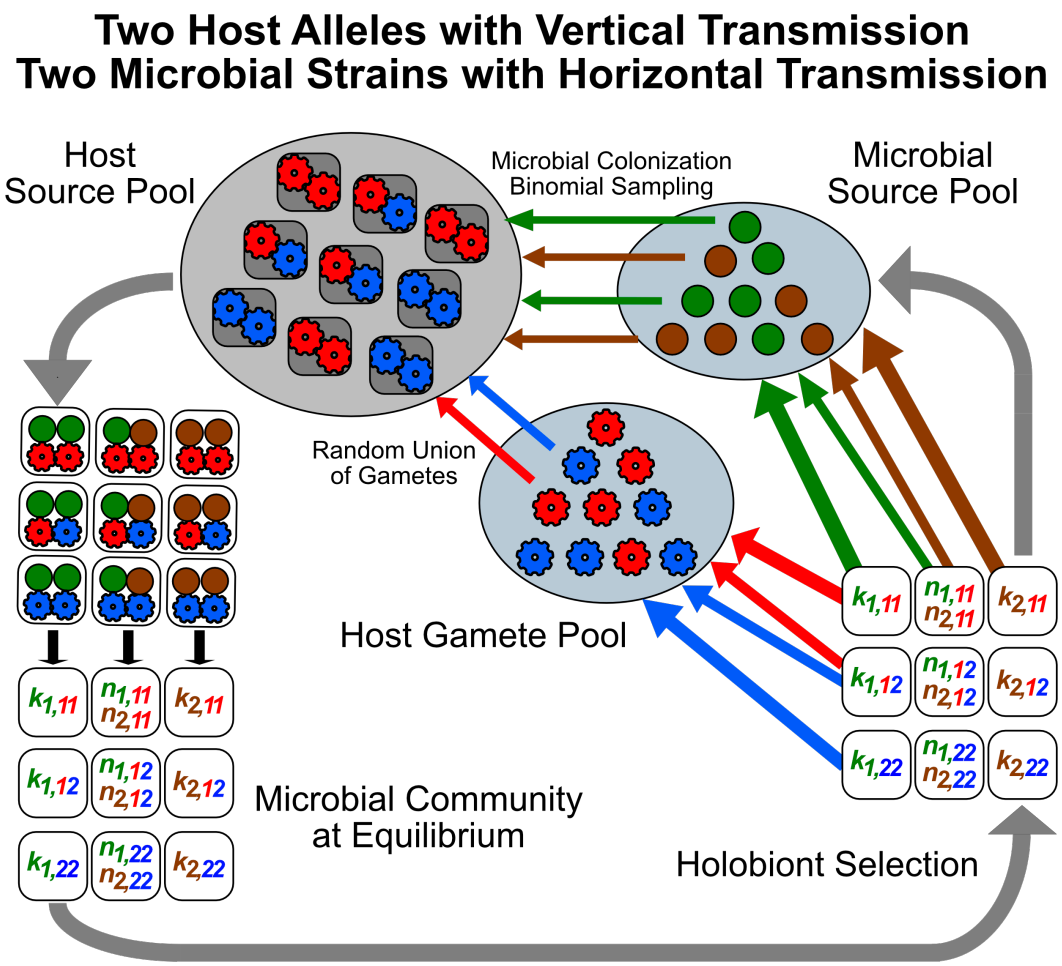


Figure 8: *Holobiont life cycle for holobiont whose microbiome consists of one or two microbial strains (green and/or brown circle) and whose nucleus consists of one allele (green or brown cog). In each generation the microbiomes for the juvenile hosts are assembled through binomial sampling from the microbial source pool. The host nuclear alleles are transmitted vertically from parent to juvenile with random union of gametes. At the holobiont selection stage, the diagram's columns indicate the strains being contributed to the microbial source pool and the diagram's rows indicate alleles being contributed to the host source pool. The microbial source pool is dilute—hologenotypes are assembled via binomial sampling of microbial strains. Generation time of microbes is short relative to host—microbiome comes to community equilibrium within each host generation. The figure illustrates how the transmission of host nuclear genes and the microbiome operate in parallel. The release of microbes to the microbe pool and the release of gametes to the gamete pool both enable genetic recombination in their respective components of the hologenome. Similarly, the statistical process of microbe colonization is the counterpart of the mating system for determining the pattern of inheritance for their respective components of the hologenome.*

So, to obtain the state variables at time $t+1$ given the state at time t, collect the previous terms. The total number of empty juvenile hosts at $t+1$ is

$$\begin{aligned}
H(t+1) &= H'''_{11}(t,k_{1,11},0) + H'''_{11}(t,n_{1,11},n_{2,11}) + H'''_{11}(t,0,k_{2,11}) \\
&+ H'''_{12}(t,k_{1,12},0) + H'''_{12}(t,n_{1,12},n_{2,12}) + H'''_{12}(t,0,k_{2,12}) \\
&+ H'''_{22}(t,k_{1,22},0) + H'''_{22}(t,n_{1,22},n_{2,22}) + H'''_{22}(t,0,k_{2,22}) \\
&= [q_1(t)^2 \ \left(p_1(t)^2 \ W_{1,11} + 2p_1(t)p_2(t) \ W_{12,11} + p_2(t)^2 \ W_{2,11}\right) \\
&+ 2q_1(t)q_2(t) \ \left(p_1(t)^2 \ W_{1,12} + 2p_1(t)p_2(t) \ W_{12,12} + p_2(t)^2 \ W_{2,12}\right) \\
&+ q_2(t)^2 \ \left(p_1(t)^2 \ W_{1,22} + 2p_1(t)p_2(t) \ W_{12,22} + p_2(t)^2 \ W_{2,22}\right)] \ H(t) \\
&= \left(q_1(t)^2 \ \overline{W_{11}}(t) + 2q_1(t)q_2(t) \ \overline{W_{12}}(t) + q_2(t)^2 \ \overline{W_{22}}(t)\right) \ H(t) \\
&= \overline{\widetilde{\widetilde{W}}}(t) \, H(t)
\end{aligned}$$

where the fitness for holobionts with A_1A_1, A_1A_2, and A_2A_2 nuclear genotypes averaged over their microbiomes are:

$$\begin{aligned}
\overline{W_{11}}(t) &\equiv p_1(t)^2 \ W_{1,11} + 2p_1(t)p_2(t) \ W_{12,11} + p_2(t)^2 \ W_{2,11} \\
\overline{W_{12}}(t) &\equiv p_1(t)^2 \ W_{1,12} + 2p_1(t)p_2(t) \ W_{12,12} + p_2(t)^2 \ W_{2,12} \\
\overline{W_{22}}(t) &\equiv p_1(t)^2 \ W_{1,22} + 2p_1(t)p_2(t) \ W_{12,22} + p_2(t)^2 \ W_{2,22}
\end{aligned}$$

and the marginal fitnesses for each allele are

$$\begin{aligned}
\widetilde{\overline{W_1}}(t) &\equiv q_1(t) \, \overline{W_{11}}(t) + q_2(t)\overline{W_{12}}(t) \\
\widetilde{\overline{W_2}}(t) &\equiv q_1(t)\overline{W_{12}}(t) + q_2(t) \, \overline{W_{22}}(t)
\end{aligned}$$

and the overall mean fitness of the holobionts is

$$\overline{\widetilde{\widetilde{W}}}(t) \equiv q_1(t) \, \widetilde{\overline{W_1}}(t) + q_2(t) \, \widetilde{\overline{W_2}}(t)$$

The notation convention is that an average taken over microbe frequencies, ie, with respect to p_1 and p_2, is indicated with an overline. An average taken over allele frequencies, ie, with respect to q_1 and q_2, is indicated with a wide tilde. A wide tilde placed on top of an overline indicates an average first with respect to microbe frequencies and then with respect to allele frequencies. An overline placed over a wide tilde indicates an average first with respect to allele frequencies and then with respect to microbe frequencies. Also, to minimize the number of diacritical marks placed about any quantity, the lowest mark is understood to refer to an average over microbiome frequencies and genotype frequencies, ie, with respect to $p_1^2, 2\,p_1\,p_2, p_2^2$ and $q_1^2, 2\,q_1\,q_2, q_2^2$ rather than with respect to p_1, p_2, q_1, and q_2.

