

ESSAY

Modeling endosymbioses: Insights and hypotheses from theoretical approaches

Lucas Santana Souza^{1,2} , Josephine Solowiej-Wedderburn^{1,2} , Adriano Bonforti^{2,3,4}, Eric Libby^{1,2} *

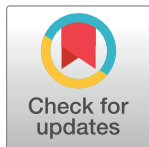
1 Department of Mathematics and Mathematical Statistics, Umeå University, Umeå, Sweden, **2** Integrated Science Lab, Umeå University, Umeå, Sweden, **3** Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden, **4** Umeå Marine Sciences Centre, Umeå University, Norrbyn, Sweden

 These authors contributed equally to this work.

* eric.libby@umu.se

Abstract

Endosymbiotic relationships are pervasive across diverse taxa of life, offering key avenues for eco-evolutionary dynamics. Although a variety of experimental and empirical frameworks have shed light on critical aspects of endosymbiosis, theoretical frameworks (mathematical models) are especially well-suited for certain tasks. Mathematical models can integrate multiple factors to determine the net outcome of endosymbiotic relationships, identify broad patterns that connect endosymbioses with other systems, simplify biological complexity, generate hypotheses for underlying mechanisms, evaluate different hypotheses, identify constraints that limit certain biological interactions, and open new lines of inquiry. This Essay highlights the utility of mathematical models in endosymbiosis research, particularly in generating relevant hypotheses. Despite their limitations, mathematical models can be used to address known unknowns and discover unknown unknowns.



OPEN ACCESS

Citation: Souza LS, Solowiej-Wedderburn J, Bonforti A, Libby E (2024) Modeling endosymbioses: Insights and hypotheses from theoretical approaches. *PLoS Biol* 22(4): e3002583. <https://doi.org/10.1371/journal.pbio.3002583>

Academic Editor: Thomas A. Richards, University of Oxford, UNITED KINGDOM

Published: April 10, 2024

Copyright: © 2024 Souza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was funded by the Swedish Research council (2018-03630 to EL) as well as the Kempe foundation grants (SMK21-0004 to EL for LSS), (JCK22-0026.1 for AB), and (JCK-2129.2 to EL for JSW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Endosymbioses are important drivers of eco-evolutionary dynamics that have the potential to forge entirely new kinds of individuals [1–4]. For example, the emergence of eukaryotes is intertwined with an endosymbiosis that would eventually evolve into mitochondria. Since gaining mitochondria, eukaryotes have demonstrated a great facility for establishing additional endosymbioses [5–7] and some may depend on their endosymbiotic associations for survival (e.g., aphids with *Buchnera* [8] and nematodes with *Wolbachia* [9]). Yet, even in associations with incredible functional integration, endosymbioses are dynamic: relationships change, partners are abandoned or swapped, and new types or levels of interactions emerge. The signatures of these past associations can persist in genomes and may influence future relationships [10,11]. It is the complexity, ubiquity, and significance of endosymbioses that make them fascinating subjects to study.

Given the pervasiveness of endosymbioses, it is perhaps not surprising that there is a wide range of empirical and experimental systems that vary across environments and taxa. If we consider systems organized by the size of the host cell, we can start with prokaryote hosts. Endosymbioses among prokaryotes are extremely rare and so, as a proxy, most attention has

been dedicated to the origin and evolution of the mitochondria within protoeukaryotes [12,13]. Within eukaryotes, there are many unicellular hosts with prokaryote endosymbionts (e.g., protists [14]), and some have been used as experimental models of endosymbioses, such as *Paramecium bursaria* [15]. Even in these relatively small organisms, there can be multiple types of endosymbionts and associated microbial communities. Similarly, large-scale organisms such as multicellular eukaryotes can have many endosymbionts and even endosymbionts that have their own endosymbionts [16]. Experimental models of these larger organisms often purposely select for relationships based on their ease of study and control, for example, the endosymbiosis between the Hawaiian bobtail squid *Euprymna scolopes* and *Vibrio fischeri* [17]. There are also many systems that lie somewhere on the spectrum of endosymbioses, such as plasmids and microbiomes that overlap with other fields of research including virology and microbial ecology. With all of this variation, it can be difficult to tease apart general features from organismal idiosyncrasies.

Mathematical modeling can serve as a useful complement to empirical techniques by allowing researchers to better understand their systems and place them in a wider context. By making the theory explicit, they can also serve as a universal language for collaboration, enabling research to be integrated from a range of disciplines [18,19] (see also [20,21] for how to interpret and integrate models with experiments), at times offering remarkable insights into biological systems [22] and simultaneously opening new fields of mathematics research [23]. Some recent reviews highlight the particular usefulness of mathematical models in evolutionary biology [24,25]. The first demonstrates how the rigorous logic of a mathematical formulation can identify the factors that facilitate the evolution of sexes or new species. The second outlines the insight that mathematical models can bring to the evolution of stress responses by integrating physiological mechanisms with an evolutionary optimality analysis. In this Essay, we consider the use of models to specifically address topics concerning endosymbioses and the types of hypotheses modeling can generate.

Integration of contrasting effects

A key question in the evolution of endosymbioses concerns the nature of the relationship between a host and its endosymbiont [26–28]. If the relationship is exploitative, then a coevolutionary arms race might ensue. If, instead, the relationship is mutualistic, then tighter integration and division of labor may evolve. Determining the nature of the interaction can thus lead to different predictions concerning the evolution of the relationship.

Empirical evidence suggests that the nature of the host–endosymbiont relationship is highly dynamic, changing across environmental conditions and time [29,30]. For example, changing the intensity of light shifted whether carrying green algal endosymbionts was costly or beneficial for its host [15]. In another interesting example, a five-year coculture evolved to change the interaction between the host and endosymbiont from initially exploitative to mutualistic [31]. A major determinant of such social evolution is the overall impact of the costs and benefits associated with behaviors, which influences the strength and mode of selection [32,33]. Since costs and benefits can change across environments and time scales, it can be difficult to determine their net effect over different contexts without using quantitative approaches.

Mathematical models can reveal surprising biological patterns when the interaction between costs and benefits is dynamic. For example, while vertical transmission was thought to be the primary way to reduce virulence in viruses, a mathematical model showed the opposite (i.e., that horizontal transmission also selects for lower virulence in viruses) [34]. Within the context of endosymbiosis research, an experiment showed that the *Dictyostelium discoideum* amoebae carrying *Paraburkholderia* bacteria intracellularly experienced a benefit in

