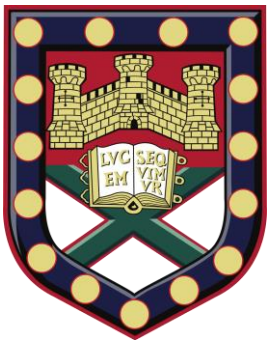
THE HOLOGENOME CONCEPT OF EVOLUTION: A PHILOSOPHICAL AND BIOLOGICAL STUDY



Submitted by **Javier Suárez** Díaz to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Philosophy

In August 2019.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……..

Abstract

The hologenome concept of evolution is a hypothesis about the evolution of animals and plants. It asserts that the evolution of animals and plants was partially triggered by their interactions with their symbiotic microbiomes. In that vein, the hologenome concept posits that the holobiont (animal host + symbionts of the microbiome) is a unit of selection.

The hologenome concept has been severely criticized on the basis that selection on holobionts would only be possible if there were a tight transgenerational host-genotype-to-symbiont-genotype connection. As our current evidence suggests that this is not the case for most of the symbiont species that compose the microbiome of animals and plants, the opportunity for holobiont selection is very low in relation to the opportunity for selection on each of the species that compose the host microbiome. Therefore, holobiont selection will always be disrupted 'from below', by selection on each of the species that compose the microbiome.

This thesis constitutes a conceptual effort to defend philosophically the hologenome concept. I argue that the criticism according to which holobiont selection requires tight transgenerational host-genotype-to-symbiont-genotype connection is grounded on a metaphysical view of the world according to which the biological hierarchy needs to be nested, such that each new level of selection includes every entity from below. Applied to hologenomes, it entails that the hologenome is a collection of genomes, and selection of hologenomes is assumed to entail cospeciation of the host with the species that constitute its microbiome.

Against that interpretation, I propose the ‘stability of traits’ account, according to which hologenome evolution is the result of the action of natural selection in a non-nested hierarchical world. In that vein, hologenome evolution does not entail cospeciation, and thus it does not require tight transgenerational host-genotype-to-symbiont-genotype connection. By embracing a multilevel selection perspective, I argue that hologenome evolution results from the simultaneous action of natural selection on each of the lineages that compose the microbiome, and on the assemblage composed by the host genome plus the functional traits of its microbiome. Hologenome selection occurs when the evolution of the traits of the microbiome result from their effects on the fitness of the host, and it can take the form of multilevel selection 1, or multilevel selection 2. In both cases, hologenome selection entails the evolution of microbiome traits, as well as evolution of the host genome, rather than cospeciation of lineages.

List of contents

List of Illustrations

List of abbreviations

***DT**: discontinuity theory of immunity*

***EM**: equilibrium model of immunity*

***HCE**: hologenome concept of evolution*

***HGT**: horizontal gene transfer*

***MLS**: multilevel selection theory*

***MLS1**: multilevel selection theory 1*

***MLS2**: multilevel selection theory 2*

***OTU**: operational taxonomic unit*

***PBI**: problem of biological individuality*

***SoT**: stability of traits*

***SoS**: stability of species*

***SoN**: stability of nucleotides*

Introduction

Introduction and motivation of the doctoral project

The doctoral project that I develop here aims at shedding light on one of the most disputed topics in theoretical biology and philosophy of biology in the last few years, the so-called 'hologenome concept of evolution' (HCE). HCE is an hypothesis about the nature of biological individuality, and specifically about the role of symbiosis in defining the (blurry) boundaries of biological individuals. Concretely, HCE posits that the entity composed by a macrobial host (animal, plant) plus the microbial symbionts of its microbiome (bacteria, Archaea, fungi, viruses) is a biological individual, that they refer to as the 'holobiont', and also a level (or unit) of selection in evolution. The statement that the holobiont is a biological individual and a unit of selection constitutes, in my view, and as I will discuss throughout the doctoral project, the main innovative thesis that the formulation of HCE puts forward. HCE defenders support their thesis by departing from two key empirical observations: the pervasiveness of macrobe-microbe symbiotic relationships, and the importance of symbionts for the life of the macrobes whose bodies they occupy. It is precisely the observation of these two empirical realities, plus a series of experimental results hard to explain with the traditional conception of 'one organism = one genome', that lead the authors to propose such a radical thesis about biological individuality.

HCE was originally proposed by biologists, but its postulates go far beyond pure biological theory. As it is a hypothesis about biological individuality, it has implications for metaphysics, ontology, epistemology, philosophy of science, and the relation between these categories. Foremost, and most intuitively, HCE raises a clear *ontological* question: if holobionts are biological individuals, how should we understand the ontology of individuals? In other

words, what makes biological individuals ‘individuals’, if symbionts are structural elements of biological individuals? And, if they are, should biological individuality be understood hierarchically, like a series of concatenated Matryoshka dolls? The ontological question immediately prompts an *epistemological* question: are there scientifically relevant criteria to determine the truth of HCE, or is it a hypothesis that cannot be empirically validated in any significant manner? Can we *really* determine whether holobionts are (or are not) biological individuals, or we can only *assume* that they are (or they are not)? What will be the impact for scientific research of assuming that holobionts are individuals? These questions are, at the same time, connected to many questions in *philosophy of science*: what is the evidence that should be demanded to accept (or reject) HCE? Is it worth pursuing a scientific avenue of research if the evidence that supports it is scarce? If HCE turns out to be based on false assumption, but it still leads to interesting pathways of research, is it possible to argue that a promising community-wide research project can be based on false assumptions?

As with most questions in philosophy, each of the questions I have just presented are not monolithic, and the answer to some subset of these questions will have immediate impact on the way of answering the others. What is more, many of these questions only arise after others have been asked, or only after others have been answered. I have come up with most of these questions during my doctoral project, precisely as a consequence of my doubts about the truth and scientific validity of HCE, which soon created an important contradictory attitude in my stance towards the hypothesis: on the one hand, I was convinced that there was a ‘core of truth’ in the theory, and I did so based, among other things, on how the assumption of the hypothesis had led, and was leading, many biologists toward really fascinating scientific discoveries that were hard to imagine had HCE not been proposed. On the other hand, I was constantly questioning the truth of the hypothesis, since mounting evidence seemed to suggest that the ‘core of truth’ that I believed to be contained in HCE vanished, and I was constantly tempted to believe that the hypothesis was simply false. Was there a way of putting order in the whole array of empirical and philosophical literature about holobionts and hologenomes, so that the

hypothesis could be consistently held, and made coherent with the accumulated empirical evidence? Especially, was it possible to elaborate a coherent conceptual explanation of the claim that holobionts are units of selection? And if so, what was the purpose of elaborating such a coherent 'story'? Another important question that revolved in my mind was how far my 'coherent story' could (and should) be from our current empirical data about the role of holobionts as units of selection. Was it possible to argue that one of the reasons why HCE was dismissed was because the authors were supporting different definitions of the holobiont, or different views about what units of selection are? I worked under the impression that this might be the case, and that the reasons for the disagreement did not have to be looked for in the empirical evidence, or in the alternative interpretations of this evidence, but also in conceptual disagreements that were prior to the collection and interpretation of the evidence. So, I thought that the best way to understand the debates concerning HCE were to explore and rethink the conceptual foundations of the hypothesis, as well as the conceptual foundations of the units of selection. Samir Okasha had previously elegantly expressed this point about the resolution of scientific disputes, especially applied to the units of selection debate:

> 'Obviously, empirical data is crucial for resolving the levels-of-selection question, as for all scientific questions; but conceptual clarity is a prerequisite too. Unless we can agree on what it means for there to be selection at a given hierarchical level, on what the criteria for individuating "levels" are, on whether selection at one level can ever be "reduced" to selection at another, on how multi-level selection should be modelled, and on whether there is always "one true fact" about the level(s) at which selection is acting, then there is little prospect of empirical resolution, however much data we collect. Focusing on conceptual questions such as these is not meant to downplay the significance of empirical data, but rather to help provide the clarification needed for addressing the issues empirically.' (2006: 2)

Like Samir Okasha, I strongly agree that conceptual reflection is a prerequisite to resolving many biological questions, including of course the

question about the levels of selection and the question about the possible role of holobionts as biological individuals and as units of selection. The empirical evidence for or against HCE will always be based on one or other conception of the units of selection, on how to individuate biological entities, on how to conceive natural selection, etc. So, it is important to reflectively speculate on what are the metaphysical/ontological assumptions that ground the support for HCE: why are HCE defenders so confident that their hypothesis is true? How do they conceive units/level of selection? How do they understand 'biological individuality'? Why are the critics so unimpressed and think, as I was once told, that HCE is merely 'bad biological practice'? Only once the different conceptions are clarified, could the debate about HCE possibly be settled and the empirical evidence evaluated. The purpose of the doctoral project is precisely to do so in order to provide a general account of how to make sense of the claim that holobionts and their hologenomes are units of selection.

Structure of the doctoral thesis

The doctoral thesis is structured in five chapters. **Chapter I** introduces HCE and traces its historical development since the hypothesis was originally formulated. **Chapter II** deals with the main criticisms that have been raised against HCE, and explains them extensively, without presupposing any background knowledge in the reader apart from the information about the hypothesis provided in the previous chapter. These two chapters play a double role in the doctoral project: on the one hand, they serve as introductory chapters, in the sense of making the reader familiar with the scientific hypothesis, as well as the difficulties it must address, that will be philosophically assessed; on the other hand, they play the role of being motivational chapters for the terms in which the HCE will be discussed from a philosophical point of view. Concretely, the two introductory chapters motivate the aspect of HCE that I will engage with in the rest of the doctoral project, which is basically the claim that holobionts and their hologenomes are biological individuals and units of selection in evolution. It is important to clarify this point because, as I said, HCE can be used instrumentally to assess many philosophical questions. However, in my doctoral

project, I will exclusively and extensively analyse one concrete philosophical and biological worry that HCE creates, namely, the question about the status of biological individuals.

Based on the principle just described, **Chapter III** examines the general debate about biological individuality, and how it has been addressed in philosophy, with a special application of some of the ideas to the notions of 'holobiont' and 'hologenome' as defined in HCE. The reader might believe that this way of presenting the debate is confusing, and that it would have been more sensible to introduce first the debate about biological individuals, as this has been extensively discussed in the philosophical (and biological) literature, and later discuss HCE in the framework of biological individuality. I disagree. To start with, I think HCE can be used as evidence for many different purposes, including answering questions in epistemology, philosophy of science, ontology, etc. Thus, it is important to present, in the first place, the hypothesis and its historical development as concisely as possible, so that the reader, while motivationally guided towards the question I aim to answer in my doctoral project, can perceive all the philosophical problems that the careful study of a scientific hypothesis triggers. Second, because I suspect that starting the thesis with the discussion of biological individuality to later move to a discussion of HCE would be somehow putting the cart before the horse, and the reader would be losing the overall motivation of the doctoral project. Biological individuality is, in itself, a very interesting philosophical topic. But the reason why it is philosophically interesting needs to be argued for, i.e. it cannot be simply assumed. Starting with the presentation of HCE, and only then moving to the discussion of biological individuality fulfils, in my opinion, the motivational gap.

Chapter IV builds on the previous chapter to argue that most criticisms to HCE presuposse a nested-hierarchical view of the biological world, according to which a holobiont would qualify as a biological individual if and only if all the species that compose the holobiont (including the species that compose the microbiome) stand in mutual relations of dependency with each other with respect to the biological *process* of interest (e.g. physiological dependence, developmental dependence, evolutionary dependence, etc.). These assumes

that the hologenome equates the sum of genomes that interact. On these grounds, many philosophers and biologists have rejected the claim that holobionts are biological individuals in any significant sense, because the relations of dependency are usually non-reciprocal between the species that interact. While the host depends on its symbionts physiologically, developmentally and, defenders of HCE argue, evolutionarily, the opposite is not usually the case, or at least it is not the case for every single species member of the microbiome. Therefore, they conclude, holobionts are not biological individuals because they are 'blurry entities' (**Chapter II**). Against this approach, I argue that the requirement concerning the reciprocity of the relations of dependency between the host and the species that compose its microbiome is unjustified, as it is grounded on a questionable metaphysical assumption about the structure of the biological hierarchy. I argue that if the requirement for a nested-hierarchy is replaced by a requirement for a non-nested-hierarchy, the hypothesis that the holobiont is a biological individuality becomes reasonable. Under the latter framework, the holobiont would not be the entity composed by the host plus the species that compose its microbiome, but rather by the host and the traits that compose its microbiome. For that reason, the traits that compose the microbiome would be coevolving both with the host genome, in virtue of the relations of dependency for certain biological processes (physiology, development, etc.), and with the species of microbes where the traits are realized, in virtue of being physically bounded to the genome of the bacterial species, which suggests the necessity of introducing a multilevel selection (MLS) analysis to study their evolution.

Finally, **Chapter V** introduces an original account of the main claim of HCE, namely, the notion that holobionts are units of selection. The chapter starts with an argument against the criticism, presented in **Chapter II**, that holobionts are not units of selection because they lack proper transgenerational transmission of the species that compose microbiome. I argue that the requirement that these authors rely on is not satisfied by canonical units of selection either, and thus, by analogy, it should not be considered a serious threat to HCE, for the condition that holobiont detractors assume is simply not necessary for an entity to be a unit of selection. Later, I present a multilevel

selection (MLS) analysis to support the thesis that holobionts are units of selection. MLS analysis conventionally includes two dimensions, multilevel selection 1 (MLS1), and multilevel selection 2 (MLS2). MLS1 occurs when the particles that compose a conglomerate engage in fitness-affecting relations, so that the final distribution of traits in the global population of particles is conditional on the relationships between the entities that form the conglomerate. MLS1 does not require that the conglomerate stands in parent-offspring relationship. MLS2, on the contrary, occurs when there is parent-offspring regression, i.e. when the conglomerate forms parent-offspring lineages. In my chapter, I first introduce a MLS1 analysis, and defend that what matters to argue that holobionts are units of selection from this perspective is to find transgenerational 'stability of traits', as opposed to transgenerational stability of species. I depart from the observation that the existence of collective parent-offspring lineages is not strictly necessary from MLS1, and argue that the only criterion that is required is that the trait-distribution in the global population of particles is conditional on the fitness-affecting interactions between the host and its microbiome. I connect this observation to our current empirical evidence and show how it suggests that holobionts are units of selection from a MLS1 perspective, at least in some cases. Secondly, I present a suggestion of how holobionts could be conceived as units of selection from a MLS2 perspective. MLS2, in contrast with MLS1, requires the existence of parent-offspring lineages of holobionts. I suggest that part of the current biological evidence could also be reinterpreted in MLS2 language, as a form of 'extended inheritance'. As in the case of MLS1, I argue that the evidence that supports that holobionts are units of selection from a MLS2 perspective also depends on conceiving the holobiont as a biological object composed by the host plus its functional microbiome. My account constitutes a suggestion of how the MLS2 model can be applied to holobionts, rather than an accumulation of empirical evidence that supports its existence.

As the reader will note, many sections in the project will be named '*A brief reflection*', '*Reflection on…*', '*A discussion of…*', etc. These sections and/or subsections are aimed at guiding the reader towards the arguments and thesis that I will be defending further on. Especially, they will serve to orientate the

reader towards the thesis that I will be arguing for in the two main chapters of my doctoral project, although they will not present the complete argument: only the reasons why I find certain positions difficult to accept, or hard to argue for. Additionally, every chapter will contain a section entitled ‘Brief summary of chapter x’, where I will summarize the main points that have been discussed in the chapter. As some of the discussions, especially the most technical ones, will be hard, I think adding a summary at the end of each chapter will help to reader to keep afresh in his memory the main points that I have discussed.

A note on the chapters

Many of the chapters from this thesis are the result of the collaboration with different people, and some of them have been either drafted or already published as papers in specialized scientific journals. These include:

- J. S. Díaz (2015): ‘El mecanismo evolutivo de Margulis y los niveles de selección’. *Contrastes. Revista Internacional de Filosofía* XX (1): 7-24.
- J. Suárez (2016) ‘Bacterial species pluralism in the light of medicine and endosymbiosis’. *Theoria. An International Journal for Theory, History, and Foundations of Science* 31(1): 91-105.
- J. Suárez (2018a): ‘The importance of symbiosis in philosophy of biology: An analysis of the current debate in biological individuality and its historical roots’. *Symbiosis* 76(2): 77-96.
- J. Suárez, and V. Triviño (2019): ‘A metaphysical account of holobiont individuality. Holobionts as emergent individuals’. Quaderns de Filosofia VI(1): 59-76.
- J. Suárez (under review): ‘Stability of traits as the kind of stability that matters. Holobionts as units of selection from a multilevel selection perspective’.

- J. Suárez (under review): ‘On the individuality of holobionts and other multispecies assemblages: A critical response to Bourrat and Griffiths’
- Á. Moreno, and J. Suárez (under review): ‘Plurality of explanatory strategies in biology: Mechanisms and networks’
- V. Triviño, and J. Suárez (under review): ‘The eco-immunity account of the holobiont. A review’
- J. Suárez, and A. Stencel (under review): ‘A part-dependent account of holobiont individuality’

Furthermore, the way of articulating these chapters is closely related to my way of thinking about two topics. On the one hand, my way of thinking of the onological structure of the biological world. This problem has sourrounded my mind since the very beginning of my Bachelor’s degree, and is probably the reason why I decided to study philosophy. On the other hand, my way of approaching the topic of scientific explanation, on which I have been working for the last few years with a close collaborator from Logos, University of Barcelona. In fact, I strongly suspect that my (probably) biased intuitions in favour of the hologenome concept of evolution are a consequence of my inclination to believe that defenders of this theory appeal to processes to explain something that cannot be explained with the same degree of precision by appealing to substances. My way of thinking about the ontological structure of the biological world is clearly a consequence of my deep interest in the work of John Dupré, whom I started reading when I was still a Bachelor student, and whose pluralism and processual ontology have clerly determined my way of approaching the problems in my dissertation. My way of thinking about scientific explanation, however, has been heavily shaped by the undeniable influence of José Díez, who has systematically insisted (in a very monistic manner) throughout his career that a good scientific explanation requires three ingredients: the existence of a regularity, the derivation of that regularity from a theoretical corpus, and the fact that the theoretical corpus from which the regularity derives is, so to speak, conceptually ‘richer’ than the phenomena that the existence of the regularity explains. Following José’s intuition, my colleague Roger Deulofeu and I have analysed three different case studies—derived from my research on symbiosis and the holobiont—that became specialized papers

and which together constituted the body of his dissertation, defended at University of Barcelona and supervised by José Díez. These papers are:

- R. Deulofeu, and J. Suárez (2018): 'When mechanisms are not enough. The origin of eukaryotes and scientific explanation'. In: A. Christian, D. Hommen, N. Retzlaff, and G. Schurz (eds) *Philosophy of Science. Between the Natural Sciences, the Social Sciences, and the Humanities*. European Studies in Philosophy of Science, vol 9. Springer, Cham.
- R. Deulofeu, J. Suárez, and A. Pérez-Cervera (2019) 'Explaining the behaviour of random ecological networks. The stability of the microbiome as a case of integrative pluralism'. *Synthese*.
- J. Suárez, and R. Deulofeu (accepted) 'Equilibrium explanation as structural non-mechanistic explanations. The case of long-term bacterial persistence in human hosts'. *Teorema*.

The chapters presented here are, however, original, and everything has been rewritten to make the whole project consistent. Nonetheless, it is important to note that most of the final content that I present here has resulted from my *symbiosis* with other authors, and that some of the original theses that I will present as a result of my research are actually the result of two, three, or even more minds thinking together.

Chapter I

‘A historical perspective on the hologenome concept of evolution’

The hologenome concept of evolution was first explicitly formulated in 2008 by Ilana Zilber-Rosenberg and Eugene Rosenberg. In this chapter, I review the history of the idea, starting from the original paper and tracing its origin both backwards and forwards, to show: first, the historical origin of the concept, as well as its relation to Lynn Margulis’ ideas about the creative power of symbiosis and its influence on evolution; second, how the idea has been modified since it was first formulated in 2008.[1]

1. The origins of the hologenome concept. A generalization from coral biology

The hologenome concept of evolution (HCE, hereafter) is a biological hypothesis about the evolution of plants and animals. Briefly sketched in Rosenberg et al. (2007a), the hypothesis was firstly developed in Zilber-Rosenberg and Rosenberg (2008) (**section 3**).[2] HCE is an hypothesis about

[1] This chapter and the next will be an extension of a work I presented in J. Suárez (2018a): ‘The importance of symbiosis in philosophy of biology: An analysis of the current debate in biological individuality and its historical roots’. *Symbiosis* 76(2): 77-96. The chapters, however, presents a more extended review of both the hologenome concept of evolution and its criticisms. I will specify in footnotes which sections of the chapter are more similar to the paper, specifying if necessary the pages where I took the reference from. The two chapters also build upon J. Jeffrey Morris (2018): ‘What is the hologenome concept of evolution’ *F1000Research* 7: 1664, doi: 10.12688/f1000research.14385.1.

[2] Whether to call it ‘hologenome *theory* of evolution’ or ‘hologenome *concept* of evolution’ is mostly an idiosyncratic decision, not necessarily connected to the scientific status of the hypothesis. In the first papers where the Zilber-Rosenberg and Rosenberg presented the idea, they referred to it as ‘theory’, whereas in later works they usually refer to it as ‘concept’. Because of this, for the rest of my dissertation, I will refer to it as ‘hologenome concept of

biological individuality, according to which the multispecies assemblage that results from the symbiotic union between an animal and/or a plant plus its symbiotic microbiota constitutes a biological individual that the authors call 'holobiont' (from the Greek, *holos*, all; *biont*, life).[3] Concretely, HCE posits that the holobiont is a *unit of selection* in evolution, with its hologenome (sum of the genetic material from the animal/plant plus the genetic material of the members of the microbiome) hypothetically getting transmitted from one generation to another. [4]

The hypothesis is a generalization of the observations previously made by the authors about corals. The proponents of HCE are well-known in the field for their work in coral biology, specifically in the *Oculina patagonica*/*Vibrio shiloi* model system, used to study the well-known phenomenon of coral decolouration (coral bleaching). In the late 90s, using the Koch postulates, *V. shiloi* had been deemed responsible for the disease affecting *O. patagonica* (Kushmaro et al. 1997). However, some analyses made a few years later showed that *V. shiloi* had mysteriously disappeared from most of the corals. This evidence suggested that corals had been able to somehow overcome the infection. Most researchers considered this fact to be puzzling for two reasons: first of all, corals do not have an adaptive immune system, which means that they cannot develop antibodies to overcome any infection, except in evolutionary timescales; second, most corals live for decades, which suggests that coral populations had not have enough time for selection to act on the coral population so that they could develop the genetic changes that would be

---

evolution', or shortly 'the hypothesis' or 'HCE'. Nonetheless, the reader might find some specific quotes that refer to the hypothesis as 'hologenome *theory* of evolution', which should always be interpreted as HCE.

[3] Before proceeding, it is important to make a conceptual clarification about the meaning of the terms that I will be using. First of all, by *'microbiota'* I will refer to the 'assemblage of microorganisms present in a defined environment'. In the case of the holobiont, the environment will be provided by the animal or plant, which acts as the host in the symbiotic relationship. Secondly, *'microbiome'* will be used to denote 'the entire habitat, including the microorganisms (bacteria, Archaea, lower and higher eukaryotes, and viruses), their genomes (i.e.; genes), and the surrounding environmental conditions' in a given environment (Marchesi & Ravel 2015: 1).

[4] It seems that Richard Jefferson had previously suggested the hypothesis in a public lecture (https://www.youtube.com/watch?v=pgL3rmZL9P0), as recognized in Bordenstein and Theis (2015: 5, Box 2). Also, in Suárez (2018a: 87, ft. 20), I suggested that Jan Sapp's concept of 'symbiome' could be understood as an equivalent to the notion of hologenome (Sapp 2003, 2004). Nonetheless, as all of this is still unclear, and more research is needed to prove the conceptual connections between the ideas, as well as the possible channels of influence, I will stick to the work that starts with Ilana Zilber-Rosenberg and Eugene Rosenberg.

required to overcome the infection. To explain the recovery from *V. shiloi* infection in corals, Reshef et al. (2006) formulated 'The coral probiotic hypothesis'. The hypothesis departs from the observation that corals contain a very diverse bacterial population symbiotically interacting with it in its tissues and mucus layers, and from this it derives the possibility that rapid shifts in this bacterial population might give rise to the acquisition of pathogen resistance. In other words, the reason why corals could get rid of the infection is not explained by appeal to coral genetics, or to coral immune system. It is a consequence of the rapid changes that coral microbiota can experience (and thus the name '*probiotic* hypothesis'), so that new species infect the corals and displace previous ones, thus generating a rapid, "Lamarckian" adaptation (Rosenberg et al. 2009).[5] And from this observation, the authors conclude: 'it is now clear that corals must be considered as symbiotic organisms consisting of the coral animal, the endosymbiotic *zooxanthellae* and a metabolically active, diverse pool of prokaryotes' (Reshef et al. 2006: 2072).

The observations concering *V. shiloi* led the authors to reflect more carefully on the general role of the microbiota in coral's ontogeny and evolution, that is, on the very nature of coral's biological individuality (Rosenberg et al. 2007a). If corals behave as the coral probiotic hypothesis suggests they do, how can we conceive their individuality? The authors noticed that the coral probiotic hypothesis was premised on a controversial assumption, namely: corals are holobionts, and their phenotypes result from the complex dynamic interactions between the coral host and its symbiotic microbiota. Because of this, changes in the coral phenotype, and thus, *coral adaptations*, can arise either from changes in the coral itself (its genomic content, the genes it expresses, etc.), *or* from changes in its microbiota. If this is correct, then the coral's microbiota is partially responsible for coral's evolutionary success and thus, the authors conclude, the environment does not select between competing corals, as it had been commonly assumed, but between competing coral-holobionts. And hence the authors propose the first explicit formulation of HCE:

[5] The use of 'Lamarckian' is here opportunistic, to keep the flow that will lead to the following paragraphs. It only aims to show that the change in the adaptive trait has been produced during the ontogeny of the organism, but see Osmanovic et al. 2018, especially 'Referee report 3: Philippe Huneman', and the authors response.

> '[T]he holobiont with its hologenome should be considered as the unit of natural selection in evolution, and microbial symbionts have an important role in adaptation and evolution of higher organisms. Therefore, microorganisms are essential not only in the health and disease of individual higher organisms, but they also are *a significant factor in species survival and evolution*.' (Rosenberg et al. 2007a: 360, Box 2, emphasis added).

This original formulation was explicitly addressed to coral biologists and was soon criticized for 'disregarding the coral holobiont'. In William Leggat, Tracy Ainsworth, John Bythell, Sophie Dove, Ruth Gates, Ove Hoegh-Guldberg, Roberto Iglesias-Prieto, and David Yellowlees (2007): 'The hologenome theory disregards the coral holobiont', the authors make the following points: first, that the Rosenberg et al. have misidentified the real causative agent of coral bleaching, since coral bleaching is not caused by any pathogen (not even *V. shiloi*), but is the consequence of broad stress response in corals; second, that Rosenberg et al. have oversimplified the complex relationship that exists between the coral and its dinoflagellates endosymbionts, which according to Leggat et al. provide a substantial set of responses and adaptive mechanisms that should be taken into account before making any general claim about the holobiont being the unit of selection; finally, Leggat et al. argue that that Rosenberg et al.'s emphasis on the importance of the microorganisms for coral biology is importantly at odds with the 'widely embraced *coral holobiont model*' (2007, emphasis added).

The three points raised by Leggat et al. were immediately addressed in Rosenberg et al. (2007b), where the authors argued none of the points raised in Leggat et al.'s paper contradicted their hypothesis. Particularly, neither the first, nor the second point would raise any problem for HCE: even if *V. shiloi* were not the causative agent of coral bleaching, and the relationship between corals and their endosymbionts are more complex than what Rosenberg et al. assumed in their first paper, this would not rule out the observation that corals did not bear

the *V. shiloi* anymore, and how this happened requires an explanation. The third point is however more problematic, and it points to an important distinction that defenders of HCE have been accused of ignoring while formulating their controversial evolutionary claims (**chapter II**). By the 'coral holobiont model', Leggat et al. refer to the model for coral disease developed in Rohwer et al. (2002), and Knowlton and Rohwer (2003), where the authors introduce the term 'holobiont' to refer to the ecological community composed by the coral, its endosymbiont *Zooxanthella*, and its microbiota, including not only bacteria but also fungi, algae, or whatever element that might be discovered (**Figure 1**). In sharp contrast with the hologenome model, the holobiont model assumes that the coral holobiont is *an ecological community*, whose components (host, *Zooxanthella* and species of the microbiota), even when they might contribute in different ways to coral fitness (e.g. helping with nitrogen fixation, as antibiotics against some pathogens, etc.), can easily move from one holobiont to another, and thus their evolutionary histories are not necessarily connected to the evolutionary history of the host, nor to the evolutionary history of the other components of the holobiont (Hester et al. 2015).

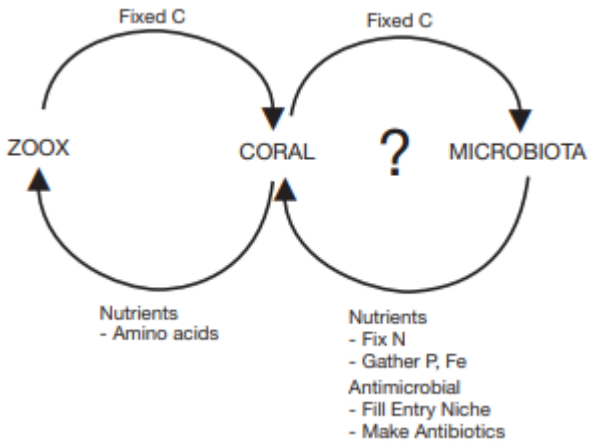


Figure 1. Rohwer et al.'s schematic representation of the holobiont. In the picture, 'ZOOX' stands for the *Zooxanthella*, coral's endosymbiont whose lost is assumed to cause coral bleaching. The question mark indicates that the exact relationship between the microbiota and the coral, as well as its function, are still unknown, although some possible roles are hypothesized. One of the points that Rohwer et al. emphasized clearly in their model is that the holobiont does not only include bacteria, but also fungi, algae or whatever unknown component that happens to reside in any of the tissues of mucus layers of corals. (From Rohwer et al. 2002: 8, Fig. 5).

The distinction between the 'holobiont model' and the 'hologenome model' raised in the dispute between Rosenberg et al. and Leggat et al. connects directly with the distinction between *holobionts* and *hologenomes*, as well as with the historical origins of the concept of 'holobiont'.[6] In the next section, I will trace back the historical origins of the holobiont concept and explain where it might have come from when it was first used by Rosenberg and Zilber-Rosenberg.[7]

2. The holobiont in its historical context. The importance of Lynn Margulis[8]

Lynn Margulis (born Alexander) was one of the most important researchers in the study of symbiosis, to which she dedicated about 50 years of her life. She is especially known for having reinvigorated the hypothesis of the symbiotic origin of eukaryotic cells, as well as for her enthusiasm about the importance of symbiosis for the maintenance of life on Earth and for its important evolutionary consequences (Margulis 1990, 1991, 1998, 2010; Sagan & Margulis 2002; Díaz 2015; O'Malley 2017; Suárez 2018a). Margulis is acknowledged as the first person to have introduced the term 'holobiont', which originally appeared in her (1990). In this work, she compares *cyclical hereditary symbiosis* with *meiotic sex* (**Figure 2**). Margulis argues that in both cases there are two entities which

[6] 'Term' and 'concept' need to be carefully distinguished at this point, since they are not coextensional. 'Term' refers to the word or expression that is used to designate one (or more) concept(s), whereas 'concept' is used to refer to a concrete idea (understood as its extension/denotation and intension/connotation) expressed by one (or more) term(s). One example can easily illustrate the distinction: the term 'bank' can be used to express the concepts of 'financial institution' or the 'slope' in a mountain. Here we have a case of one term used to refer to two concepts. On the other hand, the concept of a 'large container of hot liquids with a handle' can generally be designated by the terms 'mug' and/or 'cup'. This would be a case of one concept that can be expressed by two terms (Valdés-Villanueva 2005; García-Suárez 2011; Margolis and Laurence 2014). The distinction is important because, as I will show, the term 'holobiont' has been used differently by different authors, thus expressing different concepts. This has generated some confusion which I would like to avoid as much as possible here.

[7] From now onwards, Ilana Zilber-Rosenberg and Eugene Rosenberg will always be credited for the first original formulation and the development of HCE since, as the authors have explicitly recognized (Lamm 2018; Rosenberg and Zilber-Rosenberg, personal communication), the hypothesis has been a product of their joint work.

[8] This section is entirely based on my paper J. Suárez (2018a): 'The importance of symbiosis in philosophy of biology: An analysis of the current debate in biological individuality and its historical roots'. *Symbiosis* 76(2): Part II, sect. 2.1, pp. 86-87.

recognize each other to merge together and restart the cycle in every generation. Moreover, she hypothesizes the existence of mechanisms of mutual recognition and association which guarantee the integration of the two entities in both cases (cyclical symbiosis and meiotic sex) and, also, their subsequent dissociation, resulting in the formation of new individuals in every generation. If the entity is formed by the fusion of two haploid gametes, it will constitute a 'zygote', whereas if the entity results from the merger of two symbionts it will be a 'holobiont'. According to Margulis, both the zygote and the holobiont are new individuals and she speculates that, given the adequate recurring environmental pressures, both associations will be expected to be maintained by the selective pressure acting on the bionts/haploids (1990: 676, Fig. 3). Margulis does not, however, specify which 'bionts' should be regarded as part of the holobiont, nor does she explicitly define the term in the paper. Nonetheless, since her analogy is between cyclical symbiosis and meiotic sex, it has reasonably been assumed that she was thinking of cases of *hereditary* symbiosis (e.g. the eukaryotic cell) (O'Malley 2017).

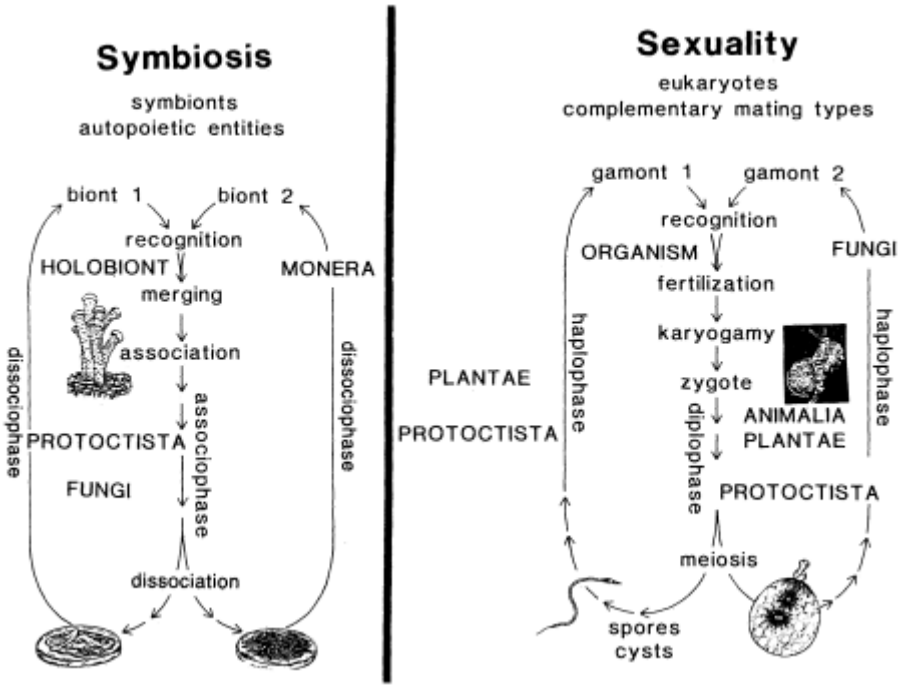


Figure 2. In this figure, Lynn Margulis compares the process of sexual reproduction with the case of *cyclical* symbiosis. She argues that, like during the process of karyogamy two haploid gametes fuse to form a diploid zygote, the process of symbiotic association can be conceived of as a process in which two previously independent bionts recognize each other and their bodies merge to establish and maintain a new individual. At a later point in time, she argues, the bionts

will dissociate in a structurally similar way than the way in which the process of meiosis gives rise to new gametes, so that the cycle can repeat itself in the next generation. (From Margulis 1990: 676, Fig. 3)

One year later, in her (1991), Margulis defines the holobiont as a 'symbiont composed of recognizable bionts', and she defines symbiosis as the physical contact between organisms of different species occurring 'throughout a significant proportion of the life history' (1991: 2, Table 1). Once more, Margulis does not explicitly spell out which bionts should be taken as constituents of the holobiont. However, she presents a definition of 'life history' as the set of 'events throughout the development of an individual organism correlating environment with changes in external morphology, formation of propagules, and other observable aspects' (1991: 2, Table 1). If one follows this definition strictly, it might be argued that the holobiont would necessarily encompass all the bionts that share their lifetime together, *irrespective of whether they are inherited or not*. However, one important difficulty must be noted here: when the definition applies to a host such as the eukaryotic cell, and a symbiont such as the mitochondrion or the chloroplast, it might be reasonable to assume that both partners share their life histories, since they are expected to have similar lifespans. But the situation is completely different if the host is an animal and the symbiont is a bacterium (e.g. aphid-*Buchnera aphidicola* association), since their lifespans are completely discordant, at least at the 'biont' or individual level. This does not mean that different *tokens* of the same *type* of bacterial species cannot interact with a host throughout the host's entire lifespan. But one must be cautious, since it seems that the entities that are being compared, if Margulis' conception of the holobiont is expected to include animal-microbe associations, are not at the same temporal and, thus, biological (in the sense in which *tempo* might affect biological properties) scale (Cáceres-Vázquez and Saborido 2018; Osmanovic et al. 2018).[9]

In any case, leaving the problem about discordant lifespans aside, and trying to be as charitable as possible with Margulis' use of the term 'holobiont', it

[9] I will come back to this problem in **chapter IV**, as I use precisely this asymmetry to criticize the hierarchical-nested view of the biological world.

seems that the conception of the holobiont that she is putting forward with her requirement of interaction during the whole life history of the bionts might be at least in part incoherent with the concept she had put forward in her previous (1990), where she seemed to be suggesting that the holobiont should exclusively include the cases of hereditary symbiosis. This is so because it might occur that symbionts that share their life history for one generation are not inherited in the next generation, if the symbionts have independent reproductive regimes. And, what is more important, even in case where the two bionts interact again one generation after another, it might happen that they do so because of a shared environment, instead of the existence of an inheritance mechanism, inheritance being understood here as 'vertical transmission' (e.g. squid-*Vibrio fischerii* symbiosis, where the bacteria are acquired from the environment every new generation).[10]

Nonetheless, despite the difficulties of interpreting Margulis' precise position about the definition of 'holobiont', it seems reasonable to consider that her initial aim was to conceive the holobiont as the entity composed by hereditary bionts. This conception becomes even clearer if one analyses the overall purpose of her 'Symbiogenesis and symbioticism', namely, to vindicate the notion of symbiogenesis as a way in which new species, kingdoms and taxa could evolve. For instance, she explicitly says that 'the highest-level taxa (…) have evolved by acquisitions of symbionts that have become *hereditary*' (1991: 11, emphasis added). I will call this conception the 'hereditary holobiont'. The 'hereditary holobiont' is also coherent with some further claims she made in her later writings (Margulis and Fester 1991; Margulis 1998, 2010; Margulis & Sagan 2001, Margulis and Sagan 2002; and also see O'Malley 2017). For instance, in one of her latest papers, where she justifies the historical role of Kozo-Polyansky in introducing the idea of symbiogenesis to biology, she argues for the necessity of genetically distinct bionts *reproducing together* for

[10] There is still another logical possibility, which seems to be however biologically impossible. It might well happen that hereditary symbionts do not share their life history with each other. For example, imagine a pair association in which one of the members—the host—is able to 'kill' its symbiont while retaining some signal that would permit its reconstitution in the precise moment of reproduction. If that case were biologically possible, it would then be a case of inherited symbionts without shared life histories.

symbiogenesis to occur. Analysing the association between eels and a specific species of shrimp (cleaning symbiosis), she argues:

> 'It is symbiosis, but not symbiogenesis. Both partners *grow and reproduce separately*. Both shrimp and eel can live separately. One sees no obvious novelty generated by this symbiosis; i.e., symbiotic physical association. The relationship between the shrimp and the eel is still a behavioral one' (2010: 1528, emphasis added)

In this vein, one might argue that, as the term 'holobiont' was introduced in comparison to meiotic reproduction, and Margulis discusses it while reflecting the importance of symbiogenesis as an evolutionary mechanism (and evolution requires inheritance to lead to adaptation), the holobiont is thus the biological individual that includes all those symbionts that are inherited together (organelles in eukaryotes, obligatory endosymbionts in insects, etc.), i.e. Margulis conceives the holobiont as the 'hereditary holobiont' (O'Malley 2017: 36, for a defence of this interpretation).

The interpretation of Margulis' understanding of holobionts as 'hereditary holobionts' is not without contestation, though. In the same volume where Margulis published her paper, John Maynard Smith proposes 'a Darwinian view of symbiosis', which he considers opposed to Margulis' conception (Maynard-Smith 1991). In his paper, Maynard Smith relates the debate about symbiosis with the well-studied problem of the units of selection[11], and embeds his discussion of symbiosis in the framework of the theory of evolutionary transitions in individuality, which he was starting to develop by that time. According to Maynard-Smith, symbiosis can be understood as an evolutionary

[11] Maynard-Smith does not use 'units of selection', but 'units of evolution', where a unit of selection is whatever entity exhibits phenotypic variation that led to multiplication of the entity within the population (thus being selected for or against), and a unit of evolution is a unit of selection that, furthermore, exhibits heredity (Maynard-Smith 1987). In contrast with Maynard-Smith, I will use 'unit of selection' as it is conventionally used, i.e. requiring heredity, variance and fitness/multiplication, and thus meaning what Maynard-Smith means by 'unit of evolution' (see Okahsa 2006; Godfrey-Smith 2009; Lloyd 2017a, 2017c: 293–297; Gontier 2010, for an analysis of the concept of 'unit of selection'). The idea of 'unit of selection' will be discussed in **chapter III**, in the context of the discussion of different ideas of biological individuality.

mechanism and interpreted in a Darwinian fashion—i.e. with the entities that interact symbiotically forming a higher-level unit of selection—*only if* the symbionts are transmitted directly during the reproduction of the host, and cannot be transmitted differently. In other words, the reproductive regimes of the symbionts must be mutually dependent (Hurst 2017, reviewed in **chapter II**, makes exactly the same point).

Maynard Smith supports his position by adducing that '[w]ith direct transmission the genes of the symbionts will leave descendants only to the extent that the host survives and reproduces' (1991: 35). The condition of 'common reproduction' is taken for Maynard-Smith as a precondition for the two symbionts to have their fitness interests aligned, which, according to standard evolutionary theory, is the only way in which two entities could be considered to evolve as a single unit. Therefore, if the reproductive regimes of the symbionts are concordant, then it will be expected that they will tend to maintain a mutualistic relation that, eventually, might make it 'reasonable to consider the association as a single unit' (1991: 38). However, Maynard Smith's argument follows, in cases of indirect transmission, this possibility is much less likely, and thus he suggests that the interacting entities should be considered as independent units (of selection).

Maynard Smith's paper is relevant in this context because he seems to be discussing Margulis' liberal views about the strength of symbiosis at generating evolutionary novelty that could be 'filtered' by natural selection. For him, the cases where symbiosis can be considered to have a substantial evolutionary impact, in the sense of affecting the 'input' of natural selection—i.e. the unit of selection, or the entity on which natural selection directly acts upon—are very limited, and probably precluded exclusively to very specific cases such as cellular organelles, as he suggests at the end of his paper. If Maynard Smith's interpretation of Margulis' views about the creative power of symbiosis were correct, then Margulis' notion of the holobiont might be interpreted not as constrained exclusively to cases of hereditary symbiosis (i.e. the 'hereditary holobiont'), such as the eukaryotic cell, but also as including symbiosis associations composed of many different bionts (which I will call the 'ecological

holobiont'). In fact, this latest view is endorsed in Guerrero et al. (2013), published two years after Margulis' death. In that paper, holobionts, considered as autopoietic (self-sustaining) units, are defined as 'integrated biont organisms, i.e., animals or plants, with all of their associated microbiota' (2013: 133, emphasis added). In that same paper, they also coined the term 'holobiome', referring to 'the assembly of genetic information contributed by the animal or plant and its associated microbiota' (2013: 134), and demanding a new look at evolution that would take into account the importance of the host genome plus the genome of its microbiota. They argued the holobiome to be a new biological entity, whose basic interacting elements would give rise to new species and, in general, new biological variety by means of the resources of both the animal or the plant, and its microbiota. At some point in the paper, the authors even endorse the theses that: (1) holobionts are subjected to natural selection; and (2) holobiomes are entities that have been selected due to their selective advantages. Even if the authors do not mention the concept 'units of selection', their paper might be interpreted as implicitly endorsing HCE, thus considering the holobiont, with its hologenome (which they refer to as 'holobiome'), a possible unit of selection in evolution.

Whether Margulis' concept of the holobiont must be interpreted as hereditary, i.e. the 'hereditary holobiont', or ecologically, i.e. the 'ecological holobiont' is not exactly relevant for the ideas I aim to present in this chapter.[12] Also, whether she considered the holobiont as a unit of selection, or she did not, is not relevant either. What matters to my purposes here is to show that Margulis was the first to propose the concept of 'holobiont' to refer to a certain kind of symbiotic assemblages with some evolutionary importance, and how the meaning she wanted to assign to the concept is ambiguous (Margulis 1998; Margulis and Sagan 2002). Importantly, this ambiguity is still today at the core of the hologenome debate.

[12] Opposing interpretations of Margulis' use of the concept of 'holobiont' might be found in O'Malley (2017) and Suárez (2018a).

2.1. Margulis' hereditary holobiont and Rohwer et al.'s holobiont model

Margulis' use of the the term holobiont might be interpreted as hereditary, at least in her first writings in the early 90s (the 'hereditary holobiont') or ecologically (the 'ecological holobiont'), as in her latest work with Guerrero and Berlanga. Setting this aside, and assuming that Margulis was using the concept of 'hereditary holobiont', what is clear is its sharp contrast with the concept of holobiont that appears in Rohwer et al.'s coral's holobiont model (Rohwer et al. 2002; Knowlton and Rohwer 2003). As argued in **section 1**, Rohwer et al.'s holobiont is the 'ecological holobiont', insofar as it refers to the whole microbial community that a given coral interacts with during its lifespan, including its endosymbiont *Zooxantella*. This established, it becomes now important to ask the following two questions: Firstly, how does the 'hereditary holobiont' differ from the 'ecological holobiont'? Secondly, why is this relevant for our purposes and, specifically, for the formulation and development of HCE?

The answer to the two questions just formulated will be given during the rest of the chapter. In the next section, I will start reviewing Zilber-Rosenberg and Rosenberg's reinvigoration of the term 'holobiont' with their formulation of the HCE, and I will try to disentangle the precise nature of the holobiont concept that their hypothesis assumes. My main argument will be that, by linking the concept of the holobiont to the concept of the hologenome, Zilber-Rosenberg and Rosenberg will be mixing Margulis' 'hereditary holobiont' with Rohwer et al.'s 'ecological holobiont'.[13]

[13] This footnote constitutes a call of caution for the reader. I have just advanced that, in my view, Zilber-Rosenberg and Rosenberg are mixing two concepts of the holobiont that are not extensionally equivalent. This should not be taken as a criticism of HCE. In fact, my whole thesis project, as announced, is a project to defend the main claim made by HCE defenders, namely: that holobionts are units of selection. Presenting the hypothesis in the way I am presenting it constitutes, in my opinion, the best way of trying to conceptualize which are the elements of the theory that are valuable and which are the elements that should be redefined to be correctly articulated.

3. Formulating the hologenome concept of evolution

3.1. The original formulation by Zilber-Rosenberg and Rosenberg

As Rosenberg et al. (2007b) explicitly acknowledge, they believe that Rohwer et al.'s holobiont model 'not only does not contradict the hologenome theory but provided some of the information on which the theory was developed.' This quote is particularly illuminating here, since it clearly shows how HCE aims to connect two apparently incoherent ideas: the 'hereditary holobiont', that might be argued to be Margulis' heritage, with the 'ecological holobiont' that is present in Rohwer et al.'s work. That connection between the two ideas is as well clear in the references that the authors mention in their very first explicit elaboration of the idea:

> 'In the hologenome theory of evolution, we suggest that the holobiont (Margulis, 1993; Rohwer et al., 2002) (the host and its symbiotic microbiota) with its hologenome, acting in consortium, should be considered a unit of selection in evolution, and that relatively rapid variation in the diverse microbial symbionts can have an important role in the adaptation and evolution of the holobiont.' (Zilber-Rosenberg and Rosenberg 2008)

Notice that, as HCE is formulated in this paragraph, it postulates, as its main claim, that the holobiont, with its hologenome, is a *unit of selection in evolution*. One obvious is question concerns what is the evidence that would be required to justify that type of claim. What are the conditions that holobionts and hologenomes would be required to satisfy to be considered units of selection in evolution, as Zilber-Rosenberg and Rosenberg claim? According to a very widely accepted conception of 'unit of selection', an entity in a population is a unit of selection if and only if it exhibits hereditary (or transmissible) phenotypic variation in fitness (Lewontin 1970; Okasha 2006; Godfrey-Smith 2009; Lloyd 2017c).[14] Taking this as their departure point, Zilber-Rosenberg and Rosenberg

[14] The debate about the units of selection is much more complex than what the assumption I will make here suggests, as I will show in **chapter III**. For the moment, it is enough to accept this vague definition as the driver of the discussions about HCE. The reader must take into account, though, that this vague assumption about the metaphysics of the units of selection will reappear

argue that their hypothesis is grounded on the following four observations[15]: first, all animals and plants interact symbiotically with a large number of microorganisms; second, the symbiotic microorganisms with which hosts interact can be intergenerationally transmitted with fidelity; third, that the interactions between the host and its microorganisms affect the fitness of the holobiont, given its environment; fourth, genetic variation in holobionts can be produced by incorporating new symbiotic microorganisms within their microbiota. Since microorganisms can respond more rapidly and using multiple mechanisms (including horizontal gene transfer) to changing environmental conditions, they offer an entire new set of possibilities for holobionts to adapt to dramatic environmental shifts, which are not limited exclusively to the genetic changes in the host's genome.