Based on these definitions, the frequency of the host nuclear alleles, A_1 and A_2, respectively, in the gamete pool at time $t+1$ as a function of both the gamete pool and microbial source pool at time t is

$$\begin{aligned} q_1(t+1) &= \frac{q_1(t)\overline{W_{11}}(t) + q_2(t)\overline{W_{12}}(t)}{\overline{\overline{W}}(t)}\, q_1(t) \\ q_2(t+1) &= \frac{q_2(t)\overline{W_{22}}(t) + q_1(t)\overline{W_{12}}(t)}{\overline{\overline{W}}(t)}\, q_2(t) \end{aligned}$$

Turning now to the microbiome, the number of strain-1 microbes in the microbial source pool at time $t+1$ is

$$\begin{aligned} G_1(t+1) &= p_1(t)^2 \left(q_1(t)^2\, k_{1,11}\, W_{1,11} + 2q_1(t)q_2(t)\, k_{1,12}\, W_{1,12} + q_2(t)^2\, k_{1,22}\, W_{1,22}\right) \\ &+ 2p_1(t)p_2(t) \left(q_1(t)^2\, n_{1,11}\, W_{12,11} + 2q_1(t)q_2(t)\, n_{1,12}\, W_{12,12} + q_2(t)^2\, n_{1,22}\, W_{12,22}\right) \\ &= p_1(t)^2\, \widetilde{w_{1,1}}(t) + 2p_1(t)p_2(t)\, \widetilde{w_{1,12}}(t) \\ &= \overline{\widetilde{w_1}}(t)\, p_1(t)\, H(t) \end{aligned}$$

where the multilevel microbe fitnesses per holobiont are

$$\begin{aligned} w_{1,1,11} &= k_{1,11}\, W_{1,11} \\ w_{1,1,12} &= k_{1,12}\, W_{1,12} \\ w_{1,1,22} &= k_{1,22}\, W_{1,22} \\ w_{1,12,11} &= n_{1,11}\, W_{12,11} \\ w_{1,12,12} &= n_{1,12}\, W_{12,12} \\ w_{1,12,22} &= n_{1,22}\, W_{12,22} \\ w_{2,12,11} &= n_{2,11}\, W_{12,11} \\ w_{2,12,12} &= n_{2,12}\, W_{12,12} \\ w_{2,12,22} &= n_{2,22}\, W_{12,22} \\ w_{2,2,11} &= k_{2,11}\, W_{2,11} \\ w_{2,2,12} &= k_{2,12}\, W_{2,12} \\ w_{2,2,22} &= k_{2,22}\, W_{2,22} \end{aligned}$$

and where the multilevel fitnesses of a strain-1 microbe per holobiont from single and dual strain microbiomes averaged over the three nuclear genotypes are

$$\begin{aligned} \widetilde{w_{1,1}}(t) &\equiv q_1(t)^2\, w_{1,1,11} + 2q_1(t)q_2(t)\, w_{1,1,12} + q_2(t)^2\, w_{1,1,22} \\ \widetilde{w_{1,12}}(t) &\equiv q_1(t)^2\, w_{1,12,11} + 2q_1(t)q_2(t)\, w_{1,12,12} + q_2(t)^2\, w_{1,12,22} \end{aligned}$$

The notation is that $w_{1,1,11}$ indicates the multilevel fitness per holobiont of a strain-1 microbe in a single strain microbiome within an A_1A_1 host whereas $w_{1,12,12}$ indicates the multilevel fitness per holobiont of a strain-1 microbe in a dual-strain microbiome within an A_1A_2 host. Thus the subscript convention is strain identity, followed by microbiome composition, and ending with host nuclear genotype. Then strain-1's marginal multilevel fitness per holobiont averaged over the nuclear genotypes is

$$\overline{\widetilde{w_1}}(t) \equiv p_1(t)\,\widetilde{w_{1,1}}(t) + p_2(t)\,2\,\widetilde{w_{1,12}}(t)$$

Similarly, the number of strain-2 microbes in the microbe pool at time $t+1$ is

$$\begin{aligned} G_2(t+1) &= 2p_1(t)p_2(t)\left(q_1(t)^2 n_{2,11}\,W_{12,11} + 2q_1(t)q_2(t)\,n_{2,12}\,W_{12,12} + q_2(t)^2 n_{2,22}\,W_{12,22}\right) \\ &+ p_2(t)^2\left(q_1(t)^2\,k_{2,11}\,W_{2,11} + 2q_1(t)q_2(t)\,k_{2,12}\,W_{2,12} + q_2(t)^2\,k_{2,22}\,W_{2,22}\right) \\ &= 2p_1(t)p_2(t)\,\widetilde{w_{2,12}}(t) + p_2(t)^2\,\widetilde{w_{2,2}}(t) \\ &= \overline{\widetilde{w_2}}(t)\,p_2(t)\,H(t) \end{aligned}$$

where the multilevel fitnesses of a strain-2 microbe per holobiont from single and dual strain microbiomes averaged over the three nuclear genotypes are

$$\begin{aligned} \widetilde{w_{2,2}}(t) &\equiv q_1(t)^2\,w_{2,2,11} + 2q_1(t)q_2(t)\,w_{2,2,12} + q_2(t)^2\,w_{2,2,22} \\ \widetilde{w_{2,12}}(t) &\equiv q_1(t)^2\,w_{2,12,11} + 2q_1(t)q_2(t)\,w_{2,12,12} + q_2(t)^2\,w_{2,12,22} \end{aligned}$$

Then strain-1's marginal multilevel fitness per holobiont averaged over the nuclear genotypes is

$$\overline{\widetilde{w_2}}(t) \equiv p_1(t)\,2\,\widetilde{w_{2,12}}(t) + p_2(t)\,\widetilde{w_{2,2}}(t)$$

The total number of microbes in the microbe pool at time $t+1$ then is

$$\begin{aligned} G(t+1) &= \left(\overline{\widetilde{w_1}}(t)\,p_1(t) + \overline{\widetilde{w_2}}(t)\,p_2(t)\right)\,H(t) \\ &= \overline{\overline{\widetilde{w}}}(t)\,H(t) \end{aligned}$$

where the microbiome multilevel mean fitness per holobiont is

$$\overline{\overline{\widetilde{w}}}(t) \equiv p_1(t)\,\overline{\widetilde{w_1}}(t) + p_2(t)\,\overline{\widetilde{w_2}}(t)$$