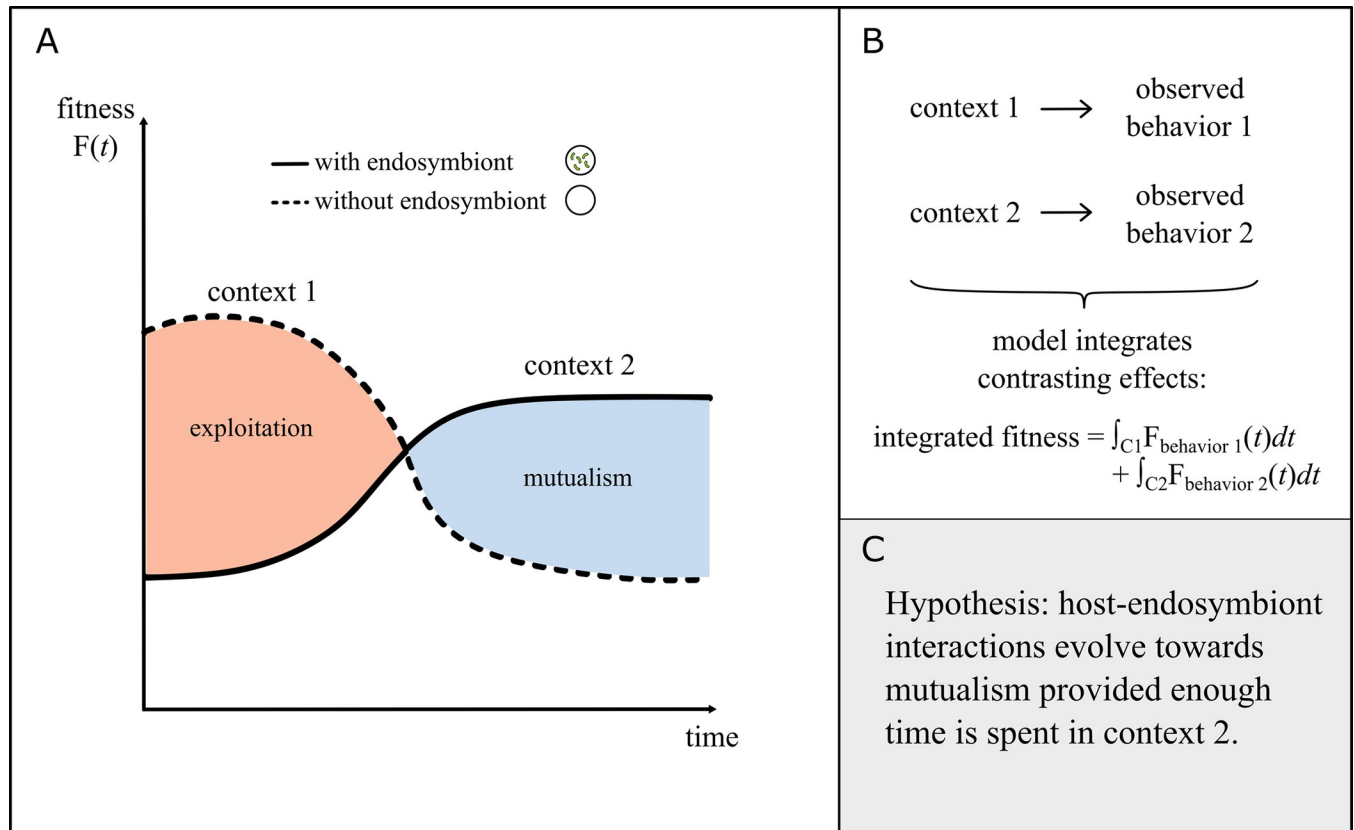


Fig 1. Integrating contrasting effects. (A) Sketch illustrating how the fitness of a host with or without endosymbionts changes in different contexts. These contexts affect whether selection favors mutualistic or exploitative relationships. Integrating over all contexts (e.g., the whole life cycle) gives a perspective on the overall behavior. (B) Schematic showing how 2 contrasting behaviors might be observed under different contexts. Integrating over the whole time period identifies the net outcome. (C) An example hypothesis generated for an endosymbiosis that appears to exhibit 2 contrasting behaviors—exploitation and mutualism—in different contexts.

<https://doi.org/10.1371/journal.pbio.3002583.g001>

nutrient-poor environments but not in nutrient-rich ones [29,35]. However, it was unclear whether carrying an endosymbiont would be beneficial in environments that switched between nutrient-rich and nutrient-poor conditions. Since the environments fluctuate, it might be reasonable to hypothesize that bet-hedging strategies would occur so that some, but not all, members of the population would carry endosymbionts. A mathematical model integrating the various costs and benefits tested this hypothesis and found that, contrary to expectation, bet-hedging strategies were rarely selected [29]. These examples highlight how mathematical models that integrate across different contexts can offer new insights into the dynamic nature of endosymbiosis interactions (Fig 1).

Revealing cross-system patterns

By employing abstractions to concentrate on processes and interactions, mathematical models effectively highlight connections between different fields and identify general themes. The typical modeling procedure entails abstracting a biological system, analyzing the model, and subsequently deducing implications for the biological system. This final stage offers an opportunity to generalize beyond the initial system. In modeling an endosymbiosis, the extent of these generalizations typically depends on how restrictive a model is. For example, a theoretical model was designed to investigate if a microbiome could induce cooperative behavior in

its host, with minimal restrictions on the microbiome itself [36]. The degree of abstraction present in this model allows its predictions to be applied to a broader range of entities that transmit between organisms, including plasmids, viruses, and multicellular symbionts.

Theoretical approaches can also identify fields that may benefit from exchange. For example, endosymbioses and certain microbial communities share similar patterns of division of labor, with both undergoing significant gene loss and evolving obligatory dependencies [37,38]. A challenge arising from such division of labor is coordination (i.e., who does what and when). If a host relies on its endosymbiont to produce energy, but sufficient energy is not provided, then the system can collapse. Understanding how coordination evolves is relevant to many areas of research in biology, including endosymbiosis, microbiomes, microbial community assembly, multicellularity, and mutualisms. These research areas all explore the ways in which selection acting on a system as a whole can lead to improved system performance, usually through some measure of fitness or function. Exchanges between these areas may help elucidate important mechanisms and common features of how multispecies systems can coordinate division of labor.

Another productive outcome of generalizing a mathematical model lies in exploring what happens when it does not apply to another system. Such failures in generalization can reveal useful distinctions that serve to organize scientific fields (Fig 2). For example, a model for