3.1.1. Evidence of microbiome transmission and microbiome-induced fitness effects

Zilber-Rosenberg and Rosenberg acknowledge that HCE requires not only the existence of interactions between the host and the symbionts (first piece of evidence in support of HCE), but that these interactions: (1) are reliably transmitted transgenerationally; (2) affect the fitness of the holobiont. If there are host-symbiont interactions but they do not have any (or they only have one) of the two shortlisted properties, then the holobiont cannot be considered a unit of selection, but a conglomerate of independent units of selection interacting ecologically with each other (and thus, HCE would be plainly false). In their original paper, Zilber-Rosenberg and Rosenberg present evidence to support the claim that the symbionts that compose a holobiont are intergenerationally transmitted with sufficient fidelity to support the claim that holobionts are units of selection (**Figure 3**). It is important to note, although in passing, that Zilber-Rosenberg and Rosenberg argue that what needs to be intergenerationally transmitted, if the holobiont can be considered a unit of selection, are the different *genomes* that constitute the holobiont. In their words:

---

in **chapter II, section 3**, when I analyse some of the criticisms that have been raised against the hypothesis.

[15] These observations, as well as the evidence that supports them, have slightly evolved with time, especially the third one, which has been reinterpreted differently.

‘The hologenome theory [concept] of evolution relies on ensuring the continuity of partnerships between holobiont generations. Accordingly, both host and symbiont *genomes* must be transmitted with accuracy from one generation to the next.’ (2008: 726, emphasis added)

Table 2. Modes of transmission of symbionts and their contribution to the fitness of the holobiont

Holobiont: Microbiota	Mode of transmission of microorganisms	Microbial contribution	References
General			
• All eukaryotes: Mitochondria	Cytoplasmic inheritance	Respiration	Margulis (1993)
• Plants: Chloroplasts	Cytoplasmic inheritance	Photosynthesis	Margulis (1993)
Invertebrates			
• Aphids: *Buchnera* sp. (primary-endosymbiont)	Via intracellular bacteria in bacteriocytes; present in ova	Provision of specific required amino acids lacking in the plant sap diet	Baumann *et al.* (1995), Wernegreen (2002), Perez-Brocal (2006)
• Aphids: Secondary endosymbionts	Via intracellular bacteria in addition to environment	Growth at high temperature; resistance to parasites	Sandström *et al.* (2001), Russell *et al.* (2003)
• Termite: Microbiota in hind gut	Feces of adult termites fed to newly hatched juveniles	Utilizable energy and carbon; nitrogen metabolism; recognition signal from odor of bacterial metabolites	Abe *et al.* (2000), Minkley *et al.* (2006)
• Anthropods/nematodes: *Wolbachia* spp.	Intracellular transmission via egg cytoplasm	Fertility and sex determination	Veneti *et al.* (2005)
• Stinkbug midgut: *Burkholderia*	Specific transmission via environment	More efficient food utilization	Kikuchi *et al.* (2007)
• Squid nidamental gland: Microbiota	Via cover of eggs originating from the gland	Protection of eggs and embryos against pathogens	Kaufman *et al.* (1998), Barbieri *et al.* (2001)
• Squid light organ: *Vibrio fischeri*	Environmental from surrounding water	Camouflage against predators	McFall-Ngai (1999)
• Corals: Microbiota	From the environment and by vegetative reproduction	Photosynthesis (intracellular algae); nitrogen fixation; protection against pathogens	Rohwer *et al.* (2002), Buddemeier *et al.* (2004), Rosenberg *et al.* (2007)
• Sponges: Microbiota	Environmental in addition to possible transmission from parent	Breakdown of complex polymers; nitrogen cycling; protection against pathogens	Webster *et al.* (2001), Hickman (2005), Taylor *et al.* (2007)
Vertebrates			
• Cow rumen: Microbiota	Physical contact with parents and via food contaminated with feces and sputum	Provision of all nutritional needs from cellulose	Dehority (2003), Russell & Rychlik (2001)
• Whale forestomach: Microbiota	Physical contact with mother	Provision of nutritional needs from chitin and other complex organics	Herwig *et al.* (1984), Olsen *et al.* (1994), Olsen *et al.* (2000)
• Human gut and mouse model: Microbiota	Via physical contact and from environment	Protection against pathogens; stimulation of immune system; angiogenesis; vitamin synthesis; fiber breakdown; fat storage	Hooper *et al.* (2002), O'Hara & Shanahan (2006), Ley *et al.* (2006a), Xu *et al.* (2007)
Plants			
• Land plants: Mycorrhiza fungi	Via seeds on ground and by vegetative reproduction	Supply of minerals from soil	Wilkinson (2001), Wang & Qui (2006)
• Nonphotosynthetic plants–fungi (some orchids)	Seeds falling on ground and by vegetative reproduction	Supply of minerals from soil and organics from other plants	Bidartondo (2005)
• Legume plants: *Rhizobium*	Environmental from surrounding	Nitrogen fixation	Stougaard (2000), Jones *et al.* (2007)
• Plant: Growth-promoting rhizobacteria	Environmental from surrounding soil	Protection against pathogens; nitrogen metabolism; acceleration of mineralization; carbon cycling; salt tolerance	Smith *et al.* (1999), Somers *et al.* (2004), Singh *et al.* (2004), Egamberdieva *et al.* (2008)
• Rice plants: *Azoarcus* sp.	From surrounding soil	Associative nitrogen fixation	Hurek & Reinhold-Hurek (2003)

Figure 3. Table presenting the original evidence of (1) rate of transmission of a host’s microbiota and (2) phenotypic effects of the microbiota in the host’s phenotype, as introduced by Zilber-Rosenberg and Rosenberg. (From Zilber-Rosenberg and Rosenberg 2008: 727, Table 2).

This is important, since this is a very specific requirement about: first, the type of inheritance that is required for units of selection; second, the nature of the

holobiont, which is here conceived as a purely genomic conglomerate of interacting species.[16]

Zilber-Rosenberg and Rosenberg do not consider the transmission of the host genome problematic, since it relies on the well-studied rules of Mendelian inheritance. What is more problematic, though, is the transmission of the genomes of microbes that constitute the host's microbiome. What are the mechanisms that ensure that the same microbes will reappear transgenerationally, to ensure the inheritance of the hologenome? Drawing upon previous work by McFall-Ngai (2002), Zilber-Rosenberg and Rosenberg distinguish direct from indirect modes of transmission of the microbes within the holobiont. The most direct case of transmission, they argue, can be found in the organelles (mitochondria, chloroplasts) of the eukaryotic cells, which are transgenerationally transmitted by cytoplasmic inheritance. A less direct, but not indirect mode of transmission occurs when the symbionts are transmitted with the reproductive cells of the host, as happens in the aphid-*Buchnera* symbiosis. Another still less direct mode of transmission occurs when the direct contact between the host and its offspring induces the passage of the microbes of the microbiota of the parents. This mode of transmission has been observed in humans, through the birth canal or through breast-feeding. The case of breast-feeding is similar to the cases of parent-offspring coprophagy, like the one observed in termites, where adult workers feed juveniles with feces. This mechanism is thought to be used to guarantee the acquisition of the hindgut microbiota by juveniles. A less direct mode of transmission occurs when the microbiome needs to be acquired horizontally, from the environment. Some cases of horizontal transmission, despite being almost indirect, can be very precise after all. One example is the *Vibrio fischeri* that bobtail squids acquire every generation to develop their light organ. Even if the transmission is not vertical, meaning that the squids need to acquire their bacteria from the environment every generation, the acquisition is very precise, and it seems that bobtail squids have developed barriers to prevent colonization from *V. fischeri* that do not emit light, suggesting that horizontal transmission does not

[16] It is important to point this here because the notion of the holobiont that I will present later will be different, and so will be the conception of the units of selection that I will develop, especially in **chapter V**, where I introduce my model of the units of selection.

necessarily prevent a faithful transmission of the microbiome that guarantees a faithful transgenerational reconstitution of the holobiont.

After they review all the possible mechanisms of hologenome transmission, Zilber-Rosenberg and Rosenberg end their section with the following highly illuminating quote:

> ‘The large varieties in modes of transmission have an interesting implication: individuals can acquire and transfer symbionts throughout their lives, and not just during their reproductive phase. This means that the parents, grandparents, *nannies*, siblings, *spouses or any organism that is in close contact with an offspring* can transfer symbionts and thereby influence the holobiont of the next generation.’ (2008: 728, emphasis added)

The highlighted part in their quote reflects an important assumption in Zilber-Rosenberg and Rosenberg’s first definition of HCE. According to the authors, the transgenerational transmission of the holobiont is a consequence of the transgenerational transmission of the host and microbial genomes (genomic view), and this transmission does not necessarily occur during the host’s reproductive phase, but it can occur during *its whole lifespan*, and be a consequence of the interaction with any member of the population (nannies, spouses, etc.), not necessarily parents. This assumption will be crucial for some of the criticisms to HCE that I will review in **chapter II**.

Part of the information included in **Figure 3** referred to the fitness influences of the symbionts in the holobiont. I will not extend this here, since I will talk more extensively about it again in **section 4** (see McFall-Ngai et al. 2013). For the moment, it is enough to say that the symbionts can have different phenotypic effects on the holobiont, including not only metabolic effects (respiration, digestion, etc.), but also behavioural effects (e.g. determining mate preference), immunological effects (e.g. protection against pathogens),

morphological effects (e.g. organ development), etc. Some of these phenotypic effects will influence the fitness of the holobiont, thus potentially generating fitter and less fit holobionts in competition with each other, as natural selection requires. Thus, according to Zilber-Rosenberg and Rosenberg, this evidence suggests that the fitness of the holobionts is comparable, which is a requirement for an entity to be a unit of selection.

3.1.2. Modes of variation among holobionts

According to Zilber-Rosenberg and Rosenberg, if evolution were exclusively based on the properties of the host's genome, evolution among animals and plants would be very slow since (a) plants and animals have relatively long generation times (i.e. the time it goes since the animal is conceived until it reproduces), (b) the only changes that get transferred to the offspring are those that affect the germ line, and (c) the number of genes required to introduce a phenotypic change is usually relatively high. This would entail the process of animal and plant evolution to be very slow. By positing that the host with its microbiome evolves as a single unit, HCE recognizes new possibilities for holobiont variation which go beyond the genetic changes within a single host (mutation, recombination, chromosome rearrangement). First of all, it also recognizes the possibility that changes within the genomes of the microbial members of the microbiome (transduction, conjugation, horizontal gene transfer) could also influence holobiont evolution. In fact, since *holobiont = host + microbiome*, holobiont variation might arise from genetic changes in any of them. Secondly, HCE also recognizes three modes of acquisition of variation that are exclusive to holobionts. These include: (a) microbial amplification and microbial reduction, i.e. changes in the relative numbers of a microbial species within the microbiota; (b) microbial acquisition, i.e. the possibility that the holobiont acquires new microbial species from its environment; (c) horizontal gene transfer between the microbes in the microbiota, or between the host and some of its microbes (**Figure 4**).

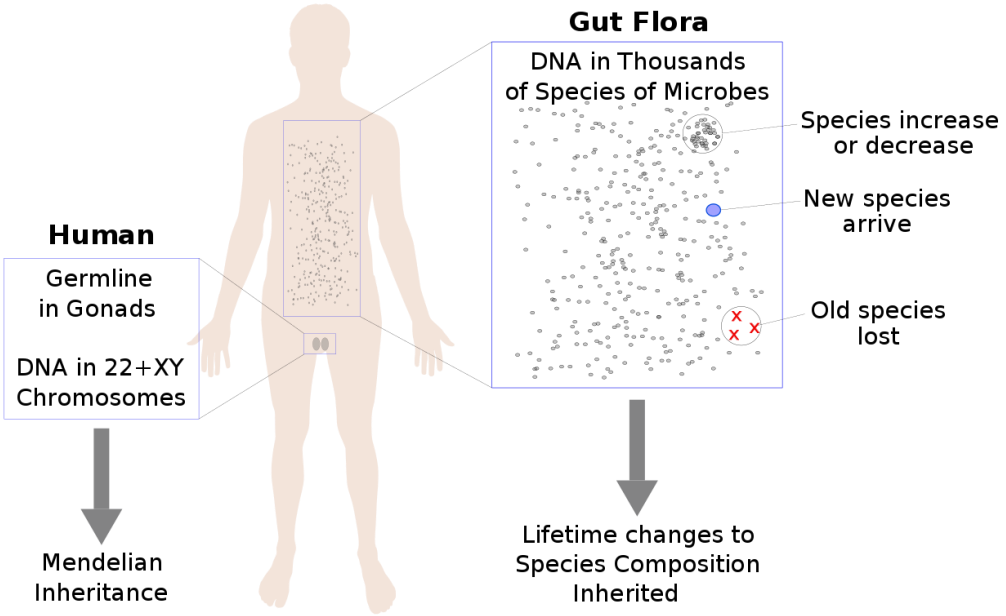


Figure 4. Modes of variation in the holobiont, and modes of transmission. The holobiont might vary as a consequence of changes in its germline cells, or as a consequence of changes in its microbiome. Variation in the microbiome includes the increase or decrease in the members of one species (microbial amplification), the arrival of new species or the loss of some species (microbial acquisition), and horizontal gene transfer among the species of the microbiome (not represented). (From Wikimedia Commons: https://commons.wikimedia.org/wiki/File:Neo-Lamarckian_inheritance_of_hologenome.svg#file)

Microbial amplification refers to the possibility that the numbers of a given population of microorganisms within the microbiome increase (or decrease) as a response to environmental changes, such as the availability of nutrients, the temperatures (including the effects of climate change), the treatment with antibiotics, etc. Zilber-Rosenberg and Rosenberg conceive the process of microbial amplification by analogy to the process of gene amplification.[17] They believe that the process of microbial amplification is well-documented and offers a wide spectrum of evolutionary capabilities to the holobiont. Importantly, the authors recognize and accept that microbial amplification might occur as a rapid response to changing environmental conditions, and thus during the ontogeny of an organism. They believe this to be a consequence of the holobiont being a

[17] 'Gene amplification' refers to a process by which the number of copies of a given gene within the genome get increased (*amplified*) without the same proportional effect occurring to other genes. For a definition check: https://www.nature.com/subjects/gene-amplification.

highly dynamic entity, whose components might shift their numbers, increasing, decreasing, or disappearing during the holobiont's lifespan.

The second mechanism is *microbial acquisition*. As animals and plants interact during their lifespan with a wide number of microorganisms from their environment, it will be expected that, eventually, some of these microorganisms will 'break the host's barriers' and, if they eventually happen to find it a suitable niche, they can incorporate to its microbiota. Zilber-Rosenberg and Rosenberg argue that the new symbiont might introduce a whole set of new genes into the holobiont which, if the environmental circumstances are appropriate, might affect the holobiont's phenotype.[18]

Finally, *horizontal gene transfer*, also called lateral gene transfer, refers to the mechanism by which bacteria can exchange genetic material with each other through plasmids, transposons, bacteriophages, etc. They believe that the microbial density in a holobiont's microbiome might accelerate the process of horizontal gene transfer, thus resulting in process of rapid evolution within the holobiont that would have not been expected 'outside' the holobiont.

3.1.3. Holobionts as interactors, hologenomes as replicators

According to Zilber-Rosenberg and Rosenberg, at least in their seminal paper on HCE, their hypothesis can be interpreted under the interactor/replicator framework of the units of selection. Following the famous distinction by David Hull (1980), an *interactor* is an entity that cohesively interacts with its environment, so that reproduction is differential, and the *replicator* is the entity whose variations are transgenerationally transmitted in a cohesive way. Replicators, but not interactors, are characterized by their 'longevity' and 'copying fidelity'; whereas interactors, but not replicators, are characterized by their high degree of 'cohesiveness', however cohesiveness is interpreted. Zilber-Rosenberg and Rosenberg interpret the interactor/replicator distinction as

[18] Zilber-Rosenberg and Rosenberg's assumption here is that, as natural selection acts on variation on phenotypes, if the new symbiont acquired affects the phenotype of the holobiont, then natural selection can act on it, as a 'unit'. This has to be taken into account at this point, since the mechanisms of microbial acquisition will be used as one of the main criticisms to HCE:

the phenotype/genotype distinction and argue that HCE entails that the holobiont is an interactor, and its hologenome is the replicator. In other words, the holobiont realizes a phenotype that can be selected for or against, and the hologenome gets faithfully replicated transgenerationally (when the holobiont is selected for), including its variants.[19]

3.1.4. Mixing the 'ecological holobiont' and the 'hereditary holobiont'

At this point, it should be clear how Zilber-Rosenberg and Rosenberg have fused the concept of 'hereditary holobiont' with the concept of 'ecological holobiont' into a new concept, represented by the claim that holobionts are units of selection, i.e. HCE.[20] On the one hand, Zilber-Rosenberg and Rosenberg argue that the hologenome is a replicator that can be faithfully transmitted intergenerationally, thus guaranteeing the sufficient degree of fidelity for natural selection to act on it as a single unit. On the other hand, they aim to include phenomena such as microbial amplification among the possible evolutionary mechanisms by which a holobiont can acquire phenotypic variation. However, the two phenomena are not on the same timescale. Ecological processes occur during the lifespan of the organism, whereas evolutionary processes extend much further, as they depend on the appearance of lineages, i.e. on the establishment of parent-offspring generations. This raises a series of biological and philosophical questions that will be raised (and partially answered) along this thesis. What are the conditions under which an 'ecological holobiont' can be considered an 'hereditary holobiont', and vice versa? Is it possible, under any possible formulation of natural selection, to consider the 'ecological holobiont' as a unit of selection? If so, what will be its properties? Do 'ecological holobionts' form parent-offspring lineages? How can the 'ecological holobiont' be individuated, and how might its individuation criteria differ from the individuation criteria of the 'hereditary holobiont'? Are the 'ecological holobiont'

[19] As I already said, many of these notions will change when HCE develops further and more evidence is gathered. For the moment, it is interesting to see what Zilber-Rosenberg and Rosenberg claimed in their seminal formulation of HCE to explore how the hypothesis has changed.

[20] Notice that the confusion is already acknowledged by Zilber-Rosenberg and Rosenberg in their first quotes, since they always refer to Margulis ('hereditary holobiont') and Rohwer ('ecological holobiont') as the original proponents of the *term*, probably ignoring the fact that the term used by Margulis and by Rohwer is definitely not co-intensional and might well not even be co-extensional either.

and the ‘hereditary holobiont’ even co-extensional in certain circumstances? When would those circumstances obtain?

Once the main problem that HCE posed has been established (i.e. mixing the ‘ecological hoobiont’ with the ‘hereditary holobiont’) two points must be noticed: first, that most of the criticisms against HCE will be based on the mixture (or confusion) of the two concepts; second, that this thesis is an effort to reconcile the two concepts, by partially relaxing the concept of ‘inheritance’ (**chapter V**).

3.2. The hologenome concept of evolution as a Lamarckian theory within a Darwinian framework

In an effort to clarify the view of evolution that HCE entailed, one year after their proposal of the hypothesis, the authors published their paper: Eugene Rosenberg, Gil Sharon, and Ilana Zilber-Rosenberg (2009): ‘The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework’. This paper is very important to fully understand HCE, since it helps in clarifying how to conceptualize the different modes of transmission and variation that are brought about in Zilber-Rosenberg and Rosenberg’s original HCE proposal. Particularly, it helps to understand which elements of their hypothesis should be characterized in a neo-Darwinian fashion and which elements should be characterized differently, as well as why they should be characterized differently.

The starting point of the paper puts the emphasis on two properties which, according to Rosenberg et al. (2009), are the core of ‘Lamarckism’: first, the principle of use and disuse; second, the principle of the inheritance of acquired characteristics. The first principle states that biological evolution will work in a way such that those organs, parts, etc. that are highly used by individuals will tend to be preserved, whereas those that are not used, or stopped being used, will be lost. The second principle states that those traits that an individual acquires during its lifetime may be inherited by their offspring. Both principles are nowadays rejected in biology, at least in the neo-Darwinian

conception of evolutionary biology, where both the principle of use and disuse and the inheritance of acquired characteristics are believed to be false for two reasons: first, because the observed variation in a population is always the result of random, i.e. non-directional, modifications in the DNA of an organism, and not a consequence of the use and disuse of one specific organ or part. For the Neo-Darwinist, evolutionary variation is 'blind'. Second, because the only inheritance that is possible is the one that affects the genotype of the germ cells of an organism; phenotypic features that result from effects that are different from modifications of the DNA of the germ cells are not inherited, and thus 'acquired characteristics' are never inherited according to Neo-Darwinism.

In which sense is HCE Lamarckian? To answer this question, it is necessary to take a look back to the mechanisms that might bring variation in a holobiont and that are additional to the conventional and *stochastic* mechanisms of sexual recombination, gene amplification, etc. that will affect both the host and its microbes (**section 3.1.2**). According to Zilber-Rosenberg and Rosenberg, these mechanisms include microbial amplification, microbial acquisition and horizontal gene transfer. In Rosenberg et al. (2009), the authors argue that microbial amplification and microbial acquisition could be interpreted as a Lamarckian aspect of the evolution of holobionts, concretely, as changes that are driven by 'use and disuse' (see also Osmanovic et al. 2018). In their view, the phenomenon of microbial amplification, as well as the phenomenon of microbial acquisition might be driven for the use that a given holobiont makes of its symbionts: if a particular species brings about a fitness-enhancing trait, then, as a consequence of the use of the trait, microbial acquisition (if the species was not previously included in the holobiont) and microbial amplification will follow; if, on the contrary, some microbes are not used, they will tend to be lost from the microbiome. The interesting aspect of this is the goal-directedness of holobiont variation, i.e. that variation among holobionts will depend on the use (or disuse) of the elements of the microbiome, which clearly fits the paradigm of the Lamarckian principle of use and disuse.

Secondly, in Rosenberg et al. the authors argue that, insofar as the microorganisms that are acquired and/or amplified during a holobiont

generation can be transgenerationally transmitted, the holobiont model also satisfies the principle of inheritance of acquired characteristics, since the changes that are brought about in one generation will tend to reappear in the next one, as a consequence of microbiome transmission (**section 3.1.1**; **Figure 3**). All this material, though, will be later 'filtered' by natural selection, and thus HCE is a Lamarckian theory, meaning that the variations and the transmission among holobionts are/can be Lamarckian, within a Darwinian framework, meaning that all the changes that appear in holobionts will later be selected for or against, depending on their influence on fitness.

3.2.1. A comment on the Lamarckian character of the hologenome concept of evolution

Rosenberg et al.'s (2009) arguments about the Lamarckian aspects of HCE need a careful look, since it does not seem as clear as the authors assume that HCE can be properly defined as a Lamarckian theory. Firstly, one of the elements that, according to Rosenberg et al. (2009), defines Lamarckism is the idea that the variation that appears in the organism is not random, or stochastic, but is brought about as a consequence of the use and disuse of the parts. In the usual textbook example, giraffes have longer necks because they tend to stretch them further and further, and thus the phenotypic variation (longer necks) is brought about by the use of the organ (see also Veigl 2017, 2019). Does HCE entail necessarily the same about holobionts, when it comes to the processes of microbial acquisition and microbial amplification? The answer is negative, for in cases where the acquisition of a positive trait leads to the amplification of the microbe that bears it, the amplification is the result of the ecological conditions of the holobiont, rather than of a goal-directedness. In this vein, it is not that clear that HCE is a Lamarckian theory.

Secondly, HCE, as originally formulated in Zilber-Rosenberg and Rosenberg (2008) (**section 3.1**) is completely opposed to the Lamarckian notion of inheritance of acquired characteristics. According to the standard interpretation of Lamarckism, the inheritance of acquired characteristics refers to the inheritance of phenotypic traits that are the result of the process of use and disuse. Neo-Darwinism explains the inheritance of traits by appealing to the

mechanisms of Mendelian transmission. HCE is not far from Neo-Darwinism, since Zilber-Rosenberg and Rosenberg argue that the genomes of the microbes that are acquired need to be transmitted. They offer a genetic interpretation of the holobiont which is closely related to the normal process of inheritance, more than it is to the notion of 'inheritance of acquired characteristics', since the key point for the latter is that the genetic modifications (if any) that bring them about do not need to be inherited.

3.3. First evidence of hologenome selection: the case of mating preferences in Drosophila melanogaster

The first non-coral evidence that supported HCE appeared in Sharon et al. (2009).[21] In their paper, they drew on a previous case of mating preference in *Drosophila pseudoobscura* reported in Dodd (1989). In his experiments, Dodd reared two groups of flies on a starch-based medium and on a maltose-based medium for more than 25 generations. Afterwards, he put the two groups together and observed that the flies reared on the starch-based media preferred to mate with these reared in the same media, whereas the opposite was the case for the flies reared in a maltose-medium. Surprisingly, despite the clear mating preference that had been developed among flies, Dodd could not find any evidence of selection for mating preference in any of the groups. He could only hypothesize that mating preference developed as correlated response to selection favouring an adaptation to a new diet.

Sharon et al. decided to replicate Dodd's experiment and take it one step further, to test whether they could have any evidence of selection on hologenomes. To do so, they divided a population of *D. melanogaster* in two groups, rearing one of the groups on a molasses medium and the other on a starch medium. They later put the two groups together and observed that a mating preference has appeared: flies reared on starch preferred to mate with flies reared on starch, whereas flies reared on molasses preferred to mate with flies reared on molasses (**Figure 5**). They observed: (1) that the mating

[21] Notice that the evidence found in corals was not about speciation, it was only about the acquisition of an adaptation, which could not be taken in any sense as an initiator of a speciation process.

preference appeared after only one generation of divergent diets; (2) that the preference was maintained for at least 37 generations. Until this point, nothing is too surprising, as Sharon et al.'s evidence is identical to Dodd's evidence. A second important step was taken when the authors decided to investigate the cause of the divergent mating preferences. Their hypothesis was that microbiome differences were underlying the diverging mating preferences in the population, a hypothesis which derives entirely from HCE. To test their hypothesis, they performed two different types of tests: first, after rearing the bacteria on their respective media, they treated them with antibiotics, observing that the mating preference disappeared; second, they did four independent infection experiments, which suggested that the flies that had lost their bacteria and for which mating preference had disappeared recovered their tendency towards a divergent mating preference if re-infected with the bacteria they had lost because of antibiotic treatment.

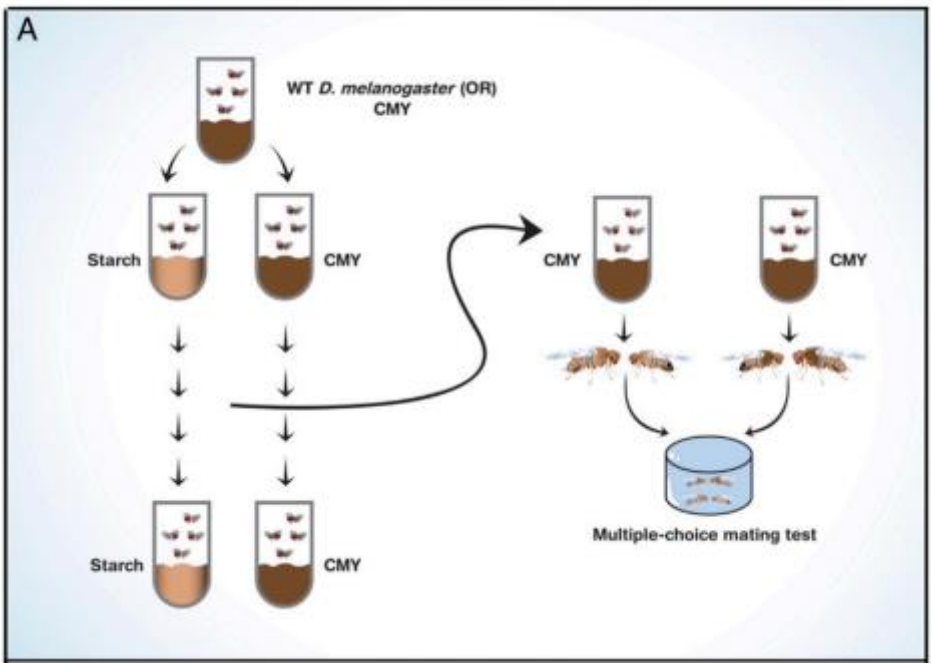


Figure 5. Description of the experimental set-up in Sharon et al.'s experiment. The population was divided in two groups and each was reared on a starch-based medium and on a molasses-based medium (CMY) for several generations (left side). Later, members from both groups were transferred from one generation into a CMY-based medium. The mating preferences were tested afterwards on 24-well plastic plates which contained four flies: two, male and female, reared on the starch-based medium, and the other two, also male and female, reared on the CMY medium (right side). (From Sharon et al. 2009: 20052, Fig. 1).

They later performed a 16S rRNA array of the bacterial communities of the flies to determine which members of the microbiota might have been responsible of the appearance of the mating preference. They observed that the group of flies reared on a starch-based medium had a substantially higher proportion of *Lactobacillus plantarum* (26%) than the flies reared on the CMY-based medium (3%), which suggests *L. plantarum* as a plausible candidate for determining the divergent mating preferences. They tested their hypothesis and observed that what might happen is that the starch-based medium induces a microbial amplification increasing the relative numbers of *L. plantarum* in *D. melanogaster*. This microbial amplification is later responsible of the observed mating preference that caused flies reared on CMY to develop a preference for other flies raised in the same medium, whereas the opposite was the case for flies reared in a starch-based medium. Therefore, the authors conclude, changes in the microbiome are responsible of the mating preferences in *D. melanogaster*, thus suggesting that the holobiont is the unit of selection.

3.3.1. Brief reflection on Sharon et al.'s results

Sharon et al. argue that their experiments supported HCE but, how true is this? In other words, is there another way of interpreting their evidence so that it is not taken as a straightforward case in favour of HCE? I suspect that, even if the results obtained by Sharon et al. suppose a strong case in favour of HCE, one must be cautious about how to interpret them, since there are different alternatives that do not necessarily match with HCE. First of all, Sharon et al. have not proven that the microbiome has an influence on the mating preferences of *D. melanogaster*. At most, they have proven that *L. plantarum* has a transgenerational influence on the mating preferences of *D. melanogaster*, under very restricted environmental conditions. This, of course, suggests that it is important to study the influence of the microbial community on a host's mating preferences, and it might also suggest that *D. melanogaster* and *L. plantarum* could evolve as a single unit under the influence of natural selection. But, unfortunately, this evidence is still far from proving that the

whole, i.e. the host plus its microbiota act as a unit of selection, which is the main claim made by HCE defenders.[22]

Secondly, there are further hypotheses which are totally consistent with the evidence found by Sharon et al. and that do not need to consider the holobiont as a unit of selection. For instance, one may argue that the CYM-based medium and the starch-based medium were different environments in different territories—each group of flies in different media with no mixture—that had the power to induce these phenotypic changes, but natural selection is not playing any role here. The effect that Sharon et al. observed is simply the result of two populations of flies, rather than of the action of natural selection on the hologenome. Under this hypothesis, it is possible to justify that this isolation could trigger, in the long term, to a phenomenon of allopatric speciation, if the genotypes of the two populations of flies start diverging as a consequence of the environmental differences. Furthermore, it is perfectly consistent to argue that even if flies from the two groups were put together afterwards to test their mating preferences, the flies would be inhabiting different environments (an environment rich in *L. plantarum* versus an environment poor in *L. plantarum*), and thus there are two populations of flies that share the same territory. Under this hypothesis, the mating preferences would be a consequence of the existence of two populations that inhabit the same territory, rather than of the action of natural selection on the flies hologenome (an effect that could trigger a sympatric speciation, if the flies genomes start diverging). However, natural seletion would not have had occurred until the genomes of the two groups of flies started diverging. Therefore, Sharon et al. would have at most proven that different environmental conditions, including changes in a host's microbiota, might trigger some phenotypic differences. But this is a far cry from the hypothesis they aimed to prove, namely: that the host and its microbiota evolve as a single unit.[23]

[22] I will provide a consistent way of explaining these results in **chapter IV, section 4.2**.

[23] Recently, Leftwich et al. (2017) revisited the experiments carried on by Sharon et al. (2009) and claimed to have found no evidence of microbiome-induced mating preferences. They also made an interesting point about the observations made by Sharon et al.: Leftwich et al. argued that, even if in some circumstances it could be possible to find cases of diet-induced mating preferences, these should be looked at carefully since, they speculated, in species with a highly flexible microbiome (like flies), diet-induced effect will always tend to be transient. Their point is relevant here since it is consistent with the 'ecological holobiont', despite being inconsistent, in

3.4. Further developments of the hologenome concept of evolution. The 'capacious' hologenome of Brucker and Bordenstein

Robert M. Brucker and Seth R. Bordenstein are two of the authors who have been most interested in HCE and have applied some of its postulates in their research, obtaining very interesting results and developing very useful notions to test the validity of HCE. Concretely, they have always been interested in the power of symbiosis to drive speciation processes, a feature of symbiosis that had been vigorously emphasized by Lynn Margulis (**section 2**), so most of their work has been oriented to test cases of speciation by symbiosis, which they analyse by taking the 'whole' (i.e. the holobiont) as the unit of evolution/selection. Their research is specifically oriented to prove the thesis that hologenomes are replicators (**section 3.1.3**). In this section, I will review their theoretical and empirical work.

3.4.1. 'Speciation by symbiosis' and 'The capacious hologenome'

Two of the most important theoretical works written by Brucker and Bordenstein include their relatively early Brucker and Bordenstein (2012a), and their (2013a). In (2012a), Brucker and Bordenstein argue that the models of speciation should include not only the study of the molecular genes or possible organelles that are responsible for speciation events, but also the microorganisms of the host's microbiome, since the latter opens a new avenue of possible speciation mechanisms that would remain unexplored if speciation is exclusively restricted to phenomena that affect molecular genes and/or cellular organelles.[24]

Their paper starts with the following reflection: first, it is clear that symbiosis is a very powerful mechanism for generating new traits (e.g. luminescence in the bobtail squid); second, it seems clear that symbiosis has played some role in accelerating evolution, i.e. in creating the conditions that

---

principle, with the 'hereditary holobiont'. For a response to Leftwich et al. (2017), see Rosenberg et al. (2018).

[24] The whole paper is concerned with the evolution of higher animals, so it is built on the assumptions of the biological species concept (i.e. the notion that species appear as a consequence of reproductive isolation, and thus reproductive isolation = speciation).

favour the appearance of genetically-based reproductive barriers. However, what would happen if the microbial species that compose the microbiome of a given host are taken as extensions of the host's genome? In other words, would our view of speciation processes change if instead of focusing on the monogenomic organism (i.e. the nuclear genes of the genome of the organism) we focus on the whole array of genes of an organism's hologenome? In the view of Brucker and Bordenstein, the obvious answer is 'yes', and not only that, but they believe that the consideration of the hologenome as the unit that 'speciates' would open the avenue for new speciation processes and the appearance of reproductive barriers much earlier than they would appear if only the changes in host's genome that led to reproductive incompatibilities are considered as potential cases of speciation. Therefore, it is the hologenome, not the host's genome, the unit that, according to Brucker and Bordenstein, should be taking as 'evolving' and, thus, driving speciation processes.

However, it is necessary to support further the claim that the hologenome is the unit of evolution. How strong is Brucker and Bordenstein's evidence to claim that the hologenome, and not the host's genome, is the entity that speciates? In (2012a) the authors refer to three sources of evidence that support HCE: first, the observation, already anticipated in Zilber-Rosenberg and Rosenberg (2008), that all eukaryotes bear microbial symbionts, i.e. the universality of eukaryote-microbial symbiosis; second, the notion of host's specificity, or the idea that the microbes that associate with an eukaryotic host are species-specific, i.e. the species of the microbiome recapitulate the phylogeny of their hosts;[25] third, the observation that the immune genes of hosts are rapidly evolving in an arms race with the microorganisms that compose the host's microbiome. If these three observations are true, then there is evidence to support the idea that the hologenome, and not the individual genome, is the entity that evolves, is naturally selected and is driven by speciation processes, as HCE postulates.

[25] This second observation, which was later called 'phylosymbiosis', as well as its importance for justifying HCE, will be explored in more detail bellow.

The question is now the following. Assuming that the three aforementioned observations are true, and thus hologenomes can be considered evolutionary units, what are the consequences for our understanding of speciation processes? Brucker and Bordenstein argue that symbiosis can drive reproductive isolation, and thus speciation, in two different ways: by generating *pre-mating isolation* conditions that reduce the gene flow between the members of a population, so that in the long-term they will become two independent populations and even two species, due to the lack of genetic exchange; by generating *post-mating isolation* conditions, or incompatibilities that arise after the offspring is produced, making it non-viable.

Cases of pre-mating isolation include behavioural preferences, including mating preferences (**section 3.3**), and ecological isolation, including phenomena like the occupation of new niches, the utilization of new nutrients, etc. Brucker and Bordenstein believe that bacterial symbionts are responsible for many of the cases of ecological and behavioural isolation, thus suggesting that symbionts can play a fundamental role in creating the conditions that prevent gene flow and thus lead to speciation processes. Cases of post-mating isolation include all the phenomena that give rise to the appearance of hybrid incompatibility. Hybrid incompatibility refers to the impossibility of hybrids being born, developing normally (hybrid inviability), or reproducing (hybrid sterility). One hypothesis to explain hybrid incompatibility is the Bateson-Dobzhansky-Muller model, according to which hybrids are sterile or, in some cases, inviable, because the epistatic interactions between their nuclear genes give rise to incompatibilities that disrupt development, or reproduction. According to Brucker and Bordenstein, considering the hologenome as a unit of selection widens the possible modes in which hybrid incompatibility can arise (**Figure 6**).

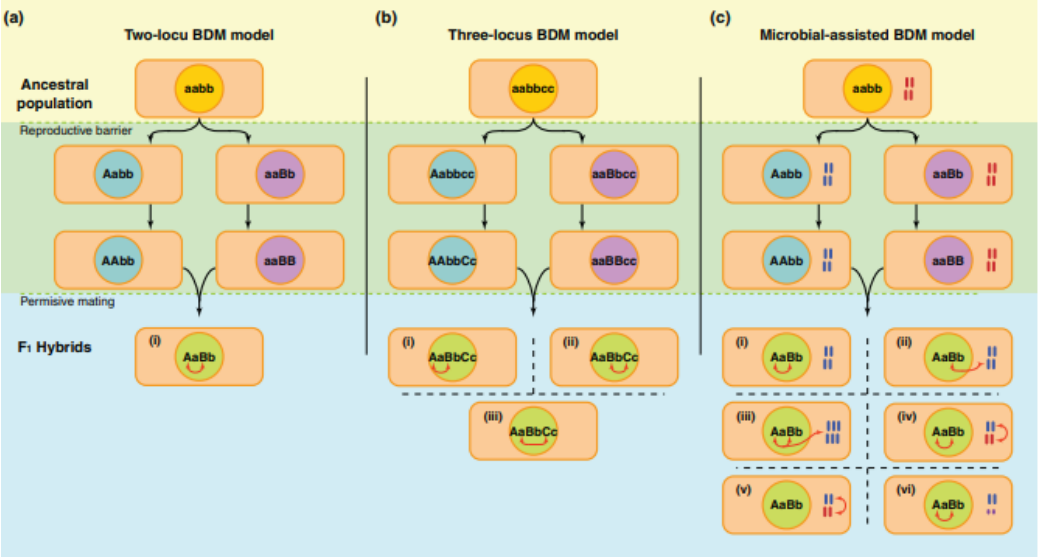


Figure 6. Schematic representation of the Bateson-Dobzhansky-Muller model of hybrid incompatibility. (a) depicts the most basic case, where the incompatibility occurs at two loci, such that alleles A and B are incompatible; (b) represents the possibility of having an incompatibility among three loci, such that A is incompatible with B and C, and B and C are incompatible with each other; finally, (c) represents a situations where the incompatibility can additionally occur between an allele and some component of the microbiome [(iii)], or exclusively between some components of the microbiome [(iv) and (v)]. (From Brucker and Bordenstein 2012a: 447, Box 2)

The mechanisms that drive hybrid incompatibilities and are known to be (at least sometimes) a consequence of incompatibilities among symbionts include cytoplasmic incompatibility (**Figure 7**), hybrid susceptibility and hybrid autoimmunity. These cases, Brucker and Bordenstein argue, suggests that symbiosis is a driver of speciation, and thus provides some support for HCE.

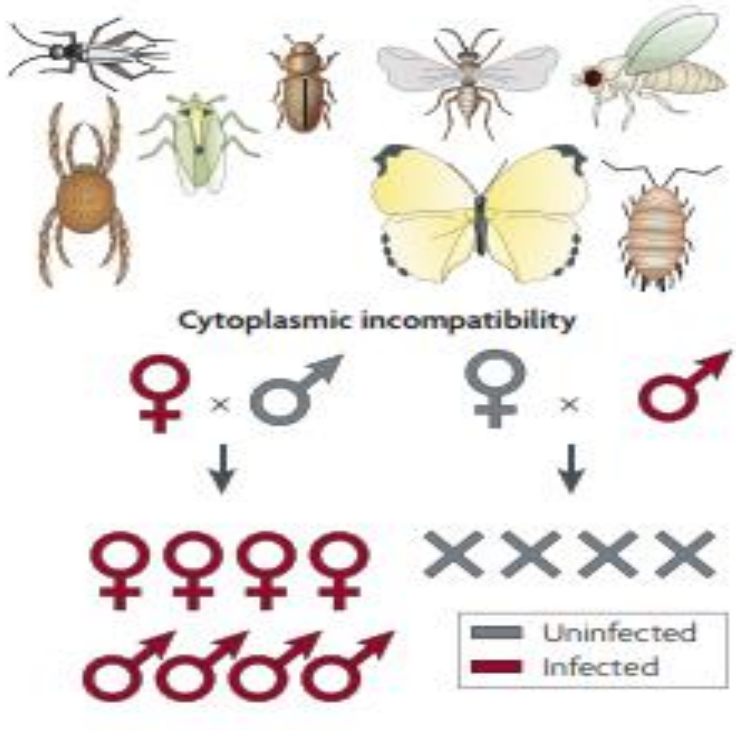


Figure 7. Cytoplasmic incompatibility among invertebrates has been shown to be caused in many cases by *Wolbachia pipientis*. As the bacteria are transmitted vertically, coupled with female reproductive cells, they disrupt the mating between infected males and uninfected females, whereas infected females can mate both infected and uninfected males. Cytoplasmic incompatibility is a method of post-mating isolation that has been shown to drive reproductive isolation in some populations. (From Werren et al. 2008: 743, Fig. 2)

Right after (2012a), Brucker and Bordenstein published (2013a), where they emphasized the importance of considering the hologenome, and not only the host's genome, as the unit of evolution. (2013a) is important because it is partially clarificatory of the meaning of the concept of 'hologenome', which had been previously introduced by Zilber-Rosenberg and Rosenberg to refer to the total among of genetic information in the holobiont (**section 3.1.**). There, Brucker and Bordenstein postulate that, among all the microbial components of a host's microbiome, there is a whole family of species-specific microbes which are acquired under the control of the host. Based on a previous study that they had conducted in *Nasonia* wasps (Brucker and Bordenstein 2011), they hypothesized that the species that compose the gut microbiome have shared part of their evolutionary history with the host, and thus represent 'an ancestral footprint of evolution' (2013a: 260). Therefore, the hologenome is the entity

composed by the host's genome, its endosymbionts (i.e. the mitochondrial genes) *and the species-specific fraction of the microbiome* (**Figure 8**).

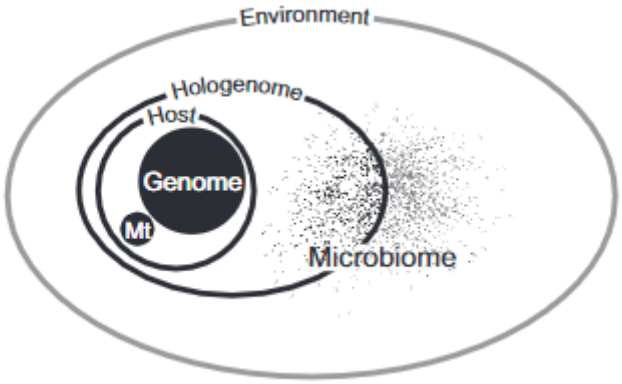


Figure 8. The hologenome according to Brucker and Bordenstein. The dark/black spots in the figure represent the total microbiome that a host interacts with. It is divided in two sets: the part of the microbiome that is considered to be a part of the environment, and the part that is considered to be part of the hologenome (to which they refer to as 'beneficial microbiome'). The division between the two sets in the microbiome is made according to whether the microbes are acquired deterministically, i.e. dictated by the host's genome, or they are the result of a pure environmental acquisition, i.e. diet, abiotic factors, etc (part of the microbiome represented in grey). (From Brucker and Bordenstein 2013: 261, Figure 1)

At this point, two important points about the hologenome, and about HCE, need to be noticed. First, Brucker and Bordenstein's definition of the hologenome is not co-extensional with Zilber-Rosenberg and Rosenberg's formulation. For the latter, the hologenome would encompass *all* the microbes of a host's microbiome. Only if this is so, it is possible to consider the phenomenon of microbial acquisition as a case of variation among holobionts, as Zilber-Rosenberg and Rosenberg aim to consider it. For the former, on the contrary, the hologenome includes exclusively the species-specific microbiome, and not the whole set of microbes of the microbiome. Again, these two concepts of the hologenome mirror the distinction introduced previously between the 'ecological holobiont' and the 'hereditary holobiont' (**section 3.1.4**), so I will refer to them as the 'ecological hologenome' and the 'hereditary hologenome'. Second, but very important, is the fact that Brucker and Bordenstein recognize that part of the species-specific microbiome is (or can potentially be) environmentally acquired (by horizontal transmission), probably by some of the mechanisms that Zilber-Rosenberg and Rosenberg had previously hypothesized. This, obviously,

creates a puzzle: it is clear, for example, that organelles, which are vertically transmitted from parent to offspring can be understood as an integrated unit together with the host's genome. Now the question is 'whether fractions of the environmentally acquired, but host associated, beneficial microbiome can be understood in a similar way' (2013a: 261).[26]

3.4.2. Introducing the concept of 'phylosymbiosis'

A key element to the discussion of HCE that was introduced by Brucker and Bordenstein was the observation that the microbes that associate with a given host are species-specific and are acquired under host control. They refer to this phenomenon with the concept of 'phylosymbiosis' (Brucker and Bordenstein 2012b). Phylosymbiosis is taken by the authors as a proxy for selection on hologenomes[27] and refers to the 'eco-evolutionary pattern in which evolutionary changes in the host associate with ecological changes in the microbiota.' (Brooks et al. 2016: 3). In other words, a phylosymbiotic pattern between two (or more) species obtains when the evolutionary patterns of the species mirror each other, such that their changes are concordantly related *in a relevant way*. Importantly, as the authors have emphasized several times, the discovery of phylosymbiotic patterns does not imply the existence of coevolution, cospeciation or even cocladogenesis among the interacting species.[28] This is because the concept 'does not necessarily presume that the members of the microbial community are constant, stable, or vertically transmitted from generation to generation' (Brooks et al. 2016: 3; see also Theis et al. 2016). This raises two important questions. First, if phylosymbiosis does not presume vertical transmission of the microbiome, what does it exactly refer to and how it can be experimentally identified? Or, in other words, what's the relevant way in which the divergent evolutionary patters need to be related so that phylosymbiosis can be identified? Second, and more importantly for the conceptual purposes of my thesis: if phylosymbiosis does not entail

[26] Brucker and Bordenstein specifically refer to the 'beneficial' microbiome in this paragraph. We will see how this definition changes in more recent conceptions of the hologenome and further developments of HCE.

[27] Notice that all this section is about selection on *hologenomes*, not about selection on *holobionts*. This distinction is crucial to understand the discussion and the type of arguments that Brucker and Bordenstein will introduce.

[28] Rosenberg and Zilber-Rosenberg (2013) have however sometimes identified phylosymbiosis with the existence of cospeciation, but this is not what Brucker and Bordenstein meant when they originally introduced the concept (Theis et al. 2016).

cospeciation, how can it be taken as a proxy for hologenome selection (**section 3.4.3**; **chapter V, section 3.3**)?