Therefore the microbe frequencies in the microbial pool at time $t+1$ are

$$p_1(t+1) = \frac{G_1(t+1)}{G(t+1)} = \frac{\overline{\widetilde{w_1}}(t)}{\overline{\overline{\widetilde{w}}}(t)}\,p_1(t)$$

and

$$p_2(t+1) = \frac{G_2(t+1)}{G(t+1)} = \frac{\overline{\widetilde{w_2}}(t)}{\overline{\overline{\widetilde{w}}}(t)}\,p_2(t)$$

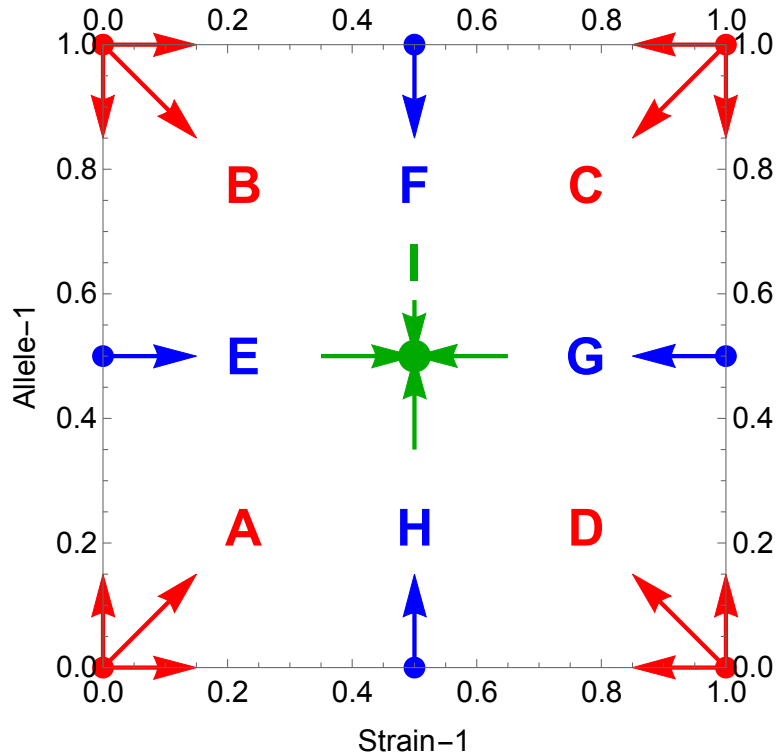


Figure 9: *Diagram of possible equilibria for holobiont evolution with two microbial strains with two host alleles. The horizontal axis is the frequency of strain-1 in the microbiome and the vertical axis is the frequency of the A_1 allele in the host nucleus. The corner equilibria are shown in red and labeled as A, B, C and D. The edge equilibria are shown in blue and labeled as E, F, G and H. The interior equilibrium is shown in green and labeled as I.*

To summarize, the equations above for $H(t+1)$, $G(t+1)$, $p_1(t+1)$, $p_2(t+1)$, $q_1(t+1)$ and $q_2(t+1)$, together constitute a dynamical system for holobiont evolution with two microbe strains and two diploid host alleles assuming (a) horizontal microbe transmission, (b) dilute microbe source pools and (c) macro time steps much longer than micro time steps. These equations correspond to those from classical population genetics for evolution at two loci with two alleles per locus. Those equations predict the trajectory in a gene pool of a two-compartment genome, one compartment per locus, whereas these equations predict the trajectory in a hologene pool of a two compartment hologenome, one compartment for the microbial genes and another compartment for the nuclear genes.

3.4.2 Conditions for Increase When Rare and Joint Microbiome-Host Polymorphism

Figure 9 shows the possible equilibria in the combined microbiome-nuclear gene system. The corner equilibria are shown in red and labelled A, B, C and D. The edge equilibria are shown in

blue and labelled E, F, G and H. The interior equilibrium is shown in green and labelled I. For the analysis, put $p_1 \to p$, $p_2 \to (1-p)$, $q_1 \to q$ and $q_2 \to (1-q)$. To understand the flow of trajectories in the p, q plane representing the combined evolution of microbial and nuclear genes, the stability of the corner equilibria are the places to investigate first.

So, start with the equation for $p(t+1)$.

$$\begin{aligned}
\overline{\widetilde{w_1}}(t) &\equiv p(t)\,\widetilde{w_{1,1}}(t) + (1-p(t))\,2\,\widetilde{w_{1,12}}(t) \\
\overline{\widetilde{w_2}}(t) &\equiv p(t)\,2\,\widetilde{w_{2,12}}(t) + (1-p(t))\,\widetilde{w_{2,2}}(t) \\
\overline{\overline{\widetilde{w}}}(t) &\equiv p(t)\,\overline{\widetilde{w_1}}(t) + (1-p(t))\,\overline{\widetilde{w_2}}(t) \\
p(t+1) &= \frac{\overline{\widetilde{w_1}}(t)}{\overline{\overline{\widetilde{w}}}(t)}\,p(t)
\end{aligned}$$

The conditions for strain-1 and for strain-2 to increase when rare in the absence of nuclear gene variation were shown earlier. These previously derived conditions can be transferred here with the proviso that the expressions for the multilevel microbial fitnesses now depend on $q(t)$ because these fitnesses are averaged over the nuclear genotype frequencies, as indicated with the wide tilde symbol. Putting $q(t) = q$, the condition for strain-1 to increase when rare is

$$p(t+1) > p(t)\big|_{p\approx 0,q}$$

which works out to be

$$E_1(q) \equiv 2\,\widetilde{w_{1,12}} - \widetilde{w_{2,2}} > 0$$

Similarly, the condition for strain-2 to increase when rare is

$$E_2(q) \equiv 2\,\widetilde{w_{2,12}} - \widetilde{w_{1,1}} > 0$$

$E_1(q)$ is the excess by which a rare strain-1 microbe out-produces a common strain-1 microbe for a given q. $E_2(q)$ is the excess by which a rare strain-2 microbe out-produces a common strain-1 microbe for a given q.

Next, turn to the equation for $q(t+1)$.

$$\begin{aligned}
\widetilde{\overline{W_1}}(t) &\equiv q(t)\,\overline{W_{11}}(t) + (1-q(t))\overline{W_{12}}(t) \\
\widetilde{\overline{W_2}}(t) &\equiv (1-q(t))\,\overline{W_{22}}(t) + q(t)\overline{W_{12}}(t) \\
\overline{\overline{\widetilde{W}}}(t) &\equiv q(t)\widetilde{\overline{W_1}}(t) + (1-q(t))\widetilde{\overline{W_2}}(t) \\
q(t+1) &= \frac{\widetilde{\overline{W_1}}(t))}{\overline{\overline{\widetilde{W}}}(t)}\,q(t)
\end{aligned}$$

The conditions for A_1 and A_2 to increase when rare in the absence of microbiome variation were shown earlier. These previously derived conditions can be transferred here with the proviso

that the expressions for the holobiont fitnesses now depend on $p(t)$ because these fitnesses are averaged over the microbiome frequencies, as indicated with the overline symbol. Putting $p(t) = p$, the condition for A_1 to increase when rare is

$$q(t+1) > q(t)\big|_{q\approx 0,p}$$

which works out to be

$$F_1(p) \equiv \overline{W_{12}} - \overline{W_{22}} > 0$$

And similarly, A_2 can increase when rare if

$$F_2(p) \equiv \overline{W_{12}} - \overline{W_{11}} > 0$$

Applying these increase-when-rare conditions to the corner equilibria yields:

Corner A, *p*=0, *q*=0

$$\begin{aligned} E_1(0) &= 2\,\widetilde{w_{1,12}} - \widetilde{w_{2,2}} = 2\,w_{1,12,22} - w_{2,2,22} > 0 \\ F_1(0) &= \overline{W_{12}} - \overline{W_{22}} = W_{2,12} - W_{2,22} > 0 \end{aligned}$$

At this corner, all the alleles are A_2 and all the microbes are strain-2. So, the first condition means that strain-1 microbes increase if twice their multilevel fitness in a dual-strain microbiome with an A_2A_2 host exceeds the fitness of strain-2 microbes in a single strain-2 microbiome with a A_2A_2 host. The second condition means that A_1 alleles increase if their fitness in an A_1A_2 host with a single strain-2 microbiome exceeds the fitness of A_2 alleles in a A_2A_2 host with a single strain-2 microbiome. A similar interpretation applies to the other corners.