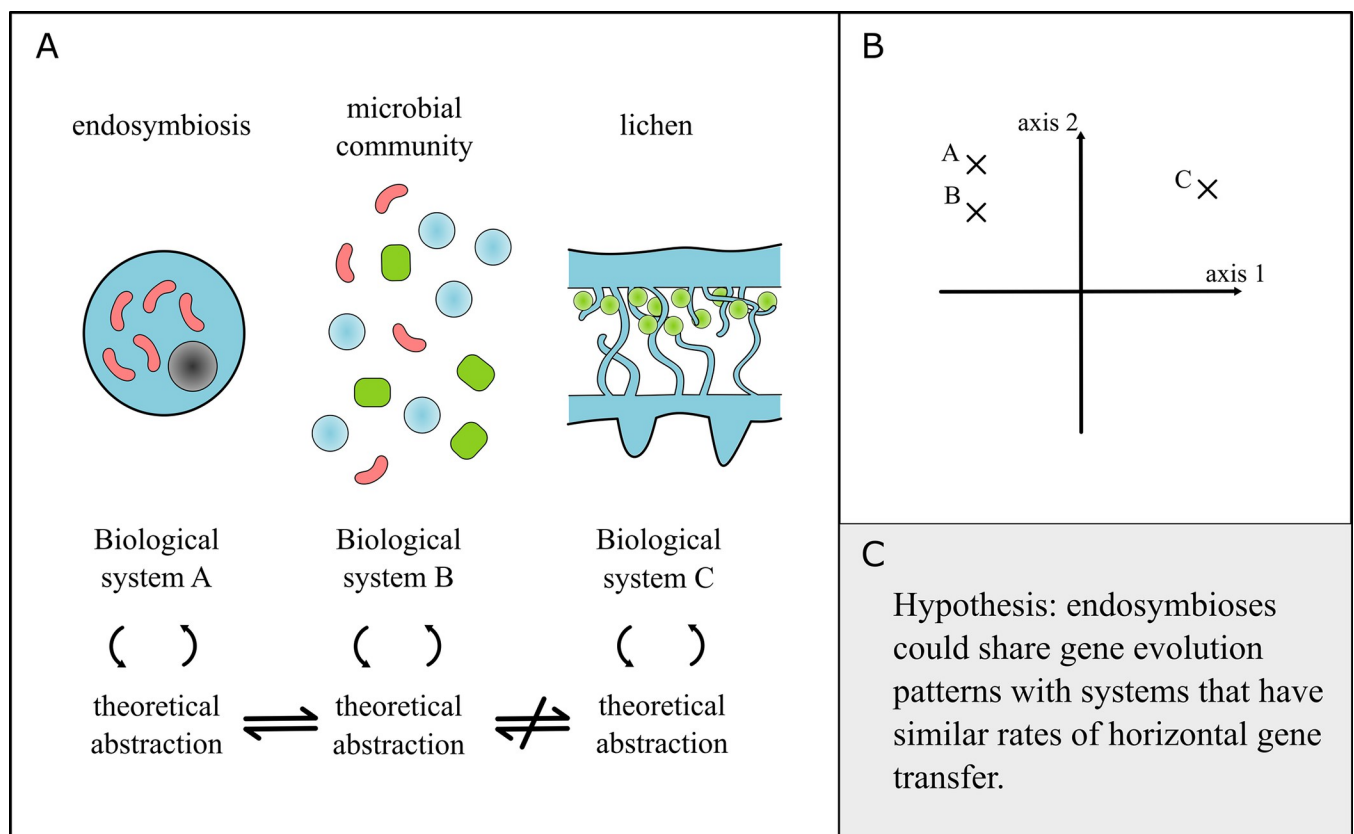


Fig 2. Broad generalizations. (A) Sketch illustrating how theoretical abstraction can be used to compare between different biological systems. For example, this can help to identify similarities in division of labor between some endosymbioses and microbial communities. Theoretical abstractions also highlight distinctions between systems (e.g., lichens have a lower frequency of horizontal gene transfer [43] than microbial communities). (B) Schematic showing how different axes can be used to separate biological systems. Here, systems A, B, and C could display similar dynamics along axis 2, but system C might be incomparable with A and B along axis 1. Moving along the axes could be interpreted in 2 ways depending on the scenario: (i) changing a parameter value, such as rate of horizontal gene transfer, shifts the system into a different regime where categorically different behaviors are observed; and (ii) entirely different models, such as mode of transfer, are required. (C) An example hypothesis generated for endosymbioses that broadly share similar rates of horizontal gene transfer.

<https://doi.org/10.1371/journal.pbio.3002583.g002>

insect–aphid symbiosis may not apply to plasmids because of the different frequencies in vertical versus horizontal transmission; this would indicate that transmission mode may be a useful axis to draw distinctions between endosymbioses. Indeed, a theoretical framework for endosymbioses organized them by the mode of transmission of endosymbionts in order to place them in the broader context of major transitions theory [2]. The value of such organization frameworks is 3-fold. First, they group systems together into categories that share similar abstractions and models, where results from one system can inform predictions about another. Second, they identify areas where empirical systems are missing, which indicates either that model systems should be developed or some constraints prevent these systems from occurring. Third, they can offer predictions on how categories evolve or transform into others (e.g., how endosymbioses may evolve into integrated units of selection/multispecies individuals) [2,39–42]. Overall, these generalizations provide a common language to explore questions spanning the spectrum of different endosymbioses.

Evaluating mechanisms

Biological systems are composed of a tangled web of interconnected components, which makes it difficult to identify the primary mechanisms responsible for a given phenomenon or behavior. To make matters worse, it is often unclear whether a crucial piece of information is missing. In such situations mathematical models can be extremely useful, providing a way to evaluate whether a set of components are sufficient to “explain” the phenomenon (Fig 3). An