A quick answer to the first question would be the following: The concordance that needs to be found to argue that there is a phylosymbiotic pattern among two or more species needs to be higher than the concordance that would be expected if the association between the host and its microbiota were due to a process of ecological filtering. Host's microbiome composition is known to be affected by different environmental factors, including diet, close physical contact with other hosts (especially sex), age, etc. If the acquisition of the microbiome were stochastic every generation, then every environmental microbe would have the same chance to colonize every host in its surroundings, and thus microbiome composition would be similar among host's colonizing the same niche, partially irrespectively of host's species. However, if the assembly is not random, then the pattern of host-microbiome concordance will be higher among hosts of the same species than among hosts of different species (i.e. intraspecific, rather than interspecific). Or, in other words, the higher the genetic differences among hosts, the higher the differences in their microbiomes, suggesting that host-microbiome assemblages are not randomly formed (**Figure 9**).

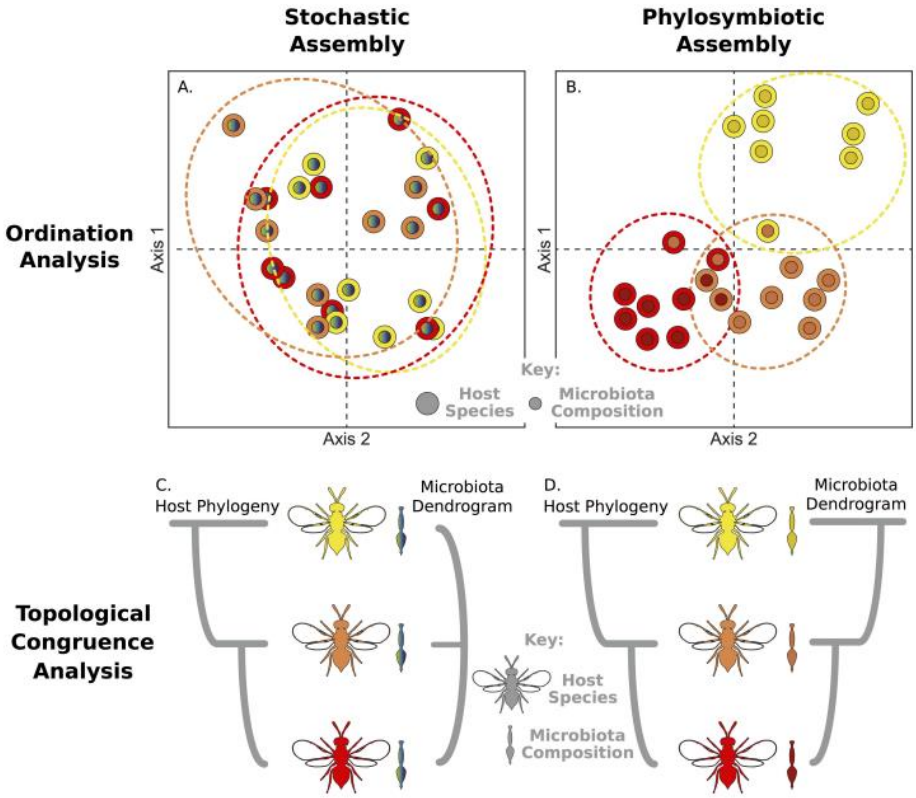


Figure 9. Representation of two patterns of stochastic vs phylosymbiotic assembly by ordination analysis and topological congruence analysis. In the ordination analysis, the dashed lines represent host-specific clustering. The divergent nature of the stochastic assembly and the phylosymbiotic assembly can be shown by comparing (A) and (B). Figures (C) and (D) compare the dendograms of a stochastically assembled holobiont versus a phylosymbiotically assembled one. It can be show how the patterns of divergence are congruent if the relationship is phylosymbiotically (D). (From Brooks et al. 2016: 3, Fig. 1)

The conceptual part of the first worry is thus solved, and the exact meaning of 'phylosymbiosis' has been made clear. Now the question is how to solve the second part of the problem, namely, how to empirically test the existence of such phylosymbiotic pattern. Is it possible to isolate all the confounding variables that might affect microbiome composition, so that the existence of phylosymbiosis can be tested? A pioneering study, previous to the introduction of the concept of 'phylosymbiosis', was done by Sebastian Fraune and Thomas C. G. Bosch (2007): 'Long-term maintenance of species-specific bacterial microbiota in the basal metazoan *Hydra*'. They analysed the microbiome community of two closely related species of *Hydra*: *H. vulgaris* and *H. oligactis*

and compared the microbiome composition (16S rRNA analysis)[29] of free-living *Hydra* with the microbial composition of *Hydra* that had been cultured and reared under laboratory conditions for more than 30 years. Their results showed that the microbiota of the *Hydra* that had been reared under laboratory conditions had preserved their composition and were significantly similar to the microbiome composition of free-living *Hydra*, in both species. Their conclusion is, thus:

> 'Taken together, these data show that both *Hydra* species select particular bacterial guilds. *H. oligactis* and *H. vulgaris* maintain these species-specific bacterial communities even when cultured under constant conditions for >30 years. Alternatively, it seems possible that the bacteria are also actively involved in selecting the host.' (Fraune and Bosch 2007: 13148)

Brucker and Bordeinstein (2011) carried out a second experiment that proved the existence of a phylosymbiotic pattern. In their experiment, Brucker and Bordenstein reared individuals from three different *Nasonia* species: *N. vitripennis*, *N. giraulti*, and *N. longicornis*. *N. vitripennis* is known to have diversified 1MY ago from the ancestor of the other two species, whereas *N. giraulti*, and *N. longicornis* are believed to have diversified from each other about 400 thousand years ago (Raychoudhury et al. 2010). The purpose of Brucker and Bordenstein's research was twofold: on the one hand, they wanted to test whether under the same rearing conditions, with controlled diets, sex, and ages, each *Nasonia* species would bear a species-specific microbiota (16S rRNA analysis), different from that of members of the other species; on the other hand, they wanted to test whether the different microbiomes recapitulated *Nasonia* phylogeny. Their results confirmed their initial hypothesis, thus suggesting that the microbiota can diversify in concordance with the diversification of the host species, i.e. in parallel to the host phylogenetic relationships and, thus, the species that compose a host's microbiota are, at least in part, specific to the host species.

[29] The fact that they analyse microbiome composition using the 16S rRNA analysis will become relevant later (**chapter II**).

Many other experiments have further supported the existence of phylosymbiosis in different animal species, including bats (Phillips et al. 2012), corals (Pollock et al. 2018), hominids (Ochman et al. 2010), sponges (Souza et al. 2017), etc. The evidence is however controversial and admits different interpretations, since it is not always clear whether all the confounding variables have been correctly isolated, whether the evidence is fully conclusive, or even whether phylosysbiosis can be given an evolutionary interpretation (Mazel et al. 2018). In any case, reviewing all the cases is outside the scope of my project. It will be important to bear in mind, though, that the concept of phylosymbiosis might have some significance for testing HCE. This leads us to the second question I had formulated above: how can phylosymbiosis be taken as a proxy of selection on hologenomes?

3.4.3. Phylosymbiosis as a proxy for selection on hologenomes

Notice that the concept of phylosymbiosis that has just been introduced and elucidated serves to demarcate which part of the microbiome can be said to integrate with the hologenome, versus which part of the microbiome must be considered a part of the environment. In other words, it introduces a criterion to make clear the boundaries that appear in **Figure 8**. This is a clear individuation criterion, which individuates the hologenome in relation to its participation in an eco-evolutionary process that generates a pattern of concordance. But, as I already explained, the observed pattern of concordance does not need to be the result of vertical transmission of the microbiome, nor necessarily indicative of a process of cospeciation, cocladogenesis, or coevolution (Brooks et al. 2016, Theis et al. 2016).[30] Then, what is exactly its evolutionary significance? Or, in other words, how can the existence of a phylosymbiotic pattern be evolutionarily interpreted?[31]

[30] 'Not necessarily indicative of a process of coevolution' does not mean that it is not indicative of one. It might be indicative of a process of coevolution, in some cases. But the key point of Brucker and Bordenstein is that it does not need to be automatically interpreted as such.

[31] See **chapter V, section 3.3** for a discussion of these points from the perspective of my own framework to study the role of holobionts as units of selection.

> The response to this question relies on the concept of 'community heritability' H^2_c, for whose existence phylosymbiosis is a proxy. In van Opstal and Bordenstein (2015), the authors suggest reconsidering the ways in which heritability measures of the microbiome had been previously carried out and concentrate instead on the notion of community heritability. Let me explain the significance of this point from the beginning. 'Heritability' is a concept from population genetics that measures the degree in which variations of a phenotypic trait can be attributed to the genetic differences among the members of the population (Visscher et al. 2008).[32] Variation in a trait can be argued to be highly heritable if and only if an important proportion of the individuals of the population that bear the trait—if it is a qualitative trait—or a concrete value for the trait—in case it is a quantitative trait—have a specific variance in their genetic makeup that can be argued to be on the basis of the trait variance. That way of thinking (the 'standard heritability' approach h^2) had been previously applied to the study of the microbiota *species composition*, where the microbiome was taken as a quantitative trait whose variation in species composition should be accounted for in terms of the genetic differences among the hosts. Thus, twin studies—the conventional method of testing heritability—were conducted to determine the influence of the host's genetic makeup and the influence of the environment in determining microbiome composition.

However, as van Opstal and Bordenstein argue, the standard heritability approach is 'unidirectional', since it presupposes that the host is the only entity determining the composition of the microbiome, and thus the composition of the hologenome. But, if the whole point of HCE was to emphasize the 'unitarian' aspect of host-microbiome interactions, i.e. the claim that the hologenome is a 'unit', in some sense of 'unit', then the standard heritability approach is misguided to understand the role of hologenomes as biological units. In their approach, a more comprehensive view is required, according to which the assembly of the host and its microbiome is driven by genotype-by-genotype interactions, i.e. by interspecies interactions. That is precisely the approach that

[32] For the purposes of the discussion here, I will not distinguish between narrow-sense and broad-sense heritability.

the study of community heritability H^2_c offers, and that van Opstal and Bordenstein propose as an alternative to standard heritability measures (**Figure 10**). Now, an obvious question arises: it is clear what standard heritability measures measure but, what does community heritability *exactly* measure? In the words of van Opstal and Bordenstein:

> 'H^2_c emphasizes that the host is part of an ecosystem and measures the extent to which variation in "whole-community" phenotype is due to genetic variation in the foundation (i.e., host) species of the community. It therefore specifies that host genetic variation will have predictable effects on microbial community assembly, in addition to having effects on specific members of the microbiome, as measured by h^2.'

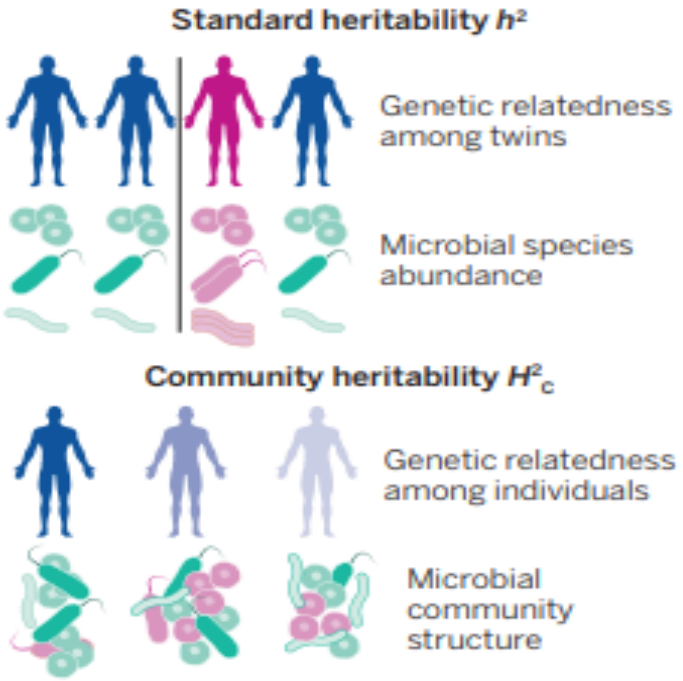


Figure 10. Comparison of the concepts of 'heritability' h^2, as it is conventionally understood in population genetics, and community heritability H^2_c, as it should be understood for the hologenome. Standard heritability measures how hosts with different genetic backgrounds might bear different microbial species in their microbiome, to determine how much of the variation in microbiome species composition can be ascribed to variation in the genetic makeup of the host (studies are usually carried out comparing the microbiomes of monozygotic and dizygotic twins). Community heritability, in contrast, measures how the differences in the genetic makeup of individuals in the population might condition the structure of their microbiomes, or whole-community phenotype. (From van Opstal and Bordenstein 2015: 1173)

Notice the important contrast that exists between H^2_c and h^2. First, whereas the later only measures the degree of *host control* on microbiome composition, the former, insofar as it only measures the relation between the phenotype of the whole-community and the genetic variation of the host, can represent three possible states: host control of microbiome assembly; host susceptibility to microbiome 'infection'; or a combination of host control and host susceptibility. Second, whereas h^2 measurements implicitly assume that microbiome composition is merely an extension of the host's phenotype, H^2_c measurements inherently incorporate the existence of non-random (i.e. not caused by a process of ecological filtering) interspecies interactions that give rise to the appearance of a whole-community phenotype. Third, whereas a high degree of h^2 for a microbial species in the microbiome underscores the possibility that the host undergoes selection to secure the transgenerational presence of that species, a high degree of H^2_c suggests that natural selection can potentially select certain forms of community assembly over other, thus potentially acting on hologenomes.[33]

The question, however, is whether the existence of community heritability of the hologenome is likely, and what are the conditions that should be satisfied if it is going to happen. Or, in other words, can we make sure that community assembly is a result of genetic factors rather than environmental factors? That question is really hard to parse, and many have suggested that the value of H^2_c is very low in most holobionts (Moran and Sloan 2015; Douglas and Werren 2016; Skillings 2016). I will review their arguments in **chapter II**.

3.4.4. How 'capacious' can the hologenome be? The case of Nasonia *wasps*

Let us grant Brucker and Bordenstein the possibility that phylosymbiosis is a proxy for the existence of selection acting on hologenomes, as well as the discovery of a phylosymbiotic relationship in *Nasonia* wasps (Brucker and Bordenstein 2011; **section 3.4.2**). The next step in their research to prove that the hologenome is a unit of selection is to show a case of speciation where the

[33] These three points about the distinction between H^2_c and h^2 will remain crucial for the rest of the thesis.

hologenome is (or can at least be considered) a relevant unit. To do so, they study the phenomenon of hybrid lethality in *Nasonia* wasps (Brucker and Bordenstein 2013b). Hybrid lethality is a post-mating (post-zygotic) mechanism of reproductive isolation that occurs when the hybrids that result from the cross-mating of two species do not properly develop, thus leading to their death before they are even born. Hybrid lethality has been traditionally assumed to be a consequence of the existence of negative genetic epistasis in hybrid genomes that do not let hybrids develop properly. However, as Brucker and Bordenstein explained (2013a) (**section 3.4.1**), there is no reason to assume that the same negative epistatic model cannot be applied to elements of the microbiome. In other words, it is possible that hybrid lethality is a consequence of the interactions between the host's nuclear genes and some elements of the microbiome, or a consequence of the interactions between the elements of the microbiome themselves.

To test this possibility in the case of *Nasonia*, Brucker and Bordenstein reared three species of *Nasonia* wasps (*N. vitripennis*, *N. giraulti*, and *N. longicornis*) under the strict same laboratory conditions to guarantee that each species mostly interacts with its phylosymbiotic community.[34] They observed that reciprocal crossings between *Nasonia* species gave rise to fertile, and diploid hybrid females in the first generation F_1. However, hybrid lethality—manifested during the larval stage—is observed in the haploid male offspring of the second generation F_2. The proportion of hybrid lethality was about 90% among *N. vitripennis* x *N. giraulti*, and *N. vitripennis* x *N. longicornis*, and about 8% among *N. giraulti* x *N. longicornis*. Brucker and Bordenstein hypothesized that the lethality in hybrids was in part a result of the presence of altered gut microbiomes in F_2 male hybrids.

Their hypothesis was then tested in two steps. First, they studied the symbiotic communities of hybrid males and non-hybrid males in F_2 and during

[34] *Nasonia* are parasitoid wasps, so all of them were reared on the same fly host *Sarcophaga bullata*, guaranteeing that all of them have the same type of diet and avoiding *Wolbachia* infections to guarantee the possibility of reciprocal crossings between the species. Also, importantly, lethality is diagnosed by the observation of larval melanisation, so that the precise moment in which it happens can be detected.

larval development, right before hybrid lethality occurs. If the microbiome plays any role in the lethality of hybrids, then hybrids and non-hybrids might bear different microbiomes. They observed that F_2 *N. vitripennis* x *N. giraulti* hybrids had a different microbiome than both of their parental species, differing both in species composition and species abundance. The major shift they observed was that the dominant bacterial operational taxonomic unit (OUT, hereafter) had changed from *Providencia* sp. IICDBZ10 (in F_2 pure species) to *Proteus mirabilis* SMBS (in F_2 *N. vitripennis* x *N. giraulti* hybrids). The shift was significantly substantial, so that some microbiome effects on lethality could be hypothesized.

Second, they studied the possible effects of the microbiome on *N. vitripennis* x *N. giraulti* F_2 hybrid lethality. To do so, they reared conventional, germ-free, and bacteria-inoculated hybrids and non-hybrids. The driving idea was the following: if hybrid lethality is a consequence of negative epistasis among the genes in the hybrid genome, then it will be expected that hybrid lethality will not be substantially different between conventionally and germ-free reared hybrids. However, if the microbiome has some influence on hybrid lethality, then it is possible that germ-free reared hybrids will recover from lethality. Indeed, their results convincingly showed a statistically significant recovery of the hybrids that had been reared in germ-free conditions, in relation to the degree of hybrid lethality among those that had been reared in the conventional environment. Secondly, their data also showed how the same levels of hybrid lethality were shown if germ-free F_2 hybrids were inoculated with *Proteus mirabilis* SMBS and *Providencia* sp. IICDBZ10 from parental *Nasonia* during their larval stage, an effect that was not observed in non-hybrid wasps (**Figure 11**), thus strongly suggesting the influence of the microbiome as one of the causative agents of hybrid lethality.

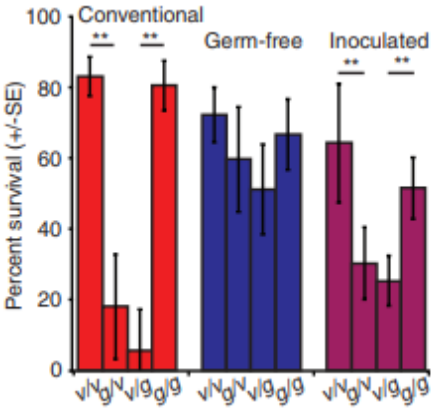


Figure 11. Comparison of the survival rate among F_2 male *Nasonia* hybrids versus F_2 non-hybid males in three different rearing conditions. The data shows how the survival rate increases significantly among the hybrids (g/v; v/g; g: *N. giraulti*; v: *N. vitripennis*) when they are reared in a germ-free environment, whereas they increase again when they are inoculated with bacteria during their larval stage. (From Brucker and Bordenstein 2013b: 668, Fig. 1C)

Brucker and Bordenstein's results strongly support the hypothesis that host-microbiome interactions can generate post-mating mechanisms that guarantee that the reproductive isolation between species is transgenerationally maintained. Furthermore, they also suggest the possible existence of an hologenomic basis underlying certain speciation events—provided that the in fact the epistatic interactions between the host genes and the microbiome were causing hybrid lethality—which supports the hypothesis that 'the hologenome [acts] as a unit of evolution [HCE], blur[ring] the line between what biologists typically demarcate as the environment and the genotype of a species' (Brucker and Bordenstein 2013b: 669).

3.4.4.1. Not so capacious, after all. Chandler and Turelli's response

Brucker and Bordenstein's experiment elicited an immediate response by James Angus Chandler and Michaelle Turelli (2014): 'Comment on "The hologenomic basis of speciation: Gut bacteria cause hybrid lethality in the genus *Nasonia*"'. In their response, Chandler and Turelli made basically two points: First, that since Brucker and Bordenstein's data did not include any evidence about specific incompatibilities between concrete bacterial lineages and the host's genome, it was not conclusive to prove the existence of hologenomic speciation. Second, that the evidence that Brucker and

Bordenstein provided to support phylosymbiosis—taken by the authors as a proxy for hologenomic speciation—was weak and inconclusive. The second criticism is highly technical and concerns the type of methodologies that are required to test phylosymbiosis, so I will not review it here. I will only concentrate on the first criticism, as it is the criticism with a broader philosophical reading.

Chandler and Turelli believe that Brucker and Bordenstein's experiment lacks the crucial test that would convincingly prove coadaptation, namely: a *phylogenetically informed* cross-inoculation experiment of gut microbes among the different *Nasonia* species. As their experimental set up was displaced, Brucker and Bordenstein proved, at most, that the gut microbiome influences hybrid lethality, but not that it is *the causative agent* of lethality. An alternative explanation of their results could be the following: the genome of hybrids is so dysfunctional due to their high degree of genetic incompatibility that any encounter with any free-living bacterial species (no matter its history of interaction with the host) will make it susceptible to lethality ('intrinsic hybrid disfunction'). But, if this interpretation of Brucker and Bordenstein's results is correct, what is the opportunity for hologenomic selection? As Chandler and Turelli put this criticism:

> 'Brucker and Bordenstein's data demonstrate that bacteria can contribute to hybrid lethality, but not because of concordant phylogenetic divergence with their hosts. The data suggest that hybrids may be generally weakened and incapable of dealing with many free-living bacteria. There are many such examples in both animal and plants. Intrinsic hybrid dysfunction is fully consistent with the (…) [notion] that host divergence leads to defective hybrids, without invoking coadaptation between hosts and their microbiota as a driver of speciation. (…) [T]he hologenomic conjecture that incompatibilities between lineage-specific, free-living, horizontally transmitted microbes contribute to speciation remains testable speculation without experimental support' (2014: 1011a)

This criticism is more important than it might seem, since it points to one of the most serious flaws of HCE: where to put the direction of the 'causal arrow'.[35] Let me explain this with a little bit more detail. Brucker and Bordenstein have shown that germ-free reared hybrid wasps do not suffer hybrid lethality, whereas conventionally reared and bacteria inoculated hybrids die. However, they have not isolated the agent that causally explains the origin of the lethality or, in other words, where the incompatibility that generates the lethality in the first place comes from. Is there any specific bacterial species, or bacterial gene, whose interactions with the host's genome are incompatible, thus producing lethality? Only if this is plausibly shown, can it be argued that the lethality is a consequence of a host-microbe incompatibility. But the evidence that Brucker and Bordenstein present is highly unspecific and, thus, inconsequential, about the mechanistic causes that produce lethality. Insofar as their data only shows a large effect of the microbiome, the direction of the causal arrow can always be reverted: what if the observed effect (lethality in conventionally reared wasps, recovery in germ-free reared wasps) is caused by the existence of a higher degree of susceptibility in hybrid genomes when compared to non-hybrid genomes? It would be precisely the existence of the susceptibility what would explain both the observed high rate of lethality among conventionally reared wasps and the lower rate among germ-free reared wasps. It turns out that for their hypothesis about *Nasonia* to be conclusive, Brucker and Bordenstein need to prove not only a large effect of the microbiome on lethality, but the specific causal role of an agent.

Brucker and Bordenstein published a response (2014) right after the paper by Chandler and Turelli was published, where they defended the interpretation of the data that they had presented in their original paper. Particularly, they disagree that their experiments are compatible with Chadler and Turelli's 'intrinsic hybrid dysfunction', but do not prove host-microbiome coadaptation. In their view, this is an inaccurate interpretation of their data, since it presupposes—as Chandler and Turelli explicitly state—that the bacteria that were inoculated to F_2 male hybrids in *Nasonia* were random species—i.e. not necessarily phylosymbiotic. However, this was not the case, since Brucker

[35] For a very similar criticism to holobiont research and the attribution of causal powers to the microbiome, see Bourrat (2018), and Lunch et al. (forthcoming).

and Bordenstein explicitly selected bacterial species that had been isolated or taxonomically found in the wasps. Thus, Chandler and Turelli's argument that 'any bacteria' can cause hybrid lethality is unjustified and, more importantly, the observation of that result would be completely irrelevant to discard adaptation. Why do Brucker and Bordenstein believe the latter to be so? Because, in their opinion, that 'holds the ambiguous equivalence that any foreign mitochondria from different animals can cause mortality' (2014: 1011b). Let me expand on the latter point a bit. The idea to grasp here is the following: to test the possibility that mitochondria are coadapted to their hosts in a way that breaks down in hybrids it is not necessary to discard the possibility that a randomly inoculated mitochondrion from any other species would have a similar effect on hybrids than the effect of a mitochondrion from the original species. It is enough to show the effect of an inoculation with a mitochondrion from the original species. And the same applies to testing the effects that certain genes might have on hybrids. In other words, Chandler and Turelli are demanding more evidence to accept a case of hologenomic speciation than the evidence they would demand for other cases (influence of a gene, influence of mitochondrion), and this seems to be unjustified from the perspective of Brucker and Borenstein. As they claim, what they are proposing is changing the 'framework for studying the basis of any reproductive isolation mechanism' (2014: 1011b). And either the new framework is accepted, or the evidence that will be demanded to prove its validity will always be biased from the perspective of the previous framework.[36]

3.4.5. Pre/post-mating isolation and the hologenome concept of evolution. A brief refection

The experiments conducted by Brucker and Bordenstein, the theoretical apparatus they introduced and the new concepts they have used to explore the consequences of HCE are probably some of the most elegant and influential steps that have been taken to gather wide support for the heavily controversial

[36] The same type of response is given to refute Chandler and Turelli's suggestion about the cross-inoculation experiment: while Brucker and Bordenstein accept that this would be additional evidence to support the hologenomic basic of speciation in *Nasonia*, they think that demanding that kind of experiment is unreasonable, since the same experiment could be demanded for every speciation gene in animals, and that type of experiments is usually not carried out.

hypothesis that natural selection can act on the multispecies consortia that hologenomes represent. However, their interpretation of the experimental results that they obtained, as well as their application of the principles of HCE to speciation is not without contestation. Specifically, their general explanation of the microbial-assisted BDM model of reproductive isolation (**Figure 6**) allows a completely different reading, closely allied to Chandler and Turelli's interpretation of their results about the hybrid lethality in *Nasonia*. The reading would be as follows: every bacterium that interacts with a host is part of its environment. As such, some interactions will be beneficial, increasing its fitness, whereas others will be detrimental, decreasing it, or even killing the host in the most extreme cases. When, then, is it legitimate to consider the holobiont as the unit that speciates/gets selected, and when it is not? Brucker and Bordenstein need to be very specific about their way of answering this question, since their hypothesis about the BDM model does not clearly screen-off *environmental* factors from *intrinsic* factors. And this distinction is essential if the dispute is about the unit of selection *qua* replicator, which is the role that hologenomes are assumed to play in the context of HCE.[37]

The existence of phylosymbiotic patterns in nature, though, could be argued to fill in this gap. The hologenome, thus, would be the ensemble of phylosymbiotic species. But notice that this could drastically reduce the 'dimension' of the holobiont, since the phylosymbiotic species that interact with a host are just a subset (variable in size) of the species that interact with a host during its lifetime. And, more importantly, proving the existence of phylosymbiotic patterns might yet not be enough to prove that holobionts are replicators because some of the phylosymbiotic species might interact within the holobiont for environmental—i.e. not intrinsic—reasons. It seems thus necessary to distinguish which elements are environmental and which elements are intrinsic if the position that will be defended is that the hologenome is a unit of selection *qua* replicator (for a long discussion of these issues, see **chapter II**).

[37] I will provide a consistent way of explaining these results in **chapter IV, section 4.2**.

4. 'Getting the hologenome concept right'. Clarifying the eco-evolutionary principles of the hologenome concept of evolution

The formulation of HCE soon gave rise to some serious theoretical criticisms, which I will review in **chapter II**.[38] Those criticisms moved the authors working under the postulates of HCE to clarify the specific meaning of their hypothesis, which gave rise to two very important pieces in the literature about HCE: Bordenstein and Theis (2015), and Theis et al. (2016).

Bordenstein and Theis's (2015) paper summarized ten principles about the holobiont and the hologenome with the purpose of making HCE clear for future research, including possible ways of evaluating the empirical validity of the hypothesis. I think what makes their contribution more valuable is that they make clear how *not* to conceive holobionts, more than they directly clarify how positively to conceive them. This is reasonable because, as the authors have expressed: 'Holobionts and their hologenomes are less entities that elucidate something per se than they are entities that need elucidation' (Theis et al. 2016: 2). Let me briefly summarize what I take to be the key elements that Bordenstein and Theis elucidate in their paper. First of all, the authors make clear that the holobiont must be considered as a unit of biological organization,[39] meaning basically that host-microbiome associations (hologenome) should not be considered as genotype-by-environment (G x E) interactions in any of its possible interpretations (i.e. the microbiome being an environment for the host, or the host being an environment for the microbiome), nor as a phenotype encoded by the host. In their view, host-microbiome associations must be conceived as genotype-by-genotype-by-environment (G x G x E) interactions. This last point should be interpreted as stating that the holobiont acts as a single unit with its environment so that certain G x G interactions can be selected for. But, what would be the conditions that G x G

[38] The reason for presenting the clarificatory papers first, and the criticisms later, is that none of the clarificatory papers directly tackled any of the criticism 'in a philosophical way', so to speak. Rather, they just stated HCE in more specific and operationalizable terms, which help to perceive the historical development of the hypothesis.

[39] Bordenstein and Theis never explicate what they mean by 'unit of organization', but from how they continue the section it might well be assumed that their definition has nothing to do with the meaning given by defenders of the organizational account of individuality (Moreno and Mossio 2015).

should satisfy so that evolution can operate on them? Basically, that their degree of temporal persistence is enough so that evolution can operate on them, and this would happen only if there is a sufficient degree of co-inheritance. Bordenstein and Theis argue this to be so based on the evidence about horizontal transmission of the microbiome reviewed in Bright and Bulgheresi (2010), Funkhouser and Bordenstein (2013), and Rosenberg and Zilber-Rosenberg (2013) (Bordenstein and Theis: Principles I and IV).

Second, following a previous statement by Rosenberg and Zilber-Rosenberg (2013), Bordenstein and Theis clarify that the holobiont is neither an organ system, nor a superorganism, nor a metagenome. These clarifications are far from trivial: Firstly, organs are conventionally individuated because they perform one function in a system; by saying that the holobiont is not an organ system, they are neglecting the possibility of considering the microbiome as an organ, thus clarifying that it can play more than one function. And, because of this, they are simultaneously indirectly clarifying that transgenerational changes in the composition of the microbiome are not commensurable with transgenerational changes in a quantitative trait. Therefore, they cannot be studied under the framework of 'standard heritability' h^2; a different way of measuring it is required. Secondly, by clarifying that the holobiont is not a superorganism, the authors are clarifying that it is a polygenomic entity, i.e. an entity necessarily composed by organisms of different species. Thirdly, by clarifying that the holobiont is not a metagenome, they are clarifying that it is not constituted by a host plus all its 'environmental microbes', but only by a subset of those. Unfortunately, they fail to define precisely which is this subset of the totality of microbes that interact with a host that should be included in the holobiont (Bordenstein and Theis: Principle II).

A clarification of this last problem is however provided in Theis et al. (2016). According to them, the holobiont is composed of the individual host and its microbial community, and the members of this microbial community 'can be constant or inconstant, can be vertically or horizontally transmitted, and can act in a context-dependent manner as harmful, harmless, or helpful' (2016: 1). Based on this definition, they define the hologenome as 'the genomes of the

holobiont at a given point of time' (2016: 1) (**Figure 12**).[40] Notice that the definition encompasses three elements. First, a temporal timescale: the holobiont includes every symbiont that interacts with a host, no matter how long their interaction lasts ('constant or inconstant'). Second, information about the modes of transmission/acquisition of the entities that compose it. Third, information about the possible effects of the microbes of the microbiome on the holobiont. Furthermore, the definition puts the emphasis on the fact that the hologenome is a spatio-temporally located *entity* which experiences/goes through (and results from) a series of eco-evolutionary *processes*, including selection, genetic drift, genetic conflict, epistasis, etc. As such, selection on the hologenome can lead to different evolutionary outcomes, including, but not restricted to, coevolution (see also Bordenstein and Theis (2015): Principles VII and VIII).

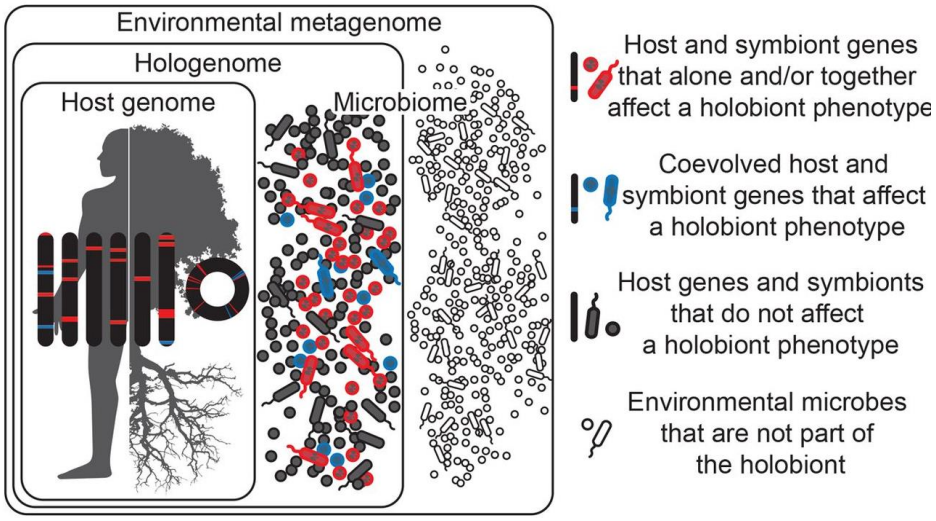


Figure 12. Schematic representation of the hologenome. The hologenome includes the total sum of genes of the members of holobiont, including the genome of the host and part of the genetic material that is contained in the microbiome (the other part would constitute an environmental metagenome). The complete genetic material of the microbiome includes a) host and symbiont genes that affect the phenotype of the holobiont but have not coevolved; b) host and symbiont genes that affect the phenotype of the holobiont and have coevolved; c) host and symbiont genes that do not affect the phenotype of the holobiont. (From Theis et al. 2016: 2, Fig. 1)

[40] This definition is different to the definition presented in Roughgarden et al. (2017), where the holobiont is defined as a host plus its symbionts, i.e. including long-term pathogens, but excluding pathogens 'which kill their host or depart in a few days'.

From the definition of the holobiont presented above follows another consequence, which is the pluralistic view of the holobiont that HCE defenders embrace. In some of the original formulations of Brucker and Bordenstein, the authors put the emphasis on the cooperative (beneficial) microbiome as the entity that, together with the host, constituted the hologenome. However, in the definition presented in Theis et al. (2016), the authors put the emphasis both on cooperation and on conflict, without restriction. In fact, that is not surprising, since defenders of the hologenome concept have always emphasized their embracing of multilevel selection theory (MLS, hereafter), the notion that natural selection targets different levels of the biological hierarchy *simultaneously*. As such, it is expected to result in contradictory effects on each level of the biological hierarchy, promoting cooperation (low conflict) in some levels, and competition in others. In principle, selection at the level of the hologenome will tend to reduce the level of conflict among the host and its microbiome, but this is not necessary to argue that the hologenome is a unit of selection, as Theis et al. emphasize, since HCE presupposes MLS. Thus, they conclude: 'Hologenomes then exist as hierarchically nested, although not necessarily integrated, levels of genomes in which all levels of selection are in play' (Theis et al. 2016: 4). This statement, as well as the rest of the clarifications about the multilevel nature of the holobiont, will be reviewed more carefully in **chapter V**.

5. The hologenome concept of evolution is a story about songs, not about their singers. A functional, but still genetic, interpretation of HCE

One of the main problems that HCE has to face is the apparent lack of transgenerational inheritance of the microbial species that compose the host's microbiome (**section 3.4.2**, **section 3.4.3**). This problem raises a question about the stability of the hologenome: does the hologenome have the sufficient degree of temporal stability so that natural selection can act *significantly* on it? Some empirical studies suggested this not to be the case, showing a high degree of transgenerational species variability (Hester et al. 2015; Taxis et al. 2015). The existence of this transgenerational variability has been interpreted by many as questioning the opportunity for significant selection on hologenomes, as I will review in detail in **chapter II, section 3**. However, this

has also produced an interesting conceptual shift, moving some authors to interpret hologenomes functionally (i.e. in terms of the genes that interact), instead of taxonomically (i.e. in terms of the species that make them up). This idea was first proposed in Taxis et al. (2015), and further developed by Lemanceau et al. (2017), Doolittle and Booth (2017), Doolittle (2017), and Doolittle and Inkpen (2018).

The notion of a functional understanding of the holobiont can be expressed as follows: the holobiont is the biological entity that *performs* a set of biological processes, including metabolic, immunological, or developmental processes, among others, in virtue of the existence of networks of functional genes whose interactions cause these processes. These processes have a high degree of transgenerational stability and can result from the interaction of different lineages (or taxa) of bacterial species, provided that the lineages that interact transgenerationally can carry out the same biological function as their predecessors—or, in other words, provided that the interacting taxa bring the same functional genes, thus giving rise to the creation of the same networks. Therefore, holobionts are units of selection *qua* interactors that lead to the transgenerational replication of different functional processes, in virtue of the replication of the functional gene-networks that have the capacity to carry out these processes.

The idea is hard to grasp, but can be elegantly captured with the following metaphor: Let's consider the biological processes that are grounded on the genetic networks as 'songs' (or as games), and let's consider the species that bear the genetic networks and interact to produce the biological processes as the 'singers' (or the players) of these songs (games). And now, let's study the evolution of songs and the evolution of their singers. For example, assume that the song is the worldwide famous '*Comme d'habitude*', first sang by Claude François. The song admits different versions, some of them slower, with a different orchestral sound, it can be sung in different languages ('My way', in English, '*A mi manera*', in Spanish, '*Precis som vanligt*', in Swedish, etc.), etc. Also, the song can be sung by different singers: Elvis Presley, Frank Sinatra, Gipsy Kings, Peter Jöback, etc. It is thus possible to talk about the evolution of

the song which, even if grounded on the singers that sing it, is somehow independent from them: the song can be sung many years after the death of some of its singers. The singers are only the interactors that guarantee the transgenerational success of the song, and of its different versions. The songs are the replicators, the entities that form the transgenerational continuous entity that evolves using its singers *opportunistically*. This is basically the idea of the song/singer model of Doolittle (**Figure 13**), which is clearly inspired by the game/players model of Taxis et al. (2015), and which the authors have summed up as follows:

> '[P]athways or more generally interaction patterns in holobionts are themselves constructed niches, created by the earliest performers of their individual steps but then setting up conditions in which very many additional taxa capable of performing the same steps (or improved versions thereof) are continually selected for (…). Because there's a song there are singers: because there are singers, there's a song. […] [And], [r]ather than seeing shared metabolisms as the products of some sort of group selection operating on individual lineages—or, in any rare mitochondria-like cases, on some hologenome—to create multilineage interactions, we imagine that such interactions already exist. Lineages evolve to carry out their steps because in each case it is selectively advantageous to individuals (*or their genes*) within those lineages to do so. There is no need to envision the independent evolution, by some onerous collective mechanism, of similar patterns in thousands of individual holobiont species.' (Doolittle and Booth 2017: 21-22, emphasis added)

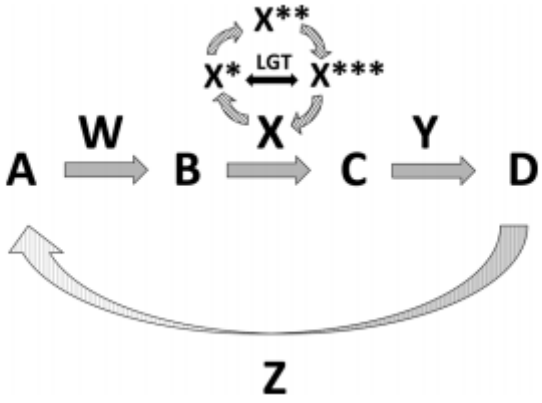


Figure 13. Doolittle and Booth's song/singer model. A-D represent the 'songs', which are the parts of a metabolic cycle, the stages in the developmental process of a host, components of a macromolecular assembly, etc. that give rise to a collective product, or *song* (A, B, C, D). W-Y are the 'singers', that is, the microbial components that are capable of producing the collective product. The singers are not required to belong to the same phylogenetic group, and are frequently interchangeable by other 'singers' (represented by X*, X**, X***), provided that those are able to make the same contribution to the collective activity. Their capability of contributing to the same cycle usually reflects the existence of lateral gene transfer between those components. Finally, Z represents the possibility that there are cyclical interactions (From Doolittle and Booth 2013: 13, Fig. 1)

The song/singers model has a clear advantage over the taxonomical conception of the holobiont, namely, that it accommodates the contradictory empirical evidence about transgenerational symbiont transmission, while keeping the one of the horns of HCE, namely that holobionts are units of selection *qua* interactors. However, by doing so, the song/singers model also gets rid of one of, probably, the most 'revolutionary' notion of HCE, namely: the role of holo*genomes* as units of selection *qua* replicators. In other words, it gets rid of the idea that natural selection operates on collections of genomes, while keeping the idea that natural selection operates *somehow* on holobionts, which would be the instruments that are used by different gene-networks to persist.

5.1. A brief reflection on the song/singer model of holobiont evolution

The song/singer model is perhaps the most elaborated and empirically adequate version of the claim that holobionts are units of selection *qua* interactors. However, it also suffers from some imprecisions: first, it is not clear how to discriminate the metabolic pathways, developmental processes,

geochemical cycles, etc. that they argue holobionts transgenerationally recreate. Is it a question of having a vague impression of which are those cycles or processes? Or is there any way of actually recognizing (or discovering) them and 'pointing them out', so to speak? Second, it is not exactly clear why or in what sense their theory is biologically relevant. HCE was introduced to explain the existence of some traits whose existence was hard (or even impossible) to explain if the individual genome was taken as the unit of selection. The song/singer model completely moves away from this explanatory context, with the sole intention of keeping the claim that holobionts are units of selection. It seems to me that, in doing so, and contrary to what they claim to be doing, they are indeed 'throwing the baby with the bathwater'. Third, the model seems to confuse cause and effect: the lineages of organisms that interact transgenerationally to form a holobiont are the effect of the existence of the metabolic pathways, geochemical cycles, etc. which use the organisms *opportunistically* for their maintenance. However, the explanation seems to work inversely, since it is hard to conceive the existence of these metabolic pathways in the first place without the lineages that interact to generate them. And, if these pathways, which are grounded on the genetic networks that instantiate them, are the replicators, then they must have been selected for as replicators precisely because the entities that instantiated them were selected for in the first place. So, what is being selected for is not the metabolic pathway itself, but the entities that interact so that those pathways obtain. And these entities are selected in virtue of these fact that these pathways confer them fitness advantages. But it is not the pathway that is selected (how can a pathway be selected?), but the set of entities that produce the pathway. Forth, and following the previous comment, the model is totally grounded in the action of single genes, which are the real focus of attention of the model, and not holobionts. The singer/song model does not really cast the holobiont as the unit of selection, but the genes that form the networks that give rise to the appearance of the metabolic pathways. In this sense, it seems to me that the model is not but another version of Dawkins' selfish gene (1976/2006, 1982a) that, as I will argue later (**chapter V, section 5**) is an obvious consequence of assuming the interactor/replicator framework for conceiving the units of selection. Fifth, but not less important, it loses track of the analogy between the genome and the hologenome, because it basically disregards any mention of the hologenome.

Proponents of the song/singer model insist that it is not useful to think of the hologenome as a coevolved entity, because in most cases it is not, but they forget the key point that defenders of HCE made: that the hologenome does not presuppose the coevolution of the species of the hologenome, only that those species will manifest a higher tendency to associate than others (H^2_c).

Despite these criticisms, here only sketched, the song/singer model is a very ingenious way of understanding the claim that holobionts are units of selection, as well as a suggestive way of making the claim coherent with our current empirical evidence. It is now time to open the discussion about the claims of HCE to other, non-evolutionary, ideas.

6. The hologenome concept as a general theory of biological individuality

The previous discussion (**section 3** to **section 5**) constituted a general review of the evolutionary claims made by the defenders of HCE, as well as their development since the hypothesis was first proposed by Zilber-Rosenberg and Rosenberg as a generalization of their observations in coral biology. This section presents a full new reading of the holobiont hypothesis, which is now conceived not as a specific theory about which units can be conceived as evolutionary individuals, but as a general theory about biological individuality. The reader might wonder in which sense it is legitimate to consider HCE as a general hypothesis about biological individuality, when the hypothesis is explicitly addressed to the role of holobionts and hologenomes as units of selection. To clarify this point, let me start by considering the following passage, taken from the introduction of Rosenberg and Zilber-Rosenberg (2013):

> ‘[The] holobiont (host + microbiota), with its hologenome (host genes + microbiome), is *a unique biological entity*, with the sum of the dynamic interactions within the holobiont giving rise to the genotype and phenotype of the organism, as we know it. The hologenome concept posits that the holobiont (host + all associated microorganisms, including viruses), being

> a unique biological entity, acts also as a level of selection[41] in evolution' (2013: viii; emphasis added)

The passage contains two different types of claim: First, that there are certain 'high-level entities' that we should call holobionts, and which act as 'a unique biological entity' (the type of biological entity not being specified); second, that these entities are *therefore* a unit of selection in evolution.[42] These claims, despite being presented as if the later were a sort of logical consequence of the former, are deeply problematic and not as easily intertwined as they assume. What is this 'unique biological entity' that Rosenberg and Zilber-Rosenberg refer to, if it is somehow different from a unit of selection in evolution? HCE will now be interpreted as a general theory about how to conceive the phenomenon of biological individuality, a thesis that I will call the 'generalized individuality thesis', in contrast to the 'particular individuality thesis' that holobionts and their hologenomes are units of selection. To that aim, I will review Dupré and O'Malley (2009), Gilbert et al. (2012), and Roughgarden et al. (2017).

Dupré and O'Malley, with their 'collaborative' view of life, were the first authors to point out, to a philosophical audience,[43] the importance of symbiosis to define one of the key aspects of multicellular forms of life, namely: their functionality as a single whole (see also O'Malley and Dupré 2007; O'Malley 2014). They introduce their point with the following reflection:

> 'Some biologists and philosophers may prefer to define multicellularity in ways derived from reflection on animals and plants, and thereby exclude these microbial communities from that category. But certainly any general

[41] Notice that 'unit of selection' and 'level of selection' are frequently used interchangeably, and so do Rosenberg and Zilber-Rsoenberg (2013).
[42] I think this is the most convincing way of interpreting the use of the paraphrase in the second sentence of the passage, as stating a sort of causal claim that connects the role of holobionts as 'unique biological entities' with their role as units of selection.
[43] Dupré and O'Malley were of course not the first to point out the importance of symbiosis and its relation to biological individuality (concretely, how fuzzy symbiosis makes some of these boundaries). These reflections started at least with Anton de Bary and have continued since. I review some of these original debates in Suárez (2018a): Part I. In this section I start with the work of Dupré and O'Malley because they are the first to point out the importance of symbiosis —and also microbes—to philosophers.

> account of the varieties of biological organization will need to take account of them and explain how they conform to concepts such [as] “multicellularity”, “invididuality” and “autonomy”. Do humans, for example, stop at their skin and have to be conceived of as tubular rather than solid in order to avoid incorporating large internal populations of gut microbes? Lederberg, with his concept of ‘symbiome’, raises the question of whether organisms are necessarily monogenomic or whether a multi- or metagenomic state is the usual state of organismal organization (Lederberg, in Hooper and Gordon 2001; Dupré and O’Malley 2007). Discussions of life and its organization have to take into account the fact that symbiotic relationships are ubiquitous and all organisms, when conceived as the functional wholes that interact with their surroundings, are multi-lineal and multigenomic.’ (2009: 11-12)

From this observation, Dupré and O’Malley derive the following conclusion: the biological world is necessarily composed of two different and non co-extensional types of entities: on the one hand, metabolic-forming entities; on the other, lineage-forming entities. They think that metabolic-wholes are multi-lineages composites, that result from the interaction of collaborative (i.e. including both competitive and cooperative) lineage-forming entities whose interactions give rise to the emergence of the functional whole that we call ‘biological individual’. And these metabolic-wholes are, for Dupré and O’Malley, the most fundamental unit of selection, something they express in the following manner:

> ‘[F]unctional entities are, rather, associations of a variety of such lineage-forming entities. A typical large eukaryote, for instance, is constituted by entities of all the kinds we have distinguished above [spatially bounded entities, evolvable entities, reproductive entities, etc.]. We might invoke here David Hull’s (1980) well-known distinction between replicators and interactors, but in a very different way from that originally supposed by Hull. Interactors, in our view, are complex systems involving the collaboration of many highly diverse lineage-forming entities. This sort of

> interactor, we also suggest, is *the most fundamental unit of selection*. This perspective has radical implications for the way we think about evolution.' (2009: 13; emphasis added)

Notice that Dupré and O'Malley specify here the meaning of one of the claims made by Zilber-Rosneberg and Rosenberg in their first formulation of HCE. Remember that, for the authors, holobionts are interactors (**section 3.1.3**). But, what are interactors? Dupré and O'Malley give a clear response to this question: interactors are the collaborative multi-lineage entities that associate to form a functional entity (and the 'unique biological entity' of Rosenberg and Zilber-Rosenberg). Nonetheless, two important points should also be emphasized: First, Dupré and O'Malley's paper does not contain any specific mention of the holobiont, nor of HCE. The most similar claim that can be found is their mention of the 'symbiome', and their reflections about the gut microbiota. So, even if their paper could have many implications for HCE, and has been used as evidence in support of the claim that holobionts are units of selection, some caution is required—especially since Maureen O'Malley has recently expressed serious doubts about the possibility that holobionts are units of selection (2016; personal communication). Second, the authors make a clear distinction between metabolic-forming wholes (interactors), and lineage-forming wholes. This is important because it can be read as a claim against the role that these 'collaborative multi-lineage wholes' are units of selection, if forming a lineage is a condition for units of selection, as in Lewontin's formulation. And, furthermore, it can be read as a claim against the definition of the hologenome as a replicator. Recall that the role of the hologenome as a replicator is a key assumption of HCE. As Dupré and O'Malley's paper is written, it is unclear how to understand that passage in terms of replicators/interactors, as well as in terms of HCE. Nevertheless, their paper makes an essential contribution by distinguishing between the notions of metabolic-forming wholes and lineage-forming wholes, which points to the non-overlapping nature of different conceptions of biological individuality.