Corner B, *p*=0, *q*=1

$$\begin{aligned} E_1(1) &\equiv 2\,\widetilde{w_{1,12}} - \widetilde{w_{2,2}} = 2\,w_{1,12,11} - w_{2,2,11} > 0 \\ F_2(0) &\equiv \overline{W_{12}} - \overline{W_{11}} = W_{2,12} - W_{2,11} > 0 \end{aligned}$$

Corner C, *p*=1, *q*=1

$$\begin{aligned} E_2(1) &\equiv 2\,\widetilde{w_{2,12}} - \widetilde{w_{1,1}} = 2\,w_{2,12,11} - w_{1,1,11} > 0 \\ F_2(1) &\equiv \overline{W_{12}} - \overline{W_{11}} = W_{1,12} - W_{1,11} > 0 \end{aligned}$$

Corner D, p=1, q=0

$$\begin{aligned} E_2(0) &\equiv 2\,\widetilde{w_{2,12}} - \widetilde{w_{1,1}} = 2\,w_{2,12,22} - w_{1,1,22} > 0 \\ F_1(1) &\equiv \overline{W_{12}} - \overline{W_{22}} = W_{1,12} - W_{1,22} > 0 \end{aligned}$$

If the four corner equilibria are unstable with trajectories leading out from them into the interior of the (p, q) plane then the interior equilibrium, I, exists. This equilibrium represents a joint polymorphism for both the microbiome and host nuclear components of the hologenome. The joint equilibrium $(\hat{p}, \hat{q})$ is found from numerically solving simultaneously for p and q in

$$\begin{aligned} p &= \frac{E_1(q)}{E_1(q) + E_2(q)} \\ q &= \frac{F_1(p)}{F_1(p) + F_2(p)} \end{aligned}$$

3.4.3 Illustration—Joint Microbiome-Host Genetic Diversity

The full model with both two microbe strains and two host alleles may be used to investigate the maintenance of genetic diversity jointly in the microbiome and host nucleus. As a illustration, return to the previous example involving a polymorphism between two microbial strains, one altruistic and one selfish. Now, suppose strain-1 is altruistic only to one genotype of host because a host with that genotype somehow facilitates microbes of strain-1. Furthermore, suppose arbitrarily that the host genotype to which strain-1 is altruistic is the A_1A_2 heterozygote genotype. This choice is made to develop a numerical example in which both the microbiome and host components of the hologenome are likely to be polymorphic. (If say, strain-1 were altruistic to a homozygote host genotype like A_1A_1 rather than to the heterozygote A_1A_2, then the A_1 allele would fix in the host component of the hologenome, thus presenting a less interesting dynamical outcome.)

Accordingly, Table 1 presents the setup. The first column of the table details all the parameter assignments. The second column details intermediate coefficients derived from these parameters. The third column details the equilibrium outcome of holobiont selection with the equilibrium frequency of strain-1 followed by the equilibrium frequency of A_1 in the top two rows. Below these are the hologenotype frequencies, $\hat{X}_{ij,kl}$ for the hologenotype with strain-i and strain-j in the microbiome and allele A_k and A_l in the host nucleus. Figure 10 shows trajectories in the (p, q) plane that approach this equilibrium from an assortment of initial conditions.

At the joint equilibrium, the altruistic strain-1 is 33% of the microbe pool and the A_1 host allele is 61% of the host gene pool. The most common hologenotypes at equilibrium are 51% two-strain microbiomes in an A_1A_1 homozygous host, 22% single strain-2 microbiomes in an A_1A_2 heterozygous host, and 20% two-strain microbiomes in an A_2A_2 homozygous host. The

Table 1: Numerical Example: Two Microbe Strains and Two Host Alleles

$W_{1,11} \to 2$	$n_{1,11} = 200$	$\hat{p} = 0.325939$
$W_{1,12} \to 3$	$n_{1,12} = 133.333$	$\hat{q} = 0.614992$
$W_{1,22} \to 1$	$n_{1,22} = 266.667$	
$W_{12,11} \to 1.5$	$n_{2,11} = 300$	$\hat{X}_{1,11} = 0.0401801$
$W_{12,12} \to 2.5$	$n_{2,12} = 333.333$	$\hat{X}_{12,11} = 0.509881$
$W_{12,22} \to 1$	$n_{2,22} = 266.667$	$\hat{X}_{2,11} = 0.171845$
$W_{2,11} \to 1$		$\hat{X}_{1,12} = 0.0503084$
$W_{2,12} \to 1$	$w_{1,1,11} = 700$	$\hat{X}_{12,12} = 0.638408$
$W_{2,22} \to 1$	$w_{1,1,12} = 900$	$\hat{X}_{2,12} = 0.215163$
	$w_{1,1,22} = 400$	$\hat{X}_{1,22} = 0.0157475$
$k_{1,11} \to 350$	$w_{1,12,11} = 300$	$\hat{X}_{12,22} = 0.199834$
$k_{1,12} \to 300$	$w_{1,12,12} = 333.333$	$\hat{X}_{2,22} = 0.0673501$
$k_{1,22} \to 400$	$w_{1,12,22} = 266.667$	
$k_{2,11} \to 400$	$w_{2,12,11} = 450$	
$k_{2,12} \to 400$	$w_{2,12,12} = 833.333$	
$k_{2,22} \to 400$	$w_{2,12,22} = 266.667$	
	$w_{2,2,11} = 400$	
$a_{12,11} \to 0.5$	$w_{2,2,12} = 400$	
$a_{12,12} \to 0.5$	$w_{2,2,22} = 400$	
$a_{12,22} \to 0.5$		
$a_{21,11} \to 0.5$		
$a_{21,12} \to 0.5$		
$a_{21,22} \to 0.5$		

microbiome in the two-strain A_1A_1 host consists of 200 strain-1 microbes and 300 strain-2 microbes. The microbiome in the single strain-2 A_1A_2 host consists of 400 strain-2 microbes. The microbiome in a two-strain A_2A_2 host consists of 267 strain-1 microbes and 267 strain-2 microbes. And similarly for the other less common hologenotypes. At equilibrium the hologenome is polymorphic for the altruistic and selfish microbial strains and for the nuclear alleles that differ in the reciprocation they supply to the altruistic microbial strain.