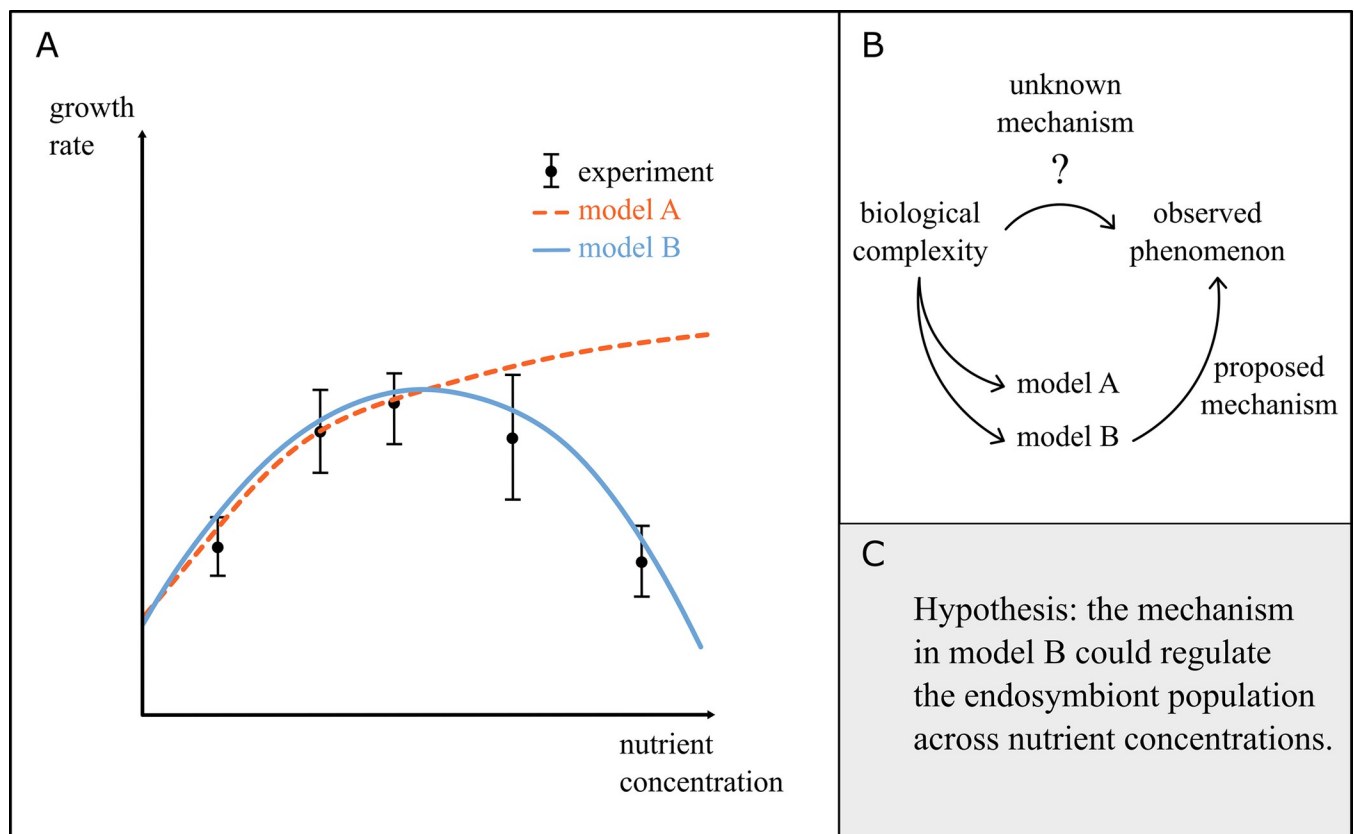


Fig 3. Identifying potential mechanisms. (A) Sketch illustrating 2 models for the growth of an endosymbiont under different nutrient concentrations with experimental results (inspired by [15]); model B is a better fit for the data. (B) Schematic showing how models can propose mechanisms for an observed phenomenon when the full details of the underlying mechanism are unknown. More than 1 model system can be explored to see when the outputs match the observations. (C) An example hypothesis generated for the regulation of endosymbionts where the primary mechanism is currently unknown.

<https://doi.org/10.1371/journal.pbio.3002583.g003>

example in endosymbiosis research is how hosts and endosymbionts navigate their relationships to maintain stability. Relationships between hosts and their endosymbionts are often regulated by a vast repertoire of chemical exchanges and physical interactions. Developing mathematical models for specific subsets of these exchanges and interactions can provide hypotheses for the primary mechanisms driving them. These hypotheses can then be experimentally explored as part of a collaborative experimentation–modeling process to identify underlying mechanisms.

When we lack empirical observations or a complete underlying molecular description, mathematical models can be used to assess whether a proposed mechanism is sufficient. An example of this in the endosymbiosis literature is a mathematical model that was developed to explain how the unicellular ciliate *P. bursaria* maintains a stable population size of its algal endosymbiont [44]. Since many of the molecular details of the underlying mechanism were missing, there were competing hypotheses for how stability could be maintained. In a combination of experiments and mathematical modeling, the authors demonstrated that a difference in the nutritional growth requirements of the host and endosymbiont was sufficient to replicate empirical observations and give rise to a stable maintenance of the population. The simple mechanism uncovered by this paper has additional properties with further implications (i.e., the mechanism functions across different growth conditions and does not require extensive coordination, which suggests it could evolve readily in different endosymbioses).

In addition to hypotheses for specific empirical systems, mathematical models can offer hypotheses concerning as yet unobserved phenomena and where to look for them. A particularly pertinent question is how host–endosymbiont relationships can become more intertwined, leading to a greater reliance on vertical rather than horizontal transmission. One model tackling this issue revealed that self-regulation would only evolve if the benefits of the relationship are sufficiently high for both the host and endosymbiont [45]. Analysis of the model also provided an explanation as to why benefits to both were necessary: when the benefit to hosts is large, then those without endosymbionts are outcompeted and lost in the population. As a result, endosymbionts effectively lose the horizontal route of transmission and stay with their hosts longer, leading them to evolve regulation of their own population size. A second theoretical study examined the costs endosymbionts face in finding new hosts and how this implicates host–endosymbiont relationship dynamics [46]. This model showed that it is more likely for endosymbionts to lose the ability to live freely and reproduce independently when host encounters are both rare and costly. Together, these results not only show when and why it might be possible for endosymbionts to control their own population, but also provide some prediction of what kinds of endosymbioses might foster this behavior.

Explore the unknown

Within the field of endosymbiosis, there is a glaring absence of endosymbioses featuring prokaryote hosts. There are only a few observed examples besides possibly the endosymbiosis that would give rise to the mitochondria [16,47,48]. This rarity is surprising when compared to the frequency of eukaryote endosymbioses and when considering the diversity and abundance of prokaryotes. There have been different qualitative arguments as to why prokaryote endosymbioses may be rare [13,49,50], but since we lack experimental systems, it is difficult to identify what factors are responsible for the rarity. Here, mathematical models can be particularly helpful in evaluating and comparing hypotheses for rarity and determining what constraints limit the successful establishment of prokaryotic endosymbioses (Fig 4).

An example of this use of mathematical modeling can be seen in a recent study that considered the role of metabolic compatibility in prokaryote endosymbioses [51]. Prior to this work,