A second key development to disentangle the meaning of the rather obscure phrase that qualifies holobionts as 'unique biological entities' came

from the work of Gilbert et al. (2012) and Roughgarden et al. (2017).[44] In their view, none of the standard criteria for individuality make sense except in the light of symbiosis: since macrobes interact during their lifespan with a wide range of microbial symbionts, it is not possible to make any definition of biological individual that excludes them as structural parts of the macrobe. Therefore, macrobes have never been independent individuals, but integrated host-microbiome ensembles (see also Dupré 2012). Let us start to show why this is so from the beginning.

Anatomical individuality

If we define biological individuality anatomically, i.e. as a structured whole where different parts work together to maintain its anatomical structure, it soon becomes clear that, for macrobes, the individual is constituted not only by the cells of the macrobe, but by the cells of the macrobe plus all the interacting symbionts. First, all macrobes interact with a wide number of microbes during their lifetime. Second, microbes occupy different compartments of the macrobe's anatomy, in many cases occupying different body parts depending on their properties. In many animals, for example, the guts are mainly composed of microbial elements, with different microbes presents in different niches, which help them in their hosts in their (usually highly inefficient) processes of digestion. In humans, the microbiome has been estimated to contain approximately the same number of cells as the body cells, with an average weight of about 0.2 kilograms (Sender et al. 2016). This led Roughgarden et al. to conclude that from an anatomical perspective, animals are holobionts, anatomically composed of host cells and bacterial cells. I will refer to this entity as the 'anatomical holobiont'.

[44] I directly combine the claims made in the paper of Gilbert et al. (2012) with the claims made in Roughgarden et al. (2017), since both papers are complementary: the first, providing the 'negative' argument that none of the dimensions of biological individuality would make sense without symbionts; the second, making the point of why host-symbiont ensembles should be considered biological individuals.

Physiological individuality

'Physiology' refers to the study of how organisms realize their normal activities and functioning, including metabolism, nutrition, and all the biochemical pathways that lead to them. The opposite of a physiological state is a pathological state. Biological entities can be individuated by attending to their physiological properties (i.e. to how they fuction normally), thus leading to the 'physiological individual', a very similar concept to Dupré and O'Malley's 'metabolic-forming' wholes, reviewed above. Gilbert et al. (2012) and Rouhgarden et al. (2018) argue that macrobes cannot be considered physiological individuals in themselves simply because their normal (i.e. non-pathological) functioning is the result of their interaction with the microbial symbionts of their microbiome. Indeed, scientific investigation has extensively proven the essential role of the microbiome for processes such as nitrogen fixation, cellulose degradation, photosynthesis, oxidation of inorganic compounds, synthesis of essential aminoacids, detoxification of poisonous chemicals, synthesis of metabolites, and so on and so forth. The authors believe that this evidence strongly supports the considerations of the microbial symbionts as essential components of the physiology of macrobes and, therefore, as a unified physiological entity, which could be called the 'physiological holobiont'. Importantly, sceptics of HCE such as Godfrey-Smith (2009, 2013, 2015) have argued that holobionts are *merely* physiological individuals, with no evolutionary significance.

Developmental individuality

A developmental individual is what proceeds 'from ovum to ovum', i.e. from a zygote to the fully 'fleshed' biological entity which has the capacity of producing new zygotes and that we refer to as 'individual' or 'organism'. In other words, a developmental individual is the entity that unfolds during a life cycle.[45] Holobiont defenders argue that no macrobe can be considered an autonomous developmental individual, since microbes are key elements in macrobe development (Gilbert and Epel 2015; Gilbert and Chiu 2015; Gilbert 2018). For

[45] In many cases, determining when the life cycle ends, and a new life cycle starts can be complex, especially in organisms with haplodiploid life cycles (see O'Malley 2016). I will put these problems aside in this discussion, though.

instance, the microbiota activates the genes that lead to the maturation of the gut and immune system in mice and zebrafish, and the same effect has been observed in the case of brain maturation in mice. The observation of these and other cases has lead Scott F. Gilbert (2018) to argue that every process of animal development is indeed a process of co-development between the macrobe and its microbiome (**Figure 14**). Therefore, by developmental criteria, macrobes cannot be considered individuals, but holobionts. I will call this notion the 'developmental holobiont'.

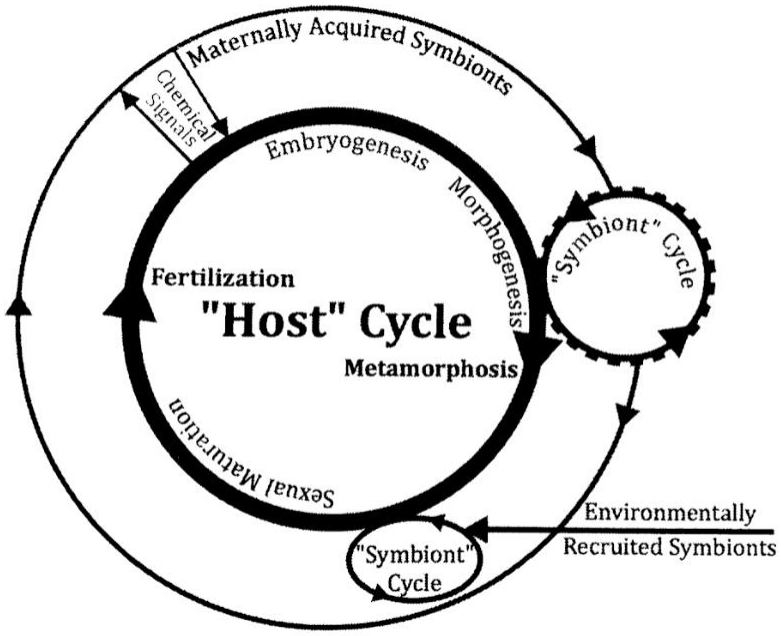


Figure 14. Gilbert's developmental holobiont. The figure shows how the traditional developmental stages of the 'host', including fertilization, embryogenesis, morphogenesis and sexual maturation include their interaction with symbionts, through different chemical signals. Most symbiont cycles are indeed shorter than the 'host' lifecycle, and, for Gilbert, symbionts can be maternally acquired or environmentally acquired. They will be considered part of the 'developmental holobiont', provided they interact with the holobiont during any of its developmental stages (From Gilbert 2018: 299, Fig. 12.1)

Immunological individuality

A classical manner of defining biological individuality is by immunological criteria, the reason why some have referred to immunology as 'the science of the self/nonself discrimination' (Klein 1982). The immunological system has traditionally been understood as a defensive system, a type of army, that

prevents the action of pathogens or any other sources that might cause damage to the 'self'. However, more recent views on immunology have seen the immunological system as the key element in defining the constitution of biological entities, i.e. as the system that defines the boundaries of the biological individual (Pradeu 2012, 2019; Pradeu and Vivier 2016). Defenders of the holobiont argue that, on the one hand, the immunological system of macrobes is created due to the active participation of their microbiome whereas, on the other, the immune system actively recruits and tolerates its symbionts, sometimes even compartmentalizing them in special structures of their bodies, such as bacteriocytes in insects. Therefore, from an immunological perspective, the holobiont is a biological individual, the 'immunological holobiont'.

6.1. A brief reflection on the 'generalized individuality thesis'

The 'generalized individuality thesis' is much more specific and, thus, more useful, than the rather unclear statement that holobionts are 'unified biological entities'. However, one might still argue that the generalized individuality thesis is orthogonal to issues of individuality, for, at most, it says that the entities that participate in physiological processes, developmental processes, immunological processes, etc. are more complex than what we think of as biological individuals (i.e. the macrobe). A version of this criticism will in fact be examined in detail in **chapter II**. For the moment, it is enough that the reader realizes of this difficulty and, especially, that she realizes that the speech about immunology, development, physiology, etc. is talk about individual biological *processes*, but it is not necessarily—at least according to some authors—talk about individual biological *substances* (Dupré 2012; Bapteste and Dupré 2013; Bouchard 2018).

Second, it is important to notice that the 'generalized individuality thesis' has been presented in relation to the individuality of macrobes (particularly, animals, although the argument can be extended to plants). In other words, the argument is basically that standard individuation criteria—i.e. the appeal to the biological processes of physiology, development, immunology—capture entities that are way more inclusive than the monogenomic entity that 'unfolds' from an original zygote. Nothing is argued about the individuality of the microbes that interact symbiotically with the host, an individuality that HCE neither questions,

nor it is required to question. HCE, conceived as a ‘generalized individuality thesis’, is a thesis about the individuality of the macrobe, not about the individuality of the microbes. This point is crucial for the argument I will develop in **chapter IV**.

7. Brief summary of chapter I: The hologenome concept of evolution as a set of distinguishable hypotheses

This chapter dealt with the historical development of the HCE, since the hypothesis was originally formulated by Ilana Zilber-Rosenberg and Eugene Roseberg as a generalization from their research on corals. The chapter has shown how the concepts used have been made more precise, how different ways of testing the hypothesis have developed and it has explained how some experiments have been carried out to test the validity of some of the claims made by the defenders of HCE. The content of the chapter can be summarized in the following key points:

1. HCE constitutes a new framework to study host-microbiome interactions, as well as a new way of conceiving these interactions.
2. The holobiont is a biological individual and a unit of selection *qua* interactor.
3. The holobiont can be conceived as a biological individual in several dimensions, including metabolically, immunologically, anatomically, and developmentally.
4. According to the standard HCE model, the hologenome is a unit of selection *qua* replicator, and it has to be modelled as a set of G x G x E interactions.
5. According to the song/singer model, hologenomes are not units of selection *qua* replicators, but holobionts are units of selection *qua* interactors.
6. Phylosymbiosis refers to the eco-evolutionary pattern by which the evolutionary changes in the host’s phylogeny associate with parallel changes in the microbiome that associates with the host.

7. Phylosymbiosis can be used as a proxy for selection on hologenomes by means of the notion of community heritability H^2_c.
8. The discovery of phylosymbiotic patterns is indicative of the action of natural selection on holobionts.
9. Changes in the hologenome can be ascribed causal responsibility in speciation events under the BDM model.

Chapter II

‘Problems with hologenomes. Three types of criticism to the individuality thesis’

The previous chapter explored the historical development of the hologenome concept of evolution—including a historical hypothesis about where some of its eco-evolutionary ideas might have come from—and sketched some of the early criticisms that had been made to its specific application to corals and *Nasonia* wasps. This chapter deals with the most important conceptual criticisms that have been raised against the hypothesis, understood now as a general thesis about the evolution of animals and plants. The chapter first discusses the problem of definition of the holobiont during the host’s lifetime (how to define holobionts as individuals ontogenetically), then connects this discussion with the evolutionary dimension of the holobiont (how to define holobionts evolutionarily), and finally discusses the problems of considering holobionts as units of selection (how can holobionts actually get selected).

1. Introduction

Since its original formulation, HCE has been subjected to several criticisms. The main criticisms it has received can be basically classified in two groups: first, criticisms concerning the logical/conceptual foundations of the hypothesis; second, criticisms about the empirical possibility of having real cases of selection on hologenomes, including criticisms about the scarce evidence supporting HCE, and the low likelihood that real events of hologenome evolution could actually happen, given our current empirical knowledge. In the literature, these different types of criticisms are not usually distinguished and are taken to lie on a continuum of problems that defenders of HCE must face for

their hypothesis to be sound. For the purposes of my doctoral project it is however crucial to distinguish the two kinds of problems, because, as already indicated, my project mostly cares about the conceptual foundations of HCE, i.e. about what might be the conception of biological individuality that might justify—if possible—such an enterprise. This does not mean, though, that the empirical problems that the application of the HCE to empirical biology presents should not have any impact on my analysis: in the end, empirical and conceptual problem *do* lie in a continuum. Nonetheless, in this chapter, I will exclusively analyse the conceptual arguments that have been made against HCE (**Introduction**).

In this chapter, I will first introduce a general problem about the definition of the holobiont, namely, the 'blurry' nature of the holobiont. Secondly, I will connect this criticism to the problem of defining holobionts as evolutionary entities. Finally, I will connect this criticism to the general problems that have been raised against the consideration of holobionts as units of selection.

2. A problem of definition. Holobionts as 'blurry' entities

A serious criticism to HCE concerns the *'blurry'* nature of the holobiont. By the problem of the 'blurry' nature of the holobiont I refer to the following: Defenders of HCE typically claim that the holobiont encompasses an animal or a plant—that plays the role of the host—plus its symbiotic microbiota (**chapter I, section 4**). However, the symbiotic microbiota of hosts is known to be very transient, contingent, and highly dependent on environmental factors such as host's diet (Hester et al. 2015; Zhang et al. 2016; Hallen-Adams and Suhr 2017; Adair and Douglas 2017). If this is so, then any hypothesis about the individuality of the holobiont will necessarily 'suffer from imprecision' (Booth 2014: 670), or so the argument usually goes, since the holobiont will not keep its identity over time (or, at least, its identity will not be the same as the identity of the host, at least not during a host's lifetime, nor is it clear how to mark different moments/stages during a holobiont's lifetime). These types of claims are present in different ways in Booth (2014), Chiu and Eberl (2016), Queller and Strassmann (2016),

and, to a lesser extent, in Bourrat and Griffiths (2018). Since all the criticisms have a very similar structure, I will concentrate here on the criticisms made in Chiu and Eberl (2016).[46]

2.1. The question of proper parts: Microorganisms are not proper parts of the host

Holobionts are composite objects, i.e. they are entities composed of different and independent parts (host, endosymbionts, microbiota) that *arguably* integrate with each other to form a higher-level object. As said in **chapter I**, defenders of HCE argue that holobionts are biological individuals, which creates the problem of explaining how the parts of the holobiont integrate with each other so that they can be considered *proper parts*. Chiu and Eberl reformulate this problem as a question of *'glue'*: what is the 'special glue' among the parts of the holobiont (i.e. microbiota and macroorganism) that allows us to consider the holobiont as a composite individual?[47]

[46] The choice of this paper as representative of claims about the 'blurry' nature of holobionts is based on three reasons. First, I consider it to be the most clearly elaborated version of the argument. Second, because the claims made in Chiu and Eberl's work seem partially different to the claims made in the other papers, which will be partially addressed in section 3 and section 4 of this chapter—concretely, Queller and Strassmann made their criticism in the context of the cooperation/conflict definition of organismality, which is the conception that I argue grounds the criticisms of the role of holobionts as units of selection (section 4); Booth's discussion is more general, and it is connected to at least two different conceptions of the role of holobionts as units of selection, which will also be discussed later; finally, I will use section 3 to discuss Bourrat and Griffith's ideas. Third, because I have drafted a critical response to Chiu and Eberl's specific account of the holobiont, together with Vanessa Triviño (University Rey Juan Carlos de Madrid), submitted to *SHPS: Part C*, which makes me much more familiar with their way of criticisms than with the criticisms made by the others. This section is importantly grounded on that draft.

[47] In this section, I have decided to introduce the problem as Chiu and Eberl do it, as the problem of the 'glue' that ties together all the members of the holobiont, since I think it is the best way of guiding the discussion. The reader will notice that there is still another logically possible way of introducing the problem, which I have decided not to consider here for argumentative purposes: the problem introduced as the question of 'glue' presupposes that there *must be* a biologically non-trivial manner of arguing that the elements of the holobiont form a relatively cohesive conglomerate, or that a biologically non-trivial criterion *needs to be* introduced for *at least* some of the members of the host's microbiota. However, there is the possibility of directly denying that there is any biologically non-trivial manner of doing so. This position is partially maintained by Queller and Strassmann (2016), who accuse holobiont defenders of basing their hypothesis solely on anatomical criteria, which are not necessarily biologically relevant. In their words 'The holobiont is defined by spatial criteria. There is no reason to believe that spatial proximity necessarily leads to functional integration' (2016: 869). 'Functional integration' should be understood here as a general reference to any relevant biological property.

To answer this question, Chiu and Eberl propose what might be called an 'immunological criterion of parthood' (ICP, henceforth), which states that 'what makes a microorganism part of an organism is not its taxonomic or functional properties, but *whether it is interconnected* with host components through the biochemical interactions of the immune system' (Chiu and Eber 2016: 822, emphasis added). ICP is grounded on two recent theories about immunity: (a) the discontinuity theory of immunity (DT, hereafter) (Pradeu and Vivier 2016) and (b) the equilibrium model of immunity (EM, hereafter) (Eberl 2016). According to DT, the immune system of an organism determines both its constituent parts (*criterion of inclusion*) and the conditions in which the organism is actively maintained in the face of constant external perturbations (*criterion of persistence*). Regarding EM, immunity should be conceived in terms of the types of reactions (signals) that are triggered as a response to the different targets that the immune system faces: intracellular signals, small extracellular signals, large extracellular signals.

Following the criteria of persistence and inclusion provided by DT, Chiu and Eberl argue that the microbiota, together with the immune system of the host, allows the persistence and individuality *of the host itself* (Chiu & Eberl 2016: 820): First, the immunity system of the host allows for the host individuality insofar as it determines its constituents. Second, throughout the immune interactions of the microbiota with the host, the individuality and persistence of the latter is guaranteed.

However, despite claiming that immunity allows to establish the individuality of the host by determining its constituents, and although the host immunity is partially activated by microorganisms, the authors do not consider microorganisms as constituents (proper parts) of the host, or as integrated with it to form a higher-level entity—i.e. a holobiont. That is where ICP enters in their picture of individuality: According to Chiu and Eberl, the microorganisms that integrate the microbiota of a host can be proper parts of the macroorganism host only in virtue of the *continuity* of host-microorganisms immune interactions. However, so their argument goes, our best current scientific evidence suggests that the interactions between the microbiota and the host are not continuous:

Insofar as most microorganisms are changeable, transient and sometimes, shared by other systems and processes, they are not continuously interacting with the immune system of the host, and therefore, they are not integrated into a single functioning and reproducing whole, even when they might occasionally have some necessary effects for the host's survival, or even when microorganisms can substitute an entire organ of a host. In their own words:

> '[T]hese fascinating cases [influence of the microbiome on the survival of the host] at most dispel the notion that macroorganisms are self-sufficient without microorganisms. They fall short from showing that holobionts are causally integrated metabolic or reproductive wholes [because] (…) [n]ot only do holobionts contain microorganisms that have negative or no effects on host phenotypes and reproduction, the selective cases of reciprocally beneficial relations are not necessarily mutualistic or cooperative' (Chiu & Eberl 2016: 821)

This last claim entails that microorganisms cannot be considered proper parts of the host. The holobiont is, thus, a 'blurry' entity, which lacks clear boundaries and clear criteria of membership/parthood and, therefore, it is not a biological individual. The host is the genuine biological individual, with its microorganisms playing the role of external resources that the host uses *opportunistically* to stay alive.

The argument presented by Chiu and Eberl shows an important flaw in HCE thinking: most HCE defenders would argue that the host plus its microbiota constitute an individual, but they fail to notice the transient nature of many of the members that compose the microbiota of a host, and how sensitive these are to variable environmental conditions. This seems to make them more similar to external resources or background conditions than to internal components of a higher-level composite object. This criticism had been advanced by Leggat et al. (2007) (**chapter I**), when they opposed the holobiont model to the hologenome model. Notice that, in both cases, the problem that the authors highlight only affects the claims made by defenders of HCE, i.e.

defenders of the hologenome model. Defenders of the holobiont model would acknowledge the transient nature of microorganisms and would in fact claim that this supports their views, insofar as they are not committed to asserting that holobionts are individuals: they only claim that they are ecological communities. In any case, the argument presented by Chiu and Eberl suggests that there are some difficulties in considering holobionts as individuals, and does so from an immunological perspective, i.e. without having to link the general idea of biological individuality with the specific concept of evolutionary individuality, whose assertion for holobionts constitutes the main claim of HCE. Now, the important question to ask is this: is Chiu and Eberl's argument a knock-down argument? Or, in other words, does the 'blurry' nature of the holobiont completely undermine the consideration of holobionts as biological individuals?

2.2. A reflection about Chiu and Eberl's criticism of the individuality of holobionts

Chiu and Eberl's problem is not about the empirical data that suggest that microbes perform really important functions for their microbes. Their concern is about the interpretation of those results as evidence for the claim that microorganisms are proper parts of the host and thus the holobiont is a biological individual. However, their interpretation of the evidence does not seem to be entirely conclusive, presenting some serious flaws that I will only summarize here. Firstly, take the case where microorganisms replace completely an organ of the host. According to Chiu and Eberl, this case should be interpreted *'at most'* as dispelling the notion that macrobes are self-sufficient by themselves. But notice that the same reasoning can be applied to every single organ of a macrobe, no matter whether it is entirely composed by microorganisms or entirely composed by host cells. In this vein, it can be argued that the heart of a macrobe is not a proper part of the macrobe, but a structure that dispels the notion that the macrobe is independent without a heart. And the same reasoning can be applied to the lungs, to the pancreas, to the liver, and so on and so forth. Chiu and Eberl's argument to discard the possibility that microbes are proper parts of the host while accepting the empirical evidence that shows their importance seems either inconclusive or very counterintuitive.

Second, the argument based on the discontinuity of the interactions seems partially inconclusive as well since, again, the same logic can be applied to the host-derived structures, giving raise to very counterintuitive results. The thesis of discontinuity can be interpreted in at least two different ways: first, talking about *tokens*, i.e. individual microbes or cells; second, talking about *types*, that is, about the interacting species or genomes. If the discontinuity thesis is interpreted as a thesis about tokens, then its results are counterintuitive: Every macrobe, to be maintained alive, requires that its cells are constantly renewed in a process known as 'cell turnover'. The existence of cell turnover entails that the interactions between the immune system and the organs of the host are, strictly speaking, discontinuous, and thus they do not constitute proper parts of the host according to ICP. On the other hand, if it is a thesis about *types*, then the thesis is inconclusive. Why should the change in types, understood as changes in the interacting genotypes, be considered so relevant for decisions about 'proper parts'? If a cell in a macrobe mutates and it spreads and starts dominating in a tissue, it would still count as a proper part of the host. And so do organs that have been transplanted from another person. Arguing that cell types can change during the lifetime of a macrobe is, thus, inconclusive to prove that they should not be taken as a proper part.

Of course, what I have just presented is a sketch of the counterarguments that can be raised against the thesis that because holobionts are 'blurry' entities, they cannot be considered biological individuals, and they would need more elaboration to be seriously considered.[48] But at least it proves that both defending and discarding any claim about their status as biological individuals presents serious problems, and neither of the hypotheses can be discarded trivially. The next section will show how the arguments about the 'blurry' nature of holobionts can be related to the arguments against the thesis that holobionts are evolutionary individuals.

[48] For an elaboration of these ideas, see my draft with Vanessa Triviño, and for a full new account that explains away Chiu and Eberl's problem, see my draft with Adrian Stencel, as well as **chapter IV** of this dissertation.

3. From the problem of the 'blurry entities' to the problem of holobionts and their hologenomes as evolutionary individuals

Bourrat and Griffiths (2018) have recently wisely connected the problem of the 'blurry' nature of the holobiont with the problem of conceiving holobionts as evolutionary individuals. According to Bourrat and Griffiths, the arguments presented by HCE defenders: (I) fail to correctly discriminate between holobionts and other random multispecies assemblages that are not holobionts and, consequently, (II) fail to support the thesis that holobionts are biological individuals, if biological individuality is understood evolutionarily.[49]

3.1. HCE arguments as 'part of the system' arguments

According to Bourrat and Griffiths, the arguments presented by HCE defenders constitute cases of what they call 'part of the system' arguments. 'Part of the system' arguments must be understood here in connection to Chiu and Eberl's concept of 'proper parts'. Bourrat and Griffiths' idea is that, no matter if the arguments about the symbionts being 'proper parts' of the holobiont are successful, they will still fail to prove that holobionts are evolutionary individuals.

'Part of the system' arguments would have the following logical structure: 'such-and-such components are essential to the functioning of some larger system, therefore those components are part of that system'. This line of reasoning, applied to holobionts (*system*) and their symbionts (*components*) might be explicated as follows:

(i) Symbionts *x, y, z* are essential for the functioning of the holobiont *H*
(ii) Therefore, *x, y, z* are parts of *H*.

[49] This section bears on my draft 'On the individuality of holobionts and other multispecies assemblages: A critical response to Bourrat and Griffiths', submitted as a response to Bourrat and Griffiths in *HPLS*. See also **chapter IV** for a response to their criticisms**.**

Notice that the argument contains an additional elliptical premise that is required to be valid, namely:

(iii) If a symbiont *X* is essential for the functioning of a holobiont *H*, then *X* is part of *H*.

Furthermore, Bourrat and Griffiths assume that when defenders of HCE defend (ii), what they are indeed claiming is that:

(iv) Therefore, *H* is a biological individual.

And, importantly, for Bourrat and Griffiths: (a) an entity is a biological individual if and only if it is a unit of selection ('biological individual = unit of selection'); (b) an entity is a unit of selection if and only if all the 'parts' of the unit have their fitness interests aligned ('unit of selection = fitness alignment'). To quote: 'the fundamental issue in identifying new levels of biological individuality should be whether some entity can function as a unit of evolution, which will depend on the fitness alignment between the partners over evolutionary timescales' (Bourrat and Griffiths 2018: 2).

Once the debate has been framed, Bourrat and Griffiths dedicate the rest of the paper to argue that (iv) does not follow from (i) and (iii), at least if (i) and (iii) are supported by the kind of evidence that defenders of HCE have used to justify them, simply because (iii) would be false. This is because the arguments use by HCE defenders are not directed to the right kind of properties that would justify considering that a system is a biological individual, but rather to other types of properties which are common to many types of relationships between organisms of different species, and even to relationships between an organism and its abiotic environment (e.g. the gravitational field; Bourrat and Griffiths 2018: 5). As they stand, the authors argue, those properties are not genuine markers of biological individuality. This leads to the second major criticism of

Bourrat and Griffiths to HCE defenders: holobiont advocates fail to provide a satisfactory theoretical criterion that allows them to distinguish between holobionts *qua* biological individuals and other multispecies assemblages *qua* consortia of interdependent biological individuals. In their view, the real criterion for determining whether a biological consortium is a biological individual is the fitness alignment among the parts that constitute it, an argument that defenders of HCE have not made. Therefore, defenders of HCE have failed in their defence of the notion that holobionts are units of selection, and a new criterion based on the concept of fitness alignment must be made explicit that satisfies two roles: (I) it allows divorcing holobionts from other multispecies assemblages, and (II) it allows supporting the thesis that (some) holobionts are biological individuals.

3.2. Bourrat and Griffiths' criticism to the 'part of the system' arguments

Bourrat and Griffiths believe that the arguments presented by defenders of HCE to support (i) are based on (false) premises, which implies that the deduction of (iv) does not work. In their view, defenders of HCE have used three sources of support for (i): metabolic (physiological) arguments, developmental arguments, and immunological arguments. The *metabolic argument*, especially emphasized by Dupré and O'Malley (2009), although also present in other formulations of HCE (Gilbert et al. 2012, 2017; Rosenberg and Zilber-Rosenberg 2013, 2016; Roughgarden et at. 2017; **chapter I**) basically states that insofar as macrobes engage in numerous interactions with their symbionts to achieve the basic goal of processing the external resources of their environment to keep the system (host + microbes) alive, the whole system should be considered a biological individual. Furthermore, in the case of Dupré and O'Malley the argument is complemented with the idea that those systems should be considered 'the most fundamental unit of selection' (2009: 13). To reach their conclusion, Dupré and O'Malley rely on the distinction between interactors and replicators (Hull 1980; Lloyd 2017a), adducing *only* that holobionts are interactors, and making clear that holobionts are systems composed of independent 'lineage-forming entities' (*sic*).

Bourrat and Griffiths do not have any concern with Dupré and O'Malley's defence of the claim that holobionts are interactors (cf. Booth 2014; Skillings 2016; Queller and Strassmann 2016). Their only concerns are: Firstly, that so stated the definition would also be applicable to other multispecies assemblages that are not holobionts; secondly, that 'although the notion of an interactor is useful in certain contexts within evolutionary theory (…), it is often of very limited value in deciding whether a multispecies entity is an individual in its own right' (2018: 11). Their first observation is supported by comparing the metabolic system of the worm *Olavius algarvensis* and its microbial symbionts, with the system formed by the seagrass meadow, bivalves of the Lucinidae family and their endosymbiotic sulphide-oxidizing bacteria, and the sulphate-reducing bacteria that reside next to the roots of the seagrass. Bourrat and Griffiths' point is that although the metabolic processes—including some extremely complex biochemical exchanges—that occur within the *O. algarvensis* and those that occur among the members of the seagrass meadow multispecies assemblage are basically the same, only the former but not the latter would be considered a holobiont by HCE defenders. Then, there would be no theoretical ground to distinguish between holobionts and other multispecies systems, apart from a basic intuition of what seems to be more an individual from our perspective.[50] And now, and always according to Bourrat and Griffiths, even assuming that holobiont advocates had a convincing argument to show why one system is a holobiont whereas the other is not, they would have only proven that holobionts are interactors. But because the notion of interactor is irrelevant for deciding whether a system is a biological individual,[51] holobiont advocates have totally missed their target, and their evidence is simply not to the point. Therefore, no matter what, HCE advocates fail to provide a satisfactory criterion to support why the metabolic argument leads to the conclusion that holobionts are biological individuals.

[50] Similar claims can be found in late 19th and early 20th century history of biology, always directed against those biologists that emphasized the integrative nature of symbiosis. Maybe the most interesting case is Pound (1893), reviewed in Suárez (2018a: 80-82).

[51] Their observation about the irrelevance of interactors for discussions about biological individuality, which I consider necessary for drawing their intended conclusion, is left unargued in their paper (for good arguments to that end see Okasha 2006; Godfrey-Smith 2009). Nonetheless, the reader should not assume that I also share their very same opinion about the usefulness of interactors for evolutionary biology.

The other two arguments that Bourrat and Griffiths discuss (developmental, immunological) have the same conceptual structure, so to keep my thesis as parsimonious as possible, I will not consider them here. Bourrat and Griffith's message is nonetheless clear: even if (some) holobionts could be proven not to be the 'blurry' entities that Chiu and Eberl, among others, have claimed them to be, proving that holobionts are interactors would not settle the question about their role as evolutionary individuals yet. What matters for an entity to be considered an evolutionary individual is that all the parts that integrate the entity have their fitness interests aligned. Insofar as this is far from being the case in holobionts, Bourrat and Griffiths argue, holobionts are not evolutionary individuals. Now that the problem of the 'blurry' nature of the holobiont has been related to the problem of evolutionary individuality (conceived as the consideration of the hologenome as a replicator), the next section will examine a bit deeper how the claims about 'fitness alignment' connect with the main claim of HCE, namely: that holobionts and their hologenomes are units of selection.

4. Some major problems with the role of hologenomes as units of selection: the problem of (the lack of) partner fidelity

The main argument against the role of hologenomes as units of selection *qua* replicators is based on the lack of proper transgenerational transmission of the entities that compose the holobiont. This thesis had been anticipated in the work of John Maynard Smith (1991): 'A Darwinian view of symbiosis' (reviewed in **chapter I, section 2**), and since the proposal of HCE, different versions of the thesis have been defended by Moran and Sloan (2015), Douglas and Werren (2016), Skillings (2016), Hurst (2017), and Bourrat and Griffiths (2018).[52] The main criticism that these authors have raised can be summarized as follows: HCE assumes that the hologenome is a unit of selection. As hologenomes are composed by a host's genome plus its microbiome, HCE entails the coevolution of the complete set of genomes. This possibility would

[52] Peter Godfrey-Smith (2013, 2015) has also made a consistent and well-argued criticism of the application of 'units of selection thesis' to holobionts. However, his arguments are based on his rather idiosyncratic definition of 'units of selection', based on his concept of Darwinian populations, so I will not consider them in this chapter, but in **chapter III** and **chapter V**.

be feasible if and only if the genomes that constitute the hologenome are inherited, i.e. their association is transgenerationally maintained (partner fidelity). Expressed in the words of Douglas and Werren:

> ‘Partner fidelity is a prerequisite for the hologenome [concept], because the host and its microbial partner(s) can only evolve as a unit if they co-occur across multiple host generations, with tight host genotype-to-microbe genotype matching.’ (Douglas and Werren 2016: 2)

By ‘partner fidelity’, Douglas and Werren mean that the genotypes that interact within a holobiont in one generation (i.e. those of the host and of its microbiota) need to be stably preserved across different generations of the host, or otherwise the hologenome cannot be considered a unit of selection. In other words, if there is a holobiont composed of a host H plus the symbiotic species S_1, S_2, S_3, the hologenome can be considered a unit of selection *qua* replicator only if the same symbiotic species S_1, S_2, S_3, re-occur across different generations of H. Otherwise, if there is no effective co-transmission of the microbiome and the host genome, selection at the level of the hologenome will be eroded, as the condition of inheritance is not satisfied. Notice, that the criticism raised by Douglas and Werren concerns the claim that the hologenome is a unit of selection *qua* replicator, so their arguments should be read that way.

Douglas and Werren’s claim follows this line of reasoning:

1. *Definition*: Holobionts are entities composed by a host plus its symbiotic microbiome.
2. *Lewontin conditions for natural selection*: For natural selection to act on a given entity in a population it is necessary that the entity (a) exhibits phenotypic variation that (b) affects its fitness and that (c) the phenotypic variation is inherited with sufficient fidelity.
3. *Partner fidelity*: The only way of guaranteeing the satisfaction of (c) among holobionts would be that both the host and the symbiotic

microbiota co-occur transgenerationally, i.e. that the host genome and the genome of the microbiome are transgenerationally co-transmitted.

4. Therefore, the existence partner fidelity is a necessary condition to claim that holobionts are units of selection, i.e. to demonstrate the existence of partner fidelity is necessary to defend HCE.

Douglas and Werren argue that partner fidelity is usually satisfied in those cases when heredity of the symbionts is direct, as it happens in cases of vertical transmission (Maynard-Smith 1991; Hurst 2017). Vertical transmission occurs when the symbionts are transmitted from parent to offspring (transovarially, through feeding, etc.) in the moment of conception, or during early development. Vertical transmission is very common in insects, which usually carry heritable bacteria, but it is normally restricted to one or two symbiont species, instead of to the entire microbiome (Werren et al. 2008; Duron et al. 2008; Osborne et al. 2009). Partner fidelity, however, is not necessarily restricted to cases of vertical transmission, as Douglas and Werren acknowledge, and it can also occur in very specific cases of horizontal transmission (Bright & Bulgheresi 2010; Shapira 2016). Horizontal transmission occurs when the symbionts need to be acquired again, from the environment, in every host generation. A well-known example and very well documented example is the case of *Vibrio fischeri* in bobtail squid, which needs to be acquired during development (Nyholm and McFallNgai 2004); this is also the case in legumes and their intracellular rhizobia, sited in their root nodules (Gage 2004). In contrast with the cases of vertical transmission, where the transmission is direct and does not require such a strong mediation, 'all horizontally transmitted symbioses require sophisticated molecular machineries to select specific symbionts from the environment' (Bright and Bulgheresi 2010: 221-222). In any case, it is important to note that in some circumstances horizontal transmission might also lead to situations of transgenerational partner fidelity.

However, Douglas & Werren (2016: 2-3, emphasis added) insist that 'even when coinheritance occurs for *a subset* of the microbial associates, it is unlikely to be so for all members of the community, and so it seems difficult to

imagine that the entire microbiome should be considered part of a "hologenome" with its host if only a subset of microbes meet the requisite conditions.' In other words, if HCE only makes sense if the totality of symbionts S_1, S_2, …, S_n that constitute the microbiome of the holobiont co-ocur across different host generations. But this case is either never fulfilled or, when it is fulfilled, it is only fulfilled for a very concrete subset of the symbionts that compose the microbiome. If this is so, then holobionts cannot be units of selection and HCE is neither a useful, nor an empirically justified framework for understanding the evolutionary dynamics host-microbiome associations.

Douglas and Werren's insistence on the necessity of transgenerational genotype-by-genotype (G x G) matching relies as much on their conception of the units of selection, as it does on their specific interpretation of the application of Lewontin's conditions to holobionts. In their view, a consortium of lower-level entities constitutes a unit of selection only if the fitness interests of every lower-level entity are aligned with each other. As holobionts are multispecies assemblages, by definition (Premise 1), the only mechanism that can guarantee that this happens is partner fidelity. This conviction leads them to express the following thought: 'For a host-microbiome association to be the unit of selection, the hologenome concept requires (near-)perfect concordance of selective interests both among the microbial partners and between the microbiota and the host. As conflicts of interest among partners increase (e.g., due to weak partner fidelity), then the host-microbiome is undermined as a single unit of selection' (Douglas & Werren 2016: 3).

As I said before, Douglas and Werren's position is not the only position in the biological (and philosophical) literature that is critical of the notion that holobionts are units of selection. A very similar position, argued on very similar grounds, is found in Moran and Sloan (2015: 6-7), who write:

> 'Heritable obligate symbioses such as those of mitochondria and eukaryotes, or of insects and obligate bacterial symbionts that provision nutrients, provide clear cases in which fitness of each party is directly

> dependent on fitness of the other. Furthermore, this codependence is maintained, in some cases, for many millions of generations through the co-transmission of genomes of both parties. […] But the majority of microbial associations of multicellular animals and plants are non-heritable. […] Across this spectrum, how can we determine if and when selection at the hologenome level has an important evolutionary impact?'

Moran and Sloan acknowledge that there are some special cases of host-symbiont associations in which partner fidelity is not intergenerationally disrupted (e.g. eukaryotic cell, obligate symbionts in insects, etc), However, their quote also reflects a clear scepticism about the opportunity of selection at the level of the hologenome. Moran and Sloan argue that many of the microbial associations that constitute the microbiome of animals and plants are non-heritable, and thus it is hard to determine the importance of selection on the holobiont. In other words, if most members of the microbiome are not inherited, is it possible to determine whether their adaptations are a by-product of selection on the holobiont, rather than a product of selection acting solely on the bacterial species? For instance, take those cases in which symbionts are horizontally transmitted. Is it possible to test whether their fitness is causally correlated with the fitness of their hosts? Even if horizontal transmission does not automatically rule out the possibility that the fitness interests of the host and the fitness interests of the microbiota are aligned, its existence creates two important problems for HCE: first, that it can result from independent selective regimes, and not from selection on the hologenome, as HCE presumes (**Figure 15**); second, that it does not seem easy to assume that *the whole* bacterial community of an environmentally acquired microbiota will necessarily have all its fitness interests aligned. They thus conclude that 'opportunity for selection at the hologenome level exists, but it may be small or insignificant relative to selection on individual interacting genomes.' (Moran & Sloan 2015: 7).

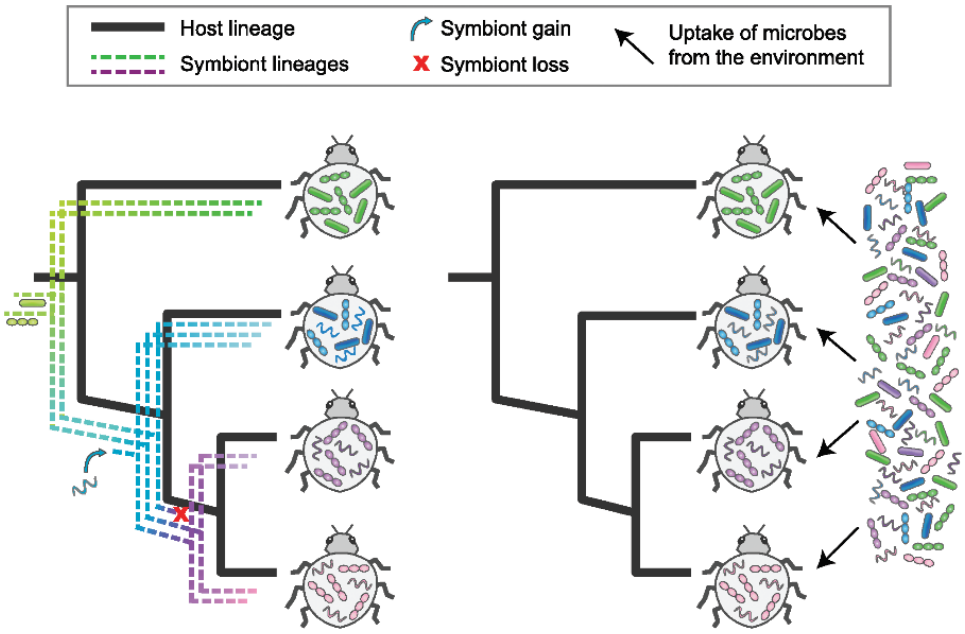


Figure 15. Two alternative explanations of the existence of phylosymbiotic patterns. The figure on the left represents the possibility that the pattern is a result of a process of host-microbe coevolution (i.e. hologenome selection). The figure on the left suggests that the process can result from a process of ecological filtering, and thus independent selective regimes affecting the hosts and its microbes. (From Moran and Sloan 2015: 4, Fig. 1)

The argument presented by Moran and Sloan adds additional support to the argument presented by Douglas and Werren. Concretely, Moran and Sloan support premise 3—the necessity of partner fidelity for selection on holobionts—with a concrete conception of biological individuality. According to this conception, a group of entities can only be considered a biological individual *qua* unit of selection if the fitness interests of all the entities are aligned, so that the interests of the entities at the lower-level does not erode (disrupt) the aggregative effects of the fitness of the entity at the higher level. In other words, a group of entities can only be considered a unit of selection if a policing mechanism that ensures that the group acts as a *real whole* is present.[53] This view about biological individuality is not original. Similar positions have been

[53] There is another way to conceive the appeal that the alienation of fitness interests has in the discourse about units of selection, and therefore in Moran and Sloan's discussion of the role of the holobiont as a unit of selection. It might be argued that, since natural selection leads to optimal evolutionary states, it is expected to cause the appearance of stable wholes. A whole will be stable if and only if there is no internal conflict that can disrupt its existence. If there are 'populations of wholes', natural selection will be expected to select for the fittest, i.e. for those with lower internal conflicts. Thus, a test to discover whether natural selection has occurred in the holobiont is to analyse the alienation of fitness interests among the members of the whole. But notice that this is an argument to test how strong natural selection is on a whole, not to test whether the whole is a unit of selection. I will make this point stronger in **chapter V**.

defended by Michod and Roze (2001), Queller and Strassmann (2009, 2016), Folse and Roughgarden (2010), Clarke (2013, 2016). The originality of the claim found in Douglas and Werren (2016) and Moran and Sloan (2015) is that they make the existence of an efficient policing mechanism in holobionts dependent on the detection of partner fidelity, because this is the way to guarantee that the fitness interests of the interacting species are aligned and, therefore, selection can occur at the level of the multispecies individual.

For partner fidelity to obtain, as I have argued, the species that compose the microbiota of the holobiont must be preserved across host generations. For the rest of the thesis, I will call this assumption stability of species (SoS, hereafter), and a specific argument against the necessity of SoS for evolution by natural selection will be presented later on (**chapter V**). For the moment, let me introduce a semi-formal definition of SoS.

> *Stability of species* (*SoS*): A holobiont in generation n+1 will be a unit of selection only if the symbionts S_1, S_2, ..., S_n, that occur within the host H_{n+1} belong to the same species as the symbionts S_1, S_2, ..., S_n, that co-occur(ed) with the host H_n.[54]

As SoS is stated here (i.e. in the works of Douglas and Werren, Moran and Sloan), it is clear that holobionts do not fulfil the strict conditions that it requires. First, empirical data about transgenerational transmission of the microbiome across different generations of a host species is scarce, limited to a very few cases, and also systematically limited to a subset of the species of the microbiome (e.g. Browne et al. 2017; Goodrich et al. 2017). Second, diet is one of the main drivers of microbial colonization and there is some evidence

[54] As John Dupré and others have noticed, the real requirement is even stronger. It is not that the species need to re-occur together, but that the *lineages* need to re-occur transgenerationally. Indeed, in my original formulation of their view, I elaborated their position as a question of 'stability of lineages'. However, I do not do this here because my claim is not that the stability of the lineages is not necessary to consider the holobiont as a unit of selection, but that not even the weaker requirement about the stability of the species is necessary, because holobiont evolution is not about the common evolution of microorganisms and the hosts that bear them, but about the evolution of hosts and the traits that their microorganisms bear. See **chapter V** for the detals.

suggesting that it partially determines the composition of the microbiota, with important shifts in species composition taking place quite rapidly related to changes in host's habits (Lang et al. 2014; Gómez et al. 2016). Therefore, if one assumes that SoS is a necessary condition to have proper inheritance—as 'inheritance' is defined in Lewontin's conditions for natural selection—and the existence of proper inheritance is required for holobionts to constitute units of selection, then holobionts are not units of selection and the hologenome concept is hollow, as Moran and Sloan have put it.

4.1. Does 'stability of species' constitute a knock-down argument against the hologenome concept of evolution? A brief reflection

The criticism just presented is perhaps the most serious of the challenges that HCE defenders must face for their hypothesis to be true. If SoS is a necessary condition for an entity to be a unit of selection, and holobionts empirically lack reliable transgenerational SoS, then HCE is patently false. But take into account that the argument depends on the acceptance of a necessary condition (SoS is necessary for units of selection), and the acceptance of a definition (holobionts = coevolved multispecies consortia). If the definition of holobiont that the critics assume is rejected, then the argument is *non sequitur*, whereas if the SoS is proven unnecessary, then the conclusion can be considered unargued. Here, I will only advance the reasons why I think the arguments just reviewed do not constitute a knockdown argument against the claim that holobionts are units of selection. A longer discussion will be presented in **chapter V, section 2**.

To start with, defenders of HCE have claimed several times that they do not assume that the holobiont is a 'coevolved' consortium. They argue that some members of the holobiont might constitute a coevolved unit, but they do not think coevolution is necessary for natural selection. Coevolution, if anything, would be a consequence of selection acting on the unit, but that is not even necessary. As Brucker and Bordenstein (2014) argued in their reply to Chandler and Turelli (2014), coevolution is usually not proven for the genes that interact within a genome, and still the 'monogenomic' organism is assumed to be the unit of selection. Thus, there seems to be no reason to believe that coevolution needs to be proven to consistently claim that holobionts are units of selection.

Critics of the holobiont would be putting the cart before the horse. They would only accept that the holobiont is a unit of selection if defenders of HCE can prove that *the action of selection* has led to the specific result of coevolution. But, species coevolution (or, rather, cospeciation) is a by-product of the action of natural selection on a consortium, so assuming it as prior to the existence of a unit of selection simply misses the point and leaves unexplained an important outcome of natural selection.

Second, related to my first worry, it seems that these authors are conflating the notions of '*response to selection*' and '*selection*' itself, a distinction that should be carefully made to gain a clear understanding about what units of selection are. This distinction follows from the breeder's equation, according to which $R = h^2*S$, where R is response to selection; h^2 is additive genetic heritability; S is selection coefficient. A unit of selection is the entity on which selection S can act, i.e. whose survival and reproduction are determined by the action of selective forces. How the entity is individuated is irrelevant for selection to act on it. Heritability is required to have a response to selection, and it is plausible to argue that, for holobionts, heritability will only be proven if there is a vertical transmission of the microbiome, i.e. SoS. But notice that heritability is completely irrelevant to address the question whether holobionts are units of selection. It is relevant to decide whether there will be a response to selection R, as well as the magnitude of this response h^2*S, but not to decide if the entity is (or is not) a unit of selection. By demanding SoS, critics of the holobiont seem to be confusing the notion of response to selection with the notion of unit of selection.