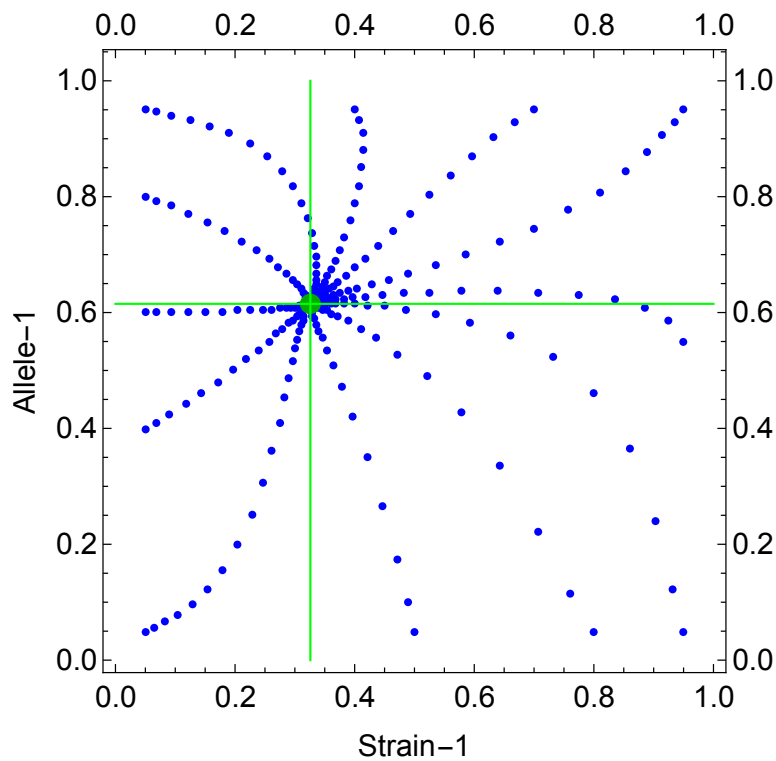


Figure 10: *Trajectories of the combined dynamics of microbe frequency and allele frequency under holobiont selection leading to a combined polymorphism representing coadaptation between the microbiome and host. Trajectories all begin at various spots near the axes and converge through time to the green dot in the interior. The equilibrium frequencies are marked with the vertical and horizontal lines in green.*

3.4.4 Direction of Joint Microbiome-Host Evolution

To analyze the direction of holobiont evolution, write the dynamics as difference equations for $\Delta p_1(t) \equiv p_1(t+1) - p_1(t)$ and for $\Delta q \equiv q_1(t+1) - q_1(t)$ yielding, after rearrangement,

$$\begin{aligned} \Delta p &= \frac{p\,(1-p)\,(\overline{\widetilde{w_1}} - \overline{\widetilde{w_2}})}{\overline{\overline{\widetilde{w}}}} \\ \Delta q &= \frac{q(1-q)(\widetilde{\overline{W}}_1 - \widetilde{\overline{W}}_2)}{\widetilde{\overline{\overline{W}}}} \end{aligned}$$

where the subscript and time index are dropped leaving $p \equiv p_1(t)$ and $q \equiv q_1(t)$. At equilibrium, the marginal multilevel fitness per holobiont of both microbe strains are equal and the marginal fitness of both host alleles are equal.

In the spirit of the previous expressions for quantities maximized by holobiont selection on a two-strain microbiome, $L_G(p)$ and on a host with two alleles $L_H(q)$, consider their product as a candidate that is maximized by holobiont selection acting simultaneously on both the micro-

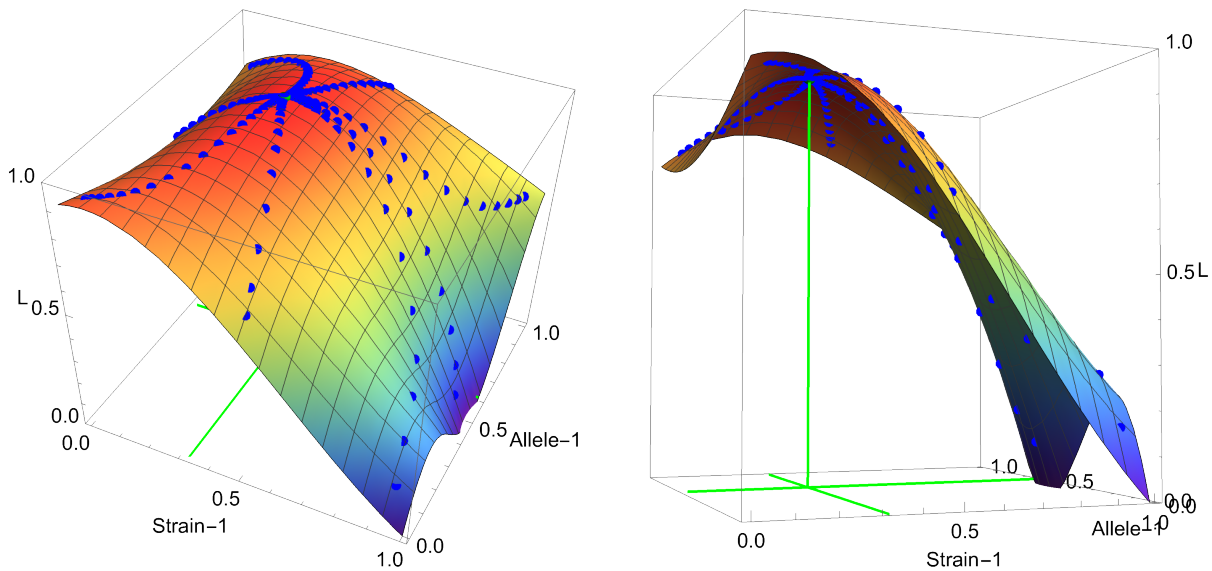


Figure 11: *The $L(p,q)$ surface representing the simultaneous degree of equality between the marginal multilevel fitness per holobiont of the microbe strains along with the degree of equality between the marginal fitness of the host alleles. The equilibrium, $(\hat{p},\hat{q})$, lies at the peak of the surface and is shown as a green dot. Trajectories of the combined dynamics of microbe frequency and allele frequency under holobiont selection are shown with blue dots from an assortment of initial conditions that start near the axes. The trajectories climb the surface and converge at the equilibrium at the top. Left, view from above the surface. Right: view from below the surface.*

biome and host nucleus:

$$\left(-(\overline{\widetilde{w_1}}-\overline{\widetilde{w_2}})^2\right)\left(-(\widetilde{\overline{W}}_1-\widetilde{\overline{W}}_2)^2\right)$$

This expression represents the simultaneous degree of equality between the marginal multilevel fitness per holobiont of both microbe strains along with the degree of equality between the marginal fitness of both host alleles. This expression may be scaled to lie mostly between 0 and 1 as follows:

$$\begin{aligned} L(p,q) &= \left(\frac{-(\overline{\widetilde{w_1}}-\overline{\widetilde{w_2}})^2}{\max[E_1(0)^2,E_1(1)^2,E_1(\hat{q})^2,E_2(0)^2,E_2(1)^2,E_2(\hat{q})^2]}+1\right) \\ &\times \left(\frac{-(\widetilde{\overline{W}}_1-\widetilde{\overline{W}}_2)^2}{\max[F_1(0)^2,F_1(1)^2,F_1(\hat{q})^2,F_2(0)^2,F_2(1)^2,F_2(\hat{q})^2]}+1\right) \end{aligned}$$

Figure 11 illustrates the $L(p,q)$ surface. It shows the equilibrium at the peak of this surface. The figure also shows trajectories for the simultaneous evolution of the microbiome and host nuclear components of the hologenome converging at the peak of the surface.