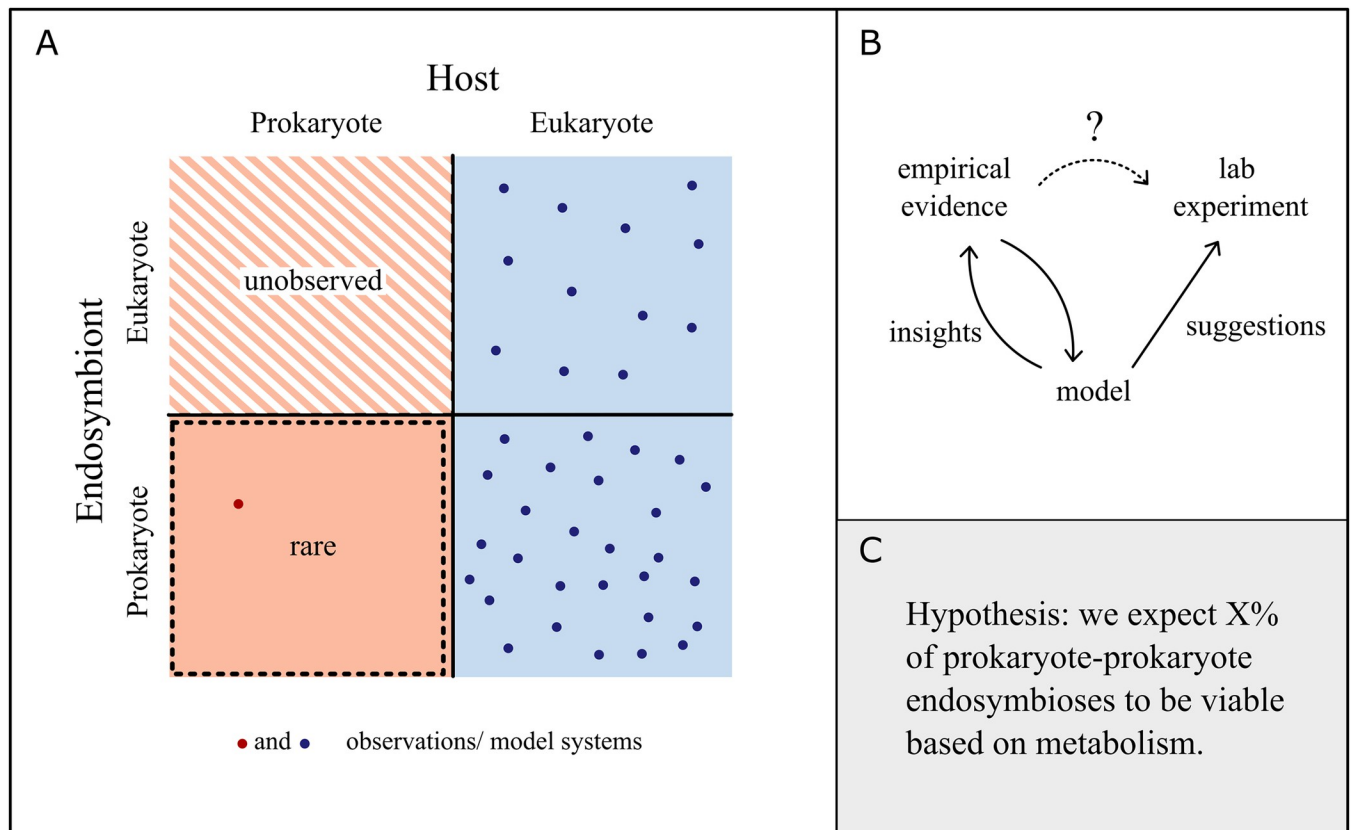


Fig 4. Exploring the unknown. (A) Several examples of endosymbioses with a eukaryote are known and well studied, but prokaryote hosts are lacking. Prokaryote hosts with eukaryote endosymbionts have not been observed and might be unfeasible due to space limitations. Prokaryote–prokaryote endosymbioses are extremely rarely observed. These endosymbioses can be explored with theoretical models, for example, to estimate the proportion of prokaryote–prokaryote pairs that could form viable endosymbioses based on metabolism (as in [51]). (B) Schematic showing how modeling can be used to give insights for empirical systems where lab experiments have yet to be devised. The model can suggest potential experimental systems, for example, by identifying species that have a high propensity to become hosts and endosymbionts. (C) An example hypothesis generated for prokaryote–prokaryote endosymbioses where experimental systems are currently unknown.

<https://doi.org/10.1371/journal.pbio.3002583.g004>

some had argued that prokaryote endosymbioses may not be viable without a prior history of coevolution (see [13] for an excellent review). For example, if an endosymbiont needs some essential compound to grow but the host cell lacks a way of transporting this compound into the cell, then the endosymbiosis would not be viable. However, it is unclear how often this type of scenario would be expected to occur in pairs of prokaryotes (i.e., to what extent is this actually a driver of the rarity). Mathematical models can be used to estimate this and thus offer a null model prediction that can later be refined or assessed by incorporating other, possibly empirical, data. In this particular case, bacterial metabolic models were randomly paired as *in silico* endosymbioses. Over half of the pairs tested were metabolically viable, though very rarely did they have higher fitness (growth rate) than their ancestors. Without empirical systems, it can be difficult to validate these quantitative models or assess their accuracy, yet they provide a null prediction that can be used to set expectations and compare with future findings.

New lines of inquiry

Mathematical models can be used to establish new lines of inquiry both in terms of scale and direction. These uses of modeling are well-illustrated in studies of corals and their

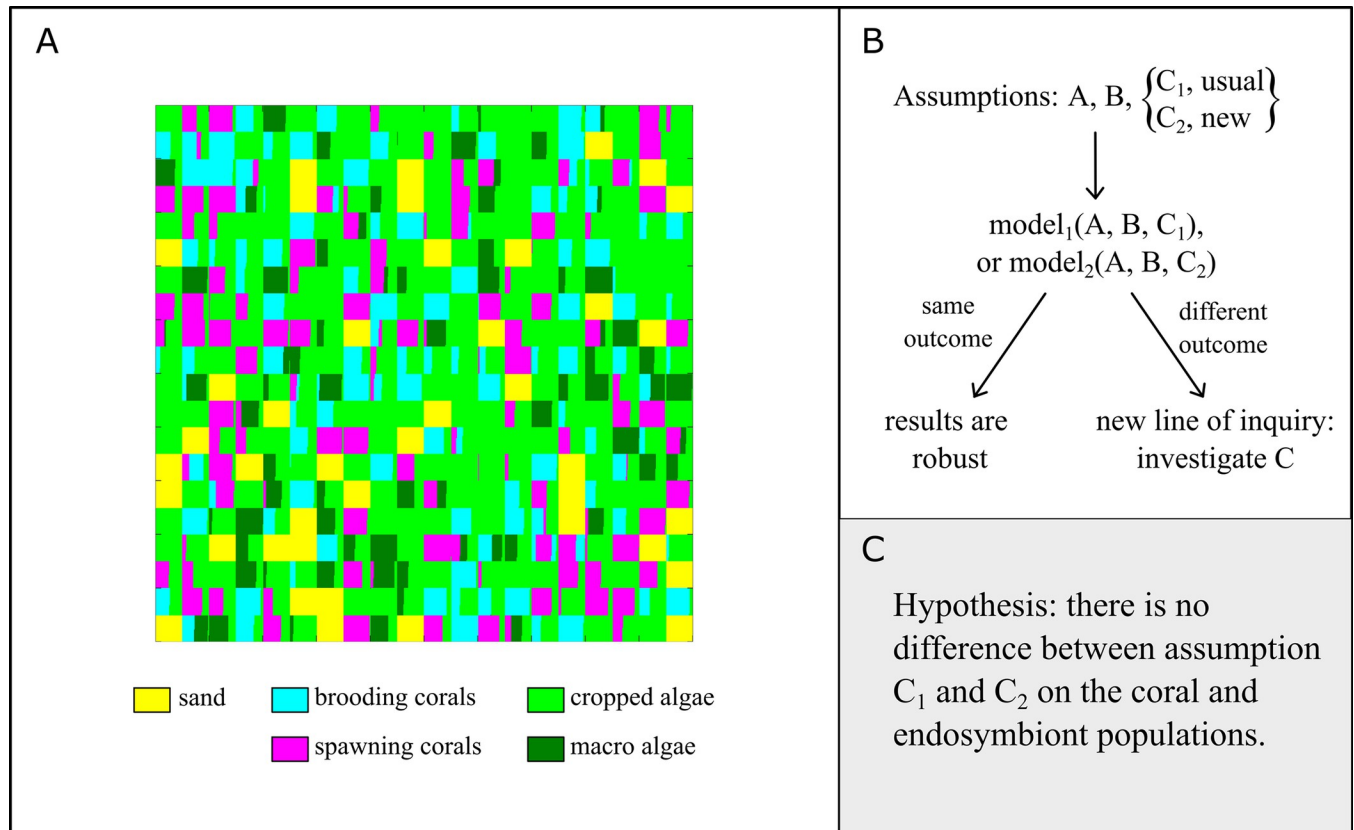


Fig 5. New lines of inquiry. (A) A representative output from agent-based model simulations of a coral ecosystem, adapted from [53]. Each grid square's color represents the relative proportions of sand, various coral species, and algal types. To enhance this model by incorporating the impact of endosymbionts on coral growth rates and mortality during stressful events, such as heatwaves [54], assumptions are required regarding the interactions between corals and their endosymbiotic algae (for example, how corals may alter their internal symbiont composition in response to such events). (B) Schematic illustrating the assumptions necessary for constructing a model. Two potential assumptions for component C are presented. By running simulations with both assumptions, we can test the robustness of the results; identical outcomes suggest that the model's results are not sensitive to the assumptions about C. Conversely, divergent outcomes indicate that further investigation into assumption C could lead to new research avenues. (C) An example hypothesis generated for a coral endosymbiosis where different assumptions can be made.