Finally, following the previous criticism, let us assume that critics of the holobiont take SoS as a necessary condition for natural selection to act on the holobiont because it is the only way of having transgenerational heredity (not heritability). Based on this interpretation, their argument is that that SoS can only (or mostly) be obtained by vertical transmission of the microbiome. Horizontal transmission, they argue, would rapidly erode selection at the level of the holobiont. But their argument is not exactly correct, as Benjamin M. Fitzpatrick (2014) has proven. Horizontal transmission, combined with a strong

action of non-additive selection[55] between the host and the symbionts (interspecific epistasis leading to linkage disequilibrium) and an appropriate population structure, can basically generate the same covariance dynamics between the host and the symbionts than the ones that are observed between the genes within a genome. Therefore, assuming that horizontal transmission will erode selection for the holobiont is equivalent to assuming that genetic recombination within one single genome will erode the action of natural selection at the level of the organism. Furthermore, as, due to the existence of abundant lateral gene transfer in the microbiome, the genetic components that guarantee the appearance of the epistatic effects that led to the linkage disequilibrium are not necessarily bound to the genetic markers used to recognise the bacerial taxa that compose the host microbiome, there can be multispecies units of selection without SoS, thus strongly contradicting the conclusions of holobiont critics (see **chapter IV** for an elaborated version of this argument).

5. Confusing sorting with selection: sorting and selection must not be equated

In the previous section, I presented the main criticisms that have been raised against HCE, conceived as a thesis about the role of the holobiont as a unit of selection. Concretely, the criticisms were raised against the thesis that the hologenome is a replicator which, as I explained in **chapter I**, is just one of the many theses about biological individuality that HCE entails. In this section, I will present a criticism that, although suggested in different ways in the critical literature about HCE, has never been explicitly presented in the way in which I will present it here. I will argue that one possible interpretation of the widespread disagreement about the role of holobionts as units of selection consists in the fact that defenders of HCE, by confusing the concepts of the 'ecological holobiont' and the 'hereditary holobiont' (**chapter I, section 3.1.4**), are confusing the notions of *sorting* and *selection*, two notions that should be

[55] In population genetics, additive genetic effects occur when the interactions between the genes are 'linear', meaning that all genes make the same contribution to the final outcome. Non-additive effects include dominance (when the action of an allele at one single locus can silence another allele in the same locus) and epistasis (when the expression of one gene at one locus depends on its genetic background, i.e. on the action of genes at different loci).

carefully distinguished in debates about the units of selection (Sober and Lewontin 1982; Sober 1984; Vrba and Gould 1986; Lloyd 1988; Okasha 2006).

Let me start by defining the two notions. In biology, and specifically, in the way in which the Modern Synthesis recasts Darwinism, evolutionary change in conceived at the level of the population. Populations are collections of causally connected individuals from the same species, amongst which there is variation. Different individuals in the population have different rates of death and birth, and as a consequence the composition of the population changes with time. This process of differential death and birth of the individuals in a population that ultimately leads to the changes in their representation of in the population that is considered to be evolutionary change is simply a process of 'sorting'. In itself, sorting is simply the process of change in biological populations, but the concept does not specify any of the causes of why this process happens: sorting is simply what results from differential death and birth.

Given this definition of sorting, an important question for biologists is thus the following: what causes the process of sorting? Or, in other words, taking sorting for granted, what causes that the population will shift in one direction (i.e. towards a specific set of changes) rather than another? Here is where natural selection enters the picture: selection one of the *causes* of the process of sorting. Concretely, it is the non-random cause of differential birth and death of the individuals in a population. Other causes, normally qualified as 'random', include the different phenomena of drift, including genetic drift. Now, an obvious question arises at this point, and it is important to tell selection and sorting apart, namely: in what sense is natural selection a non-random cause of sorting? Here, a quote from Elisabeth S. Vrba and Stephen J. Gould is particularly illuminating:

> 'Selection encompasses those interactions between heritable, emergent character variation and the environment that cause differences in rates of birth or death among varying individuals' (Vrba and Gould 1986: 219)

Notice that the definition includes three elements: first, the localization of a focal level which is distinct from the environment; second, the localization of variation in character (phenotypic variation) at that level; third, the heritability[56] of that variation in character. If these three properties are localized, then selection can arguably be understood as a cause of sorting, i.e. as a plausible cause for the differential rates of birth and death of the individuals in a population. However, and importantly, the biological world is usually considered to be hierarchically organized. That is to say, it is possible to find different focal levels where one can observe individuals that satisfy the three properties included in Vrba and Gould definition. Thus, the observation of sorting in one level of the biological hierarchy can always be the result of a process of selection acting on the upward or on the downward level, which generates a random (i.e. not due to selection) distortion in the level under study (Sober and Lewontin 1982; Vrba and Gould 1986; Lloyd 1988; **Figure 16**). Therefore, sorting and selection should not be conflated.

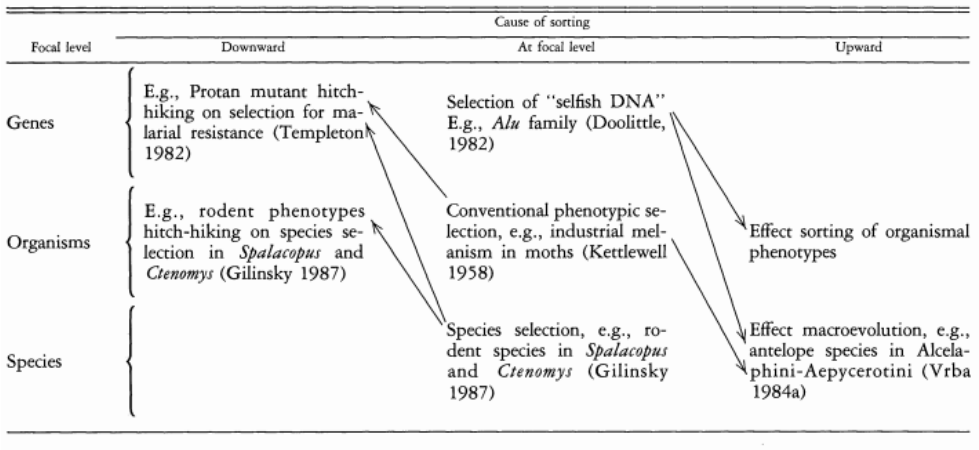


Figure 16. Table suggesting the relationship between selection at a focal level with its upward and downward effects on the sorting of the entities at the upper and lower-levels. The table shows how what might look as selection at one level (e.g. effects of sorting on the organismal phenotypes) is just a result of the selection at a lower level (e.g. selection of 'selfish DNA'). (From Vrba and Gould 1986: 220, Table 1)

[56] Heritability should not be understood in its technical meaning in population genetics, but rather as a way of expressing the possibility of being inherited.

Take the following, and conventional example of selection: the case of selection for darker phenotypes—industrial melanism—in *Biston betularia* moths due to industrial pollution (Kettlewell 1958). In this case, the population is formed by two phenotypes: darker and brighter moths. In a non-polluted environment, darker months are less successful, because their pigmentation makes easier to identify by predators, and thus more susceptible to be eaten *due to* their phenotypic characteristics. However, in a polluted environment, the opposite is the case, and thus the trait distribution in the population shifts due to the higher susceptibility of brighter moths to be eaten. Selection is acting on a very specific trait (pigmentation), and thus might generate a distortion in the genetic makeup of the population (downward level) or on the macroevolutionary pattern of moths (upward level). However, the upward and downward level effects are mere sorting, which is due to the effect of selection on the focal level of moths. The changes in the genetic distribution, as well as the changes in the macroevolutionary patterns of moths (if any) are the distortions that result from the causal effect of natural selection on a different level of the biological hierarchy. Therefore, to prove selection at one level it is indispensable to prove that the change in the phenotypic character occurs at that level, that it can be inherited, and that it *causally* leads to the differential death and birth of the individuals at that level.

How to apply the distinction between sorting and selection to the case of HCE? I think it is possible to argue that defenders of HCE are confusing the two notions and are therefore taking cases of mere sorting as cases of selection or, in other words, confusing an upward effect with a real cause acting at a lower-level in the biological hierarchy. First, it seems unquestionable that holobionts, defined as a host plus its microbiome, have differential rates of birth and death. For instance, if holobionts are the metabolic-wholes of Dupré and O'Malley (2009; **chapter I, section 6**), it seems easy to argue that some holobionts live longer than others, and some holobionts produce more offspring (i.e. more metabolic-wholes) than others. Second, it seems at least plausible to think that holobionts have different phenotypic characteristics, and their different rate of birth and death can arguably be attributed to these differential phenotypic characteristics. But now the crucial question to decide whether selection can

possibly act at the level of the holobiont is whether inheritance can be detected, so that phenotypic features have a way of getting fixed, if they are advantageous. Hologenome defenders argue that this is possible, on the basis that the hologenome is the replicator. Hologenome detractors, on the other hand, argue that it is not possible to tell apart selection from sorting in the case of holobionts: how to distinguish cases of genuine host-microbiome coevolution and cases of ecological filtering? Assuming that the holobiont can be singled out as a genuine focal level that can be distinguished from the environment (which is something that some detractors of the holobiont would neglect, **section 2**), what are the criteria to tell apart cases of sorting from cases of genuine selection, i.e. differential rates of death and birth *due to* heritable phenotypic characteristics? The argument here would be that, since this is not empirically possible, then the holobiont cannot be singled out as a unit of selection, and thus HCE is empirically false. Defenders of HCE are thus conflating the indirect effects of selection on the differential rates of birth and death of the holobiont with the direct effects of selection on the individual species that interact in the holobiont and that produce sorting of holobionts as a byproduct.

5.1. But should sorting and selection really be distinguished? A brief reflection

The argument presented above seems a knock-down argument against HCE and seems even more strict that the argument based of SoS that I presented in **section 3**. The case I have made here is not that SoS is not detected among holobionts. It is that even if it were, the authors cannot really tell apart which are the cases of genuine selection where the holobiont is the focal level, versus cases of sorting, where the species that compose the microbiome are the focal level. It is just another (more elaborated) way of expressing Moran and Sloan's worry that you can never tell apart cases of convergent host-microbiota phylogeny and cases of mere ecological filtering. I think reflecting their worry in terms of the distinction between sorting and selection is important, because it reflects one of the most genuine empirical worries to consider when selection for holobionts is being studied.

However, is this worry legitimate? From a conceptual point of view, it seems important to consider that not all causes of differential births and deaths of the individuals in a population are due to selection for their characteristics. Also, it is empirically important to single out the mechanisms, or causal agents, that are producing the selection, in case its effects are detected. But it is not so clear that the possibility of the effects of selection at the lower-level should be taken *simpliciter* as a knock-down argument against selection at the upward level, because doing so might bear the risk of a regression *ad infinitum*. In a hierarchical-nested view of the world, it is true that holobionts are made of different species, but so is it true that genomes are made of different genes, genes are made of different molecules, molecules are made of different atoms, and so on and so forth. If the possibility of distortion is taken 'too seriously', one might end up with no level where the effects of selection can be studied. There must be one level where one can stop, and the reason for stopping at that level must be somehow independent from the possibility of distorting effects from the lower level, or no empirical research could be pursued. Godfrey-Smith suggests the idea that we must start '*afresh*' at each level where we aim to study individuality and the levels of selection, and Okasha suggests that we must break up with the necessity of a nested hierarchy in biology. I will pursue these two thoughts for holobionts in **chapter IV** and **chapter V**. For the moment, it is enough to say that the argument based on the distinction between sorting and selection must not be taken as an argument that rules out the possibility that holobionts are units of selection, and thus the possibility that some formulation of HCE is true. The purpose of the project is precisely to figure out *which formulation* of HCE is true, and in virtue of what it can be considered true.

6. Brief summary of chapter II: Three levels of criticism to holobionts and the hologenome concept

In this chapter, I have introduced the main criticisms that have been raised against HCE, including one last criticism (the confusion of the ideas of sorting and selection) that I came up with. The main ideas that have been presented in the chapter can be summarized as follows:

1. Holobionts are ‘blurry’ entities. Because of the transient nature of many microorganisms during the life of the holobiont, holobionts lack a sufficient degree of ontogenetical temporal stability to be studied as biological individuals in any relevant sense.
2. None of the arguments given by HCE defenders to support the thesis that holobionts are units of selection is to the point, since they fail to capture the essence of units of selection.
3. Holobionts lack transgenerational stability of the species that compose their microbiota, insofar as they lack partner fidelity, so hologenomes cannot be units of selection *qua* replicators.
4. Defenders of HCE are confuting the notions of sorting and selection. Some of their arguments prove that holobionts are units of sorting, while failing to prove that selection is the cause of this differential sorting.

(This page is intentionally left blank)

Chapter III

‘The concept of biological individuality: A review with an application to the discussion about the role of holobionts as biological individuals’

The previous chapters introduced the hologenome concept of evolution, and its problems, as a debate about the nature of biological individuality. The guiding question was whether holobionts could be considered biological individuals in any significant sense, particularly emphasizing their role as units of selection. But, of course, deciding whether holobionts are or are not biological individuals depends on what one means by ‘biological individual’. This chapter reviews recent debates on the (many) meaning(s) of *biological individuality*, as well as the relationships between different characterizations of the concept. I will put special emphasis on the discussion about the conditions that a biological individual has to satisfy to be considered a unit of selection, as I take this to be the main claim of the hologenome concept of evolution.

1. The never-ending problem of biological individuality

My discussion in the previous chapters made clear that the main problem for HCE defenders is to offer a coherent account of biological individuality[57] that

[57] The concepts of “biological individuality” and “organism” are sometimes used interchangeably (e.g. Clarke 2013: 413, ft. 1), sometimes the later is taken as a subset of the former (e.g. Lidgard and Nyhart 2017b: 18), sometimes are taken to be non-synonymous, although sometimes generating partial overlaps (e.g. Godfrey-Smith 2013; cf. Pradeu 2016a, 2016b), etc. To avoid confusion, I will use the concepts of “biological individual” and “organism” interchangeably, and I will always specify the nature of the biological individual I will be talking, e.g. *metabolic* individual, *anatomical* individual, *developmental* individuals, *organisational* individual, *evolutionary* individual, etc.

accounts for the individuality of the holobiont, including its role as a unit of selection. The problem that this creates to HCE defenders, however, is not exclusively reduced to them. It connects with a classical and well-known debate that has caught the attention of biologists, historians and, of course, philosophers, for many centuries (see Lidgard and Nyhart 2017a, for the first effort to offer an integrative approach; reviewed in Suárez 2018b): What is a biological individual? How can biological individuals be demarcated and distinguished from the environment that surrounds them? How can the parts of an individual be discriminated from the background conditions that sustain its existence, without being part of it? One of the reasons why this debate has caught (and still catches) the attention of so many researchers is that, while the concept of individuality is crucial in biology (biology deals with individuals), biologists cannot agree on its most fundamental meaning, and this generates confusion and conflict. Following Ellen Clarke (2010, 2013), I will refer to this lack of general agreement as *the problem of biological individuality* (PBI, hereafter). PBI arises from the combination of the following observations: (1) biological research needs a concept of biological individuality, i.e. a way of discriminating biological individuals, since biological individuals are the entities that biology studies (Ruiz-Mirazo et al. 2000; Pepper and Herron 2008; Moreno and Mossio 2015); (2) there are different criteria to define biological individuality; (3) these criteria generate non-overlapping classifications of biological individuals (Wilson and Barker 2013; Lidgard and Nyhart 2017b); (4) these non-overlapping classifications generate epistemological problems about how to study the entities that biological research is interested in studying; (5) therefore, there is a PBI which generates the following question: how can one *arbitrate* over the different definitions that are to a certain extent accepted in the current literature?[58]

In this chapter, I will examine the PBI with a particular emphasis on its relation to the problems of individuation that arise for HCE. I will start reviewing

[58] There is always the possibility, embraced by some authors (e.g. Wilson 2000), to deny that the PBI needs to be arbitrated, arguing that biology can do very well without a clear concept of individuality. I think this is a legitimate option, but adopting it would still require explaining the PBI—even if it is to explain it *away*. As the purpose of this chapter is only to illustrate that there is a *problem*, no matter whether it has or not (serious) epistemological consequences, I will ignore these positions.

different criteria that have been used to define biological individuality, to show that they generate a non-overlapping classification of biological individuals. Second, I will present in depth the problem of the units of selection. Following Samir Okasha, I will distinguish between *synchronic* and *diachronic* approaches to the units of selection question, and I will analyse both approaches in detail. In the case of the synchronic approach, I will put the emphasis on the interactor/replicator framework and its application to the case of holobionts. I will put forward one hypothesis about the nature of the interactor category, which is inspired by Kevin de Queiroz's 'resolution' of the species problem (de Queiroz 2005a, 2005b, 2007): all the criteria of biological individuality that have been offered accord with different biological processes in which biological individuals can participate and that, taken together (i.e. if an individual satisfies *all of them*), characterize the interactor. In this sense, interactors are highly evolved entities that appear because of their simultaneous participation in different biological processes.[59] In the case of the diachronic approach, I will put the emphasis on the MLS theory, distinguishing between multilevel selection 1 (MLS1, hereafter) and multilevel selection 2 (MLS2, hereafter). I will argue that the MLS approach is applicable to a level of biological individuality provided that the entities individuated at that level satisfy, *at least*, one of the criteria of biological individuality aforementioned—i.e. provided that they participate at least in one biological process. This said, I will defend (here and in **chapter V**) that the MLS approach is more appropriate than the interactor/replicator framework to study whether holobionts are units of selection.

2. Different criteria for defining biological individuality and their non-overlapping nature

The PBI has lots in common with the problem of species in biology and philosophy of biology.[60] As it occurs for the category of *species*, for which biologists and philosophers agree that there are different definitions that

[59] The ideas about the interactor that I present in this section have come to my mind thanks to the thoughtful discussions I have had with Álvaro Moreno, as well as other members of the IAS Research Group, University of the Basque Country (Spain), especially Mark Canciani. No less important in the development of these ideas has been the role of Çaglar Karaca.

[60] I have analysed the problem of species, applied to the case of bacterial species, in my paper J. Suárez (2016): 'Bacterial species pluralism in the light of medicine and endosymbiosis'. *Theoria* 31(1): 91-105.

generate non-overlapping classifications of the biological world (cf. Ruse 1987: 238, for a notable exception), so it occurs for the case of biological individuals. There are nowadays dozens of different criteria to determine whether a biological ensemble is a biological individual—instead of a 'random' aggregation of biological stuff—and these different criteria generate non-overlapping classifications of the world. Or, in other (metaphysically more neutral) words, each of these different criteria entail the existence of distinct ways of 'lumping' matter together (each criterion corresponding to one legitimate way of lumping, and each lump corresponding to one biological individual), so that: (1) some lumps cut across each other, and (2) some ways of lumping matter will dictate that there are two (or more) individuals, on occasions where others will dictate that there is only one, and vice versa. Interestingly, and that is the reason why the PBI is indeed a *problem*, as Clarke (2010: 313-315; 2013: 413-418) has convincingly argued, the existence of different criteria generate methodological disputes, and discarding some definitions as inadequate can sometimes overlook some important biological phenomena.[61] To illustrate the non-overlapping nature of the definitions, and the problems that this generates for biological practise, let me first offer some candidate criteria for deciding how to delimit the boundaries of biological individuals (adapted from Clarke 2010; Wilson and Barker 2013; Lidgard and Nyhart 2017b: 19-21, Table 1.1.)[62] [63]:

- *Physical boundedness*: a biological individual is a physically bounded, spatially discrete unit. This criterion is similar, although not equivalent, to the notion of 'Anatomical individuality' which I first introduced in **chapter I** (Gould and Lloyd 1999; de Sousa 2005).

[61] Notice that 'criteria' and 'definition' are not synonymous, although the use here might mistakenly suggest so. To be clear, I use 'definition' to refer to the *criterion* (if one) or *set of criteria* that are used to characterize the necessary and sufficient conditions that the entities of an ensemble must fulfil to be 'lumped' together as a biological individual, and criterion to refer to any of the biological mechanisms that could be included in the definition.

[62] The problem I will consider here is not how to reidentify one and the same organism/individual over time, but how to determine whether we should count one or two individuals. Notice that these two questions, although related, are not metaphysically equivalent (Pradeu 2010; Wolfe 2010).

[63] The list is not, does not aim to be, exhaustive. I will only suggest some criteria that have been offered in the literature to illustrate the points that will connect the discussion in this section with the rest of sections in this chapter.

- *Genetic homogeneity*: a biological individual is the ensemble composed by the set of genetically homogenous cells (Dawkins 1982a; Santelices 1999).
- *Development from a fertilized zygote*: sexually (diploid) distinguishable sex cycle, excluding other means of propagation (Janzen 1977; de Sousa 2005).
- *Realization of a life cycle*: having a demarcated life cycle, irrespective of the means of propagation/generation of the entity that realizes the life cycle (Rainey and Kerr 2011; Wilson and Barker 2013)
- *Immunological integration*: existence of an immunological system that delimitates the boundaries of the system, protecting it from external attacks and recognising the entities that are 'parts' of it and maintain its functionality through time. This criterion is similar to the notion of 'Immunological individuality' which I introduced in **chapter I** (Pradeu 2010, 2012; Tauber 2016; Gilbert and Tauber 2016).
- *Germ/soma specialization*: existence of a division of labour that clearly demarcates reproductive cells (i.e. cells that will contribute to the formation of the zygote in the future) and somatic cells (i.e. cells that compose the body of the individual) (de Sousa 2005; Godfrey-Smith 2009, 2013; Folse and Roughgarden 2010).
- *Bottleneck*: originated from a narrowing (usually in the form of a unicellular stage) of the material constituents that follows some form of propagation (Godfrey-Smith 2009, 2013, 2015).
- *Reproduction*: originated from a distinguishable form of reproduction, normally following certain parameters that define the correct way of reproducing and tell reproduction apart from growth (Godfrey-Smith 2009).
- *Functional integration*: existence of cohesion among the parts, usually defined in terms of shared metabolism (Sober and Wilson 1989, 1994, 1998; Dupré and O'Malley 2009; Folse and Roughgarden 2010; Moreno and Mossio 2015).
- *Organisational closure/ Autonomy*: ensembles whose components together realise a regime of 'regulated closed agential emergent organisation' (Moreno and Mossio 2015).

- *Reduced conflict/ High cooperation among the parts*: ensembles whose parts interact in a harmonic manner (Queller and Strassmann 2009, 2016; Folse and Roughgarden 2010).

The criteria just offered are not exhaustive, and some of them might even be argued to be specifications of others (e.g. the autonomy approach is a way of making 'functional integration' precise). However, they will serve to illustrate the points I want to make for the purposes of this chapter. First of all, on a general ontological level, it is clear from the list above that some definitions are exclusively given on the basis of the satisfaction of *one* criterion ('A biological individual is an entity that is X', where X is a criterion), whereas others include *several* criteria that might be satisfied either conjunctively ('A biological individual is an entity that is X, *and* Y', where X and Y are criteria), or disjunctively ('A biological individual is an entity that is X, *or* Y', where X and Y are criteria), or both ('A biological individual is an entity that is X, *or* Y, *or* (Z, *and* Z', and Z'')', where X, Y, Z, Z', Z'' are criteria). For instance, Godfrey-Smith appears three times, Moreno and Mossio appear twice, and Pradeu and Santelices appear only once. This is because for some authors, an entity is a biological individual if and only if it has certain properties or participates in certain biological processes (e.g. it reproduces in the appropriate way, which means it forms bottlenecks, or has a clear germ/soma specialization). For others, on the other hand, the requirements are different and thus their definitions include only the satisfaction of one property or the participation in a specific type of biological process. In this sense, some definitions are 'harder to satisfy' than others, i.e. they are more demanding. I will come back to this point at the end of this chapter.

Second, and partially a corollary of the first point, notice that some definitions are 'absolute', whereas others might accept the existence of degrees of biological individuality. According to the former, there is no possibility of vagueness: either an ensemble is a biological individual, or it is not. According to the latter, some ensembles will be biological individuals *absolutely*, while others will be individuals only to a very low degree. A canonical defender of the first position is Janzen. For him, a biological individual is the entity that derives

from a sexually formed zygote, and only that. Either an ensemble satisfied that, and thus it is a biological individual, or it does not, and thus it is a clone, or a part, of another biological individual. On the other hand, Queller and Strassmann defended the view that some ensembles are more 'organismal' than others, depending on the degree in which the parts cooperate (or compete) with each other (**Figure 17**).

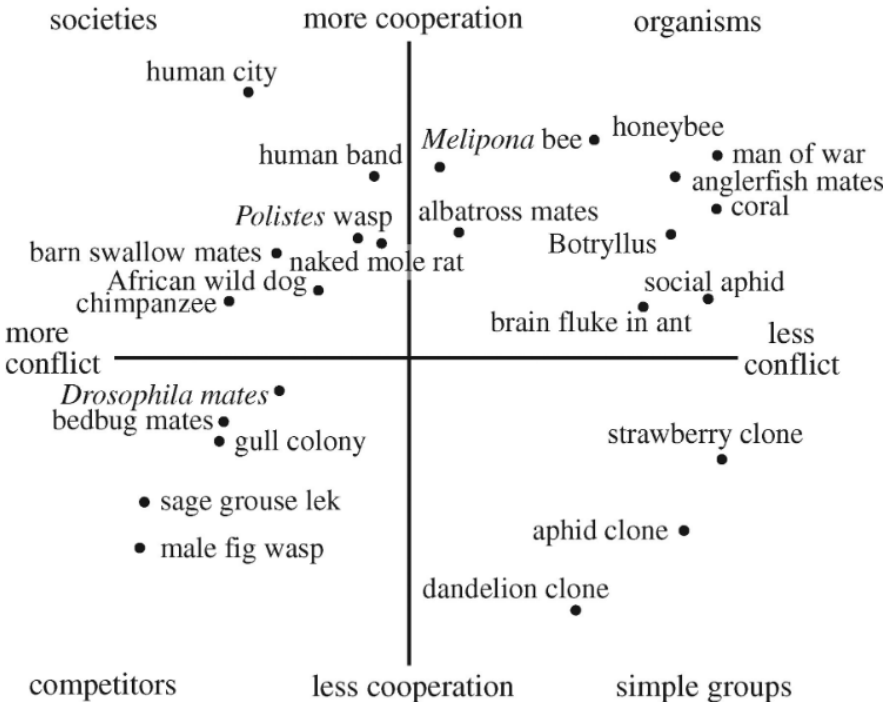


Figure 17. The higher the level of between-part cooperation and the lower the level of between-part conflict, the higher the score in 'organismality'. The figure presents a classification of some 'lumps' of matter and their degree of 'organismality' according to these criteria (From Queller and Strassmann 2009: 3145, Fig. 2)

The same is true of Godfrey-Smith, who has a rather idiosyncratic view of the problem. On the one hand, Godfrey-Smith believes that there are different ways of significantly individuating biological 'lumps' of matter, and thus distinguishes *organisms* (functionally integrated wholes) from *Darwinian individuals* (units of selection). On the other, he accepts that Darwinian individuals, which are things that *reproduce correctly*—i.e. according to any of the modalities of reproduction that Godfrey-Smith considers as 'correct'—come by degrees, depending on how many mechanisms they possess that guarantee that their reproductive regime is clearly demarcated/ distinguishable (**Figure 18**). This classification

allows him to distinguish between paradigmatic Darwinian individuals and aggregates that approximate—more or less—to the paradigm.[64]

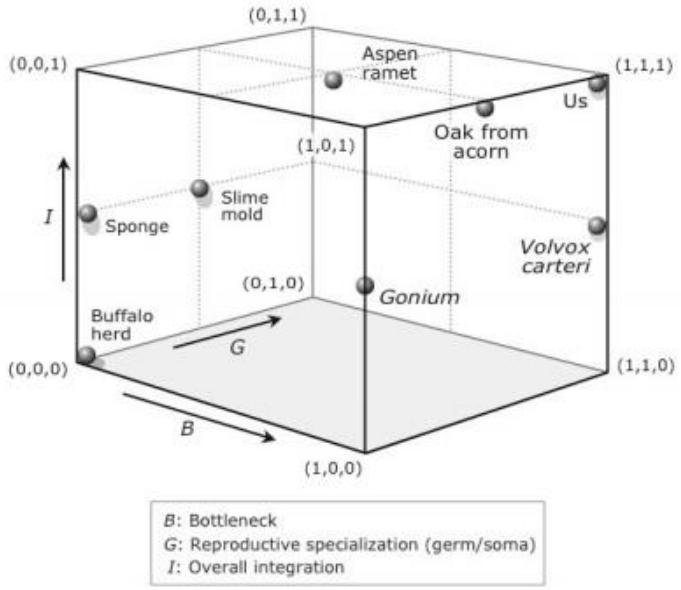


Figure 18. Three mechanisms to determine whether a collective reproducer is a paradigmatic Darwinian individual or not. Depending how many of the mechanisms a collective possesses, it will score higher in Darwinian individuality, with the higher score constituting the case of 'paradigmatic Darwinian individuals'. Notice that, according to this criterium, individuals come by degrees, with paradigmatic individuals like us, and paradigmatic non-individuals, like buffalo herd. In between, cases like the slime mold, an aspen ramet or a sponge. (From Godfrey-Smith 2009: 95, Fig. 5.1)

Third, as in the case of species, the different definitions do not classify biological 'lumps' in the same manner. Take, for example, the case of *individuals**[65] with a haplodiplontic biological cycle (for an analysis of this type of reproduction, see O'Malley 2016; **Figure 19**).[66] According to sexual definitions

[64] Surprisingly, one of the criteria that he considers is 'overall integration', which is the criterion he uses to characterize what he calls 'organisms'. For a sustained criticism to his model see Sterner (2015) and Stencel (2016).

[65] I will use '*individual**' to avoid pre-judging their status as biological individuals, since the point of the section is to show how different definitions offer different verdicts.

[66] Notice that I say 'haplodiplontic' biological cycle, not 'haplodiploidy', which refers to the mechanism of sex determination in some insects. While the former means that one organism goes through a haploid stage during its life cycle, the later refers to the species where members of one sex (females) are diploid individuals, whereas members of the other (males) are haploid individuals. Of course, these criteria of individuality according to which the haploid life stage is

of biological individuality, the gametophyte, no matter how well developed it might be, is just one of the stages of development, i.e. one temporal stage of the life cycle of the biological individual. But there is only *one* individual, in different stages if its life cycle. On the other hand, for those that consider that the biological individual is the entity that realizes a life cycle, *both* the gametophyte and the sporophyte are biological individuals, thus there are two biological individuals, and not only one, as the defender of the sexual criterion of individuality assumes. And the same would be true for the defender of the bottleneck view, since haplodiplontic *individuals** experience two bottlenecks during their life time, one for each part of the cycle (for an excellent exposition of the type of disagreements, see Clarke 2010, 2013).

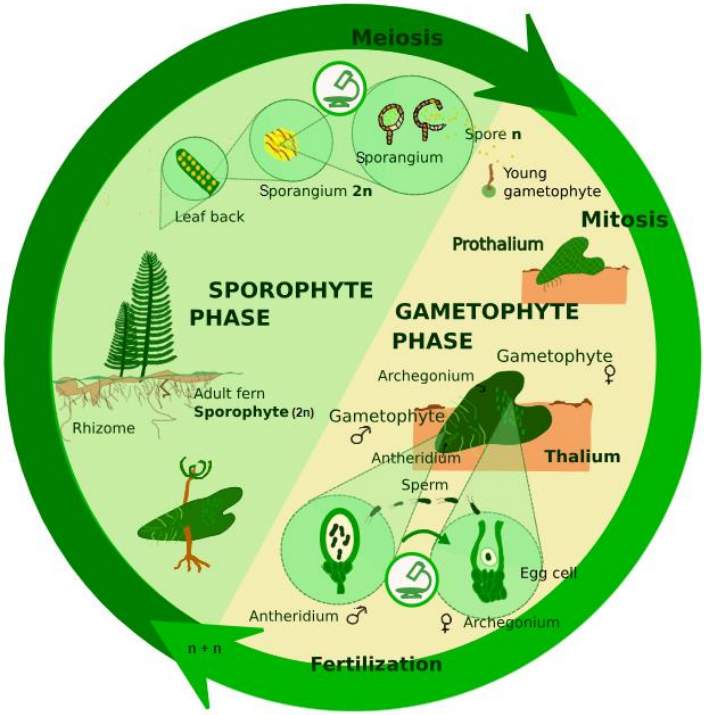


Figure 19. Haplodiplontic life cycle. The *individual** that develops during the gametophyte phase is haploid (*n*), whereas the one that develops during the sporophyte phase is diploid (*2n*). Notice that both cycles require respective bottlenecks, and that during both lifecycles a complete and fully formed *individual** appears. (From Wikimedia Commons: https://upload.wikimedia.org/wikipedia/ commons/f/f7/Diplohaplontic_English.svg).

just a stage of one individual make it hard to conceive haploid individuals *qua* individuals in haplodiploid species. However, I will not consider these cases here.

Fourth, and no less important, these disagreements have sometimes important practical consequences, that should not be left unnoticed. To explain this, take the case of aphids, and the verdicts about how to count new aphids that the sexual criterion and the bottleneck view render. As it is known, aphids can develop and grow to reach an adult stage both from fertilized and from unfertilized eggs. Defenders of the sexual view consider that only those aphids that appear from a sexual fusion of gametes must be counted as new individuals—the others would be 'clones', according to e.g. Jenzen (1977)—whereas defenders of the bottleneck view would count both aphids that develop from fertilized and unfertilized eggs. Why is this problematic? As Ellen Clarke (2013: 416-418) has convincingly argued, counting individuals one way or the other might overlook cases of selection. For instance, imagine a case in which one of the clones (i.e. the aphid resultant from an unfertilized egg) acquires a mutation and thus triggers a process of within-clone selection. In this scenario, the mutant can cause a process of informational decay (e.g. if the mutant expands it mutation to the 80% of the clone population), a process that cannot possibly be appreciated by the scientist who, as Jenzen, is exclusively concerned with the counting of the sexually produced aphids. Therefore, as Clarke summarizes:

> '[D]ifferent counts of the number of individuals in existence can lead to different measures of the fitness of their traits, and thereby lead the scientist to generate a different figure for total evolutionary change than she would have done if she had counted aphid insects, instead of aphid clones. This undermines our ability to understand how natural selection acts on wild populations, as well as our efforts to conserve and manage our natural environment. This is the sense in which the biological counting problem is real. There is a genuine multiplicity of concepts of the organism and this multiplicity is potentially damaging to the discussion and application of evolutionary theory.' (Clarke 2013: 417)

Finally, different ways of defining—and counting—biological individuals are sometimes guided by the type of problems that the scientists are investigating. It is not the same to be concerned about how the immunity system reacts to an

external ‘invader’, or how it can overcome an internal dysfunction (e.g. a tumour), than to be concerned with how an organism develops into adulthood, or how an organism reproduces. The three are significantly important biological processes that require adequate tools to be studied and handled carefully. However, these different processes can lead to different ways of ‘lumping’ the biological matter, and thus to different responses about how biological individuals must be counted (Pradeu 2016a, 2016b; DiFrisco 2017). This observation needs to be taken into account, as it is, in my view, part of the essence of the importance of the PBI.

In conclusion, the point I wanted to make in this section was simply that there is an intrinsic connection between our way of defining biological individuals, the problems we aim to study, and the set of criteria we use to study them. But this analysis leaves open an important question: how is the PBI related to the debate about holobionts and the nature of HCE?

2.1. Brief reflection on the problem of biological individuality and its importance for the hologenome concept of evolution

The PBI is fundamental to understand some of the issues raised by HCE, as some people have already emphasized (Gilbert et al. 2017; Stencel and Wloch-Salamon 2017; Suárez 2018a), simply because the debate about holobionts puts tension on some conceptions of biological individuality. If defenders of HCE are right, there is a genuine set of biological processes where the entity to be individualized is the multispecies community. Which this set of biological processes might be, it is a difficult question to address here, and I suspect it is highly relative to the individual that is being studied (see **chapter IV**). In any case, it is important to notice three points, that result from my analysis in the previous section: first, that the choice of one or other criterion of biological individuality depends on the problem that the scientists want to address; second, that different criteria respond to different, but relevant, biological processes, and thus different criteria capture some dimensions of the biological individual; thirds, that the use of different criteria is legitimate, if they are connected to the appropriate type of questions—i.e. if they serve to give responses to a biologically relevant question. These three points are, and will

be, fundamental for my position about HCE, and about the role of holobionts as units of selection. Thus, the lesson to derive from this discussion is that the choice of a criterion of biological individuality is (and must be) connected to the kind of problem that one is addressing.[67]

3. Biological individuals as units of selection

The reader might have noticed, at this point, that the list I have just presented in **section 2** lacks a fundamental criterion, one that many take as *the* criterion, to faithfully demarcate biological individuals. I am referring to the notion of units of selection, and specifically to the idea that biological individuals *are* units of selection and *only* units of selection, i.e. entities or ensembles that participate in the selection process (Clarke 2013). There are two reasons for introducing the PBI first, in a rather general way, not specifically addressing the issue of the units of selection. First, that I will propose to incorporate different dimensions of the different criteria of biological in the notion of the units of selection I will favour in this chapter; second, that I agree with Ellen Clarke that what defines a biological individual is its capacity to participate in a selection process, and that this capacity is multiply realizable.[68]

Let me start this section by noting a very basic point about natural selection. What is natural selection about? Or, in other words, what does natural selection *explain*?[69] In its original Darwinian formulation, natural selection was aimed at explaining the existence of *adaptations* in the living world. That is to say, it explained why *species* are apparently 'built' in such a way that they seem as if they had been specifically designed to do what they do—i.e. to feed how they feed, to reproduce how they reproduce, etc..[70] Darwin's idea to explain why

[67] Different versions of this position have been defended by Lidgard and Nyhart 2017b, Love and Brigandt 2017. I won't discuss how my version differs from theirs, as it is outside the scope of this thesis.
[68] Although, as the reader will see, I disagree with her favoured view of natural selection (i.e. the conditions that a biological ensemble must satisfy to arguably participate in a selection process).
[69] Most of the exposition that I introduce here is based on Ariew and Lewontin (2004): 'The confusions of fitness'. It is, however, partially modified, as what I present is my own interpretation of Darwin's ideas.
[70] Remember that Darwin's book was entitled *On the Origin of Species*, not 'on the origin of organisms.'

this was so can be summarized as follows: (1) species are collections of organisms with different phenotypic characteristics (natural properties); (2) the different phenotypic characteristics among the organisms in a species make some of them fit the environment better than others (fitness to the environment); (3) there is a 'struggle for existence', i.e. organisms compete with each other to survive; (4) those organisms that fit the environment better will survive longer, and the phenotypic characteristics that allow them to survive longer will spread ('survival of the fittest')[71]; (5) as a consequence, the species composition will change over time (changes in representation in the population), and thus (6) the fittest species will tend to dominate the world (generalized adaptation). This can be summarized in the following simple schema

> (S) Natural properties → Fitness to the environment — [Struggle for existence] → Survival of the fittest → changes in representation in the population → generalized adaptation[72]

Notice that, in (S), 'fitness to the environment' is presented in abstract terms, and in relation to the natural properties of the fitness bearer. What is the fitness that (S) assumes exists between the natural properties of an entity and the environment? To understand well Darwin's concept of 'fitness' one must note what Darwin aims to explain by appealing to the 'survival of the fittest'[73], namely: the existence of adaptation, or the generalized belief that living creatures seem to be designed by a sort of engineer to do what they do. Thus, in this sense, Darwin's theory of evolution as a consequence of the survival of the fittest is only an explanation of why living creatures seems to be designed, i.e. of why they bear adaptations. In this sense, the concept of 'fitness' entailed by his theory takes the higher reproductive rates of the fitter living creatures as a proxy to explain the survival of the fittest—i.e. it is a way

---

[71] Notice that Darwin uses 'survival of the fittest' as a (more adequate) synonym of 'natural selection'.

[72] 'Struggle for existence' is just the reason why organism's fit to the environment leads to the survival of the fittest. But it is neither caused by the fitness of an organism to its environment, nor is it the cause of the survival of the fittest. It is only a background condition that one must assume.

[73] Notice that the phrase 'survival of the fittest' was originally introduced by Spencer in 1864, and later used by Darwin after a suggestion from Wallace. See https://en.wikipedia.org/wiki/Survival_of_the_fittest. Accessed 29th May 2019.

in which the fittest can guarantee their survival and thus the fact that the properties that make them fitter spread in the population. But notice that *fitting the environment* is logically prior to *surviving* (including reproducing the type that makes it fit the environment). I wanted to emphasize this point here to contrast the notion of fitness that Darwin uses with the notion that appears in population genetics, where fitness is *the consequence* (and not *the cause*) of the survival of the fittest (and thus, of the higher reproductive rates) (see Ariew and Lewontin 2004, for an excellent exposition of this point). This point is important for the argument in support of HCE I will give in **chapter V**.

Second, notice that the existence of an *entity* whose natural properties (or phenotype) *fit* the environment, i.e. an entity that acts as fitness bearer, is implicitly assumed, but *its nature* is left unexplained, in Darwin's schema.[74] Which is this entity? The debate about the units of selection is, precisely, the debate to determine what these units can be, how they can be defined, how they can be individualized, and how they can be empirically studied. The question is partially scientific—can we empirically discover these entities?—although inherently philosophical, insofar as it concerns *causality*, i.e it is a question about the entity on which causality operates in the biological process of evolution by natural selection (**chapter II, section 5**) . Insofar as it concerns causality, it creates problems related to how causality must be understood, and how it must be applied in this case. Determining which the entity (or entit*ies*, as we will see) on which natural selection causally acts, has worried biologists and philosophers for more than one century. I will review the main approaches that have been offered to answer the question.

3.1. Lewontin's original formulation

The original formulation of the debate about the units of selection, as this debate has been understood in the last half century, is conventionally credited

[74] In this paragraph I am motivating the debate in Aristotelian terms: a property (*idion*) must inhere in an entity (*ousía*), or there is no property!

to Richard Lewontin (1970): 'The units of selection'.[75] His paper starts as follows:

> 'The principle of natural selection as the motive force for evolution was framed by Darwin in terms of a "struggle for existence" on the part of organisms living in a finite and risky environment. The logical skeleton of his argument, however, turns out to be a powerful predictive system for changes at all levels of biological organization. As seen by present day evolutionists, Darwin's scheme embodies three principles (…):
>
> 1. Different individuals in a population have different morphologies, physiologies, and behaviours (phenotypic variation).
> 2. Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
> 3. There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).
>
> These three principles embody the principle of evolution by natural selection. While they hold, a population will undergo evolutionary change' (Lewontin 1970: 1)

Interestingly, Lewontin's formulation is very abstract, as well as quite general and broad in scope, two features that Lewontin himself recognises. First, his three principles are applicable to *individuals*—Lewontin conventionally refers to 'multicellular organisms'—but they would be applicable to 'any entities in nature that have variation, reproduction, and heritability' (Lewontin 1970: 1), a point he proves clearly in the rest of his paper by applying the principles to organelles, cells, gametes, etc, i.e. to different levels in the biological hierarchy. Second, the meaning of the key terms in the principles is left completely open. In other words, neither a particular mechanisms of inheritance is presupposed by Principle 3—'[n]o particular mechanism of inheritance is specified, but only a

[75] The reader might be puzzled that I start with a paper from the 70s, provided that I said that the debate has concerned philosophers and biologists for more than one century. Nonetheless, starting with Lewontin's paper is the conventional practise in philosophy, since it is the moment when the collaboration between biologists and philosophers started to flourish, and the debate became *a debate*, properly speaking, and not a mere disagreement in opinions (Lloyd 2017c). I will thus follow that convention here.

correlation in fitness between parent and offspring' (Lewontin 1970: 1)—nor the reason(s) why different phenotypes contribute differently to different generations (Principle 2) is given.

The generality of Lewontin's principles is such that he believes that they could (and must) be applied not only to the convential elements that he describes in his hierarchy, but also to higher level entities, such as multispecies communities:

> 'At yet higher levels, the species and *the community, natural selection obviously must occur*. Species evolve to survive in a certain environmental range, and if the environment should suddenly change, some species will become extinct but others will survive. *The same is true of communities* whose stability of composition depends upon the interaction among their constituent species.' (Lewontin 1970: 15, emphasis added)

The last point is remarkable for at least two reasons. First, because it shows that Lewontin's self-conception of his framework for the units of selection is such that it applies to *every imaginable* level in the biological hierarchy, without exception. Second, because it can be applied to holobionts, as holobionts are communities of species. Let me expand on the second remark. In Lewontin's framework, an entity *x* is a unit of selection if and only if *x* exhibits variability, fitness differences and heritability. If communities of species, that is holobionts, are units of selection—which is something Lewontin seems to be claiming in the paragraph above—then they must exhibit the three properties, *in one way or another*. And remember that Lewontin's formulation of the properties is merely functional; that is to say, it is not tied to any particular mechanism that realizes the properties. Therefore, the task for the 'Lewontinian' is to discover which are the mechanisms that communities of species—or holobionts, in our case of interest—have that make them susceptible of materially realizing the three

properties, and thus being units of selection.[76] But notice that for the Lewontinian, the fact that communities of species are units of selection is an *a priori* fact, whose *a posteriori* (empirical) consequnces would be to discover the mechanisms that materially realize the properties that make the community a unit of selection. Therefore, and this argument is crucial for my doctoral project, and of course for defending HCE from the Lewontinian perspective: a) it is legitimate to claim that a conglomerate (superorganism, holobiont, etc.) is a unit of selection; b) a consequence of the claim would be trying to discover the biological mechanisms that allow the conglomerate to realize the properties that make it a unit of selection; c) some of these properties must be new or unknown; and, thus: 1) it is illegitimate, from this perspective, to argue that a conglomerate is *not* a unit of selection because it does not realize any of the mechanisms conventionally attributed to the units of selection (e.g. germ/soma specialization; functional cohesion; bottlenecks; etc.); 2) the most rational way of proceeding is trying to discover the mechanisms (some of them possibly new, unknown before the entity was proposed as a unit of selection) that the conglomerate possesses and that allow it to act as a unit of selection.[77] A clear implication of this is that there must exist a 'diagnostic criterion' for identifying units of selection that is previous to the discovery of the mechanisms that allow the entity (or conglomerate) to satisfy Lewontin's conditions.[78] I will come back to this point in **section 3.4**.

[76] To avoid confusion, by 'mechanisms' or 'material realization', I refer to biological properties such as having bottlenecks, having germ/some specialization, etc. (specified in **section 2**), which are considered as *realizers* of biological individuality, or realizers of properties such as inheritance. The claim thus is that holobionts *must* have some of these realizers, even if they are different to those conventionally known and described in the literature on biological individuality.

[77] Notice that this argument can be perfectly made against Godfrey-Smith's notion of Darwinian individuality, as Sterner (2015) has already noted, although in a different fashion than the one proposed here.

[78] Of course, one might wonder how we can decide what counts as a biological assemblage, and what does not, as well as whether this already entails a view about biological individuality. But notice that this would just generate a chicken-egg problem. It is clear, though, that, as it occurs in the case of species, there are some basic 'diagnostic criteria' that everyone would accept, e.g. the whole set of 'stuff' that occupies a position <*x, y*> at time *t*, and that seems to have a defined boundary. I will assume, thus, that there are basic diagnostic criteria that everyone would agree upon to determine what counts as a biological assemblage, and what does not.

At this point, a clarification of Lewontin's framework is required. In his framework, he presents the three principles that he takes to be independently necessary and jointly sufficient for an entity to qualify as a unit of selection and refers to them as: (1) existence of phenotypic variation; (2) existence of differences in fitness derived from the differences in the phenotypes; (3) heritability of fitness. I want to call the attention to the combination of conditions (2) and (3), as he seems to be conflating the notion of 'fitness', with the notion of 'phenotypic trait' (**section 3**). On the one hand, the requirement, in (2), that individuals with different phenotypes 'score' differently on fitness appeals to the differences in survival and *reproduction*. And, what is more acute, when Lewontin re-takes his formulation again across the paper, he refers simply to the concept of 'reproduction' to talk about the differences in fitness. On the other hand, in (3), Lewontin seems to be using a different notion than fitness, when he claims that fitness must be heritable. To start with, because he argues that Darwin's theory explains the *distribution of phenotypes* in a population (i.e. why the organisms of a species have the phenotype they have, and which ones are getting 'selected'). Second, because for this type of explanation to be accurate, what needs to be heritable is not the difference in the capacity to reproduce (reproduction rate), but the phenotypic feature that is undergoing positive selection. And, importantly, the two notions are not equivalent, as Godfrey-Smith has demonstrated: fitness does not need to be heritable for the traits in the population to be heritable, and only the latter are what matter to argue that there is evolution by natural selection going on in a population (Godfrey-Smith 2009: 17-27). For this reason, I will rather use a different formulation of the general Lewontinian principles that makes the difference between (2) and (3) explicit:

> 'A sufficient mechanism for evolution by natural selection is contained in three propositions:
>
> 1. There is variation in morphological, physiological, and behavioural traits among members of a species (the principle of variation).
>
> 2. The variation is in part heritable, so that individuals resemble their relations more than they resemble unrelated individuals and, in particular, offspring resemble their parents (the principle of heredity).