Holobiont selection does not lead to a peak of overall adaptation for the holobiont. Figure 12 shows the surface for the overall holobiont mean fitness, $\widetilde{\overline{W}}(p,q)$. The equilibrium is indicated as

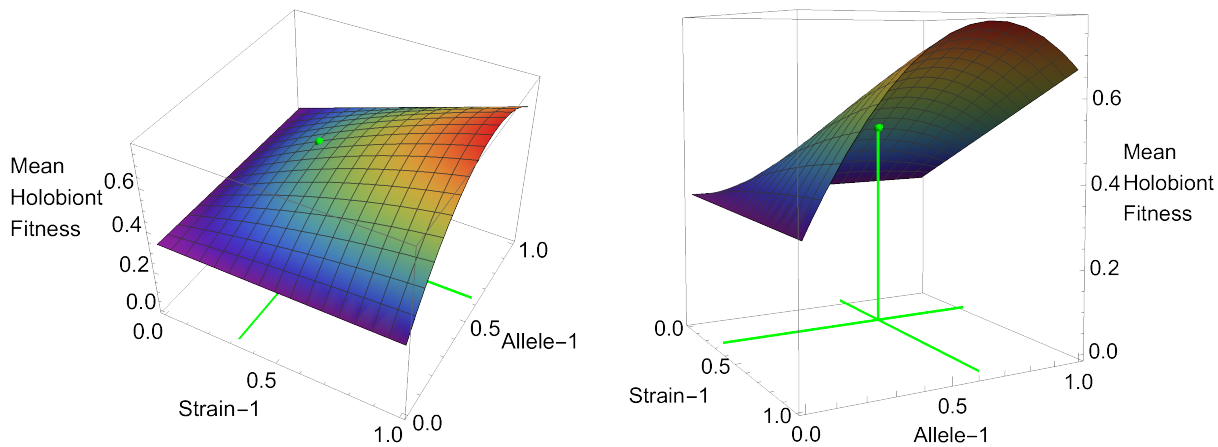


Figure 12: *The $\overline{\overline{W}}(p,q)$ surface representing overall holobiont mean fitness. The equilibrium, $(\hat{p},\hat{q})$, is shown as a green dot. Left, view from above the surface. Right: view from below the surface.*

a green dot on the surface. This dot does not lie on any peak of the surface. From the standpoint of maximal holobiont fitness, the best outcome would be for the microbiome to consist entirely of the altruistic strain-1 and for the nuclear alleles to yield the heterozygotes that benefit from strain-1 altruism. However, this outcome does not result from holobiont selection.

The difference between holobiont selection maximizing $L(p,q)$ and not $\overline{\overline{W}}(p,q)$ points out that from the standpoint of the host, a gene obtained via a microbial strain is "not as good as" the same gene present in its own nucleus. Classical natural selection with multiple genes at one locus leads to an outcome that maximizes the mean fitness. But if those same genes are embedded in microbes, then the evolutionary dynamics do not generally result in an equally adaptive outcome.

The hologenome is best compared with a multiple locus genetic system rather than a one locus multiple allele system. Classical natural selection with multiple loci generally does not result in maximizing mean fitness unless the linkage between the loci is tight by being close to one another on the same chromosome. The hologenome and a two-locus genetic system both contain two components. Just as selection operating on two locus genetic system does not generally maximize mean fitness, selection operating on a combined microbiome-host genetic system should not be expected to maximize mean fitness either.

On the other hand, although a microbial gene might not be "as good as" a nuclear gene for the host with respect to realizing an optimal adaptive outcome from selection, a microbial gene might be beneficial to the host for other compelling reasons. From a host's perspective, a microbial gene is a "contract gene", so to speak. A host presumably can jettison the gene if not useful, and can assimilate a replacement if needed. The facultative employment of zooxanthellae

by corals under differing light and temperature conditions comes to mind. In contrast, recruiting a gene to join the nucleus represents a serious commitment by the host, like the granting of tenure passed on to the host's offspring over many generations and perhaps in perpetuity. The topic of hologenome architecture is a conceptual counterpart to that in classical population genetics concerning how natural selection sculpts the genetic system by modifying linkage relationships with chromosomes and modifying mutation rates with DNA repair mechanisms.

3.5 Strategies of Coadaptation

3.5.1 Definition of a Holostable Strategy (HSS)

Consider the hologenotype equilibrium point at which both the frequency of allele A_1 equals 1 and the frequency of strain-1 equals 1, corresponding to the corner equilibrium C in Figure 9. Next, consider a set of possible alternative alleles and a set of possible alternative microbial strains, a so-called mutant hologenotype. Then ask under what conditions this hologenotype equilibrium, C in Figure 9, is stable to invasion by any alternative nuclear allele and/or alternative microbial strain taken from the sets of possible alleles and strains—the opposite of the condition for increase when rare. That is, the (A_1, strain-1) hologeotype is stable if the mutant (allele, strain) hologenotype cannot increase when rare.

If allele A_1 is fixed in the host nucleus and strain-1 is fixed in the microbiome, then alternative alleles and strains *cannot* increase when rare if

$$\begin{aligned} W_{1,11} &> W_{1,12} \\ w_{1,1,11} &> 2\,w_{2,12,11} \end{aligned}$$

In words, a hologenotype consisting of A_1A_1 paired with the strain-1, will be called "holostable" if *both* the fitness of an A_1A_1 homozygote with a single strain-1 microbiome exceeds the fitness of an A_1A_2 heterozygote with a single strain-1 microbiome *and* the multilevel fitness per holobiont of a strain-1 microbe in a single-strain microbiome with a A_1A_1 host exceeds twice the multilevel fitness per holobiont of a strain-2 microbe in a two-strain microbiome with a A_1A_1 host. A holophenotype that leads to such a hologenotype will be called a "holostable strategy" (HSS). The HSS is stable to invasion by alternative phenotypes both ecologically and evolutionarily. An HSS might be expected to result from holobiont selection.

3.5.2 Criteria for an HSS

Consider a trait controlled by the host, x, and another trait controlled by the microbe, y, from a set of possible traits. The host fitness and the multilevel microbe fitness depend on both these

traits. Write the fitnesses as functions of these traits and unpack the multilevel fitness into its components as

$$W(x,y)$$
$$k(x,y)\,W(x,y)$$

A focal strategy, $\hat{x},\hat{y}$, for a host and its microbe is an HSS if it is unbeatable by any mutant strategy, x_m, y_m, for a mutant host and microbe. According to the conditions above for a non-invasible hologenotype, the conditions for the focal strategy, $\hat{x},\hat{y}$, to be an HSS are

$$\begin{aligned} W(\hat{x},\hat{y}) &\geq W(x_m,\hat{y}) \\ k(\hat{x},\hat{y})\,W(\hat{x},\hat{y}) &\geq 2\,n(\hat{x},\hat{y},y_m)\,W(\hat{x},y_m) \end{aligned}$$

for any x_m, y_m from the set of possible traits, where $n(\hat{x},\hat{y},y_m)$ is the mutant's population size in a two-strain microbiome given that the host's trait is $\hat{x}$ and the other microbe's trait is $\hat{y}$. The $n(\hat{x},\hat{y},y_m)$ stands for the mutant strain's equilibrium population size from the Lotka-Volterra competition equations or from any other applicable model of population dynamics within a two-strain microbiome.