<https://doi.org/10.1371/journal.pbio.3002583.g005>

dinoflagellate endosymbionts, which provide an exceptional system for studying endosymbiosis because there is abundant data in terms of spatial variation of coral communities, as well as a rich historical record of environmental conditions and coral growth. This data has led to the development of intricate models that explicitly define the many mechanisms and processes that affect individual organism behaviors and their interactions with the environment. In particular, agent-based models excel in capturing heterogeneity and spatial interactions among individuals and other components of the system, crucial for accurately modeling complex ecosystems like coral reefs [52–54]. Because these models are very complex, they are often calibrated and validated by comparing their hindcast predictions with historical data, thereby improving their reliability for future projections about coral health in response to a changing climate [55]. Given the relevance of such historic data, recent efforts have focused on extending the temporal timescale further back to procure more robust data [56], creating lines of inquiry for endosymbiosis that consider timescales beyond human lifetimes.

Both creating a mathematical model and analyzing it can uncover hidden assumptions and reveal new directions to explore (Fig 5). The process of writing down a model, such as an agent-based model, forces one to make many explicit decisions on how agents behave and

within what space they interact. Reflecting on these decisions can identify unknown details and outstanding questions. For example, in [54], the authors extended a spatially explicit agent-based model for corals [53] to assess the long-term benefits of switching from thermally sensitive to thermally tolerant symbionts after heat waves. In developing the model, the authors faced the decision of whether to include symbiont reversal to the original composition after some time or to keep the switch from sensitive to tolerant symbionts unidirectional. Including symbiont reversal would require also knowing the relationship between the dynamics of symbiont switching and coral growth. The authors concluded that the empirical data to characterize this was insufficient at the time and assumed the switch from sensitive to tolerant symbionts was unidirectional. This is a clear example of a situation in which the implementation of a model forces the authors to make certain assumptions explicit (the reversibility of symbiont switching) and points to new lines of inquiry (how to characterize the switching dynamics).

In some cases, there is not enough evidence to distinguish between competing assumptions in a model. Experiments may help determine which assumption is better supported, but conducting experiments for every modeling assumption is not cost-effective. By analyzing a model using each assumption, we can explore whether the competing assumptions actually lead to different outcomes. To see how this works, we consider the earlier case of the assumption concerning symbiont switching. Models could be run with different assumed relationships between coral growth and symbiont switching dynamics as a kind of sensitivity analysis. If a wide swath of dynamic relationships produce the same behavior, then this would suggest that while the exact relationship is unknown, it may not be as important to investigate as other parameters. If instead different dynamic relationships produce different behaviors, then this would indicate that this parameter should be investigated. Since computations tend to be cheaper and faster than experiments, this methodology can determine which of the unknowns in a model are the most important to learn.

Frameworks for integration

With the variety of empirical systems and theoretical approaches available, we also need ways of combining them into common frameworks to address central questions on endosymbioses. This comes with challenges, such as simplifying complex problems, ensuring realistic results, identifying influential variables, and requiring iterative refinement. Although in principle there are many ways of doing this (e.g., Bayesian models), an interesting way that may be relevant for questions in endosymbiosis research is the use of Fermi estimates. Fermi estimates simplify complex problems by dividing them into smaller, more manageable parts and using educated guesses. Possibly the most famous example of a Fermi estimate is the Drake equation, which tries to estimate the number of radio-communicative civilizations in the Milky Way [57,58]. This involves estimating factors such as the rate of star formation, the fraction of those stars with planets, and the likelihood of life developing intelligence.

To see how Fermi estimates may be useful in endosymbiosis research, one could consider the question of why prokaryote endosymbioses are rare. This question could be addressed by estimating the likelihood of establishing a prokaryote endosymbiosis, which in turn could be broken down into the following terms: the total number of interactions between different prokaryotic species; the fraction of those encounters that lead to one cell getting inside another; the fraction of those newly created endosymbioses that can reproduce; and the fraction of those viable endosymbioses that persist long enough to fix in a population. Though we might not know the precise value of any of these terms, we can make educated guesses for some by using existing empirical data. For example, we can estimate the first term by multiplying the

average number of interactions per species [59] with the number of prokaryote species [60]. Where data is unavailable, theoretical approaches can provide approximations for some of the gaps; for example, metabolic models have been used to estimate the fraction of newly created endosymbioses that are viable [51]. Of course, this approach to estimation raises all sorts of possible issues and omissions; however, addressing these factors and determining how to incorporate them is part of the value of such a formulation.

Conclusions

In this Essay, we have explored how mathematical models can serve as useful tools in endosymbiosis research, showcasing the types of hypotheses mathematical models can generate and how they can be used to complement empirical approaches. We purposely selected examples representing a diverse set of models and model systems to highlight the breadth of possibilities in terms of the utility and application of mathematical models. To get a sense of the types of questions within endosymbiosis research that mathematical modeling may be particularly well-suited to address, we outline some key questions in Box 1. Many of the modeling examples considered in this Essay address one of these questions (e.g., coral endosymbiosis models often address the first question, concerning the effects of environmental conditions). Yet, there are some questions that have been relatively unexplored. For example, the third question poses how an endosymbiosis may respond to an additional element such as another endosymbiont or a virus. This question is relevant in terms of plastid acquisition, whereby eukaryotes have gained additional endosymbionts following the mitochondria. Regardless of the extent to which these questions have been considered, the tremendous diversity of endosymbioses means there is plenty of room for exploration, both building on existing models and in new directions.

Box 1. Topics in endosymbiosis fit for mathematical modeling

1. **Dynamic environments and evolving relationships:** Do changing environmental conditions alter the costs/benefits of endosymbioses and their evolution along the mutualism–exploitation spectrum?
2. **Coordination of reproduction:** How do endosymbioses ensure consistent, synchronized reproduction of host and endosymbionts in the face of possible conflicts over resource use and selection to maximize growth rate?
3. **Third party influence:** How do endosymbioses respond to the introduction of a third species such as another endosymbiont or a virus?
4. **Division of labor:** In what ways can endosymbioses partition labor (e.g., energy transformation or metabolism) to gain sustainable, synergistic benefits?
5. **Comparative effects of spatial organization:** When does the specific spatial arrangement of endosymbioses offer different outcomes than symbioses between partners that are not arranged in a nested architecture?
6. **Harmonious pairings:** Are there certain types of species that are more likely to produce a successful endosymbiosis than others?
7. **Long-term horizon:** What factors influence whether an endosymbiont remains a long-term partner, evolves into an organelle, or deteriorates until lost or replaced?

Certainly mathematical models have their limitations, and for many relevant questions in endosymbioses it is better to interrogate an experimental system or empirical data. But there are some tasks for which mathematical models are particularly well suited. Moreover, the act of producing a model can be informative in and of itself, because it requires explicitly formulating assumptions and identifying what necessary information is missing. For now, there is a distinction in many fields of biology, such as endosymbiosis, between modelers and empiricists/experimentalists. They often collaborate, but owing to their different methodologies and academic backgrounds, they sometimes occupy different spaces within the same field. In this way, modelers exist as a type of endosymbiont within their host biological fields, seeking inspiration and exciting questions from the complexities and mysteries of biological systems. While there are costs to having modelers, hopefully the net exchange is positive, with modelers providing useful generalizations, identifying key mechanisms, and offering new insight and lines of inquiry.