> 3. Different variants leave different numbers of offspring either in immediate or remote generations (the principle of differential fitness).' (Lewontin 1985: 76)

Notice that in this new formulation the possible equivocation with the ambiguity of the concept of 'fitness' vanishes: evolution by natural selection can be detected whenever there is (1) variation in character; (2) heritability *of this variation*; and (3) fitness differences among variants. There are two characteristics of the new formulation, though, that should not be left unnoticed. First, in the new version, the conditions are introduced as *sufficient* for natural selection to occur, but not as *necessary*. Second, the point about parent-offspring resemblance, which in the original version only appeared in 3 ('fitness is heritable') and was expressed as the necessity of a correlation, is reinforced in the new version. In fact, this point is doubly reinforced in the new version, as Lewontin complements it with the following remark: '[i]f variation exists but is not passed from parent to offspring, then the differential reproductive success of different forms is irrelevant, since all forms will produce the same distribution of types in the next generation' (Lewontin 1985: 76). These two features are fundamental because they provide a good guidance for further debates about the units of selection, and they provide good guidance for approaching HCE defender's claim that the holobiont is a unit of selection. Overall, Lewontin's approach can be abbreviated as follows: for an entity to be a unit of selection it is sufficient, *but not necessary*, that the entity exhibits heritable variation in fitness, i.e. that there is fitness-mediated transgenerational parent-offspring resemblance with respect to type. If these conditions are satisfied, then evolution by natural selection will occur.

3.2. The interactor/replicator framework

Soon after Lewontin provided his original formulation of the 'recipe' that units of selection must satisfy, Richard Dawkins (1976/2006) published his well-known book *The Selfish Gene*, where he proposed his gene-centred view of the evolutionary process and suggested a new way of conceiving the units of selection controversy, which in certain sense specifies Lewontin's framework. Dawkins' view of the evolutionary process was soon made more precise by

David Hull (1980, 1981, 1988), and became the standard view among biologists and philosophers of biology for the next two decades (Dawkins 1982a, 1982b; Brandon 1982, 1988; Sober and Lewontin 1982; Sober 1984; Mitchell 1987; Lloyd 1988, 1992, 2001, 2017a; Sober and Wilson 1994).[79] Let me put a little bit of context here, before I introduce Dawkins and Hull's framework, since it will be useful to fully comprehend why the authors decide to suggest their specific type of approach to address the controversy about the units of selection.[80]

Before Dawkins published *The Selfish Gene*, there was an agitated debate among evolutionists about which was the entity or entities that natural selection could 'select', i.e. which was the level of the biological organisation at which selection was more effective.[81] During the 50s and the 60s, it was common among biologists—who mainly followed Darwin's observation about the possible action of natural selection 'for the good of the group' (1871, *The Descent of Man*)—to interpret some types of animal behaviour as adaptations that benefitted the group or species that the animal belonged to. Such general appeal to group benefit was explicitly endorsed by V. Copner Wynne-Edwards in his major work (1962): *Animal Dispersion in Relation to Social Behaviour*, where he acknowledged the general existence of between-group selection in nature, and showed no objection in postulating it as one of the main causes of evolution and the existence of adaptations. Wynne-Edwards's views comprised a forceful attack on the Darwinian orthodoxy according to which the individual organism is the main unit of selection, i.e. it is the main unit on which adaptations evolve. Wynne-Edwards views, however, were soon criticized by Maynard-Smith (1964) and, especially, by Williams (1966). Williams' criticism was based on the existence of a conceptual distinction between '*adaptations*', and '*fortuitous benefits*': in his view, adaptations look at the *past history* of the trait, whereas fortuitous benefits look at its future. Wynne-Edwards confuses the

[79] A notable exception to this trend is developmental systems theory (Oyama 1988; Griffiths and Gray 1994, 1997).

[80] As in most philosophical debates, distinctions and alternative frameworks to conceive the same process are usually introduced to overcome one problem or set of problems that philosophers see as 'urgent' and in need of a solution. The same is true about the units of selection controversy, so introducing the context is fundamental, and prior, to introducing Dawkins and Hull's interactor/replicator framework.

[81] The reconstruction is based on Okasha (2001): 'Why won't the group selection controversy go away?', *Br J Philos Sci* 52: 25-50, and in Gould (2002): *The Structure of Evolutionary Theory*, pp. 544-556.

two notions and, thus, his arguments about the widespread existence of group selection are just misguided.

Let me explain Williams' idea more succinctly, because his criticism is fundamental to understand the debate about the units of selection. According to Williams, a trait is said to be an adaptation for doing *p* if and only if the trait evolved because there was selection for the trait, and there was selection for the trait precisely because having the trait promoted doing *p* (Sober 1984: 208). Applied to the context of group selection, a trait can be argued to be an adaptation for the group if and only if there was group selection for having the trait (Sober 1993: 85). However, and this is the key argument to be found in Williams' criticism of Wynne-Edwards, the fact that one trait now benefits the group of entities that bear that trait does not mean that the trait was selected in the first place *because* it benefitted the group. Notice the causal language that appears here: to say that a trait evolved because it conferred benefits to the group whose individuals bear the trait entails that one is able to tell a causal story of why the trait is there, a causal story that appeals to the benefits that the trait confers to the group. As I said in **chapter II**, natural selection provides a causal explanation of why a trait exists, so the existence of a causal story that can be told is fundamental to talk about group adaptations.[82] Williams' point against Wynne-Edwards is, thus, that what he regards as traits that have been *selected for* the good of the group might benefit the group *now*, but this benefit: 1) might not be the reason why the trait exists on the first place (i.e. it does not have the right kind of causal history); 2) might be the consequence of *kin selection* or *inclusive fitness theory*, which is presented as an alternative to group selection that would explain the existence of the same type of traits that group selection explains, without appealing to a benefit for the group (see Hamilton 1964a, 1964b, for the concept of 'kin selection').

[82] There is a whole debate about what natural selection *really* explains (Stegmann 2010; Birch 2012; Díez and Lorenzano 2013; Mogensen 2016), or about whether explanations that appeal to natural selection are causal or rather statistical (Walsh et al. 2002; Brandon and Ramsey 2007; Lewens 2010; Huneman 2012). I will not enter in any of these disputes here, as that is outside the scope of this thesis. I will only appeal to the explanatory nature of the notion of 'natural selection' for illustrative purposes.

About 1), Williams made the following point: the individuals that compose a group have generally shorter living times than the groups themselves; if this is so, then the opportunity for between-group selection will systematically be eroded by the opportunity for between-individual selection. Why? Because for natural selection to be an efficient driver of evolution, it is necessary that the degree of variation within the object under consideration is low, so that the possible effects that a specific variation *v* might have on the fitness of the objects that bear *v* allow these objects to outcompete other objects in the populatin that do not bear *v*. However, if each object in the population is constantly varying due to selection at the lower level, then natural selection is not possible. This condition, he believes, is not satisfied among groups, since they experience constant variation and, therefore, the opportunity for group selection is rare, compared to the opportunity of individual selection. Concerning 2), Williams develops the following line of reasoning: According to kin selection, some traits (e.g. offspring caring behaviour) that would in principle seem like group adaptations, are in reality individuals adaptations, because the individuals that evolve those traits will be in a clear fitness advantage over those that do not. Importantly, for an adaptation to be counted as a case of kin selection, it is necessary that the trait evolved to benefit those that are alike, i.e. to benefit the relatives (offspring, or tightly connected relatives). Those cases, detractors of Wynne-Edwards argue, are not to be counted as cases of group selection and, furthermore, most cases of what Wynne-Edwards counts as group selection are indeed cases of kin selection.

Given the context aforementioned, it is now time to introduce Dawkins' solution to the tension between group selection and kin (or individual) selection, which was just the tension between Wynne-Edwards and Williams. It must be noted, although in passing, that by the moment when Dawkins introduces his solution, this was seen more as a 'challenge from below'—and Wynne-Edwards' model was interpreted as a 'challenge from above'—in the sense that it also seemed to question the role of the organism as a unit of selection. However, this perception was somehow misleading, since one of his points consisted precisely in defending that the organism is *also* a unit of selection, in at least one of the two different meanings that the concept can adopt (Dawkins

1982a; see Wilkins and Bourrat 2018: section 3). Let me introduce his solution to the tension by quoting Dawkins' own words in the 'Introduction to the 30th Anniversary Edition' of his opus magnum, where he re-evaluates the importance of *The Selfish Gene* for the debate about the units of selection:

> 'I should perhaps have gone for *The Immortal Gene. The Altruistic Vehicle* (…). Perhaps it would have been too enigmatic but, at all events, the apparent dispute between the gene and the organism as rival units of natural selection (a dispute that exercised the late Ernst Mayr to the end) is resolved. There are two kinds of unit of natural selection, and there is no dispute between them. The gene is the unit in the sense of replicator. The organism is the unit in the sense of vehicle. Both are important. Neither should be denigrated. They represent two completely distinct kinds of unit and we shall be hopelessly confused unless we recognize the distinction.' (Dawkins 1976/2006: ix)

This paragraph introduces Dawkins' popular view about the evolutionary process. The evolutionary process is, for him, the process by which natural selection causes the evolution of two special types of entities: on the one hand, the gene, which is the unit of selection in the sense of being a replicator; on the other, the organism, which is the unit of selection in the sense of being a vehicle. But, what does this jargon exactly mean? And, more importantly, what is the significance of introducing two units in a process which originally involved the existence of only one?

Dawkins believes that we must carefully distinguish between: on the one hand, the units that have the properties that a successful unit of selection must have; and, on the other hand, the larger units that these original units must form to foster their own success. Dawkins refers to the first units as *replicators*. Replicators are the entities of which copies are made, and that share three basic properties: longevity, fecundity, and copying-fidelity. Why are these three properties necessary to determine the degree of success of the replicator? The answer to this question is simple: if the phenotypic effects of an entity are going

to outcompete the phenotypic effects of another, then the entity must be long-lived enough so that the effects can accumulate, a lesson we learnt from Williams. Longevity, fecundity, and copying-fidelity guarantee that this is so. The unit that fulfils these three properties is the replicator, and the (selfish) gene is for Dawkins the paradigmatic case of a replicator.[83] On the other hand, Dawkins also recognises the existence of a different type of unit of selection that he calls the *vehicle*. The importance of introducing the vehicle and its necessity as a unit of selection becomes clear in the following passage taken from Stephen J. Gould:

> 'I find a fatal flaw in Dawkins' attack from below. No matter how much power Dawkins wishes to assign to genes, there is one thing he cannot give them –direct visibility to natural selection–. Selection simply cannot see genes and pick among them directly. *It must use bodies as an intermediary*.' (Gould 1977: 24, emphasis added)

Notice the problem that Gould is pointing at. It might well be true that replicators are long-lived, that they have a high-degree of copying-fidelity and that they are fecund. However, they are invisible to natural selection, which can only see and, thus, *proximately select*, phenotypic differences among individuals. But notice that this is a consequence of the fact that replicators tend to appear combined with each other forming chromosomes, cells, organisms, multispecies communities, etc. (Lewontin's hierarchy). All these structures are not replicators: they are vehicles that: 1) are directly visible to selection, and 2) their success ultimately determines the rate of success of the replicators that compose them. In Dawkins' words: 'Vehicle selection is the differential success of vehicles in propagating the replicators that ride inside them.' (Dawkins 1982b). At this point, two things become clearer: first, that part of the disagreement about the units of selection turns out to be a semantic question, due to the fact that different researchers mean different concepts by using the

[83] Notice that the replicator is a conceptual tool, the name to refer to the unit of selection that satisfies the three properties, and the gene is the biological entity that satisfies the properties. In Dawkins' framework, however, other entities, such as memes, can also act as replicators and, what is more interesting, due to the empirical nature of the category, it is perfectly conceivable that other entities apart from genes could also be replicators, provided they satisfy the desired properties.

same expression ('units of selection' is polysemic);[84] second, Dawkins' observation in the 30th years Edition of *The Selfish Gene*, where he claims that a more adequate title would have been *The Immoral Gene. The Altruistic Vehicle*.

Let me start from the beginning. In which sense is the dispute about the units of selection semantic? Dawkins gives the following example: 'The controversy about group selection versus individual selection is a controversy about whether, when we talk about a unit of selection, we ought to mean a vehicle *at all*, or a replicator.' (Dawkins 1982b). In Dawkins' view, Wynne-Edwards, as well as other defenders of group selection, are referring to the vehicles, insofar as groups are in part responsible for the success of the replicators that make up the group. On the other hand, Williams, as well as other detractors of group selection, usually mean the 'replicator' when they talk about the unit of selection, and thus strongly oppose Wynne-Edwards' hypotheses about group selection.[85] Thus, the dispute between group selection and individual selection is partially a semantic dispute, as the authors mean different concepts by the expression 'unit of selection'. However, the dispute is only *partially* semantic, since sometimes they also disagree about which entity is the vehicle of selection, and there is a genuine empirical debate, as Dawkins reminds us: 'the organism and the group of organisms are true rivals for the vehicle role in the story, but neither of them is even a *candidate* for the replicator role. The controversy between "individual selection" and "group selection" is a real controversy between alternative vehicles' (Dawkins 1976/2006: 254-255).

Second, Dawkins believes that a more appropriate title for his book would have been *The Immoral Gene. The Altruistic Vehicle*. Why? Because in Dawkins' view of the evolutionary process, the real beneficiary of evolution, the

[84] A semantic question, but not a verbal dispute. There is a genuine semantic disagreement about the meaning of 'unit of selection', the disagreement is not merely about what to call the different units of selection.

[85] It is not strange that Dawkins explicitly acknowledges that the properties that he attributes to replicators, and that he had polemically expressed in *The Selfish Gene* by demanding the immortality' of replicators (1976/2006: chapter 3; see also 1982b), derive from the properties that Williams attributes to the units of selection.

real unit that has the right properties and that pursues its own interests is the replicator—which is empirically the 'gene', conceived as 'a piece of chromosome which is sufficiently short for it to last, potentially, for *long enough* for it to function as a significant unit of natural selection' (Dawkins 1976/2006: 45)— whereas vehicles are there altruistically, just to serve the interests of the replicators that make them up. Importantly, in Dawkins' view, the replicators make their vehicles, and the replicators that survive longer are those that make better vehicles, as the vehicles are the entities that compete which each other to pass on their replicators. Using the conventional distinction between proximate and ultimate causation (Mayr 1961), it could be argued that the replicators are the ultimate beneficiaries of the evolutionary process, whereas the vehicles are only the proximate causes that determine the success of the former.[86]

Dawkins' view of the evolutionary process is very controversial, and has been subjected to a lot of criticism, that I will not examine here. But one important criticism that I will examine came from David Hull, who made a fundamental contribution by clearing up the notions of *replicator* and *vehicle* that Dawkins had originally introduced (Godfrey-Smith 2000: 404-407). For Hull (1980, 1981, 1988)[87], the process of natural selection encompasses two different types of entities: on the one hand, an entity whose *structure* persists throughout evolutionary time; on the other, an entity that has the capacity of interacting with others and, thus, causes differential replication. According to Hull, these two types of entities had not been clearly defined by Dawkins, and even if he had introduced the notion of replicator and vehicle, the distinction was not clear enough to make sense of these two roles. The two entities introduced by Hull came to be known, respectively, as the *replicator* and the *interactor*. To quote:

[86] Lisa Lloyd would disagree with referring to the replicator as the beneficiary of the evolutionary process, because in her view of the units of selection, there are not two, but four different questions (Lloyd 1992, 2001, 2017a). I think she is right that the question about which entity is the replicator and the question about which entity is the beneficiary of the evolutionary process are not necessary the same question. However, this does not refute the claim that Dawkins mixes the two categories in one, and that's the reason why I introduce Dawkins' replicator by calling it the 'ultimate beneficiary'.

[87] It is conventionally accepted to use the definition of the replicator that Hull uses in his (1980). However, I think the definitions that he gives in his excellent (1988), is far more precise, as it eliminates the weird requirement of 'directedness' from the definition. I will thus take (1988) as my referent.

> ‘[I]n an effort to reduce conceptual confusion, I suggest the following definitions:
>
> *replicator* – an entity that passes on its structure largely intact in successive replications
>
> *interactor* – an entity that interacts as a cohesive whole with its environment in such a way that this interaction *causes* replication to be differential
>
> With the aid of these two technical terms, selection can be characterized succinctly as follows:
>
> *selection* – a process in which the differential extinction and proliferation of interactors *cause* the differential perpetuation of the relevant replicators.’ (Hull 1988: 408-409).

I now need to explain more clearly what Hull was getting at with this distinction. To do so, let us go back to Gould’s criticism of Dawkins. According to Gould, the main problem with Dawkins’ replicator-centred account was that genes (or replicators, see fn. 83) were not directly visible to the process of natural selection. And this was a non-salvable obstacle for the replicator account. Remember again a point I made in **chapter II**, and that has been repeated several times: natural selection is a causal process. Evolutionary change can be driven by many different types of processes; those different processes generate a sorting among the interacting entities. Among those processes, selection is *the causal* process. Relegating natural selection to the persistence of replicators simply misses the main point. Hull’s vision of natural selection, on the contrary, overcomes this difficulty: natural selection is a story of *both* replicators *and* interactors. The replicators are the structures that persist through evolutionary time (structural unit). The interactors, on the other hand, are the cohesive wholes that causally determine how successful these structures are (functional unit). As Hull expresses his thoughts:

> ‘If an entity is to function as a replicator, it must have a *structure* and be able to pass this structure on to successive generations of replicators. (…)

> Although replicators might be part of functional systems, they themselves need not be functional systems.
>
> Interactors must exhibit structure but toward quite different ends – they must be able *to cope with their environments*'. (Hull 1980: 318, emphasis added)

At this point, I want to make a clarification. What is at stake in this debate? Why does it seem so necessary to discriminate between structural units (replicators) and functional units (interactors)? Here, again, a bit of history will help. When Dawkins wrote *The Selfish Gene*, junk DNA had recently been discovered. As its name suggests, junk DNA is a portion of the DNA in an organism with no defined biological function, i.e. a branch of DNA that is not coded into a biologically functional protein.[88] In those days, to account for the maintenance of junk DNA was paradoxical: if natural selection tends to favour variants that increase the fitness of their bearers, and junk DNA does not produce any kind of biological function, then it does not increase the fitness of its bearers and is expected to be selected *against*. Dawkins was interested in explaining why junk DNA was so common, and thus he came up with the idea that the real beneficiary of selection, the real and more genuine unit of selection, was the structure that was able to copy itself, at zero cost, and guaranteeing its survival throughout the cost for a third entity, i.e. the selfish gene. The selfish gene has the interesting peculiarity of getting copied despite producing no phenotypic effect at all, and mostly thanks to the phenotypic effects produced by other non-selfish genes that it happens to 'live with' within a genome (Doolittle and Sapienza 1980). The existence of junk DNA is thus Dawkins' inspiration to formulate his distinction between replicators *qua* structural entities and vehicles *qua* functional entities, which Hull would later reformulate as a distinction between replicators and interactors.

Notice, thus, that the questions I have asked in the previous paragraph get an immediate response here. First, it is necessary to discriminate between

[88] The exsistence of junk DNA is now questionable for the fact that a branch of DNA has no biological function does not mean that it does not have biochemical activity, i.e. junk DNA might be transcribed, but the proteins that are produced fail to have a biological function. See Doolittle (2013a).

structural and functional units of selection because it is necessary to distinguish the units that get selected due to their own interactions with the environment (i.e. the phenotypes, and thus the genes that code for phenotypic properties), and the units that get selected because they are good at making copies of themselves. [89] Furthermore, the existence of junk DNA strongly suggested, at least for Dawkins, that the real winner of the evolutionary process was the entity that made copies of itself, and that any (possible) phenotypic effect of this entity would be, at most, instrumental: it would simply be a way of getting copied better. But evolution was not just a tale of making good phenotypes, but a tale of using those that make good phenotypes to get copied better, without taking the effort of making the good phenotype yourself. These points will be fundamental, because most of the problems that gave rise to the formulation of the distinction between replicators and interactors will re-appear in the discussions of HCE (**chapter II**), and my response to them will be very similar to the responses that have been given in the past (**chapter V**).

Having settled the interactor/replicator framework, a fundamental question arises: how is this framework connected to Lewontin's hierarchical view of the units of selection? Does the interactor/replicator framework entail that the biological hierarchy is just composed by two units of selection, and that all the debate that Lewontin (and Wynne-Edwards, and Williams, etc.) started about the role of the organism, the superorganism, or the group, as the unit of selection must be forgotten? The exact answer to the first question is unclear, as Samir Okasha (2006) and Godfrey-Smith (2009) have discussed at length in their work. Nonetheless, both the first and the second question might be replied to as follows: once the replicator/interactor framework was formulated, it became widely assumed that the replicators were, empirically speaking, the genes, conceived in Dawkins' structural sense of 'an entity capable of making copies of itself' and, thus, they became recognised as the entities that

[89] It is questionable that 'being good at making copies of onselef' entails 'getting selected', as Dupré has told me, because the ones that are selected are those that have 'adaptive phenotypic consequences'. I think he is right about pointing that out. However, it is conceptually possible to argue that since different genes have different 'copying capacities', their capacity to make copies of themselves is indeed an adaptive phenotype of the gene. This seems to me the most charitable interpretation of Dawins and Hull's view of natural selection, at least.

accounted for the phenomenon of heredity.[90] Therefore, the debate about the units of selection became a dispute about which entity is the interactor; in other words, to put it in Robert Brandon's terminology, the scientific relevant problem became to identify 'the hierarchy of interactors'. Thus, the interactor/replicator framework somehow subsumes Lewontin's view under the proviso that all heredity will be carried out by replicators, and thus one should not worry about hereditary relations when she is trying to discover which entities play the role of the interactor. The debate about the units of selection, thus, became reformulated as the debate about the entities that interact cohesively so that, *as a consequence*, replication is differential.

I have highlighted '*as a consequence*' for one simple reason: what became relevant during the 80s and the 90s of the 20th century was to discover at which level natural selection *causally* acts. And remember that in Hull's definition of the interactor and the replicator, the only entity that plays a genuine causal role in evolution is the interactor. This is thus another argument to support my view that the entity that played a philosophically significant role for those decades was the interactor, and the replicator was just seen as a 'lucky survivor' of the process of selection. Importantly, many of the philosophers that proposed elaborated accounts to cope with the problem of causality did so thinking about interactors (see Brandon 1985; Lloyd 1988). And, indeed, as I already said in **chapter II**, the main issue about the units of selection is how to distinguish sorting from selection, that is, how to distinguish a random effect of sampling from a causal effect derived from having a fit phenotype. I will come back to these issues later, now it is time to revise the main problems that the interactor/replicator framework has.

[90] Of course, this opinion needs to be put into context, since some people have also argued that there were both a hierarchy of interactors *and* a hierarchy of replicators (e.g. Brandon 1988), under the assumption that each interactor requires the existence of a specific replicator that accounts for the parent-offspring similarity at that level of the biological hierarchy. Brandon's view is nonetheless not the standard one, and in fact the most popular view of the replicator recognizes two entities that can play this role: the gene and the meme, the last being the unit of cultural heredity.

3.3. Problems with the interactor/replicator framework

As I said at the beginning of the last section, the interactor/replicator framework developed by Hull and Dawkins was the dominant view about the units of selection for about two decades. However, after some powerful critiques, that framework became widely questioned, and has been slowly abandoned. What are the main problems that the interactor/replicator framework created, and that has moved researchers to abandon it? I will present three different types of criticism that, in one way or another, have been formulated in the literature. The first one, which is partially derived from Gould (2002), and refined by Godfrey-Smith (2009), concerns the empirical correctedness of the view of the selection process that the interactor/replicator framework presumes. Secondly, I will introduce a criticism partially derived from James Griesemer (2000a, 2000b), and the *major transitions in evolution* research project, which questions the necessity of the replicator as a category, on the basis that it leaves too much unexplained. Finally, I will introduce an 'Ockham razor's' type of criticism, which basically states that the replicator/interactor framework multiplies the units of selection beyond necessity, postulating the existence of two entities for a process that could perfectly be explained by postulating the existence of one.

First of all, in his extensive (2002), zoologist Stephen J. Gould strongly opposes the interactor/replicator framework based on his strong opposition to the concept of 'replicator'. According to him, while it is true that the concept of the 'replicator' captures the role of the gene as a very important entity in the evolutionary process, he thinks that the role of replicators in natural selection is completely irrelevant. Let me present his argument with his own words:

> 'Replication identifies a valid and important criterion for the crucial task of bookkeeping or tracing evolutionary change; but replicators cannot specify the causality of selectionist processes, which must be based upon the recognition and definition of interactors with environments' (Gould 2002: 72)

Notice what it is at stake here. To do so, it is useful to go back to a point I made in **chapter II**, about the distinction between sorting and selection, as well as to the point made at the beginning **section 3** of this chapter, concerning the two different notions of 'fitness' that Lewontin and Arriew (2004) pointed to. The point I made there is that natural selection is a question of finding causes, i.e. of finding the reason why biological individuals have the traits they have and that make them fit their environments. And, precisely, the problem with population genetics, replicator-based models of natural selection is that, at most, they can track the effects of selection on the genetic pool. But this is a far cry from identifying the actual causes that operated in the population on the first place, and that in the end caused that distribution (see also Sober and Lewontin 1982). The primary problem that makes Gould suspicious of the interactor/replicator framework is, thus, that it conflates the effect of selection with its actual cause, i.e. it conflates sorting with selection.

But this is not the only problem that Gould sees with the interactor/replicator framework. There is a more fundamental problem that affects the conceptual necessity of replicators for selection processes. Remember that, in the interactor/replicator framework it is generally assumed that the genes play the role of the replicator, insofar as they are capable of making copies of themselves and thus maintaining their structure for the longest time. However, there is a historical and, I would say, empirical paradox in this account. When Darwin first presented his theory he was completely unaware of the existence of entities that could perfectly copy their structure, i.e. he was completely unaware of the existence of genes. He simply formulated his theory on the basis of his empirical observations on artificial selection experiments, as well as some field observations, both of which were completely disconnected from any knowledge about the material basis of heredity. This criticism against the interactor/replicator framework is relevant for two reasons. First, because it makes the valid and substantial point that what is empirically necessary to 'presume' the action of natural selection is simply parent-offspring resemblance, however this is understood; second, because it puts the focus on why the transmission of a subset of particles from parent to offspring, while empirically relevant, is nothing but a historical contingency of the course of evolution. It is

perfectly plausible that evolution would have been different, and that natural selection could act on individuals with a completely different hereditary basis to the one living creatures happen to have (see also Godfrey-Smith 2009: 36-39). In fact, the existence of other non-genetic channels of inheritance, like epigenetic inheritance, or cultural inheritance, put stress on the notion of the replicator and, importantly, question its validity as a scientifically relevant category (Jablonka and Lamb 2005; Pontarotti 2016).

Secondly, a substantial argument against the interactor/replicator framework derives from the work on the major transitions in evolution and, especially, from the work of James Griesemer (2000a).[91] The best way of appreciating this criticism is, I think, going back to the original definitions of the replicator and the interactor. According to them, the replicator is an entity with a high degree of *copying-fidelity*, insofar as it has the capacity of getting its structure preserved, and the interactor is a highly *cohesive* entity that interacts with its environment as a single unit. The paradox of these definitions is this: are not the capacity of making highly accurate copies of oneself and the property of being a highly cohesive unit already highly *evolved* biological characteristics whose existence needs to be explained by natural selection? Notice the problem here. Let us assume that the interactor and the replicator are the only units of selection that exist *now*. If this is so, then evolution by natural selection will only occur in the empirical objects that happen to have the features of the interactor and the replicator. However, what becomes paradoxical is that the very existence of these objects needs to be explained by the same principles that explain why they evolve by natural selection *now*.[92] That is, we need natural selection to explain why certain objects became, say, 'paradigmatic' units of selection. But this immediately puts the cart before the horse, since it is at least conceptually perplexing to explain the existence of the units of selection by the action of natural selection on objects that are not themselves units of selection!

[91] A similar criticism appears in many other works, most notably in the work on group selection by Wilson and Sober (1989), and Sober and Wilson (1994, 1998).

[92] Notice that when I use the expression 'evolve by natural selection' I mean evolve *in response to selection*. That is, evolve according to the formula $R = h^*S$, which I introduced and explained in **chapter II**. In other words, it is about which entities evolve according to the breeder's equation.

The tradition of thinking of the levels of selection *diachronically*, that is to say, the study of the evolution of the canonical units of selection that we know nowadays (i.e. the real biological entities that play the roles of interactors and replicators respectively), is conventionally traced back to the original work by Buss (1987). Buss, however, did not exactly think that his work supposed a challenge to the interactor/replicator framework *per se*, but a framework for tracking a different type of question: on the one hand, the identification of the interactors and the replicators will give us information about the entities in which biological adaptations will appear; on the other hand, the study of the origins of individuality will provide information about how new units of selection (i.e. units that can evolve adaptations) came to exist in the first place. In this sense, the two questions, although related—identifying how new units appear would be somehow conceptually prior to the discovery of how the current units can evolve their own adaptations, are conceptually different (Fontana and Buss 1994; Griesemer 2000a; Okasha 2006). However, Buss's view of the problem is not completely correct. Because new units of selection will only evolve if there is a group adaptation that evolves so that two previously extant units become one.[93] But this process, even if seen as a transition in individuality, is the process of *evolving an adaptation* that fuses two previously extant units into a single unit, and thus it is not true that the questions about the entities that become new levels of selection *and* the entities that evolve adaptations can be conceptually separated. This objection puts a fundamental tension on the replicator/interactor framework, since it makes clear that, although it might be a good framework for identifying some candidate units that would presumably evolve adaptations, it falls short of being a sufficiently general view of the process of natural selection.

Finally, there is a substantial problem that results from the appreciation of the previous two problems, in combination with Ockham's principle of parsimony. By separating the properties of 'interacting with the environment' (i.e. bearing fitness) and 'standing in hereditary relations', and attributing them

[93] Which these group level adaptations are is, of course, a question whose answer depends on the author who studies the problem. I will later consider the view of Ellen Clarke, who argues, following Michod (1999) and Queller and Strassmann (2009), that the key adaptation is the appearance of a policing mechanism. See **section 3.3.1.1**.

to two different types of entities, the interactor/replicator framework multiplies the type of biological entities to study to two instead of one. That multiplication would be legitimate if and only if the role played by the two types of entities were: (1) completely necessary for natural selection to happen; and (2) the entities that play these two roles were completely independent from each other. In other words, the postulation of two types of entities that participate in the process of natural selection would be necessary if and only if the type of properties whose existence is explained by appealing to natural selection were so radically apart from each other that it would be impossible to imagine that a single entity could play both roles (i.e. heredity and fitness bearing). Otherwise, Hull and Dawkins would be multiplying the number of entities beyond natural necessity, which contradicts the principle of parsimony dictated by Ockham's razor. Personally, I think the intuitive appeal of parsimony is, at least, philosophically questionable and, at most, completely unjustified from an empirical point of view: why should nature be parsimonious? In any case, setting aside possible methodological discussions about general philosophical principles, this type of criticism to the interactor/replicator framework is a way of making a valid point. It is not just the case that the interactor/replicator framework separates two types of properties and links them to two different types of entities. In addition to that, it partially blocks every type of research on every other mechanism of inheritance that is not channelled by replicators, thus limiting importantly the concept of inheritance. And, in this sense, the strict separation of the two types of entities might turn out to be epistemically harmful and ontologically mistaken.

Let me make a thought experiment to show why this issue is of vital importance, and especially acute for the topic of my dissertation.[94] Imagine that we discover a foreign planet where there are creatures like us, that reproduce and produce offspring, that these offspring resembles their parents on average more than they resemble the offspring of any other random member of the same population, and so on and so forth. The only substantial difference between these creatures and us is that when they reproduce they do not pass any structural particle to their offspring, but those are created *de novo* following

[94] In **chapter V** I will elaborate on this with the real-world example of holobionts. But for the moment, a thought experiment will illustrate this point more abstractly and I think more clearly.

a mysterious recipe that every member in the population knows. Now imagine, for the sake of the argument, that the repetition of this process several times gives rise to different generations and, with time, those individuals that are better suited to their environment happen to be dominant in the population. I have no doubt we would say that the process that gave rise to the dominance of the fitter types in our imaginary planet was natural selection. Now suppose we believe that natural selection only occurs when we have two entities, the interactor and the replicator, as in Hull and Dawkins framework. It is easy to see what the interactors would be in our thought experiment, as I made it easy when I said that the foreign planet was inhabited by 'creatures *like us*'. So, our duty as scientists is finding out the replicators. Can we identify the particles that account for the features we perceive in the population and that lead to the dominance of the fitter type? Actually, we cannot, because in my example the parent-offspring similarity was due to the capacity of the parent to follow a recipe and agglutinating previously extant material so that the offspring would resemble them. For the defender of the replicator/interactor framework, thus, the observed similarities are not a consequence of selection, but of a different, and unknown biological process. But this conclusion is highly counterintuitive. Would not the duty of the biologist be to look for whatever substance or process (in this case, the mysterious recipe and the materials taken to make each offspring) is responsible for the observed parent-offspring similarity, no matter how different this might be from canonical replicators? And, importantly, would it not be more ontologically accurate to say that inheritance, in this foreign planet, is channelled through the mysterious recipe? I suspect everyone would answer to that question affirmatively, which suggests that replicators are definitely not necessary for natural selection.

The thought experiment just explained is in line with the first criticism against the interactor/replicator framework, as this was explained by Gould. But in my case, the experiment has a different purpose. I wanted to show that any argument based on the claim that 'there are no replicators' to justify that natural selection is not occurring at one biological level is simply misguided, as it is

simply a case of inappropriate reification (or 'misplaced concretenes').[95] In general, the discovery of the entities that account for the observed similarities will always be an *a posteriori* task. The way of discovering new levels of selection, thus, cannot consist in identifying both the interactor *and* the replicator, but rather in identifying transgenerational similarity relations that need to be accounted for. Thus, in this case, the Ockham's parsimony principle must be applied for, as I have proven, not doing so would not only not benefit, *but even damage*, scientific research.

3.3.1. Brief reflection. Is there a way out for the interactor/replicator framework? A proposal to reconceive the problem of the units of selection

The three types of criticism I have just presented put a serious challenge to the interactor/replicator view of the selection process. However, I think that defenders of the framework were correctly pointing at something that no one before them had pointed to, namely: that there is a genuine distinction to be made between the set of structural properties that guarantee (even if only contingently) that a concrete pattern of similarity reappears, and the functional properties that make that pattern environmentally efficient (i.e. efficient in competition with other patterns), and thus increase its probability of transgenerationally reappearing. Or, in other words, there is a distinction to be made between the *phenotypic properties* that determine the fitness of a biological individual *vis-à-vis* other individuals, and the *hereditary basis* of those properties. In canonical units of selection it might happen that there is a one-to-one equivalence between the processes that instantiate these properties (one interactor / one replicator), but it must be noticed that the two are logically distinct and, thus, it is possible to imagine the former without the latter, and vice versa. Precisely, I suspect that what the framework of the major transitions in evolution entails is that the two properties can be biologically disconnected, that this is so in many cases, and that a transition occurs precisely when both properties get tightly intertwined with each other.

[95] As presumably there must be a process of inheritance, but this does not need to consist in the transmission of particles. Thanks to Dupré for reminding me of the fallacy of 'misplaced concreteness' to define what occurs in this scenario.

This said, I want now to propose a way out for the interactor/replicator framework, i.e. a possible way of rethinking it so that it can be applied extensively in evolutionary biology, and thus can still be used as a model.[96] This way out, as announced in **section 1**, mirrors Kevin de Queiroz's resolution of the species problem;[97] so let me start by presenting it. First, what is the species problem? Briefly described, the species problem consists in that existence of different, scientifically relevant, and non-overlapping criteria for determining when a new species has appeared, and thus for deciding when two species, instead of one, exist.[98] I will assume, but not argue, that there is such a problem, since it is not strictly relevant, as I will only use de Quieroz's solution as an inspiration for my own solution to the PBI, which I have already proven to be a real problem.[99]

In his (2005a), Kevin de Queiroz notices that despite the enormous divergence among different species concepts (i.e. different criteria for determining species membership and, thus, for determining when there are two instead of one species), all the definitions share a basic commitment: 'species are segments of lineages at the population level of biological organization' (de Queiroz 2005a: 196). The only disagreement among different concepts concerns not the basic meaning of the category, but the *diagnostic criteria* that are used to determine whether and when a concrete segment can be considered a separate species or not. Let me explain this more succinctly, by making use of one example. For some biologists—e.g. ecologists like van Valen—the decision of whether a conglomerate of biological individuals can be divided into one species or must be divided into two separate species depends on whether the individuals share the same adaptive niche. In their opinion, if

[96] The reader might wonder what my motivation in is trying to 'force' a way out for a framework that simply does not work. But my motivation has been clearly stated in the previous paragraph: I think that Hull, Dawkins, and others were pointing to a distinction that needs to be made, and that it is useful.

[97] Interestingly, B. Sterner has also used de Queiroz's resolution to propose a very original account about how to individualte population lineages (Sterner 2017). I only came across with his paper after I had written this section, and even finished the whole draft of the thesis, so I had not enough time to think about how his ideas and mine can be related to each other.

[98] To be clear, this is just one of the species problems, namely, the problem of deciding between competing criteria for species membership. But that's not the only philosophical problem that has been considered under the name of 'the species problem'. For a brilliant exposition of the different types of species problems see Dupré (1993: 37-59).

[99] Solid arguments that motivate the existence of the species problem can be found in Dupré (1993), Hey (2001), Richards (2010), and Suárez (2016).

two groups of biological individuals live in different ecological niches, then they are two species. On the other hand, for other biologists—e.g. Paterson—the criterion of the niche is not valid: what matters, and what must be considered fundamental in deciding whether the two groups constitute one single species or two, is whether they can recognise each other as potential mates.

Notice that the existence of these two criteria can generate conflicts in deciding whether two groups belong to the same species, or to two different species (**Figure 20**).

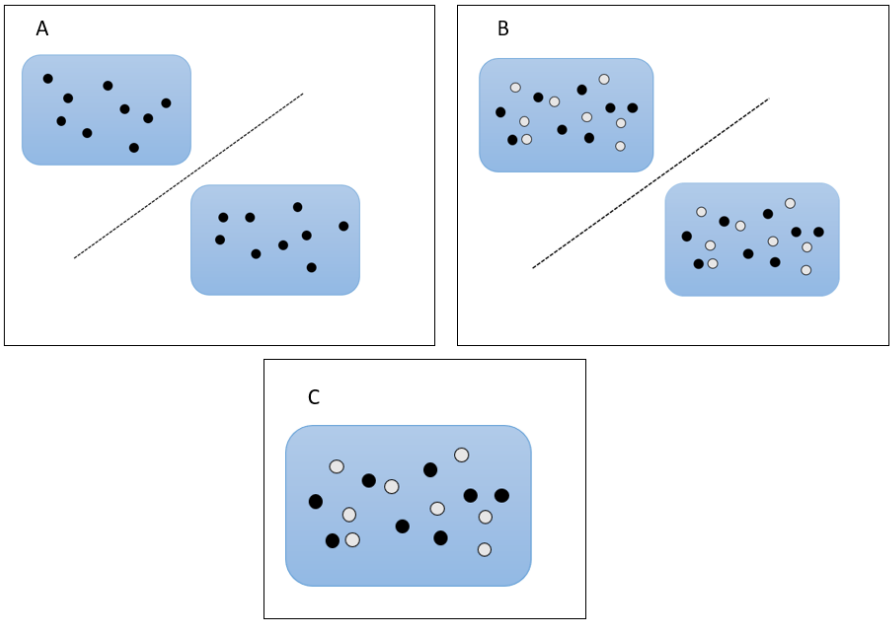


Figure 20. Three possible scenarios where different species criteria would provide different verdicts. Each blue box represents a niche; the dashed line between two blue boxes represents a geographic obstacle that isolates the individuals from each niche from the individuals of the other niches; finally, the spots represent individuals. By convention, the individuals represented in black can only recognise and mate with individuals in black, and the same is true for individuals represented in white. Notice that, in A, there would be two species according to van Valen, but only one according to Patterson, whereas in C the opposite is the case. In situation B, both van Valen and Paterson would recognise the existence of two species, but they would delineate them differently.

Why does this situation occur? De Queiroz's response is simple, but very convincing: speciation is a process during which the members of a group

became dissociated and form two different groups. The process occurs slowly, and it is realized by different types of subprocesses: at some point, the two groups might start living in different niches; later in time they might stop recognising each other as potential mates; still later, they might stop generating fertile offspring; and so on and so forth. There is no biological or metaphysical reason to expect that all the subprocess that drive speciation would occur at the same time: on the contrary, it is probably a more plausible biological hypothesis that the processes will occur in different times. There will probably be a moment in time in which all the subprocess will have occurred, and thus it will be clear that the two groups form two different species. However, in between, there will be a period in which only some of these subprocesses will have occurred, and thus different biologists will give different responses to the question about how many species are there (**Figure 21**).[100]

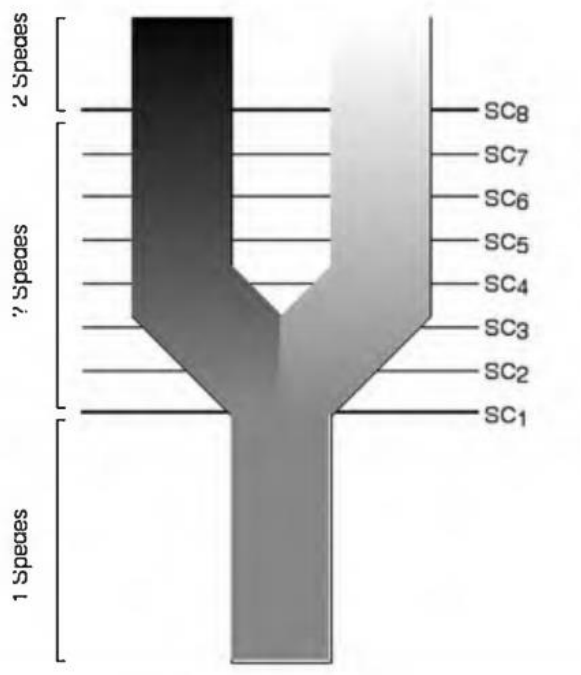


Figure 21. De Queiroz's schematic representation of the process of speciation. SC1-SC8 represent different subprocesses that can be used as diagnosis criteria for determining the existence of two species instead of one. At some point, everyone will agree that the two groups constitute two separate species, whereas at another everyone will agree that there is one single

[100] Notice that this solution is not very different from what John Dupré (1993: 47) has suggested. Nonetheless, where Dupré wonders why it is not legitimate to assume that there are different 'ways of being a species', if different species have different origins (conceived as different criteria), de Queiroz provides a common manner of harmonizing these different 'ways of being a species'.

species. In between, the decision will depend on the subprocess that the biologist is interested in studying. (From de Queiroz 2005a: 204, Fig. 3)

The question now is, how is de Queiroz's resolution to the species problem relevant to think about the PBI and, particularly, to defend the interactor/replicator framework? I think the answer to this question lies, precisely, in the processual nature of biological individuals (Dupré 2012; Nicholson and Dupré 2018). Remember that the point of de Queiroz was that speciation was a process that included a series of subprocesses. The same logic applies to biological individuality: biological conglomerates are individuated always in virtue of the biological processes they participate in (Bouchard 2018). This is clear when one reads the different criteria of biological individuality that I presented in **section 2**: all of them refer to one type of process (physiological, immunological, developmental, etc.) and, importantly, the PBI arises precisely because the entities that are individualized in virtue of their participation in certain types of processes do not necessarily participate in all the other processes that are used to define biological individuality. In **section 2**, this was illustrated with the case of clonal individuals but, of course, it generalizes more, applying to absolutely every candidate for being a biological individual, and for every diagnostic criteria that can be applied for its identification. Now the question is: is it true that, like in the case of de Queiroz's resolution of the species problems, all criteria of individuality share something in common that has remained unnoticed due to the pervasiveness of disagreement over different diagnostic criteria? And, if so, how is this relevant for the interactor/replicator framework?

The answer to the first question is, in my view, affirmative. Indeed, all the different diagnostic criteria of biological individuality share a very important feature in common: biological individuals are the entities that participate *as a whole* in at least one relevant biological (sub)process of interest and, thus, the entities that need to be studied—manipulated, preserved, cured, compared, counted, etc.—if we want to gain knowledge about the biological (sub)processes in question. Thus, each criterion of biological individuality gives us a way of diagnosing when two entities become one, in the sense of realizing

a biological (sub)process together. In some cases, some ‘conglomerates’ participate as a unique object in more than just one single (sub)process, and thus we could arguably say that they are a ‘paradigmatic individual’, to use Godfrey-Smith’s terminology. However, it is not less true that any of the criteria point to interesting (sub)processes that deserve empirical treatment, and thus I suggest that it is reasonable to accept that a conglomerate should be considered a biological individual if it participates in at least one of these processes (**Figure 22**). Notice, however, that this does not mean that because it participates in at least one of these processes, and thus it would count as a biological individual according to my definition, then it needs to be studied as a single individual with reference to all the other biological (sub)processes. That would be simply mistaken and might obfuscate understanding, at least as much as not considering the conglomerate a biological individual if it participates in at least one process would obfuscate understanding. I will elaborate on this more in **chapter IV**.

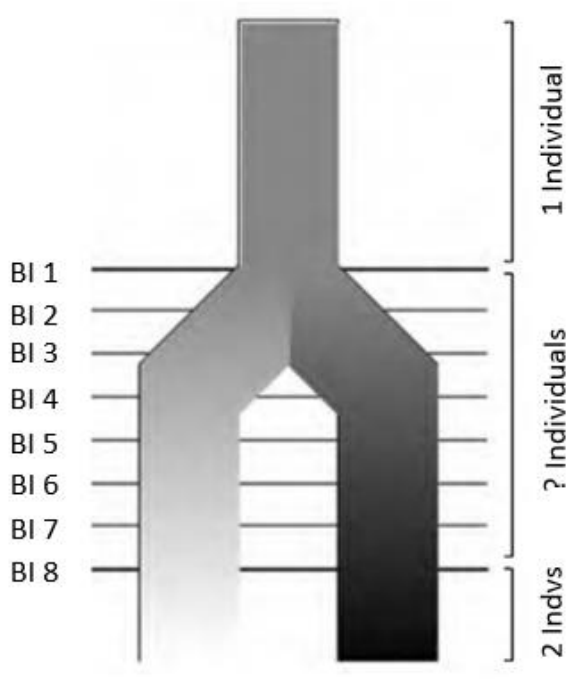


Figure 22. The *processes* of biological individuality. In contrast with the case of species, in which the key question was deciding if there are two rather than one, for biological individuality the opposite happens: the question is deciding when there is one rather than two. In the figure, B1-B8 refers to different candidate criteria. If two individuals meet the eight criteria, then we would rather say that there is only one, whereas if they only meet some, but not others, then the answer will depend on the criteria we adopt. (Adapted from de Queiroz 2005a: 2004, Fig. 3)

The answer I have given to the first question does not however reply to the second question. How is my definition of biological individuality relevant for the interactor/replicator framework of the units of selection? For one simple reason: Because the interactor/replicator framework strictly separates two properties—inheritance and interaction with the environment—whose separation is crucial to get a better understanding of the units of selection. My way of thinking biological individuality saves the interactor/replicator framework insofar as it also separates the two properties, as they are different biological (sub)processes that, thus, account for different dimensions of biological individuality. In my account, a conglomerate would count as a unit of selection provided that it can be shown to participate in at least one significant biological (sub)process of interaction with the environment (the role of the interactor). Once this is so, then, it becomes necessary to look for the set of factors that might be involved in the hereditary process, i.e. in the process that has made the two entities join and participate together in this biological (sub)process. These factors might well be replicators, but they do not need to be, as I will argue in **chapter V**. In this sense, my account takes a lesson from the interactor/replicator framework without falling prey to its vices, namely: leaving too much unexplained.[101]

Notice, thus, that my requirement for a conglomerate to qualify as a unit of selection is very unrestrictive. As I will explain, inheritance needs to be found at some point, but it must always be widely conceived. Why do I want to have such an unrestrictive view of natural selection? Because otherwise we run the risk of leaving too much unexplained. Natural selection is normally used as the *explanans* of many evolutionary innovations. It makes sense to limit its explanatory scope so that it does not become a trivial, quasi-tautological *explanans*. This is precisely the spirit of the works of Gould, Vrba, Brandon, Sober, Lloyd, Okasha, and others. However, I suspect that the limitation must always be dependent upon: 1) the *explanandum* we aim to account for; 2) the

[101] I suspect that my account of when to argue that the entitites at a level are a unit of selection roughly corresponds to what Jantzen (2017), has mathematically argued. He argues that concentrating on whether the objects at one level have certain *properties* is question begging, and the debate should concentrate on whether the objects have certain *dynamics*, which he elaborates mathematically. I think his way of making the point is a mathematical way of demonstrating what I have argued only verbally.

existence of alternative *explanans* for the same *explanandum*; and 3) how much of the *explanandum* can be assumed to be random (i.e. due to non-causal factors). It seems to me that a key element of my account is that it allows developing alternative explanations for the same phenomenon and which would probably lead, for instance, to the discovery of new channels of inheritance. Of course, it might happen that for some of the conglomerates that are considered biological individuals there is not a common selective explanation that accounts for why they bear the traits they bear. But in assuming that such explanation could in principle exist, we would have gained a deeper understanding of the phenomenon. I have made this point in most of my published—and my yet unpublished—papers, which I have mentioned in the **Introduction**, so I won't repeat the same points here. In any case, an exploration of the alternative I suggest here will be developed in **chapter V** for the specific case of holobionts and HCE.