The first condition for an HSS is satisfied if the focal x as a function of the focal y, denoted $\hat{x}(y)$, is obtained by maximizing W with respect to x for a given y,

$$\hat{x}(y) = \max_x W(x,y)$$

because any x_m other than $x_m = \hat{x}(y)$ for a given y, leads to a W that is less than $W(\hat{x}(y),y)$. Thus, $\hat{x}(y)$ obtained in this way is the host component of the HSS for the given y.

The second condition for an HSS is not so straight forward. The condition indicates that the focal microbe's multilevel fitness must exceed twice the mutant microbe's multilevel fitness. So the HSS focal y must be such that for any mutant y_m the focal microbe multilevel fitness equals twice the mutant multilevel fitness if $y_m = y$ and the focal microbe multilevel fitness exceeds twice the mutant multilevel fitness if $y_m \neq y$. Therefore, the condition for the microbe's HSS, $\hat{y}$, is obtained by maximizing the following expression with respect to y for any y_m

$$\hat{y} = \max_y \left(k(\hat{x}(y),y)\,W(\hat{x}(y),y) - 2\,n(\hat{x}(y),y,y_m)\right)W(\hat{x}(y),y,y_m)) \;\; \forall\; y_m \in [0,1]$$

Then, $\hat{y}$ can be tested by various mutant y_m to ensure that the focal y cannot be invaded by any y_m.

The focal $\hat{x}$, which is the host's component of the HSS, is then obtained by evaluating $\hat{x}(y)$ at $y = \hat{y}$. In this way, both the host and microbe components of the HSS are calculated. Whether $\hat{x},\hat{y}$ actually exist as a solution to these two HSS conditions depends on the tradeoffs among the host and microbe phenotypes as specified in a phenotype submodel.

3.5.3 Example Phenotype Submodel

To illustrate how to develop a phenotypic submodel, consider an example concocted for an idealized coral-zooxanthellae holobiont. Let x be the "altruism effort" by the microbe toward the host and y the "altruism effort" by the host toward the microbe. The altruism effort by the microbe refers to the cost of the transfer of photosynthate from the zooxanthellae to the coral, and the altruism effort by the host refers to the cost of the infrastructure the coral provides for the zooxanthellae. Let the product, $x\,y$, be the "altruism synergy" between the microbe and host. This quantity indicates how much both parties are working together to provide assistance for one another.

The altruism synergy is taken to be the fraction of the host's base fitness, W_o, that it forgoes to provide infrastructure to the microbe and also the fraction of the microbe's base carrying capacity, k_o, that it foregoes to provide energy to the host. Accordingly,

$$\begin{aligned} W(x,y) &= (1 - x\,y)\,W_o + \ldots \\ k(x,y) &= (1 - x\,y)\,k_o + \ldots \end{aligned}$$

The costs of the altruism from the host to the microbe and from microbe to host are only part of the picture. The other part is the benefit each receives from their altruistic synergy.

So, credit the host with the benefit that the microbe is supplying, $c_1\,x^{c_2}\,y^{c_3}\,k_o$. The coefficient, c_1, converts units of microbe carrying capacity to those of host fitness. Both x and y are raised to powers, $c_2 < 1$ and $c_3 < 1$, to describe diminishing returns to scale for the effort. Furthermore, $c_2 < c_3$ to indicate that the effect of the host's helping the microbe must do a "round trip" to be experienced by itself, whereas the effect of the microbe's helping is only discounted once because it results from a "one-way trip".

$$W(x,y) = (1 - x\,y)W_o + c_1\,x^{c_2}\,y^{c_3}\,k_o$$

Similarly, credit the microbe with the benefit that the host is supplying to it, $c_4\,x^{c_5}\,y^{c_6}\,W_o$. The coefficient, c_4, converts the units of host fitness to those of the microbe carrying capacity and $c_5 < 1$ and $c_6 < 1$ indicate the decreasing returns to scale for the efforts. Moreover, $c_6 < c_5$ to account for the different discounting between round trip and one-way contributions of altruistic effort. Hence the equation for the strain-1 carrying capacity in a A_1A_1 host becomes

$$k(x,y) = (1 - x\,y)k_o + c_4\,x^{c_5}\,y^{c_6}\,W_o$$

These two equations constitute the coral-zooxanthellae submodel to be analyzed. The task now is to determine the values of altruistic efforts, $\hat{x}$ and $\hat{y}$, that comprise the HSS.

Table 2: Numerical Example: Parameters for Host-Microbiome Holostable Strategy (HSS)

$W_o \to 1$	$c_1 \to 1/1000$	$c_4 \to 300$
$k_o \to 1000$	$c_2 \to 3/4$	$c_5 \to 1/4$
$a \to 1$	$c_3 \to 1/4$	$c_6 \to 3/4$

3.5.4 Example Computation of an HSS

Table 2 presents parameter values for a numerical example.

For the host HSS condition, the focal host fitness is

$$W(x,y) = (1 - x\,y) + x^{3/4}\,y^{1/4}$$

which is a $\cap$-shaped function of x for a some values of y. Setting the partial derivative of $W(x,y)$ with respect to x equal to zero and solving for x yields

$$\hat{x}(y) = \frac{81}{256\,y^3} \quad \text{for } \; 0.68142 \leq y \leq 1$$

The $\hat{x}(y)$ is a monotonically decreasing function of y over the interval, $[0.68142, 1]$—this is the HSS value of the host's trait, x, for a given value of y.

Turning now to the microbe HSS condition, the first term is the focal microbe multilevel fitness is $k(\hat{x}(y), y), W(\hat{x}(y), y)$. The microbe carrying capacity is

$$k(x,y) = 1000(1 - x\,y) + 300\,x^{1/4}\,y^{3/4}$$

Substituting $\hat{x}(y)$ for x in $k(x,y)$ yields

$$k(\hat{x}(y), y) = 1225 - \frac{10125}{32\,y^2} \quad \text{for } \; 0.68142 \leq y \leq 1$$

which is a monotonic increasing function of y over $[0.68142, 1]$. Similarly, substituting $\hat{x}(y)$ for x in $W(x,y)$ yields

$$W(\hat{x}(y), y) = 1 + \frac{27}{256\,y^2} \quad \text{for } \; 0.68142 \leq y \leq 1$$

which is a monotonic decreasing function of y over $[0.68142, 1]$. The product of these, which is the focal microbe multilevel fitness, is then

$$k(\hat{x}(y), y)\,W(\hat{x}(y), y) = \left(1225 - \frac{10125}{32\,y^2}\right)\left(1 + \frac{27}{256\,y^2}\right) \quad \text{for } \; 0.68142 \leq y \leq 1$$

which is a monotonic increasing function of y over $[0.68142, 1]$. In particular, $y = 1$ is the microbe trait that maximizes the focal microbe multilevel fitness. Furthermore, as noted above, $k(\hat{x}(y), y)$ is also maximized at $y = 1$ and any $y \neq 1$ leads to a lower k.