Author Contributions

Conceptualization: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Formal analysis: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Funding acquisition: Eric Libby.

Investigation: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Methodology: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Project administration: Eric Libby.

Resources: Eric Libby.

Supervision: Eric Libby.

Validation: Eric Libby.

Visualization: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Writing – original draft: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

Writing – review & editing: Lucas Santana Souza, Josephine Solowiej-Wedderburn, Adriano Bonforti, Eric Libby.

References

1. Rafiqi AM, Rajakumar A, Abouheif E. Origin and elaboration of a major evolutionary transition in individuality. *Nature*. 2020; 585:239–244. <https://doi.org/10.1038/s41586-020-2653-6> PMID: 32879485
2. Estrela S, Kerr B, Morris JJ. Transitions in individuality through symbiosis. *Curr Opin Microbiol*. 2016; 31:191–198. <https://doi.org/10.1016/j.mib.2016.04.007> PMID: 27131019
3. Radzvilavicius AL, Blackstone NW. The evolution of individuality revisited. *Biol Rev Camb Philos Soc*. 2018; 93:1620–1633. <https://doi.org/10.1111/brev.12412> PMID: 29575407
4. West SA, Fisher RM, Gardner A, Kiers ET. Major evolutionary transitions in individuality. *Proc Natl Acad Sci U S A*. 2015; 112:10112–10119. <https://doi.org/10.1073/pnas.1421402112> PMID: 25964342

5. Hackett JD, Yoon HS, Butterfield NJ, Sanderson MJ, Bhattacharya D. Plastid endosymbiosis: sources and timing of the major events. *Evolution of Primary Producers in the Sea*. Academic Press; 2007. p. 109–132.
6. Kneip C, Lockhart P, Voß C, Maier UG. Nitrogen fixation in eukaryotes—new models for symbiosis. *BMC Evol Biol*. 2007; 7:55. <https://doi.org/10.1186/1471-2148-7-55> PMID: 17408485
7. Nowack EC, Melkonian M. Endosymbiotic associations within protists. *Philos Trans R Soc Lond B Biol Sci*. 2010; 365:699–712. <https://doi.org/10.1098/rstb.2009.0188> PMID: 20124339
8. Shigenobu S, Wilson AC. Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. *Cell Mol Life Sci*. 2011; 68:1297–1309. <https://doi.org/10.1007/s00018-011-0645-2> PMID: 21390549
9. Kaur R, Shropshire JD, Cross KL, Leigh B, Mansueto AJ, Stewart V, et al. Living in the endosymbiotic world of Wolbachia: a centennial review. *Cell Host Microbe*. 2021; 29:879–893. <https://doi.org/10.1016/j.chom.2021.03.006> PMID: 33945798
10. Wilson AC, Duncan RP. Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. *Proc Natl Acad Sci U S A*. 2015; 112:10255–10261. <https://doi.org/10.1073/pnas.1423305112> PMID: 26039986
11. Jackson R, Henry LM, Wurm Y. Evolution: The legacy of endosymbiosis in ants. *Curr Biol*. 2020; 30: R1385–R1387. <https://doi.org/10.1016/j.cub.2020.09.023> PMID: 33202242
12. Martin WF. Symbiogenesis, gradualism, and mitochondrial energy in eukaryote origin. *Period Biol*. 2017; 119:141–158.
13. Zachar I, Boza G. Endosymbiosis before eukaryotes: mitochondrial establishment in protoeukaryotes. *Cell Mol Life Sci*. 2020; 77:3503–3523. <https://doi.org/10.1007/s00018-020-03462-6> PMID: 32008087
14. Husnik F, Tashyreva D, Boscaro V, George EE, Lukeš J, Keeling PJ. Bacterial and archaeal symbioses with protists. *Curr Biol*. 2021; 31:R862–R877. <https://doi.org/10.1016/j.cub.2021.05.049> PMID: 34256922
15. Lowe C, Minter E, Cameron D, Brockhurst M. Shining a Light on Exploitative Host Control in a Photosynthetic Endosymbiosis. *Curr Biol*. 2016; 26:207–211. <https://doi.org/10.1016/j.cub.2015.11.052> PMID: 26748854
16. Von Dohlen CD, Kohler S, Alsop ST, McManus WR. Mealybug β -proteobacterial endosymbionts contain γ -proteobacterial symbionts. *Nature*. 2001; 412:433–436.
17. McFall-Ngai M. Divining the Essence of Symbiosis: Insights from the Squid-Vibrio Model. *PLoS Biol*. 2014; 12:e1001783. <https://doi.org/10.1371/journal.pbio.1001783> PMID: 24504482
18. Robinaugh DJ, Haslbeck JM, Ryan O, Fried EI, Waldorp LJ. Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspect Psychol Sci*. 2021; 16:725–743. <https://doi.org/10.1177/1745691620974697> PMID: 33593176
19. Zuk M, Travisano M. Models on the runway: how do we make replicas of the world? *Am Nat*. 2018; 192:1–9. <https://doi.org/10.1086/697508> PMID: 29897802
20. Möbius W, Laan L. Physical and mathematical modeling in experimental papers. *Cell*. 2015; 163:1577–1583. <https://doi.org/10.1016/j.cell.2015.12.006> PMID: 26687351
21. Grainger TN, Senthilnathan A, Ke PJ, Barbour MA, Jones NT, DeLong JP, et al. An empiricist's guide to using ecological theory. *Am Nat*. 2022; 199:1–20. <https://doi.org/10.1086/717206> PMID: 34978962
22. Jungck JR. Ten equations that changed biology: mathematics in problem-solving biology curricula. *Bios*. 1997; 23:11–36.
23. Reed MC. Mathematical biology is good for mathematics. *Not Am Math Soc*. 2015; 62:1172–1176.
24. Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, Stern CA, et al. Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. *PLoS Biol*. 2014; 12:e1002017. <https://doi.org/10.1371/journal.pbio.1002017> PMID: 25489940
25. Taborsky B, English S, Fawcett TW, Kuijper B, Leimar O, McNamara JM, et al. Towards an evolutionary theory of stress responses. *Trends Ecol Evol*. 2021; 36:39–48. <https://doi.org/10.1016/j.tree.2020.09.003> PMID: 33032863
26. Frank SA. The origin of synergistic symbiosis. *J Theor Biol*. 1995; 176:403–410. <https://doi.org/10.1006/jtbi.1995.0208> PMID: 8538218
27. Law R, Dieckmann U. Symbiosis through exploitation and the merger of lineages in evolution. *Proc Biol Sci*. 1998; 265:1245–1253.
28. Roughgarden J. Holobiont evolution: Population theory for the hologenome. *Am Nat*. 2023; 201:763–778. <https://doi.org/10.1086/723782> PMID: 37229712