3.3.1.1. Comparison of my 'resolution' of the problem of biological individuality with Ellen Clarke's solution

One might now object that my 'resolution' of the PBI is indeed no different from what Ellen Clarke proposes in her (2013). There, Clarke argues that there are many different ways of being a biological individual (being a biological individual is multiply realizable), but despite this, biological individuality can be defined in terms of two necessary and sufficient functional conditions: A biological individual is an object that possesses *simultaneously* at least one mechanism that limits its capacity to experience selection from below (policing mechanism) *and* at least one mechanism that increases its capacity to be selected vis-à-vis other objects of the same type (demarcation mechanism) (adapted from Clarke 2013: 427). As both policing and demarcation mechanisms are multiply realizable, and will very likely be realized differently in different objects, then Clarke believes that her definition is sufficiently general to define biological individuals. And, in addition, she believes that the existence of these mechanisms provides a good way to spell out more clearly Lewontin's criteria for being a unit of selection (**section 3.1**).

While I believe that Clarke's definition is on the right track, I also believe it is too restrictive. First, what would happen if an object—call it "*object**"—just exhibits one of the two mechanisms that she believes are necessary and sufficient for being a biological individual? That the *object** simply would not qualify as one. But, again, that option seems to leave too much unexplained. Let us assume that we have a real-world example of an *object** which is formed by two more basic primitive entities linked to each other by a policing mechanism (say, one entity coerces the other if it does not cooperate). Now, let us assume for the sake of the argument that the *object** happens to evolve a demarcation mechanism. What would be a good explanatory account of why the *object** evolved the demarcation mechanism? Probably that appeals to natural selection *for* the *object**. But, for Clarke, that would not be possible, because by definition *object** is not a unit of selection. Then the explanation would be that there was selection for the entities that compose *object**, and thus sorting at the *object** level *caused** the appearance of the demarcation mechanism. Why *caused**, and not simply "caused"? Because, remember, sorting at one level is just a mere consequence of causation at a different level (the lower level), so sorting at one level is, *by definition*, never causal (**chapter II, section 5**). But this is paradoxical, because a demarcation mechanism is a group level adaptation, or it is nothing. That is, it is an adaptation that appears *at the level of the object**, and thus, necessarily, there must have been selection for *object**, according to what Sober and Wilson (2011) call Williams' principle (from Williams 1966). Therefore, Clarke's definition, by considering that the two criteria are both necessary and jointly sufficient leaves too much biological complexity unexplained.

As a follow up from this criticism of Clarke's view, it becomes clear that, while I share her view about the multiple realizability of biological individuals, I cannot share her views about the definitional criteria to consider that an object is a biological individual. My definition is, thus, less stringent than hers. And, as I will develop extensively in the next two chapters, the criteria we should accept as individuating criteria do not even need to be considered either policing or demarcation mechanisms in any sense.

3.4. The diachronic view of natural selection: Multilevel selection theory

Until now I have presented two approaches to the topic of the units of selection that have been characterized as *synchronic* approaches: they are oriented to determining the characteristics that a biological object must satisfy in order to evolve in response to natural selection *now* (Okasha 2001, 2006). However, this type of approaches fails to explain how the objects that are units of selection *now* came into existence *in the first place*. To overcome this difficulty, a *diachronic* approach, i.e. one that explains how these objects came into existence, is needed. The task of developing a diachronic approach, while important for many decades, acquired a special sense of urgency in the mid 90s, connected to the research on the major transitions in individuality (Maynard-Smith and Szathmáry 1995). That was simply because the study of the major transitions in individuality made clear that:

> '[T]he levels-of-selection question is not simply about identifying the hierarchical levels(s) at which selection *now* acts, (…) but about identifying the mechanisms which led the various hierarchical levels to evolve in the first place.' (Okasha 2006: 16)

Why? Because as I explained before, the question of how new objects with the properties of a unit of selection evolved requires knowing how natural selection made possible that the mechanisms that confer these very properties to these objects evolved in the first place. Or, in other words, a selectionist story of how the mechanisms which guarantee heritable variance in fitness (to use Lewontin's formulation) in certain biological objects needs to be told. To accomplish this task, neither Lewontin's hierarchy of objects, nor Hull/Dawkins' hierarchy of interactors and replicators can play any significant role, because it is their very existence that needs to be explained in the first place. It is in this context that MLS theory becomes especially relevant (Damuth and Heisler 1988; Okasha 2001, 2006; Kramer & Meunier 2016). The framework of MLS is thought to account for the evolution of particles that are nested within collectives, as well as for the evolution of collectives, by appealing to: (1) the evolutionary (Lewontinian) properties of the particles within the collective, and (2) the evolutionary (Lewontinian) properties of the collectives. MLS offers thus

a diachronic perspective to the study of the units of selection which encompasses two different types of scenarios: MLS1, in which the particles are the focal units whose level of reproduction is calculated, and the collective is taken to be the unit composed of the entities that engage in fitness-affecting interactions with each other; MLS2, in which the collectives are the focal units whose number of offspring is tracked (**Figure 23**; Sober & Wilson 1998; Okasha 2006; Clarke 2014). I will now outline more extensively the essence of the MLS approach to natural selection.

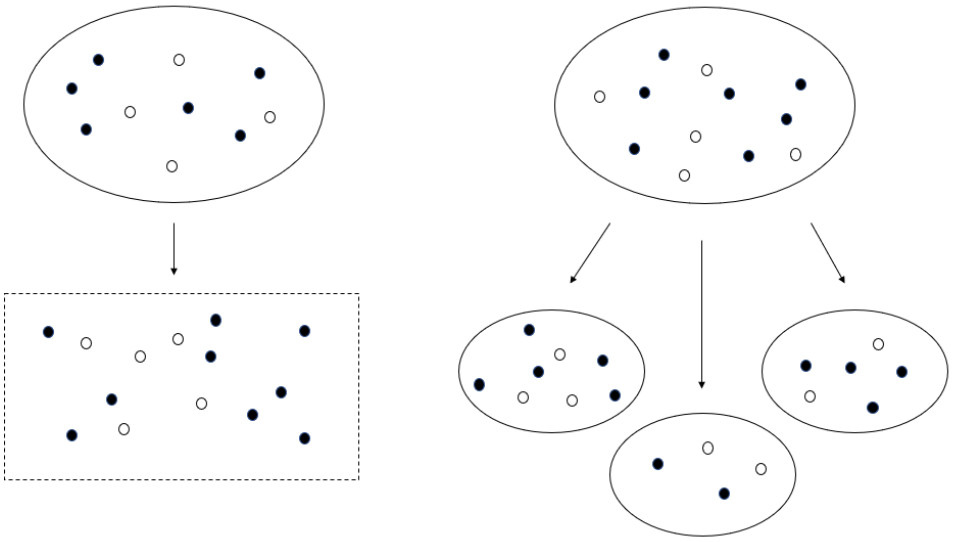


Figure 23. Difference between MLS1 (left hand) and MLS2 (right hand). The ellipses represent the collectives; the coloured and non-coloured circles represent particles with different traits; and the dashed rectangle represents the global population of particles (which is not itself a collective) produced by a collective. In MLS1, what matters is how many particles the collective releases to the global population of particles, and thus collective parent-offspring relations are not tracked. In MLS2 on the contrary, what is measured is the number of offspring collectives that a single collective produces. (From Suárez, under review, Fig. 1).

Let me start explaining MLS from the beginning.[102] The first question is: when is a MLS framework preferred to a single-level model of evolution? When there is a suspicion that a set of particles engages in a fitness-affecting relationship with

[102] My presentation of MLS will be *qualitative*, in the sense that I will not make use of the mathematimatical aparatus derived from the Price approach to MLS, since I think it is neither needed to understand the essence of MLS, nor required to understand the application of MLS to HCE. For a brilliant exposition of MLS that combines quantitative and qualitative analysis, see Okasha (2006: 40-75).

each other that, somehow, modifies its evolutionary success. In other words, when the evolution of the collection of particles is different from what their evolution would be like if they did not engage in fitness-affecting interactions with each other, and, *importantly*, this outcome is a consequence of their engagement in fitness-affecting interactions with each other—remember that natural selection is supposed to provide a causal story of why certain objects bear the traits they bear, so it is essential that the 'distorted' outcome occurs *as a consequence* of the interactions. Therefore, MLS will be applicable to every set of particles that engages in fitness-affecting interactions with each other that alter their expected evolution. And, additionally, it will be required in these cases precisely because the evolution of the particles will be conditional upon the collectives they form.[103]

Secondly, what is the difference between MLS and one-level scenarios? In contrast with single-level scenarios, MLS requires considering the evolution of each particle in the context of the whole collection of particles, i.e. in the context of the global population; and, additionally, MLS requires considering the evolution of each collective in relation to other collectives in the population. To do so, a MLS framework requires two types of counting: first, how much do particles reproduce *within* the collective (that is, how selection acts *within* collectives); second, how much collectives reproduce and form new collectives (that is, how selection acts *between* collectives). Therefore, in any MLS framework, two different types of scenarios must be taken into account: one where the particles are the focal unit (MLS1), and one where the collectives are the focal unit (MLS2).

Thirdly, what has to be looked for in each MLS scenario, that is, in each case of MLS1 and MLS2? Basically, the extent to which the particles satisfy Lewontin's conditions for evolution by natural selection. That is to say, to which extent the objects manifest heritable variation that is causally connected to their fitness, so that the objects that possess a concrete form of variation are fitter

[103] Technically speaking, MLS theory would be perfectly applicable to one level scenarios, providing one considers that the level of reproduction—and thus, of possible cross-level influence—of the objects below and above the focus level is zero.

than the objects that possess a different type of variation, and this variation is heritable (i.e. there is parent-offspring regression). The challenge, in a MLS framework, is how to conceive these properties at two different levels or, in other words, to establish which is the object that must manifest these properties. Are they the particles, each of them individually? Or is it rather the collective? Or are they both? In the MLS framework, a strict separation between particle-level selection and collective-level selection is required, and that is precisely the difference between the two types of MLS: MLS1 (particle-level selection, selection *within* collectives), and MLS2 (collective-level selection, selection *between* collectives). Importantly, in MLS, the collective is a unit of selection in both scenarios, as I will explain in detail.

Let me start with MLS1. In MLS1, the entities whose heritable variation in fitness is measured are the particles *within* the collective. In this sense, MLS1 particle-fitness will simply mean that number of offspring particles that a given particle produces, and MLS1 collective-fitness will be measured in virtue of how many particles each collective produces and releases to the environment so that they can form new collectives in the next generation.[104] Let me describe this process in detail.

- MLS1 particle-level fitness. In MLS1, the collective provides *exclusively* the population structure where the particles evolve. The key point that MLS1 emphasizes is that the particles bear the traits they bear, and they evolve how they do, because they engage in significant fitness-affecting interactions within a collective, and part of their evolution will depend on the response of the other particles within that very collective. That is to say, the particles within the collective are mutually responding to each other (competing, cooperating, etc.) in a way that affects their evolution, i.e. that affects how their trait distribution will change over time.

[104] By ‘fitness’ I will always mean ‘reproductive output’, i.e. entity whose offspring is ‘counted’. Nonetheless, there might be other ways of counting the fitness of entities, as some people have emphasized (e.g. Bouchard 2008, 2011; Triviño and Nuño de la Rosa 2016). MLS would be, in principle, also applicable in these cases, although I will not discuss here how this could be done.

- MLS1 collective-level fitness. Furthermore, in a MLS1 scenario, collectives are expected to expel part of their particles to the environment, and these particles are expected to mix up and make new collectives. In the 'mixing up' stage, the competition of particles with each other to form new collectives is essential to understand particle evolution. Why? Because those collectives that survive longer will have more chances of releasing their particle types to the environment, and thus the probability that new collectives that are composed of the same particle types as the particles types that are more abundant in them increases. In other words, collective selection in MLS1 means that some collectives survive longer than others in virtue of the particle types they are composed of and, as a consequence, their longer survival causes an increase in 'the parent-offspring resemblance *in the global population of particles*' (heritability condition) (Okasha 2006: 72, emphasis added).

There is one lesson that can be extracted from this analysis, and that results from the combination of two theses: First, in MLS1, the entities whose reproduction rate is measured are the particles, not the collectives; second, in MLS1, the fitness of the collectives is measured by measuring how they contribute to the global population of particles, and thus how they will affect the population of particles in subsequent generations, and not by how many offspring collectives they leave.[105] Therefore, in MLS1, the collective, despite not being counted as the 'reproducing' unit, is as much a unit of selection as the particles are, simply because the longer the survival of the collective, the more the collective will affect the global population of particles in subsequent generations.

[105] Notice that the decision of not counting the collectives as the entities that reproduce, i.e. that form parent-offspring lineages, is rather ideosyncratic. It is mainly conditioned by the fact that, in some MLS scenarios, we simply do not know how to recognise parent-offspring lineages. Okasha (2006: 51-52) uses the slime mould as an example of this problem, and argues that in some cases it might be better to apply MLS1 simply for the epistemological difficulty of knowing whom is parent of whom. But the reader must keep in mind that this is just based on our lack of knowledge of how to trace lineages for some 'weird' biological objects, but it is not a conceptual necessity. This point will gain its significance in **chapter V**.

Let us go now to MLS2. In contrast with MLS1, the focal unit whose reproduction rate is measured in MLS2 is the collective, and thus what has to be analysed is the competition *between* collectives in forming parent-offspring lineages. In other words, the type of parent-offspring regression that matters in this case is 'collective parent / collective offspring' regression. What needs to be studied in MLS2, thus, is how much a collective offspring resembles its collective parent. Again, I will now describe this process in detail.

- MLS2 collective-level fitness. In MLS2, the objects whose fitness is measured, that is, the objects for which parent-offspring regression matters are the collectives. Thus, to apply MLS2 to a collective it is necessary to find a way to discover parent-offspring lineages, and to measure parent-offspring regression or resemblance with respect to traits under investigation. In other words, it is necessary that the collectives satisfy the Lewontin's conditions at least to a certain extent.

- MLS2 particle-level fitness. Measuring particle-level fitness in MLS2 is a complex issue, as the translation between levels is not as immediate as in the case of MLS1, where MLS1 collective-level fitness was interpreted simply as the average of MLS1 particle-level fitness. That is because in MLS2 particle characters'—that is, particle phenotypic traits—will also be selected within each collective, and their response to selection will depend on the magnitude of parent-offspring resemblance of each particle for the trait under study. This could eventually, *although not necessarily*, have effects on the collective character, for example by biasing collective-level reproduction. However, and this is important, this does not need to be taken into account in most cases: in most occasions, MLS2 will lead to two independent evolutionary responses, which will be measured in different units (how many offspring collectives vs. how many offspring particles), and that will, at most, bear a contingent relation to each other.

Now I have described the difference between MLS1 and MLS2, let me summarize the key points of MLS: First, both in MLS1 and MLS2, the collective

is a unit of selection; what changes in these two scenarios is the object whose parent-offspring relations are tracked, but this does not rule out the fact that the collective is a unit of selection in *both* scenarios. Second, that reproduction (in the conventional sense of parent-offspring regression) at the focal level is not necessary for arguing that a collective is a unit of selection; it is sufficient that the existence of a collective of entities engaged in fitness-affecting interactions biases the evolution of the particles, so that changes that were not expected to occur had the collective not existed are indeed produced—and, conventionally, these changes are called group-level adaptations, as they are selected *because* there are groups. Third, that a transition in evolutionary individuality is precisely the process by which there is a shift from the MLS1 scenario to the MLS2 scenario; for a transition in individuality to happen, the collective needs to evolve biological mechanisms to guarantee that: (1) there are parent-offspring lineages; (2) there is parent-offspring resemblance at least to certain extent.

4. The holobiont through the lens of multilevel selection theory

Now I have examined the debate about the biological individuality and the problem of the units of selection, it is time to connect the two conceptions of the units of selection (synchronic and diachronic) to the issues raised by HCE. I will start this section with a quote from an old paper of mine, which reflects and, to my knowledge, makes explicit for the first time, the main point about holobiont evolution that I will advance here, and develop in detail in **chapter V**:[106]

> '[E]very time that there is a process of symbiogenetic fusion (…), natural selection will act *simultaneously* in at least two levels. On the one hand, inside the new organism, which I have called "symbiome" [holobiont], where some symbionts and some hosts will be naturally selected in virtue

[106] The paper I am referring to is J. S. Díaz (2015): 'El mecanismo evolutivo de Margulis y los niveles de selección'. *Contrastes* XX (1): 7-24. The paper, which summarizes the main point I developed in my undergraduate dissertation, focuses exclusively on the *most provocative* work by Lynn Margulis (see Suárez 2018a, and this thesis, **chapter I**), and not on the hologenome tradition that starts after the original publications of Zilber-Rosenberg and Rosenberg. The conclusion, as well as the argumentative strategy, is nevertheless also valid for HCE. Notice, however, that the terminology I use is slightly different from the terminology used by HCE advocates, e.g. I use 'symbiome' instead of 'holobiont'; or I use the *ad hoc* terms 'intraorganismic' and 'supraorganismic'.

> of their respective traits and whether they are (or not) beneficial in their new environment. On the other hand, in the relationship which is established between the symbiome [holobiont] and its environment.
>
> Making the hypothesis more precise, I will coin two terms to refer to these two levels of selection. First, the *intraorganismic* level of selection, which refers to what occurs *within* the symbiome [holobiont]. In this level of selection, the symbiome is the environment where selective processes occur, and both the symbionts and the host will be the units of selection. (…) [In line with this], we can assume that natural selection will favour the perpetuation of those that will increase their fitness within the symbiome [holobiont] […]
>
> Second, the *supraorganism* level (…). This level refers to what happens *outside* the symbiome [holobiont], that is, in the relation between it and its proximate environment. In this environment, there will be prey, predators, climate conditions, etc. that will foster the survival of some supraorganisms and the death of others. In this level, (…) the symbiome [holobiont] is the unit of selection (…). [and thus] natural selection will foster some fitness advantages for the symbiome [holobiont], which are determined by its interaction with the proximate environment and will occasionally go against the intraorganismic "advantages."' (Díaz 2015: 15-16)

That paragraph, even if focused mainly on Margulis' ideas about the role of symbiosis and symbiogenesis in evolution, encapsulates what I take to be the main lesson of HCE. Or, in other words, it expresses faithfully the conception of natural selection that underlies HCE, namely: a diachronic, multi-level perspective. Notice that the point in the paragraph is not to say that holobionts—or the 'symbiome', as the expression I used in that paper, following Jan Sapp (2003, 2004)—are units of selection *qua* interactors or *qua* replicators. The holobiont is in many cases neither cohesive enough, nor has the hologenome the high degree of copy-fidelity that is demanded for replicators.[107] But this is irrelevant. The lesson to take from HCE is that natural

[107] Of course, in some species it might well happen that this is not true, and that it is possible to argue that the hologenome is a replicator with the same degree of fidelity transmission as genomes. But notice that this does not need to be the norm.

selection must be thought of as MLS—i.e. it must be understood diachronically—at least if one aims to make sense of certain evolutionary innovations that multicellular organisms possess. And also, and importantly, the paragraph expresses why the diachronic perspective does not neglect the possibility that selection processes also occur *within* the holobiont, and not only *between* holobionts[108]: From a diachronic perspective, the force of selection acts in both units, 'forcing us to accept a *plurality* of simultaneous levels [of selection]' (Díaz 2015: 17).[109]

Why is this so relevant and so novel? The diachronic approach to the units of selection is common, and has been applied before, as I showed in **section 3.4**, so there seems to be no particular 'novelty' in HCE: it just demands that we apply an approach that is available and has been applied before. This is correct, but only to certain extent: while it is true that MLS is not a 'weird' approach to the units of selection, the approach is normally applied to cases where a transition of biological individuality *has already occurred*, not to cases where it is arguably *occurring now*. This makes sense, as the result of a transition in individuality is hard to predict: will the entities overcome the difficulties of transitioning, or will they rather go back to the original state, the one previous to their 'coming together'?[110] The point of HCE, however, is that that worry, while legitimate, should not preclude us from studying holobionts as units of selection from a MLS perspective, on the risk that ignoring that possibility could mask many cases of real selection. In this sense, HCE is novel, as it encourages us to study transitions that, presumably, are still happening (see **chapter IV** for an ontological justification of the application of MLS to holobionts).

I will not review here the implications of applying a diachronic view to the study of holobiont evolution. It is enough to settle the problem, and frame it

[108] What in the quoted paragraph I call '*outside* the holobiont'.
[109] The necessity of the interaction of the two levels, as well as the necessity of recognising that the outcome of selection would result from their interaction, was presented in a figure that supported the explanation (Díaz 2015: 16-17, Fig.1).
[110] I use 'coming together' because I am assuming the transition has not been completed, and thus I cannot say that the entities go to the stage 'before the transition has occurred', even when that would be the most logical sentence.

adequately, since my model will be presented and defended in detail in **chapter V**. But the consequences of the presentation of HCE in terms of MLS I have just sketched here should at least be anticipated, even if still not argued for. The consequences are threefold: first, if MLS is the right way of thinking about HCE, i.e. about the role of the holobiont as a unit of selection, then the arguments against HCE that claim that holobionts are not interactors and hologenomes are not replicators are automatically invalid—as are invalid, of course, the arguments in support of HCE which are framed under the interactor/replicator framework. Second, that there is no possibility of saying that HCE would have negative epistemological implications, like ignoring the possibility of selection *within* the holobiont. So expressed, the criticism is clearly misguided: MLS theory precisely emphasizes the existence of selection *within* the holobiont. Third, that there is an urgent need to think about holobiont evolution from a MLS, and this perspective must overcome the species-specific understanding of the holobiont. Remember that I argued that a MLS could be applied to any level provided that at least one biologically significant process occurs at that level (**section 3.4**). However, I have not presented the way in which the entities at the level should be individualized. Saying how this is so in the case of holobionts is not a simple task to accomplish, and that is why it deserves independent treatment.

5. Brief summary of chapter III: The problem of biological individuality meets the holobiont

In this chapter, I have introduced and discussed extensively the PBI, with special emphasis on the problem of the units of selection, as I take it as the main problem that is at stake in the discussions about the HCE. The key points that I have presented in the chapter can be summarized as follows:

1. There is a genuine problem of biological individuality, which derives from the facts that: (a) biological objects are individuated in virtue of the biological (sub)processes they participate in, and (b) the objects individuated by these (sub)processes are not necessarily co-extensional.

2. One criterion of biological individuality is considering whether the object studied is a unit of selection.
3. There are two main approaches to the topic of the units of selection: (a) a synchronic approach (which objects participate in the process of selection *now*); (b) a diachronic approach (hot the objects that participated in the process of selection have *evolved*).
4. The interactor/replicator framework is a way of approaching the question of the units of selection synchronically.
5. The interactor/replicator framework has many problems, especially for the commitments it makes about the existence of two highly evolved entities.
6. Multilevel selection theory is the conventional way of presenting the diachronic approach.
7. Holobiont thinking, in its HCE version, entails that some types of biological entities must be studied through the lens of multilevel selection theory *now*, and thus multilevel selection does not need to be applied exclusively to how some biological objects *evolved* in the past.

Chapter IV

‘Starting afresh at each level. Towards a hierarchical-non-nested view of the biological world’

This chapter derives a consequence of the processualist perspective of biological individuality that I introduced in chapter III, to provide a reply to the arguments by Chiu and Eberl, and by Bourrat and Griffiths (chapter II). Drawing upon the case of gut development via *Bacteroides fragilis* in mice, and the study of the life cycle of *B. fragilis*, I make three points. First, I argue that the application of process-based individuation criteria to symbiotic assemblages reveals that the dependency relations among the entities that constitute the collective are non-reciprocal. Whereas the host in the collective depends on the symbionts for realizing the developmental, immunological, and physiological processes that characterize it, the dependency relation is not necessarily mutual and, thus, while the holobiont appears to be an individual from the host-perspective, it turns out not to be one from the symbiont-perspective. Second, I take issue with Chiu and Eberl, and Bourrat and Griffiths. They assume that biological individuality requires the existence of a co-dependency between the host and its symbionts, and thus they reject the notion that holobionts are biological individuals. I argue that their position derives from view of biological individuality, which is grounded on the assumption that the biological world is hierchical *and* nested, which leads to an incorrect account of the ‘boundaries’ of the holobiont. I defend that, since the process-based criterion for individuating biological entities is not committed to a hierchical-nested view of the world, it can explain why holobionts are biological individuals and units of selection, as well as it can

explain away the concerns expressed by Chiu and Eberl, and Bourrat and Griffiths.[111]

> *"All our knowledge is linked and governed by metaphysics; metaphysics is the network that makes cohesive every piece of matter we interact with. But this network, and its nodes, is embedded within our ordinary consciousness under different states of affairs (…). The universal edges of the network were neither made explicit, nor became elements of our reflection."*
>
> G.W.F. Hegel (1807) *Phänomenologie des Geistes*, my translation

1. Introduction

In the previous chapter, I argued that there are many different and non-co-extensional ways of individuating biological entities, a phenomenon that I referred to as the PBI. There, I also proposed a processual and pluralistic picture for conceiving individuality, according to which a collection of entities can be considered a biological individual if the collective participates in a significant biological (sub)process as a whole (metabolic, developmental, immunological, etc.) (**chapter III, section 3.3.1**). In this chapter, I want to use the pluralistic and process-based approach presented there to question the standardly assumed hierarchical *and* nested picture of the biological world. In my view, the problem that HCE entails for the debate about biological individuality and the units of selection is, precisely, that it questions this

[111] This chapter is almost completely based on my work with Adrian Stencel: Javier Suárez and Adrian Stencel (under review): 'A part-dependent account of holobiont individuality'. There is, though, a substantial difference, for, as Adrian does not share my intuitions about HCE, in our draft we do not commit ourselves to any of the evolutionary implications that I draw here, and that I will develop in detail in the next chapter.

hierarchical-nested view in many ways.[112] According to the latter, the biological world would be like a set of Matryoshka (Russian) dolls, such that each new level would contain absolutely every entity that belongs to the level that is immediately below, as well as the hereditary basis of the entities in lower levels. In this sense, a superorganism would be a collection of organisms, which would be a collection of cells, which would be a collection of genes, etc. Applied to the holobiont, this view entails that the holobiont is a collection of genomes, or organisms that are member of different taxa. This type of structure of the world is also very popular among those who have worked on the topic of the units of selection (Lloyd 1988; Gould 2002). However, as Okasha (2006, 2011)[113] has noticed, precisely in the context of discussing the units of selection, this seems to be a non-argued *a priorism*, and biological processes are perfectly compatible with a hierarchical-non-nested conception (**Figure 24**). To quote:

> '[O]n one standard conception of multi-level selection theory, higher levels of selection arise via fitness-affecting interactions between lower-level units. Though models of this process *often assume* a nested hierarchy [hierarchical-nested view, in my terminology], the underlying causal mechanism *does not require* nesting; it could work equally well with overlapping groups of lower-level units.' (Okasha 2006: 44, emphasis added)

[112] And, importantly, defenders of the HCE have not noticed this consequence sufficiently well.

[113] Notice that in his (2011), Okasha only discusses the non-nested nature of phylogenetic hierarchies as elaborated by cladists, but not of the hierarchies generated in the debate on the evolutionary transitions in individuality, which according to him *needs to be* (*is*) strictly nested. Concerning the later, Okasha only argued that the way of *ranking* the individuals across levels should be free, and thus we should eliminate disputes about whether the new level is a group, a superorganism, etc. In this chapter, I will only argue for the possibility of classifying physiological and developmental individuals hierarchically but in a non-nested way, and in the next chapter I will apply the same logic to the MLS scenario, thus extending Okasha's framework to a domain where he has not extended his ideas yet. Nonetheless, as Lisa Lloyd has told me (personal communication), applying these ideas to MLS is not equivalent to apply them them to the context of the evolutionary transitions in individuality, as MLS is applicable to more scenarios than the scenario of the evolutionary transitions in individuality, as this was conceived by Maynard-Smith and Szathmáry. For similar hypothesis about the non-nested nature of biological hierarchies see e.g. McShea and Simpson (2011).

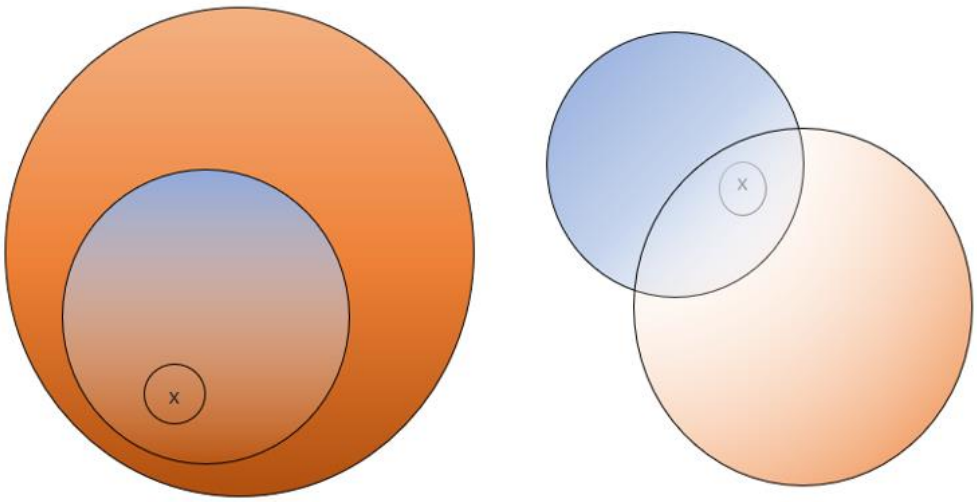


Figure 24. Comparison of a nested (left) and a non-nested (right) hierarchy. In a nested hierarchy, if *x* belongs to two higher level units (blue circle, orange circle), then one of the levels must necessarily contain mereologically all the elements that belong to the other level. In a non-nested hierarchy, on the contrary, it is possible that *x* belongs to two levels, one higher than the other (in this case, represented as 'wider' circles), without the necessity that the lower level is completely included in the higher level. As the figure also shows, in a nested hierarchy, levels do not overlap, but are subsets of each other. On the contrary, in a non-nested hierarchy levels can intersect, and thus *only parts of the lower level*, rather than the whole lower level, are a subset of the upper level.

In this chapter, I will argue that the precessual view of biological individuality that I presented before contradicts the hierarchical-nested view of the biological world in a way that is perfectly consistent with Okasha's (2006) intuition. The account is based on a contrastive case study that suggests that the holobiont is a biological individual from the perspective of the host, whereas it is not from the perspective of symbionts that compose the host's microbiome (as these are individuated according to16S rRNA analysis).[114] In other words, the type of biological (sub)processes that the host participates in requires the inclusion of the microbiome as an essential component (that is, as a part of the process, and thus as part of the individual that the process individuates), whereas the same is not usually the case if one is studying the biological (sub)processes that each of the symbiotic species that compose the host's microbiome engages

[114] Elsewhere I have speculated that 16S rRNA bacterial species classification might not be very appropriate for the type of classifications that are needed in symbiosis research (Suárez 2016). This section is thus, in part, a further development of that idea.

in. This analysis provides a particular view of the way in which the microbiome must be individuated. In my account, the microbiome must not be individuated according to the taxa that compose it (hologenome = collection of genomes), but rather functionally, according to the role that these taxa play, and the traits that allow the taxa to play these roles (hologenome = host + traits encoded by the microbiome). Building on that claim I argue that evaluations of the individuality must always start *afresh* at each level. Each level must be considered a genuine biological level independently of some possible problems that could theoretically arise 'from below' or 'from above', and thus ascriptions of individuality might generate a cross-classification of the biological hierarchy (Karaça 2019, for a similar argument).[115] I will argue that holobionts are a canonical example of this type of cross-classification, as there are important differences between the kind of biological (sub)processes that a host engages in, and the kind of biological (sub)processes that a bacterial taxa in the host microbiome engage in. Applied to evolution, my account entails that hologenome selection results from the *coevolution* of the host genome and the traits that compose the host microbiome, rather than from the *cospeciation* of the host with the bacterial taxa that compose its microbiome (**chapter V**).[116]

Simulteneously, this chapter constitutes a critical response to the arguments presented by Chiu and Eberl according to which holobionts are very 'blurry' objects to be considered biological individuals (**chapter II, section 2).** I argue that my way of accounting for biological individuality overcomes their problem. On the one hand, it explains why holobionts are perceived as 'blurry' objects by some authors (basically, because they assume a hierarchical-nested view of the biological world that leads them to hold a mistaken interpreted of HCE). On the other hand, it explains why the blurry nature of holobionts is not

[115] In line with what I argued in **chapter III, section 3.1**, the key ideas are that first, the processes that cause the appearance of the phenotypic properties must be studied apart from the hereditary basis of these properties, and, second, that the study of the processes that cause the appearance of the phenotypic properties at different levels generate a cross-classification of the individuals in the biological world, even when some of these levels would theoretically (in a Matryoshka-looking world) incorporate the individuals from below.

[116] I am making a distinction here between the notion of 'cospeciation', which means 'species-species coevolution', and the notion 'coevolution', which does not necessarily mean 'species-species coevolution', but it could also mean 'trait-trait coevolution'. In my view of the holobiont, the holobiont is a unit of selection because it creates the conditions that make the existence of 'host genome-microbe trait' coevolution biologically possible, rather than 'host species-microbe species coevolution'.

problematic to assert their individuality (for the existence of biological (sub)processes that generate a cross-classification of the biological world necessarily gives rise to a 'blurry' perception of some of these individuals). At the same time, the chapter constitutes a partial response to Bourrat and Griffiths' criticism to the 'part-of-the-system' arguments (**chapter II, section 3**).

The chapter will be driven by assuming that the biological (sub)processes of interest, and thus the ones that will determine the 'boundaries' of the biological individual, are development and physiology (**chapter III, section 2**).[117] Bearing this in mind, I will proceed as follows. First, I will present some research on the development and physiological behaviour of the gut in mice that, I argue, shows the necessity of taking the holobiont seriously from a developmental and a physiological perspective (the host / macrobe perspective). Building on that, I explore the evolutionary implications of that observation. Second, I present some recent evidence about the mode of life of *Bacteroides fragilis*, which is the bacterium that happens to be essential for mice's gut development and physiology (the symbiont / microbe perspective), to suggest that the taxa is physiologically and evolutionarily independent from the holobiont.[118] Drawing on these two cases, I argue that the commitment to a hierarchical-nested view of the biological world is invalid, as the latter would entail a wrong interpretation of the boundaries of the biological individual (in this case, of the holobiont). Finally, I show how this overcomes the problem of the 'blurry' nature of the holobiont.[119]

---

[117] Keep in mind that, in my account, this means that these processes will determine the phenotypic properties of interest and thus, indirectly, where the hereditary basis of these processes must be investigated.

[118] I will use 'macrobe perspective' as synonymous to 'host perspective', and 'microbe perspective' as synonymous to 'symbiont perspective', because in the case of the holobiont, the host is conventionally a macrobe, and the symbionts are microbes. Nonetheless, I am conscious this is not necessary in other cases of host / symbiont relationships. But as I will make my case for holobionts, I will take them as synonymous.

[119] A cautionary note for the reader. This chapter is about *metaphysics*. It is about how the biological world would look like if our current empirical evidence about how some biological (sub)processes occur was taken seriously. I will make a case about how 'hierarchy' and 'individuality', conceived as abstract metaphysical concepts, must be understood. In this sense, the biology, while important, is partially irrelevant, because the claims aim to be much broader.

2. The host perspective – when the holobiont is the biological individual

To prove the case that holobionts are biological individuals from the perspective of the host,[120] I will study the influence of the gut microbiome in the development and physiological maintenance of mice. Concretely, I analyse the role that the microbiome plays in the normal maturation and maintenance of the gastrointestinal tract, as this has been widely documented in the scientific literature. I make two points. First, that from the point of view of developmental and physiological processes, the microbiome is an essential component of mice, and thus it is part of the biological individual (i.e. mice are holobionts). Second, that therefore mouse evolution will strongly depend on the composition and on the evolution of its microbiome, to the extent that the 'evolving object' is the holobiont.

2.1. The case of gut development and gut physiology in mice

In mice (as it is also the case in humans, and many other mammals, particularly ruminants), the gastrointestinal tract presents a very high density of bacterial species from different phyla, the species of the families Bacteroides and Firmicutes being the most predominant. One of the main characteristic of these microorganisms is that they are innocuous in that environment, which does not entail that they are completely innocuous for their host; on the contrary, most of them will become very virulent in the case of 'bacterial infiltration', that is, if they happen to cross the intestinal mucus layer and travel to any other tissue or organ. Interestingly, several studies that compared gut development in germ-free mice with gut development in conventionally raised mice have shown that the microbiome plays a fundamental role in two essential processes for the survival of mice: first, the maturation of the intestinal layer; second, its maintenance. Germ-free mice, in contrast with conventionally raised mice, show

[120] The word 'perspective' should not be understood in any anthropomorphic sense. By using 'perspective', I will only refer to the element of the group whose individuality is being studied. In this sense, 'host / macrobe perspective' means that I am studying the individuality of the host, whereas 'symbiont / microbe perspective' means the opposite. In this chapter, I concentrate exclusively on the study of these two perspectives because the problem is to discover whether the study of some biological (sub)processes from the perspective of any of them would justify the consideration of the holobiont as a biological individual, and whether the consideration of the holobiont as a biological individual is equally applicable to all the taxa that belong to the symbiotic assemblage.

a considerable reduction in the thickness of the mucus layer of their gut, which is combined with its denaturalization, that is, with an alteration in its normal physiological properties. It has been hypothesized, and some current empirical evidence supports this hypothesis, that these two alterations result from the essential role that gut microbiome plays in inducing the differentiation of RORγt+NKp46+ natural killer (NK)-like cells. These are a very specific type of lymphocyte that is known to induce the construction of the gastro-intestinal tract (Zheng et al. 2008; Sanos et al. 2009; Jakobsson et al. 2015).

It is important to note here that these empirical findings support the hypothesis that the microorganisms that compose the gut microbiome are at least partially integrated with the mouse, a point I will develop extensively later. Why is that? Because it is not only that these microorganisms must simply be *present* during the process of maturation of the mucus layer. They are also required to interact with their host in sophisticated ways (by producing very concrete biomolecular components) to induce the differentiation of the lymphocytes that would later induce the construction of the gastro-intestinal tract. Furthermore, for gut maturation to occur normally, the interaction between the host and the microbiome is temporally restricted; that is to say, it must necessarily occur during the developmental stage, or otherwise the process is not possible, not even if the inducers are provided at some point later in time (Olszak et al. 2012). As in every other homeorhetic process, this observation suggests the importance not only of the existence of a process of interaction, but also that this interaction is mediated by the right kind of components, and that it occurs in a particular moment during development. The maturation of the mucus layer in the gut of mice requires thus, first, that the appropriate set of interactions between the mouse and its microbiome are triggered; and, second, that these interactions are accompanied by a series of consecutive co-responses which alter the nature of the mouse, as well as the nature of the microbiome, to make both objects ‘partners through development’ (Gilbert 2017). If these interactions fail to be triggered, maturation is not possible and, thus, mice are not expected to survive in their standard environmental conditions.

In addition to the role it plays in the maturation of the intestinal barrier, the gut microbiome in mice is also hypothesized to play an essential role in maintaining the intestinal homeostatic equilibrium. It does so by balancing the fragile equilibrium that exists between pro- and anti-inflammatory T cells (another type of lymphocyte), which are the cells that actively maintain the intestinal barrier. How does the microbiome contribute to the maintenance of this homeostasis? It has been experimentally proven that the level of iNKT pro-inflammatory cells is higher in gnotobiotic (= germ-free) mice than it is in conventionally raised mice. This disproportion has been shown to correlate with a higher incidence of colitis among gnotobiotic individuals (Smith and Garret 2011). Furthermore, it has also been experimentally shown that germ-free mice need to be colonized at birth, or their level of iNKT pro-inflammatory cells never reaches an equilibrium level, which suggests that the process of gut maturation is strongly linked to the processes that maintain the homeostasis of the gut. Otherwise, if mice are for instance colonized when they are adults, their homeostatic levels are never restored (Olszak et al. 2012). This evidence suggests that the microbiome is essential not only for gut development, but also for keeping its homeostasis; and, second, that the processes of gut development and gut homeostasis are intimately connected with each other in ways that still need to be investigated.

Drawing on the previous observations, recent research driven under very specific lab conditions has allowed the identification of the bacterial *taxa* that are responsible for the pro-inflammatory and anti-inflammatory cell responses that have been documented in mice. Round et al. (2011) have shown that *B. fragilis* is in control of some of these processes, since it has the capacity of activating some of the pathways—particularly, the toll-like receptor pathways—that suppress the inflammatory response. Concretely, *B. fragilis* suppresses the pro-inflammatory response of T-helper 17 (Th17) cells, while simultaneously inducing the differentiation of $CD4^+$ T cells into regulatory-T cells (Treg) (see also Mazmanian et al. 2008). These two processes—fundamental for gut colonization and, thus, for gut maturation—are mediated by polysaccharide A, which is one of the most common compounds in the bacterial wall of *B. fragilis* (Round and Mazmanian 2009).

This last point is especially remarkable to understand the role that the microbiome plays in the physiology of the host, and the level of interconnectedness that exists between the host and the microbiome. The microbiome will trigger the appropriate physiological responses in mice if its interaction occurs during a very precise moment of mouse development and if *the right type of elements composes it*. That is, if some, among the whole set of elements that compose the microbiome, have the capacity to induce the appropriate responses in the host, so that cell differentiation occurs, and the pro-inflammatory response of Th17 cells is supressed. The question now is to determine 'the right types of elements'.

These observations however shows the intrinsic connection that exists between the host and its microbiome, as well as why the microbiome cannot simply be conceived as a 'background' condition. Let me go a bit deeper into this observation. One standard argument against the consideration of the microbiome as a part of the biological (sub)process(es) that the host participates in is that the 'part-of-the-system' arguments do not work (**chapter II, section 3**). According to that criticism, the symbionts can be conceived as 'background conditions' that are required for the biological (sub)process to take place, in the same sense as some inorganic elements (oxygen, polysaccharides, or even the gravitational field, to use an example that Bourrat and Griffiths use) are sometimes conceived. That interpretation, while legitimate, is misguided, because the continuity that exists between the microbiome and the host's genome is larger than the continuity that exists between the microbiome and any of the elements that are conventionally considered 'background' conditions. I will explore this point in what follows.

2.2. The continuity between the host and the microbiome: the microbiome is more than a mere background condition

In what sense does the evidence about gut development and gut maintenance support the existence of continuity between the host and the microbiome, and how does this continuity support the individuality of the holobiont? First, I will

argue that the entities that compose the microbiome, in contrast with standard background conditions, play a role as *evolutionary agents*—in the sense in which genes are evolutionary agents, i.e. as entities that are subject of the evolutionary process—in the biological (sub)processes of development and physiology and, in this sense, it is possible to tell them apart from background conditions.[121] Second, I will connect this claim to the discussion about the units of selection by arguing that, insofar as the elements of the microbiome play a role as agents in these (sub)processes, they do therefore play a fundamental role in the evolution of the host.

2.2.1. The microbiome as an 'evolutionary agent'

The evidence in the previous section shows that, during the development of the gut in mice, the microbiome plays a fundamental role via the action of *B. fragilis*. Notice that the role played by *B. fragilis* in this process depends on how it interacts with some elements of the immunological system of the host and, more precisely, it is strongly connected to the existence of an appropriate peer-communication mediated through polysaccharide A (Round and Mazmanian 2009). Importantly, *B. fragilis* plays a similar role in this process to the role that a gene (or set of genes, if the process is controlled by more than one gene) of the host could play. To show why, let us assume the following scenario. Suppose that a gene that gets expressed within undifferentiated $CD4^+$ T cells would trigger their differentiation of the into Treg. If this were shown, it is possible that we could deactivate the expression of that gene (or set of genes) under strict laboratory conditions. Call that gene (or set of genes) *XYZ*. If we did so, we could probably argue that *XYZ* plays a basic role in the development of mice, i.e. that it enters the process of development. Notice, though, that *XYZ* deactivation would not stop gut development. It would only 'make it worse' as the gut that will be denaturalized. Importantly, for what we know about gene expression, it can be asserted that a necessary *and sufficient* condition for mice to have a 'normal' gut is not only that they bear *XYZ* within their genome. As

[121] Notice that the practice of endogenizing background conditions and making them part of the system that one aims to explain, is common in evolutionary biology, as Okahsa has argued (2018). Importantly, to stop seeing a phenomenon as a background condition and start seeing it as part of what needs to be explained, it is first necessary to find a way of thinking of the phenomenon holistically. That is what I argue needs to be done with the microbiome, in virtue of the evidence currently available.

important as bearing *XYZ* is that *XYZ* expressed itself during the stage of development when it is required. A mouse whose *XYZ* gene fails to express will have a failure in gut development. Because this is so, we would say that *XYZ* is a part in the process of development in mice and, thus, it is part of the biological individual that is developing. *XYZ* is hence a difference-maker in the process of mouse development, and thus it is part of the system that 'unfolds' during development.

The same logic that applies to *XYZ* in my hypothetical scenario applies to the polysaccharide A in the wall of *B. fragilis*. Polysaccharide A is a difference maker for gut development, despite nor belonging to the host genome, but to its microbiome. This is so because all the properties that I argued *XYZ* would have in my hypothetical scenario are now ascribed to polysaccharide A. Indeed, the lack of polysaccharide A in the walls of *B. fragilis* would have the same effect in mice than the lack of *XYZ* in the host genome. The fact that *XYZ* was a physical part of the host genome, whereas polysaccharide A is not, does not invalidate the fact that polysaccharide A is part of the developmental individual insofar as it is a difference maker in development. And, thus, polysaccharide A is part of the developmental individual.

If my analogy works, then I would have proven that the 'part-of-the-system' arguments invoked by defenders of HCE are *legitimate*, thus contradicting the thesis defended by Bourrat and Griffiths (**chapter II, section 3**). However, they work differently to what some defenders of HCE have assumed. First, because if in my thought experiment *XYZ* would be 'part-of-the-system' in virtue of its participation in development, and in reality polysaccharide A plays the role of *XYZ*, then what should be considered 'part-of-the-system' is polysaccharide A, including the gene (or set of genes) that are responsible for its expression in *B. fragilis*, rather than *B. fragilis*. Second, because if my argument is correct, then the way of setting the boundaries of the holobiont is radically different from the way that defenders of HCE tend to favour, and that Bourrat and Griffiths have argued against. In the process-based view of biological individuality that I favour, the microbiome would consist in the set of traits that are involved in the relevant biological (sub)processes.

Assuming that these traits *need to be* localized in the genome of the host, or otherwise will not be part of the process, is an unjustified apriorism that commits a fallacy of misplaced concreteness (cf. Stencel and Proszewska 2017 for a similar argument). [122]

2.2.2. The microbiome as an evolutionary agent for the host. A response to Bourrat and Griffiths

How does the evidence just presented connect to the claim that holobionts are units of selection? A full argument to support that idea will be developed in **chapter V**, but I will now outline the master lines of my response, and connect them to the case study I have just presented. Notice that my argument will be connected to the view about the units of selection I presented in **chapter III, section 3.3.1**, where I argued that the set of entities that participates as a unit in a significant biological (sub)process must be considered a unit of selection (at least from a MLS perspective), or otherwise the evolution of some traits would be unexplained. Let me now apply that reasoning to gut development in mice to argue that the host is not the unit of selection, but the holobiont is, and thus some of the most significant traits involve in mice development are the result of group selection.

First, $CD4^+$ T cell differentiation is a basic stage in the development of mice and, more importantly, our current evidence suggests that it is mediated by the action of polysaccharide A in *B. fragilis*. Let me make a thought experiment now, which partially follows the thought experiment from the previous section. Imagine that the process of $CD4^+$ T cell differentiation is mediated by *XYZ*. And, also, imagine, for the sake of the argument, that every single member of a population of mice bears *XYZ* within its genome. If this were so, then there would be one legitimate biological question about *XYZ*, namely, how the process by which *XYZ* was selected worked. A plausible answer would be that *XYZ* conferred a fitness advantage to its bearers, and thus, mice with a functional copy of *XYZ* in their genomes would produce more offspring than

[122] Notice that this is response to Bourrat and Griffiths because it was my personal decision (for different reasons, see footnote 46) to analyse their arguments against the claim that the holobiont is an individual. Nonetheless, notice that my comments also refute the arguments given by Booth (2014), and Queller and Strassmann (2016).

those without it, and hence in the long term all the members in the population would have a copy of *XYZ*. This would be a conventional selective explanation. But notice that the same type of explanation is possible if the members of the population of mice had been able to transgenerationally 'kidnap' an *entity*[123] from their environment that produced the same effect without being integrated in the genome of the mice. Could not mice, say, 'externalize' the entity (resource)[124] that is needed to carry out their development? If this happens, is there any solid reason to reject the idea that the factor that is externalized belongs to a common unit on which natural selection can act?