Table 3: Numerical Example: Results for Host-Microbiome Holostable Strategy (HSS)

$\hat{x} = 0.316406$	$W_o = 1$	$W(\hat{x}, \hat{y}) = 1.10547$
$\hat{y} = 1$	$k_o = 1000$	$k(\hat{x}, \hat{y}) = 908.594$
$\hat{x}\,\hat{y} = 0.316406$	$w_o = 1000$	$w((\hat{x}, \hat{y}) = 1004.42$

The second term in the microbe HSS condition is the mutant microbe multilevel fitness which is the product of the mutant population size and the host fitness with the mutant microbe. The mutant individual is assumed to differ from a focal individual *only* in x and y, and not in other ecological dimensions. Hence the competition coefficient, a, between a focal individual and a mutant is one, $a = 1$. According to the Lotka-Volterra competition equations, if $a = 1$, the focal microbe and the mutant microbe cannot coexist ecologically in the microbiome—which of these has the higher k excludes the other with the smaller k. Now, if the mutant $y_m \neq 1$, then the mutant k is less than the focal k, and therefore the mutant is ecologically excluded, resulting in $n = 0$. Therefore, the microbe HSS condition reduces to maximizing solely the first term, which is the focal microbe multilevel fitness, because the second term is zero. And as has been seen, the first term is maximized at $y = 1$, which is the HSS value of y. And therefore, the HSS value of x is $\hat{x}(1)$ which equals $81/256 = 0.316$.

In summary, the HSS is the pair $\hat{x}$ and $\hat{y}$, that jointly maximize the host fitness, $W(x, y)$, and the microbe multilevel fitness, $w(x, y) = k(x, y)\, W(x, y)$.

In the numerical case here, Table 3 records the HSS and its implications. At the holostable strategy, the hosts contribute an altruistic effort to 0.316, the microbe an altruistic effort of 1, and the altruistic synergy is 0.316. Compared to a non-altruistic base line, the reciprocal altruism raises the host fitness from 1 to 1.11, the microbe carrying capacity drops from 1000 to 908.1, and the microbe multilevel fitness rises from 1000 to 1004.4. Thus, the coadaptation between the components of the holobiome raises the fitnesses of both the host and microbe components above their base levels with no altruistic efforts. A mutant with any other values of host and microbe altruistic efforts cannot displace a host and microbe with these HSS values of altruistic effort.

In this example, the HSS results from an appealing condition: the joint maximization of the host fitness and the microbe multilevel fitness. It is presently not clear if this criterion for an HSS extends to any phenotype model, or is restricted to the model that has been analyzed. The possibility of a general theorem for an HSS maximization criterion awaits further work.

Phenotype models might be devised in the future for other strategic situations. For example, the reverse of this analysis would be to investigate reciprocal exploitation in which x might represent the effort by a pathogen to exploit the host and y the effort by the host to retard the pathogen or even to extract from it some useful byproducts. Indeed, a systematic study of

strategies with $++$, $+-$, $-+$ and $--$ interactions between microbes and host, all in a holobiont setting wherein the microbe and host's interests are partially aligned, also awaits further work.

4 Discussion

Some new theoretical concepts to highlight: (1) the symmetry between the random union of gametes for nuclear genes resulting from the mating system and the random assortment of microbes resulting from the host-colonization process as illustrated in Figure 8, (2) the definition of a multilevel fitness for microbes that combines both within-host and between-host success, and (3) the holostable strategy (HSS) for an evolutionarily and ecologically stable holobiont.

Holobionts do not destabilize the concept of an individual in biology. The status of a holobiont as an individual derives from the status of the host. What is problematic about a coral as an individual is that its polyp is part of a colony of polyps, not that it harbors zooxanthellae. What holobionts do destabilize is the concept of inheritance as a lineal rather than a collective process.

5 Acknowledgments

I thank Erol Akçay, Lynn Chiu, Scott Gilbert, Javier Suárez, and Jeremy Van Cleve for helpful comments on the manuscript.

References Cited

Babcock, R. C., G. D. Bull, P. L. Harrison, A. J. Heyward, J. K. Oliver, C. C. Wallace and B. L. Willis. 1986. "Synchronous Spawnings of 105 Scleractinian Coral Species on the Great Barrier Reef." *Marine Biology* 90:379–394.

Bordenstein, S. R. and K. R. Theis. 2015. "Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes." *PLoS Biol.* 13(8):e1002226.

Brown, J. S. and T. L. Vincent. 1987. "Coevolution as an Evolutionary Game." *Evolution* 41:66–79.

Carmona, D., C. R. Fitzpatrick, and M. T. Johnson. 2015. "Fifty years of Co-evolution and Beyond: Integrating Co-evolution from Molecules to Species." *Mol. Ecol.* 24:5315–5329.

Damuth, J. and L. Heisler. 1988, Alternative Formulations of Multilevel Selection." *Biology and Philosophy* 3:407–430.

Dieckmann, U., and R. Law. 1996. "The Dynamical Theory of Coevolution: a Derivation from Stochastic Ecological Processes." *J. Math. Biol.* 34:579–612.

Fenner, F. 1953. "Changes in the Mortality-Rate due to Myxomatosis in the Australian Wild Rabbit." *Nature* 172, 228–230 doi:10.1038/172228a0

Foster, K. R., J. Schluter, K. Coyte, and S. Rakoff-Nahoum. 2017. "The Evolution of the Host Microbiome as an Ecosystem on a Leash". *Nature* 548:43–51.

Gilbert, S. F., J. Sapp, and A. I. Tauber. 2012. "A Symbiotic View of Life: We Have Never Been Individuals." *Q. Rev. Biol.* 87:325–341.

Hirose, M. and M. Hidaka. 2006. "Early Development of Zooxanthella-containing Eggs of the Corals, *Porites cylindirica* and *Montipora digitata*: The Endodermal Localization of Zooxanthellae." *Zoological Science* **23**:873–881.

Jennings, E. C. 2019. "A Holobiont Characterization of Reproduction in a Live-bearing Cockroach, *Diploptera punctata*." Phd. diss. Retrieved from https://etd.ohiolink.edu/.

Margulis, L. 1991. "Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis." In *Symbiogenesis and Symbionticism,* edited by L. Margulis and R. Fester, 1–14, Cambridge: MIT Press.

Maynard Smith, J. 1964. "Group Selection and Kin Selection." *Nature* 201:1145–1147.

Mayo, D. and N. Gilinsky. 1987. "Models of Group Selection." *Philosophy of Science* 54:515–538.

McFall-Ngai, M., M. G. Hadfield, T. C. Bosch, H. V. Carey, T. Domazet-Loo, A. E. Douglas, N. Dubilier, et al. 2013. "Animals in a Bacterial World, A New Imperative for the Life Sciences." *Proc. Natl. Acad. Sci. (USA)* 110:3229–3236.

Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford UK: Oxford University Press.

Roughgarden, J. 2020. "Holobiont Evolution: Mathematical Model with Vertical *vs.* Horizontal Microbiome Transmission." *Philosophy, Theory and Practice in Biology*, 12:2. http://dx.doi.org/10.3998/ptpbio.16039257.0012.002.

Trench, R. K. 1993. "Microalgal-Invertebrate Symbioses: A Review." *Endocytobiosis Cell Res.* 9:135–175.

Zilber-Rosenberg, I. and E. Rosenberg. 2008. "Role of Microorganisms in the Evolution of Animals and Plants: the Hologenome Theory of Evolution." *FEMS Microbiology Reviews* 32:723–735. https://doi.org/10.1111/j.1574-6976.2008.00123.x