29. Scott TJ, Queller DC, Strassmann JE. Context dependence in the symbiosis between *Dictyostelium discoideum* and *Paraburkholderia*. *Evol Lett*. 2022; 6:245–254. <https://doi.org/10.1002/evl3.281> PMID: 35784451
30. Keeling PJ, McCutcheon JP. Endosymbiosis: The feeling is not mutual. *J Theor Biol*. 2017; 434:75–79. <https://doi.org/10.1016/j.jtbi.2017.06.008> PMID: 28624393
31. Jeon KW. Development of cellular dependence on infective organisms: micrurgical studies in amoebas. *Science*. 1972; 176:1122–1123. <https://doi.org/10.1126/science.176.4039.1122> PMID: 5035468
32. Van Vliet S, Doebeli M. The role of multilevel selection in host microbiome evolution. *Proc Natl Acad Sci U S A*. 2019; 116:20591–20597. <https://doi.org/10.1073/pnas.1909790116> PMID: 31548380
33. Okasha S. *Evolution and the levels of selection*. Clarendon Press; 2006.
34. Lipsitch M, Siller S, Nowak MA. The evolution of virulence in pathogens with vertical and horizontal transmission. *Evolution*. 1996; 50:1729–1741. <https://doi.org/10.1111/j.1558-5646.1996.tb03560.x> PMID: 28565576
35. Mather RV, Larsen TJ, Brock DA, Queller DC, Strassmann JE. *Paraburkholderia* symbionts isolated from *Dictyostelium discoideum* induce bacterial carriage in other *Dictyostelium* species. *Proc Biol Sci*. 2023; 290:20230977. <https://doi.org/10.1098/rspb.2023.0977> PMID: 37464760
36. Lewin-Epstein O, Hadany L. Host–microbiome coevolution can promote cooperation in a rock–paper–scissors dynamics. *Proc Biol Sci*. 2020; 287:20192754. <https://doi.org/10.1098/rspb.2019.2754> PMID: 32075531
37. Morris JJ, Lenski RE, Zinser ER. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio*. 2012; 3:e00036–e00012. <https://doi.org/10.1128/mBio.00036-12> PMID: 22448042
38. Pande S, Kost C. Bacterial unculturability and the formation of intercellular metabolic networks. *Trends Microbiol*. 2017; 25:349–361. <https://doi.org/10.1016/j.tim.2017.02.015> PMID: 28389039
39. Queller DC, Strassmann JE. Beyond society: the evolution of organismality. *Philos Trans R Soc Lond B Biol Sci*. 2009; 364:3143–3155. <https://doi.org/10.1098/rstb.2009.0095> PMID: 19805423
40. Queller DC, Strassmann JE. Problems of multi-species organisms: endosymbionts to holobionts. *Biol Philos*. 2016; 31:855–873.
41. Libby E, Ratcliff WC. Lichens and microbial syntrophies offer models for an interdependent route to multicellularity. *Lichenologist*. 2021; 53:283–290.
42. Andersson R, Isaksson H, Libby E. Multi-Species Multicellular Life Cycles. In: *The Evolution of Multicellularity*. CRC Press; 2022. p. 343–356.
43. Tunjić M, Korać P. Vertical and horizontal gene transfer in lichens. *Period Biol*. 2013; 115:321–329.
44. Iwai S, Fujiwara K, Tamura T. Maintenance of algal endosymbionts in *Paramecium bursaria*: a simple model based on population dynamics. *Environ Microbiol*. 2016; 18:2435–2445. <https://doi.org/10.1111/1462-2920.13140> PMID: 26625979
45. Uchiumi Y, Ohtsuki H, Sasaki A. Evolution of self-limited cell division of symbionts. *Proc Biol Sci*. 2019; 286:20182238. <https://doi.org/10.1098/rspb.2018.2238> PMID: 30963950
46. Nguyen PL, van Baalen M. On the difficult evolutionary transition from the free-living lifestyle to obligate symbiosis. *PLoS ONE*. 2020; 15:e0235811. <https://doi.org/10.1371/journal.pone.0235811> PMID: 32730262
47. Yamaguchi M, Worman CO. Deep-sea microorganisms and the origin of the eukaryotic cell. *Jpn J Protozool*. 2014; 47:29–48.
48. Wujek DE. Intracellular bacteria in the blue-green alga *Pleurocapsa minor*. *Trans Am Microsc Soc*. 1979; 98:143–145.
49. Martin WF, Tielens AG, Mentel M, Garg SG, Gould SB. The physiology of phagocytosis in the context of mitochondrial origin. *Microbiol Mol Biol Rev*. 2017; 81:e00008–e00017. <https://doi.org/10.1128/MMBR.00008-17> PMID: 28615286
50. Bhattacharya D, Archibald JM, Weber AP, Reyes-Prieto A. How do endosymbionts become organelles? Understanding early events in plastid evolution. *Bioessays*. 2007; 29:1239–1246. <https://doi.org/10.1002/bies.20671> PMID: 18027391
51. Libby E, Kempes CP, Okie JG. Metabolic compatibility and the rarity of prokaryote endosymbioses. *Proc Natl Acad Sci U S A*. 2023; 120:e2206527120. <https://doi.org/10.1073/pnas.2206527120> PMID: 37071674
52. Bozec YM, Mumby JP. Australian Institute of Marine Science. Supplementary report to the final report of the coral reef expert group. S7. Coral reef models as assessment and reporting tools for the Reef 2050 Integrated Monitoring and Reporting Program—a review. 2020. Available from: <http://hdl.handle.net/11017/3570>.

53. Mumby PJ. The impact of exploiting grazers (Scaridae) on the dynamics of Caribbean coral reefs. *Ecol Appl*. 2006; 16:747–769. [https://doi.org/10.1890/1051-0761\(2006\)016\[0747:tioegs\]2.0.co;2](https://doi.org/10.1890/1051-0761(2006)016[0747:tioegs]2.0.co;2) PMID: 16711060
54. Ortiz JC, González-Rivero M, Mumby PJ. Can a thermally tolerant symbiont improve the future of Caribbean coral reefs? *Glob Chang Biol*. 2013; 19:273–281. <https://doi.org/10.1111/gcb.12027> PMID: 23504738
55. Hughes TP, Barnes ML, Bellwood DR, Cinner JE, Cumming GS, Jackson JBC, et al. Coral reefs in the Anthropocene. *Nature*. 2017; 546:82–90. <https://doi.org/10.1038/nature22901> PMID: 28569801
56. Kamenos NA, Hennige SJ. Reconstructing four centuries of temperature-induced coral bleaching on the Great Barrier Reef. *Front Mar Sci*. 2018; 5:283.
57. Drake FD. Project ozma. *Phys Today*. 1961; 14:40–46.
58. Tarter J. The search for extraterrestrial intelligence (SETI). *Ann Rev Astron Astrophys*. 2001; 39:511–548.
59. Kodera SM, Das P, Gilbert JA, Lutz HL. Conceptual strategies for characterizing interactions in microbial communities. *iScience*. 2022; 25:103775. <https://doi.org/10.1016/j.isci.2022.103775> PMID: 35146390
60. Locey KJ, Lennon JT. Scaling laws predict global microbial diversity. *Proc Natl Acad Sci U S A*. 2016; 113:5970–5975. <https://doi.org/10.1073/pnas.1521291113> PMID: 27140646