The answer to these two questions is affirmative. In fact, the case described in this section proves that this can be so. But, more importantly, in **chapter I**, I reviewed part of the evidence that supported, according to HCE defenders, the thesis that holobionts are units of selection. There, I avoided to make any claims about how to interpret their position from the perspective of the debates about the units of selection, because I had not presented my analysis of the concept of biological individuality yet (**chapter III**). However, now I am in a condition to make a stronger claim about the relevance of some of the evidence gathered by HCE defenders. Their evidence points to a *proxy* for detecting the existence of a common (host + microbiome) unit of selection because if a host necessitates the active intervention of its microbiome to realize a biological (sub)process, then it is likely that the object that is being naturally selected is not the host itself, but the host plus *some of the traits* of its microbiome. If the microbiome had not been part of the units of selection the evolution of the biological (sub)process would have been radically different. Ignoring the role that the traits of the microbiome have played in the evolution of these processes would mask the evolution of some of the adaptations that sustain the (sub)process. Or, alternatively, they would be explained as a result of drift, despite the connection between these traits and the fitness of the biological object. Therefore, if we (*now*) regard these traits as adaptations—and taking into account that an adaptation is a feature that has been selected *for*

[123] 'Entity' should not be equated with 'microbial taxon' here, as I will make the claim that the entity is the trait that plays the functional role irrespectively of the microbial taxa that carries it.
[124] I use 'externalize the resource' meaning 'not integrating the factors responsible for its appearance within the genome'.

(**chapter III, section 3.3**)—then it *must* have been selected for the good of the whole, i.e. for the good of the holobiont. And this is only possibly if the host plus its microbiome are conceived as a unit of selection.[125]

Notice that my response dispels the concerns raised by Bourrat and Griffiths about the differences between the microbiome and a background condition, such as the gravitational field. 'Background conditions' refers to a set of physicho-chemical factors that are essential for the life of organisms and/or that partially determine that an individual has the propiertes that characterize it. One standard example would be oxygen in aerobic organisms. The main difference between the physicho-chemical interactions between a host and its microbiome, and the physicho-chemical interactions between an organism and its background conditions is that the former, but not the latter, depend on the expression of some traits on the host and on its microbiome, as I have argued in this section. Therefore, even while the interactions between the host and its microbiome are directly mediated by physic-chemical elements (and it could not be otherwise, unless one believe in the existence of non-material objects), these physic-chemical elements appear and are transgenerationally maintained as a consequence of the traits that are coded in the microbiome, and that are responding to these interactions.

Second, my response also explains why Bourrat and Griffiths' disregard of the holobiont is based on an ungrounded premise. In their view, the only entities that must be considered 'part-of-the-system' are those that are coded in the species genome. While it is true that this compartmentalization grants the existence of a higher degree of transgenerational fidelity in the transmission of the traits, it says nothing about their common history of selection. Conceptually speaking, the compartimentalization could just be a historical accident in the evolution of species. The fact that *many* phenotypic features of a species can be explained by appealing to how the traits in the genome have been selected does not mean that this *must* be so for every phenotypic feature. Conceptual necessity cannot be predicated from biological contingency. Indeed, the case I

[125] I will not make any appeal to MLS here, but I obviously mean unit of selection 'from a MLS perspective'. See **chapter III, section 4**; and **chapter V**.

have reviewed here suggests that, if it is sound to consider that the processes of development in mice requires the intervention of microbiome traits, then it is conceptually sound to extend the 'boundaries' of the unit of selection beyond the genome, to the hologenome. In that vein, it can be said that the evolutionary history of *mice* (i.e. the evolutionary history of the *host*, or the *macrobe*) cannot be understood unless some of the traits it bears are relativized to its hologenome. And this is what defenders of HCE have been claiming for several years, as I showed in **chapter I**.

3. The symbiont perspective – when the holobiont is not the biological individual

In the previous section, I argued why the host's perspective requires the holobiont 'lens', that is, why the evolution of the host needs to be studied as a story of how the holobiont evolves. An immediate interpretation of the position I held in the previous section would be to consider that, insofar as mice need to be considered holobionts for their interactions with *B. fragilis*, then *B. fragilis* must necessarily be part of the holobiont, and thus mice and *B. fragilis* will be cospeciating. However, I think that interpretation of the case, while intuitive, is not sufficiently fine-grained, because it would mask some of our current knowledge about the species *B. fragilis* that suggests that mice and *B. fragilis* are cospeciating. In this section, I will review part of this evidence and derive some consequences about the boundaries of the holobiont. Concretely, I argue that *B. fragilis* participates in a series of biological (sub)processes at a different temporal scale than the (sub)processes that define the boundaries of the holobiont and, thus, *B. fragilis* is not part of the holobiont.

3.1. The life mode of B. fragilis*: some basic considerations*

B. fragilis is a bacterium and, in discussing the notion of biological individuality, this feature has particular relevance, because bacteria are, in general, unicellular organisms.[126] Insofar as bacteria are unicellular organisms, there are

[126] Of course, the case of biofilms is an important exception, because in that case it seems that the individuality of the bacterium gets 'diluted' in the individuality of the 'multicellular' biofilm (Ereshefsky and Pedroso 2013, 2015, 2016).

two key features of their biological nature that cause that the kind of biological (sub)processes that define the boundaries of their individuality are given at a different temporal scale than the kind of biological (sub)processes than define the boundaries of multicellular organisms, and this has important biological consequences.[127]

B. fragilis is a Gram-negative bacillus, that is, it is a rod-shaped bacterium, whose body-boundaries are limited by a peptidoglycan layer and a lipid membrane. Their diameter oscillates between 0.5 and 0.8 μm, and their longitude between 1.5 and 9 μm. *B. fragilis* does not form spores, and it does not have any elements that would allow motility (flagella, cilia, etc.). Despite its immobility and its reduced size, *B. fragilis* engages in a series of biological (sub)processes. It engages in a set of physiological processes that maintains each bacterium alive, it develops to a certain extent by multiplying its body structures, and it reproduces by binary fission. All these processes occur within the boundaries of its cellular membrane and its bacterial wall (being channelled by a process of exchange of matter and energy with its external environment) and occur in very short periods of time. The temporal scale at which these biological (sub)processes occur is dictated by the reduced size of the bacterium, and its short life cycle. Due to these two features, the bacterium can only interact with some of the elements of its immediate surroundings, and thus its possibility of generating an extensive network of interactions with other biological individuals is strongly limited.

In contrast with *B. fragilis*, multicellular organisms are composed by a wide variety of cell types (epithelial, muscular, etc.). Each of the cells of a multicellular organism, taken individually, engages in its own biological (sub)processes. However, the multicellular individual that is formed by these different cell types is given at different temporal scales. Its boundaries are necessarily differently, because of the different type of biological (sub)processes that the multicellular engages in. There are necessarily ontological differences between *any* multicellular individual and *any* unicellular

[127] I have chosen *B. fragilis* because it is the species that triggers development in mice. But notice that my arguments should generalize to most microorganisms.

individual, even when the unicellular individual only acquires its properties because it is a zygotically-derived part of the multicellular individual (e.g. specialized cell types).

The biological (sub)processes that *B. fragilis* engages in and that determine its individuality can occur in a wide spectrum of ecological niches. The presence of *B. fragilis* is well documented in hosts that have a very different evolutionary history than mice. For instance, it is abundantly found in humans (Mazmanian et al. 2008), bears (Sommer et al. 2016), camels (Paul and Dey 2015), and pigs (Tajima and Aminov 2015), among others. It is widely acknowledged that the species composition of each of these holobionts is different (Ley et al. 2008), which suggests that the range of environment where *B. fragilis* can grow and reproduce is very diverse. This is important for it entails that the individuality of *B. fragilis*—conceived in relation to the biological (sub)processes that it engages in—does not depend on its relationship with any specific species of host, or any other specific bacterial species. Of course, its ecological range depends on some of its biological properties, and thus some of the environments will be more suitable than others for realizing its biological functions and forming a population. However, even while that is so, none of these factors will necessarily determine the phylogenetic evolution of *B. fragilis*, as I will argue in what follows.[128]

3.2. The discontinuity between the symbiont and the holobiont: why B. fragilis *is not a part of the evolving holobiont*

My argument above supported the thesis that the species *B. fragilis* is not part of the evolving holobiont, despite the influence that it has in mice development. Furthermore, even in the case that its influence in mice development had been a result of slection, that would not necessarily entail that *B. fragilis* evolves as a member of the holobiont. Let me explain these points more clearly. In **section 2.2.3** I had introduced the notion of interchangeability, i.e. the thesis that it is possible that a holobiont realizes the same capacities (i.e. bears the same phenotypic traits), despite the fact that the species that compose its microbiome

[128] Notice that I will later connect the evolutionary independence of the bacterial lineage to the fact that these are individuated according to their 16S rRNA composition.

are not the same. I argued there that it is enough that the species that form the host's microbiome have the functional capacity of acting as 'agents', i.e. they have the capacity of functionally expressing the traits that the holobiont needs to develop. Because different bacterial species usually carry similar information (factors, genes, etc.), this situation is to be expected, and this creates an apparently strange asymmetry: while the host is a unit of selection *qua* a holobiont (i.e. host + microbiome are the unit of selection for at least some phenotypic traits), the same is not necessarily true for the species that compose the host microbiome. In other words, while the phylogeny of the host is inseparable for the evolutionary history of the holobiont, the phylogeny of the microorganisms that compose the host's microbiome is not necessarily linked to the evolutionary history of the holobiont. The case of *B. fragilis* revised in this section reflects the biological reasons why this is so.

First, the ontogeny and reproduction rate of each bacterium is such that their *tempo* of evolution is quicker than the *tempo* of the holobiont, so that every evolutionary change that might be triggered as a consequence of living within a concrete holobiont can always be explained by appealing to the evolutionary history of the bacterial lineage, rather than appealing to the evolutionary history of the holobiont. In other words, the biological (sub)processes that a bacterium engages in are necessarily not co-extensional with the biological (sub)processes that the holobiont engages in. This is so even though the microbiome *functionally defined* participates in the same biological (sub)processes that the host. This situation is especially remarkable for those cases when a bacterium can live in more than one environment, such as *B. fragilis*. Insofar as the bacterium can live in different holobionts, and even in different populations that are not necessarily linked to any host, it is unreasonable to assume that the evolutionary changes that could occur in *B. fragilis* were a result of its cospeciation with any of the hosts it interacts with. In other words, any evolutionary change that the bacterium could experience and that would benefit a concrete holobiont can perfectly be explained as a 'fortuitus benefit' that makes the bacterium apt to survive in a concrete environment. As such, this fortuitous benefit will be expected to disappear as soon as bacteria change their environment.

Second, but related. As the most widely accepted criterion for determining bacterial phylogeny is the evolution of 16S rRNA, precisely due to its low rate of evolution compared to other components of the bacterial genome, the evolution of those lineages of microorganisms whose interaction with a specific host lineage is contingent is not expected to be linked to the evolution of the host lineage. In other words, host lineages and bacterial lineages are not expected to cospeciate. If the changes in the bacterial lineages are required to affect the 16S rRNA to be considered phylogenetically significant, then the interaction between a bacterial species and a host species must be much longer than the time that most host-symbiont interactions last. Only some cases of obligatory host-symbiont interaction that have become evolutionarily established (e.g. aphid-*B. aphidicola*) would satisfy this criterion. These are the cases that, in **chapter I**, I referred to as the 'hereditary holobiont'. Notice, however, that the point I am making in this section is not that the holobiont should be defined as the 'hereditary holobiont'. I am only posing some problems to including the bacterial *species*—as these are defined according to the 16S rRNA criterion—that compose the microbiome of a host within the definition of the holobiont; that is, within the holobiont *as an independently evolving entity*. Or, to use the vocabulary that I employed before, I am posing problems to the notion of the 'ecological holobiont', on the basis of my review of the ecological ranges of *B. fragilis*. Bearing this in mind, in the next section I will present my account of the boundaries of the holobiont as a mid point between these two extremes. This will only be possible by getting rid of the commitment to a hierarchical-nested view of the biological world.

4. Biological individuality must be ascribed afresh at each level. Towards a hierarchical-non-nested view of the biological world

The cases I examined illustrate the existence of two non-co-extensional responses to the question about the individuality of holobionts, depending on the perspective that one adopts. On the one hand, I argued that the host perspective requires considering the holobiont as the biological individual—and the unit of selection—because the biological (sub)processes that define its

individuality include some traits of the microbiome as part of the system (*contra* Bourrat and Griffiths, **chapter II**). On the other hand, I have argued that the symbiont perspective does not require considering the holobiont as the biological individual, for symbionts are usually shortlived unicellular organisms that can live in a wide range of ecological niches. In this sense, the fact that a concrete bacterial taxon resides in the tissues of a host can be the result of *chance*, rather than a result of host-microbe cospeciation. I connected that claim to the fact that bacterial phylogeny is usually traced by 16S rRNA analysis, a gene that was chosen, precisely, for its low rate of mutation/evolution. In this vein, I argued that the symbiont perspective clashed with the host perspective for both generate apparently contradictory responses about the individuality of the holobiont. Is there a way of solving this inconsistency? Or, in other words, is there a reason (a 'hidden' metaphysical assumption) why this inconsistency appears? And, if so, can we explain away the inconsistence by getting rid of the 'hidden' metaphysical assumption?

Before giving a reply to the questions, I want to go back to the distinction between the interactor and the replicator that I presented in **chapter III, section 3.2**. I had argued that, while it was true that the distinction between the two categories was problematic (**chapter III, section 3.3**), it captured the essence of a very important distinction that should be kept for subsequent discussions about the units of selection. That distinction was the difference between the *phenotypic properties* that determine the fitness of a biological individual *vis-à-vis* other individuals, and the *hereditary basis* of those properties. I said that the biological (sub)processes that lead to the realization of these properties are very heterogeneous. Furthermore, these (sub)processes can be materially realized at different temporal scales in the biological hierarchy, and this can potentially generate a multiplicity of biological individuals, each at a different timescale. Drawing on that distinction, plus the lessons derived from the cases I reviewed here, I now argue that the apparent contradiction derives from the assumption that the biological hierarchy needs to be strictly nested, and that this nestedness must be traced according to the entities that can participate in Mendelian relations of inheritance. These two assumptions entail a view of the hologenome that equates it with the sum of genomes (i.e. the collection of

organisms that are member of different taxa) that compose the symbiotic assemblage. According to it, the holobiont would be a unit of selection if and and only if all the taxa that compose it cease to be units of selection in their own. Or, in other words, the holobiont would only be a biological individual if all the microbial taxa that compose its microbiome were, so to speak, *under the service of the holobiont*, i.e. if they evolved according to the necessities of the host, rather than according to their own evolutionary interests—i.e. they should evolve to reduce conflict with the host. But, is there any way of judging when such type of evolution has occurred?

From the perspective of those who assume a hierarchical-nested view of the biological world the answer to the last question is obvious. A set of species would have evolved as a single object when the set of dependency relations that make the species participate in the same evolutionary process are reciprocal (i.e. when the species are cospeciating). And this general view would apply to every biological (sub)process that is used to define individuality[129] (physiology, immunology, development, etc).[130] In this vein, a holobiont would be a biological individual if and only if *all the taxa that are structurally included in the whole* are necessarily involved in the (sub)process. If some of the parts of one taxon (e.g. the 16S rRNA gene that is used to characterize the taxa of the microbiome) are not involved in that process, then the holobiont is not a biological individual. (**Figure 25**). Elwick (2017) provides an elegant argument suggesting that the hierarchical-nested view of the biological world has historically been—and still is—the standard metaphysical conception of biological individuality. Furthermore, both parties in the HCE dispute assume the nestedness of the biological hierarchy, irrespectively of whether they concive the holobiont as the 'ecological holobiont' or as the 'hereditary holobiont', and irrespectively of the fact that it contradicts hologenome selection.

[129] Remember that, in the view I am putting forward, a biological individual can only be individuated in virtue of the set of processes it participates in.
[130] Notice how this assumption is the same that underlies Maynard-Smith's criticisms to the view of Margulis about the role of symbiosis in evolution (**chapter I, section 2**).

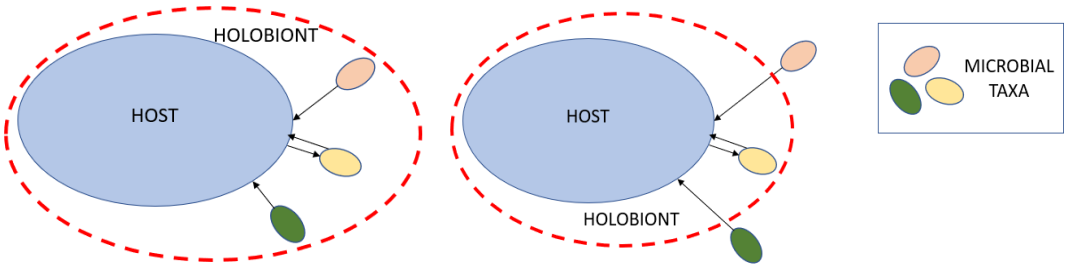


Figure 25. Simplified view of the dependency relations in a holobiont, and its boundaries, according to the hierarchical-nested view. The arrows stand for dependency relations with respect to a biological (sub)process. According to the hierarchical-nested view, only the host and the yellow taxa would constitute a biological individual, as their dependency relations suggest that they are cospeciating (right figure, 'hereditary holobiont'). However, the 'ecological holobiont' (left figure) would not be a biological individual, for even if the host depends on each microbial species, only the yellow species depend on the host, and thus the holobiont is merely a community of species evolving independently from each other.

The cases I analysed in this chapter shows why the hierarchical-nested view of the biological world is mistaken, as it obliges us to decide either that holobionts are individuals (and, thus, that all the species that compose the host's microbiome are necessarily cospeciating), or that holobionts are not biological individuals (and, thus, that except in rare occasions, the species that interact evolve independently from each other). The problem is that neither of these two extremes seems the right one, as I have shown, because the very nature of symbiosis leads to scenarios in which the relationships of dependency are not established between taxa, but *between the traits that the taxa bear*, especially when the biological (sub)processes that are being studied are biological (sub)processes that definife the individuality of the host. My hypothesis is that the holobiont is a unit of selection because the host genome is constantly coevolving with some of the traits of its microbiome, whereas the taxa (as they are determined by the 16S rRNA analysis) that compose the host microbiome are in most cases evolving independently from the host genome. If my hypothesis were true, it would create a dilemma about the unit of selection for some of the traits of the microbiome for, while these traits are physically connected to the bacteria taxa that bear them (they are parts of the bacterial genome, and thus they reproduce with the bacteria), they would be coevolving with the host genome.

I believe that my way of describing holobiont works, first, because our current biological evidence suggests that complex niche specializations of a host are usally a consequence of hologenomic evolution. Second, because it explains why the dependency relations between the host and the taxa that compose its microbiome are usually non-reciprocal. I will make this point more vivid in **chapter V**, where I will explain how these facts would work evolutionarily by using the tools of MLS. For the moment, though, it is enough to point out that some of most salient evolutionary innovations of certain animals depend on adaptations in their microbiome (Mendoza et al. 2018). In other words, some specializations (e.g. dietary specializations) that require strong cumulative selection could not be explained if the adaptations were looked for exclusively in the host genome. The functional contributions of the microbiome need to be investigated too, or it is not possible to explain why the host has the specific lifestyle that it has. Furthermore, some of these niche specializations have been argued to underlie the existence of speciation events is some animal clades. This suggests, first, that hologenomic evolution *must* have played a role in the evolution of some animals and, importantly, that it has been a result of the coevolution between the host and *the traits of its microbiome*, rather than a coevolution between the host and the *taxa that compose its microbiome*.[131]

Notice that, if my account of the boundaries of the holobiont were correct, then it would dramatically question the validity of the hierarchical-nested view of the biological world, especially when it comes to evaluate the hereditary basis of the phenotypic properties that biologists are interested in explaining. I propose to adopt instead a hierarchical-non-nested metaphysics for thinking about biological individuality, where hereditary relations are realized in different ways, and even sometimes between traits that are not bound to the same genome, but to different genomes. Possible conflicts between levels appear thus in virtue of the non-nestedness of the hierarchy. But its non-nested nature allows explaining why the functional microbiome is evolving with the host genome (as

[131] Notice that this suggests that symbiosis could play a role in speciation events, as Brucker and Bordenstein (2012a) had hypothesized. However, notice that its role is very different from the role they assumed, as in the view I am putting forward hologenomic incompatibilities would be a consequence of trait incompatibilities, rather than of species incompatibility. See **chapter 1, section 3.4.1**.

some evidence suggests) while the bacterial taxa are evolving independently from the host (as critics of HCE insistently point out). (**Figure 26**).[132]

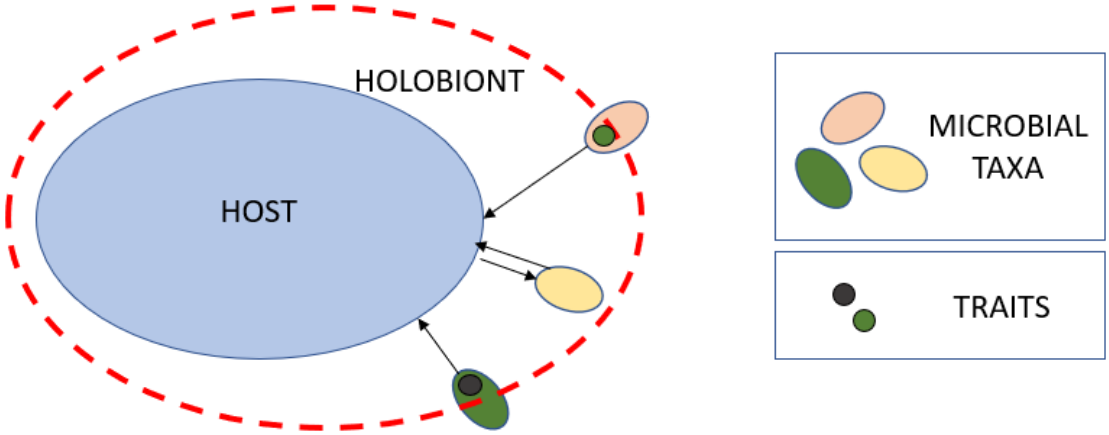


Figure 26. Simplified view of the boundaries of the holobiont in a non-nested-hierarchical world. The arrows stand for dependency relations with respect to a biological (sub)process. The non-nested-hierarchical view of the biological world entail that the holobiont is a biological individual that includes the host and the functional parts of the bacterial taxa it interacts with, rather than the whole taxa, except for the yellow taxon as it establishes relations of codependency. According to this view, the individuality of the host, and thus its evolution, is a consequence of its interaction with the traits of the bacterial taxa of its microbiome, while simultaneously the evolution of these traits is partially conditional on the evolution of the bacterial taxa that bear them. Under my interpretation of HCE, these traits part of two independently cospeciating entities. First, they are part of the host, despite the lack of physical connection to *its* genome. Second, they are part each of the microbial taxa in virtue of being physically bounded to their genomes, and thus reproduce with them. In this sense, holobiont evolution entails a non-nested-hierarchical view of the biological hierarchy.

My reason for defending that the biological hierarchy is non-nested is intimately connected to my process-based view of biological individuality. In a process-based view, ascriptions of individuality *must* start afresh at each level, and must be established depending on the (sub)processes of interest, and the temporal scale that is being studied (which will also constrain the (sub)processes). As this is so, different timescales and/or the different biological (sub)processes will probably generate a non-overlapping classification of biological individuals, in which the phenotypic properties of the individual and the hereditary basis of this

[132] My proposal is not unwarranted, but a way of fleshing out the type of metaphysical committments about the world that HCE entails.

properties are realized differently. And thus, the hierarches *of entities* generated in virtue of these (sub)processes would necessarily be non-nested. Applying the same logic that John Dupré applies to the case of species, I argue that if different biological individuals participate in different processes, then it is not expected that we must have a unique, and perfectly overlapping classification of biological individuals, nor a unique and perfectly overlapping classification of relations of inheritance.[133] Importantly, the point is ontological. Because if different biological individuals have different ontological properties derived from the different processes they engage in, then it is expected that they will evolve differently, and possibly in conflict with each other. As this is so, and some traits are most likely evolving as a consequence of two opposing selective forces, the perspective of MLS provides a perfect framework to model the evolution of holobionts.

4.1. The non-nested hierarchy and the 'blurry nature' of the holobiont

Once the commitments of a non-nested-hiearchical view of biological individuality have been presented, it becomes easy to explain away the problem of the 'blurry nature' of the holobiont. To recapitulate, the problem of the 'blurry nature' of the holobionts dereives form the observation that the symbiotic microbiome (defined in virtue of the species composition) of any hosts is usually very transient, contingent, and highly dependent on environmental factors such as host's diet. Because of this, any hypothesis about holobiont individuality will necessarily 'suffer from imprecision' (Booth 2014: 670), because the holobiont will not keep its identity over time (**chapter II, section 2**). Notice that, as it is formulated, the problem immediately vanishes from the perspective of my approach to HCE. Of course, it is expected that the species that compose the microbiome of the holobiont will change over the host's lifetime, as it is expected that each of the zygotically-derived cells will change. It would be surprising, though, that the *functions* of the microbiome would change; that is, that the microbiome would not remain functionally stable over time. But, of course, as neither holobiont thinking, nor HCE entail necessarily that there is

[133] Dupré applies a similar logic to the case of lineages, in his (2017).

cospeciation, changes in the species composition of the microbiome are irrelevant to discard any of the holobiont hypotheses.[134]

In my paper with Vanessa Triviño we even went further to argue, *contra* the specific case of Chiu and Eberl (2016), that they seemed to be neglecting several facts that strongly suggested the existence of a conceptual and biological continuity between host individuality *qua* monogenomic multicellular, and the individuality of the holobiont *qua* polygenomic multicellular.[135] Particularly, we pointed out that, first, they neglected the possibility of considering functional integration—or what I have called in this chapter 'dependency'—a correct criterion for determining the existence of a biological individual. This, we argue, is a strange decision, because, as they explicitly acknowledge, in some cases microorganisms substitute a whole organ of the host. It strikes us as surprising to claim that even in these cases the microorganisms are not part of the host, as Chiu and Eberl do, and seems to us rather an ungrounded decision, unless they apply the same criterion for *every* organ of the host.[136] Second, but related, they seem to defend a strange view about the concept of 'proper part', i.e. about the necessary conditions to determine when a type belongs to a compound (that is, it is *a part* of the compound), and when it does not. In their view, a necessary condition for a type to belong to an ensemble is that at least one token of this type is always interacting with other parts of the ensemble. But, as we see the situation, this criterion is very weird, because it would entail that proper parts of an object cannot change their type during the time that the object exists. This criterion would exclude things such as transplanted organs, which seems paradoxical. If someone gets her heart replaced by the heart from a donor, the new heart is indubitably a part of herself!

---

[134] Notice that the criticism based on the notion that either HCE is a hypothesis about cospeciation, or it is simply non-meaningful, is a common place in the literature, as I showed in **chapter I** and **chapter II** (e.g. Maynard-Smith 1991; Leggat et al. 2007; Chandler and Turelli 2013; Hester et al. 2015; Moran and Sloan 2015; Douglas and Werren 2016; Hurst 2017). Obviously, as the non-nested-hierarchical view of the biological world entails that we have to start afresh, *at each level*, cospeciation is not necessary for holobiont evolution.

[135] I follow the distinction made by Dupré (2012).

[136] Of course they could always say that their criterion is strictly genetic, but this would not solve the problems at all, due to the fact that most living creatures are indeed chimeras.

These two points seem paradoxical, but I think they are a consequence of Chiu and Eberl's commitment to the nestedness of the biological hierarchy. If one renounces to the nested-hierarchical view of the biological world, their problems are easily explained away. The authors are disregarding the facts that, first, there is a distinction to be made between the host perspective and the symbiont perspective, and second, that evaluations of individuality need to start afresh at each level. Since they disregard both facts, their arguments tend to emphasize how different the symbiont is from the host, and how this challenges thinking of the holobiont as a biological individual. But in doing so, they are throwing the baby out with the bathwater. Instead of focusing on the continuity between the multicellular host and its symbionts (host perspective), the authors focus on the discontinuity that exists between the symbionts and the host (symbiont perspective), and thus their criticism would automatically apply to other elements that are clear parts of the host. As a consequence, their account automatically 'dissolves' the individuality of the host, which becomes immediately de-individualized. Had the authors started afresh at each level, their problems would have been different, and thus the picture of holobiont individuality that would have emerged from their application of DT and the ICP would have been very different.

4.2. Reinterpretation of Sharon et al. experiment and Brucker and Bordenstein's experiment in the light of the non-nested-hierarchy

It is important to see how the experiments carried out by Sharon et al. (2009) (**chapter I, section 3.3**) and Brucker and Bordenstein (2013a) (**chapter I, section 3.4.4**) acquire their entire significance and justification when they are reinterpreted under the definition of the holobiont I am putting forward. Remember that, for the authors, their cases supported HCE. In the first case, because mating preferences in *Drosophila* had been consistently correlated to some components of the microbiome; in the second, because they could prove that there was a correlation between the microbiome and hybrid lethality in wasps of the genus *Nasonia*. The reader might remember that both cases raised similar concerns: why should we consider that the new microbiome is a part of the evolving holobiont instead of a 'new environment' where the 'old' host develops so that, in the end, it will trigger host evolution? An obvious

answer could be as follows. Assume, for the sake of argument, that the traits of the microbiome that determine the outcome of these processes could persist transgenerationally due to their fitness effects on the host, and no other changes in the host genome take place. In that case, could we possibly explain *causally* the persistence of any of these traits? The intuitive answer for those who are committed to the claim that natural selection is only effective on genomes would be negative for natural selection cannot 'select' anything *in the host genome*. But there seems to be something wrong in this answer because it seems that there must be *a causal basis* for this process. How can this tension be resolved?

A first answer would consist in arguing that the 'ecological holobiont' is the unit of selection, and thus a cospeciating entity such that the 'ecological hologenome' is its replicator.[137] But this answer would be obviously misguided, because it is not the case that all the species that a macrobe hosts during its life time get transgenerationally transmitted. A second answer, which is the one I am putting forward in this chapter, consists in renouncing to the nestedness of the biological hierarchy. If one accepts the claim that the holobiont is a unit of selection, and rejects the necessity of host-microbiome cospeciation, then one can argue that the host and its *functional microbiome* have been selected together, as a unit. In that vein, and combining this claim with the committment to MLS, I argue that the host has not changed its environment in any of the experiments, but rather its functional microbiome. This change in the functional microbiome triggers the two phenotypic traits observed, mating preference in *Drosophila* and hybrid sterility in *Nasonia*. At the same time, none of the bacterial taxa that compose the microbiome of the host has been selected as a unit with it. They have rather changed their environment and, thus, acquired some fortuitous benefits. In **chapter V** I will explain how this type of response can give a satisfactory solution to the HCE controversy.

[137] Notice that the problem is saying that the hologenome is the replicator, not saying that it is a unit of selection.

5. Brief summary of chapter IV: The non-nested nature of the biological hierarchy and the holobiont

In this chapter, I have presented a view of the biological hierarchy that, relying on some empirical evidence on mouse development and contrasting it with the life mode of *B. fragilis*, entails an important breakdown with standard accounts. I have argued that standard accounts of biological individuality fail to appreciate the individuality of holobionts because they assume that holobionts could be biological individuals only if the relations of dependency between all the species that compose the microbiome of the host are reciprocal. I have argued that such type of view is grounded on the belief that the hierarchy of objects that compose the biological world needs to be strictly nested, like a set of Russian dolls. I argued that such assumption is a misguided apriorism that first, disregards the distinction to be made between phenotypic properties of a whole, and the hereditary basis of these properties; and second, it disregards the possible existence of hierarchies of biological objects that are made in relation to different temporal scales in cirtue of the processes that are being studied and are, thus, non-overlapping. Drawing on this observation, I have argued that:

1. Ascriptions of biological individuality must start *afresh* at each level, without considering the possible disruptions that might come from 'below'.
2. In symbiotic assemblages, the host perspective, and the symbiont perspective must be always separated when biological individuality is being ascribed, as they will be in most cases species that are evolving independently
3. Because of the different temporal scales between the host and the symbionts, the host perspective entails that the holobiont (host + *functional* microbiome) is the unit of selection, whereas the symbiont perspective entails that the symbiont species, rather than the holobiont, is the unit of selection.
4. Therefore, the holobiont is a unit of selection from the host perspective, and its evolution must be studied from the perspective of a MLS framework.

Chapter V

‘Stability of traits as the kind of stability that matters in units of selection: An account of the role of the holobiont as a unit of selection from a multilevel selection perspective’

This chapter builds on the non-nested-hierarchical view of the biological world and presents a conceptual model to test the hypothesis that holobionts are units of selection. I start reconsidering the arguments against the claim that holobionts are units of selection presented in chapter II and argue that they are misguided, since they misidentify the target of selection, as they reduce the debate of the units of selection to the debate about identifying the replicators. Second, drawing on this criticism, and the necessity of applying a multilevel selection framework for thinking the role of holobionts as units of selection, I present a multilevel selection approach to think about HCE. I first elaborate a conceptual framework to test the hypothesis that holobionts are units of selection from the MLS1 perspective, and I later elaborate a model to interpret the claim that holobionts are units of selection from a MLS2 perspective. Concerning MLS2, I suggest an alternative definition of ‘inheritance’ that would serve to test whether my model applies, and how often, in real biological systems.

1. From the non-nestedness of the biological hierarchy to the multilevel selection perspective

The previous chapter provided a philosophical story about the metaphysical properties of biological individuals and what this entails for the biological hierarchy. Concretely, I argued that: 1) ascriptions of biological individuality must start afresh at each level because 2) the biological world is hierarchical and non-nested. I grounded my claims on recent observations about the physiology and development of mice. Even though the chapter contained some suggestions about the role of holobionts as units of selection, I did not properly develop an account of what holobionts would look like as units of selection. This is problematic because, first, HCE is the claim that holobionts *are* units of selection (**chapter I**) and, second, my project is to develop an account of what they would look like as units of selection (**Introduction**). To overcome that gap, this chapter articulates a new conception of how the holobiont could be conceived as a unit of selection from a process-based perspective to biological individuality. To do so, it is essential to keep in mind three theses that I have developed, and that justify the account of HCE that I elaborate in this chapter:

> *Thesis 1*. If a biological object participates in a significant biological (sub)process, then it is minimally a unit of selection (**chapter III, section 3.4**).
>
> *Thesis 2*. The role of the holobiont as a unit of selection needs to be thought from the perspective of a MLS framework (**chapter III, section 4**)
>
> *Thesis 3*. The holobiont is a unit of selection in virtue of the association between the host species and its functional microbiome (**chapter IV**)

The result of conjoining the three theses is the MLS framework I will elaborate here.[138] But before doing so, I need to to show why the last family of arguments against the role of the holobiont as a unit of selection is invalid.

[138] Notice that what I will present is just a 'sketch' of a possible endogenization in terms of evolutionary biology of the type of knowledge gathered by defenders of HCE.

2. 'Stability of species' is not a valid criterion to determine whether holobionts are units of selection.

In **chapter II, section 4**, I presented what I took to be the main criticism that had been raised against HCE, namely, that the species that compose the microbiome of a host are not transgenerationally co-transmitted together with the host (i.e. there is no partner fidelity), and thus this invalidates the claim that the holobiont is a unit of selection. I said there that, while it is true that the species that compose the microbiome of a host are not transgenerationally co-transmitted with the host, the validity of the argument depended on the assumption of SoS:

> *Stability of species* (*SoS*): A holobiont in generation n+1 will be a unit of selection only if the symbionts S_1, S_2, ..., S_n, that occur within the host H_{n+1} belong to the same species as the symbionts S_1, S_2, ..., S_n, that co-occur(ed) with the host H_n.[139]

Building on my discussion about the topic of the units of selection in **chapter III, section 3**, I am now going to argue that SoS is an incorrect criterion to characterize the units of selection for it is based on the replicator/interactor framework, and a hierarchical and nested view of the biological world, and it is thus prey of the very same problems that affect that account of evolution. Second, based on what I suggested in **chapter III, section 4**, I will argue that HCE must be interpreted from the perspective of MLS, and not from the perspective of the interactor/replicator framework, and thus the criticism against HCE based on SoS is completely invalid.

Let me start from the beginning. A criticism to HCE on the basis of SoS could only be consistently held if the same type of criterion of inclusion—perfect transgenerational transmission—would be applied to every single level of the

[139] See footnote 54.

biological hierarchy where selection at the lower level could disrupt selection at the higher level. One such case would be a multicellular organism. When could one argue that a multicellular organism is a unit of selection? Assuming, for the sake of the argument, that all cells in a multicellular organism share the same genome (so that the match one-organism/one-genome is satisfied), and assuming that a genome is a collection of nucleotides (the minimal entity of which copies are made), a multicellular organism would be a unit of selection if and only if there is a perfect transmission of nucleotides. Call that 'nucleotide fidelity'. Following the same scheme of reasoning against HCE I presented in **chapter II, section 4**, one could argue that:

1*. *Definition*: Multicellular organisms are entities composed by the collection of nucleotides that constitutes their genome.
2*. *Lewontin conditions for natural selection*: For natural selection to act on a given entity in a population it is necessary that the entity (a) exhibits phenotypic variation that (b) affects its fitness and that (c) the phenotypic variation is inherited with sufficient fidelity.
3*. *Nucleotide fidelity*: The only way of guaranteeing the satisfaction of (c) among multicellular organisms would be that the whole collection of nucleotides of their genome co-occurs transgenerationally.
4*. Therefore, the existence of nucleotide fidelity is a necessary condition to claim that multicellular organisms are units of selection.

One could even condense these four premises in one single criterion similar to SoS, namely: stability of nucleotides (SoN, hereafter):

Stability of nucleotides (*SoN*): A holobiont in generation n+1 will be a unit of selection only if the nucleotides N_1, N_2, ..., N_n, that appeared in the multicellular organisms M_{n+1} are the same as as the nucleotides N_1, N_2, ..., N_n, that appeared in the multicellular organism M_n.

But to take a criterion as SoN as a necessary condition for having a unit of selection would be very problematic, for it would simply rule out the possibility that *any* organism that reproduces sexually is a unit of selection! And that would obviously be an unacceptable response for any biologist. Interestingly, Benjamin Fitzpatrick has mathematically demonstrated a very similar point to the one that I am making here. Like me, he also believes that the disanalogy between sexual reproduction and microbiome transmission is not as abrupt as some critics of HCE assume, at least in the case of horizontal transmission of the microbiome. To quote: '[t]he conclusion that horizontal transmission rapidly erodes extra-genomic associations is equivalent to the conclusion that recombination rapidly erodes associations between genes within a genome' (Fitzpatrick 2014: 1). Nonetheless, in contrast with my criticism to SoS, Fitzpatrick's argument against critics of HCE still assumes that SoS is a necessary condition for claiming that holobionts are units of selection. He only points out that, where critics of HCE perceive a sharp discontinuity (SoN in sexual reproduction versus SoS in holobiont 'reproduction'), there is indeed a continuity, and SoS can be maintained by *some means* that critics of HCE have not thought about (horizontal transmission in a structured population or in a population where selection is acting). How does my criticism of SoS differ from Fitzpatrick's? I think the real problem with SoS is not empirical, but *conceptual*. Insofar as SoS is simply a way of restating SoN at a different level of the biological hierarchy, it has the problem of reducing the question about the units of selection to the question about which entity is the replicator. And, as I argued in **chapter III, section 3.3**, the replicator/interactor framework of the units of selection is simply invalid to describe some basic dimensions of the debate about the units of selection.[140]

The way in which I have reasoned that there is an immediate conceptual move from SoS to SoN can be criticized as a sort of straw man argument. Indeed, defenders of SoS would argue, the logic that requires the existence of SoS for holobionts does not apply to the genome of multicellular organisms, and

[140] In fact, some critics of HCE have excplitly stated that all that the debate on units of selection should be about is about replication, and since HCE defenders have for the moment failed to prove that hologenomes are replicators, then their theory is simply wrong. This is, for instance, the position that comes out from reading Skillings (2016: 886, ft. 7), or from reading Bourrat and Griffiths (2018).

thus accepting SoS does not necessarily commit them to accept SoN. Why? Because it could be argued that there is a sharp difference between, say, paradigmatic multicellular organisms and holobionts. How could this difference be captured? Both Moran and Sloan (2015) and Douglas and Werren (2016) argue that the key distinction between these types of 'assemblages' is that in the case of holobionts there is no 'fitness alignment' among the entitites that compose the microbiome of the host, whereas the same is not true for multicellular organisms, where it is possible to identify different mechanisms that guarantee that the fitness of the objects that compose the whole is *de facto* aligned (e.g. bottlenecks, germ/soma specialization, etc.). As a consequence of the lack of similar mechanisms in the holobiont (for there is no SoS, and SoS would either be the only possible mechanism, or the precondition for these mechanisms to evolve), the different partners that compose it will enter in a permanent 'arms race'—for the microbes could 'escape' from the microbiome to a different environment—that would erode any possibility for natural selection to act on the collective.

I suspect their reasoning is partially correct, and I think these authors have actually made a good point. If there are no, say, 'policing mechanisms' that guarantee that selection at the upper level is not disrupted by selection at the lower level, then all the changes that are attributed to selection on the upper level might indeed be mere sorting that results from selection at the lower level (**chapter II, section 5**). Nonetheless, after the discussion I introduced in **chapter III** and **chapter IV**, we are now in a position to see why this line of criticism, as persuasive as it might sound, is completely irrelevant to refuting HCE. First, it departs from a wrong concept of 'unit of selection', fot it assumes the paradoxical situation that the existence of 'fitness alignment' is prior to the existence of the object that is being selected. But, as I argued, this would leave the existence of 'fitness alignment' unexplained, and for sure biologists would demand a causal explanation of why the parts of a concrete conglomerate have their fitness interests aligned. To have such a story, we need to think of the units of selection from a MLS perspective. Second, it mischaracterizes the concept of 'holobiont' that is used in the HCE. As I argued, the concept of holobiont does not demand the existence of cospeciation, at least as species

are defined according to 16S rRNA criterion. HCE only demands that the host and its microbiome coevolve to a certain extent, and this coevolution can take sometimes the form of cospeciation, or it can also take the form of coevolution between the host and some of the rapidly evolving elements of the microbiome. In that vein, individuality has to be ascribed afresh at each level, and instead of thinking which are the elements of the lower level that could disrupt selection at the upper level, we must think about how to conceive the action of natural selection at the concrete level we are investigating. [141]

3. The holobiont as a unit of selection from a MLS1 perspective

I will now present a MLS1 perspective to study the evolution of holobionts, and to conceive their role as units of selection. First, I will motivate the introduction of the MLS1 perspective to think about holobiont evolution by relying on two recent biological studies that, I argue, implicitly rely on a MLS1 conception of the role of the holobiont as a unit of selection. Second, I will introduce my MLS1 model for the evolution of the holobiont.

3.1. Toxin exposure and vampire bats

The first case study I consider derives from Osmanovic et al.'s (2018) population genetics model of a symbiotic community. In their study, the authors evaluate how a host and its vertically-transmitted population of bacteria jointly adapt to an environment that is polluted with a toxic agent. The details of the model are irrelevant, but the results are significant for the point I will make in **section 3.2**. First, Osmanovic et al. observed that the stress created by the exposure to the toxin can be alleviated by changes in the traits of the host's associated bacterial community, and that these occurred during a single host generation—thus referred to as 'Lamarckian adaptations', see **chapter I, section 3.2**. Their discovery was, therefore, that for a concrete environmental pressure, a host-microbiome system could overcome it by selection on the traits of the microbiome, an opportunity that is not feasible in germ-free macrobes.

[141] Notice that this is a consequence of the metaphysical conclusion according to which the biological world is structured in a non-nested hierarchy, so that different levels overlap with each other.

Notice that this type of response depends on the different timescales on which the traits of the microbiome and the traits of the host can evolve and, thus, on the different scales on which selection for each of their traits is possible. Importantly, selection for microbiome traits determines selection for hosts (and for host's traits) within a single host generation: hosts whose microbiomes evolve to overcome the toxin exposure will survive longer than those that fail to do so and, thus, will release more microbes with the appropriate traits to cope with the toxin and will produce more hosts with their very same traits.

Second, but related, Osmanovic et al. study how selection acts transgenerationally on the host-microbiome system, assuming that there is vertical co-transmission of both elements. They observe that, over multiple host generations, there is a substantial reduction in the physiological stress caused by the toxin, which results from an increase in the total amount of detox secreted by the microbiome.[142] That leads the authors to conclude that the selection of hosts is 'based on a collective property of the bacterial community' (Osmanovic et al. 2018: 8). In other words, assuming vertical co-transmission, selection will favour those hosts whose microbiomes evolve the traits that allow them to cope with the environmental challenges posed by the toxin. The authors do not analyse, though, what would happen in host-microbiome systems where there is no host-microbiome co-transmission. Nonetheless, they speculate the following: 'Horizontal transmission is not expected to compromise the acquisition of toxin tolerance, but rather *to promote sharing of the benefits with offspring of other hosts*' (Osmanovic et al. 2018: 8, emphasis added). The reason for their conclusion is obvious: in evolutionary host-scales, only these hosts that acquire toxin tolerance will survive.

The second case comes from the study on the evolution of vampire bats carried out by Mendoza et al. (2018). In their research, they explicitly argue to be applying a hologenomic approach to study the evolution of obligate hematophagy (blood-sucking diet) in the common vampire bat (*Desmodus rotundus*). Their research is presented in two stages. First, the authors show

[142] Another population model of holobiont evolution that relies on vertical transmission can be found in Roughgarden et al. (2017).

that, whereas a taxonomic analysis of the gut associated microbiome in the common vampire bat shows the phylogenetic influence of the host in the composition of the microbiome (with more proximity to the taxonomic composition of carnivorous and insectivorous bats than to the taxonomic composition of frugivorous bats), this similarity masks a striking dissimilarity at the functional level. In their words: 'While there is little differentiation between the functional gut microbiomes of carnivorous, insectivores [*sic*] and frugivorous bats, the common vampire bat functional microbiome is almost completely distinct, and exhibits the least intra-species variation between the samples' (Mendoza et al. 2018: 661). The data is relevant because it suggests that the vampire bat microbiome might harbour some functions that are specialized to the extreme type of diet that vampire bats have, which means the microbiome could play a key role in the evolution of their lifestyle.

Second, the authors analyse whether their last observation was true by determining which of the adaptations—in their research, 'traits undergoing positive selection'—that allow vampire bats to cope with the challenges posed by hematophagy were located in the bat genome, and which were located in its microbiome. Mendoza et al. observed that many of the traits of the bat microbiome that were undergoing positive selection were causally involved in coping with some of the challenges posed by sanguivory, including both nutritional and non-nutritional challenges (**Figure 27**). Thus, they conclude:

> 'It is clear from our results that the common vampire bat has adapted to sanguivory through a close relationship between its genome and gut microbiome. (…) We showed that extreme dietary specializations, such as that of the common vampire bat, provide a comparative framework with which to tease apart the relative roles of genomes and microbiomes in adaptation. In conclusion, our study illustrates the benefits of studying the evolution of complex adaptations *under a* [sic] *holobiome framework*, and suggests that vertebrate adaptation studies that do not account for the action of the hologenome may fail to recover the full complexity of adaptation' (Mendoza et al. 2018: 664, emphasis added).

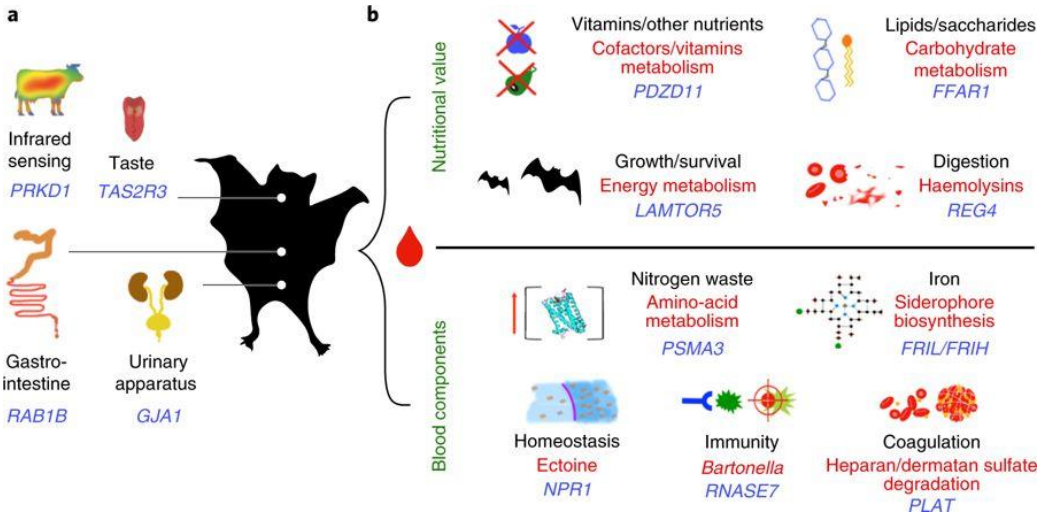


Figure 27. Mendoza et al.'s 'hologenomic' vampire bat. The figures indicate the dietary challenges that the common vampire bat has overcome. a) Adaptational contributions to sanguivory accountable to genomic changes alone (blue labels). b) Adaptational contributions to sanguivory within a hologenomic context (blue labels for host genes; red labels for gut microbial traits). (From Mendoza et al. 2018: 661, Fig 2).

In the next section, I will review the exact implications of this type of research for the debates about the units of selection, and for how the main claim of HCE should be conceived.

3.2. Stability of traits as the kind of stability that matters from a MLS1 perspective[143]

The research just presented has shown the usefulness of considering the holobiont as the unit of selection to uncover the evolution of complex lifestyles and thus indirectly discover the origin of some adaptations.[144] Notice that the

[143] Part of this section is based on my paper J. Suárez and V. Triviño (2019) 'A metaphysical approach to holobiont individuality: Holobionts as emergent individuals'. *Quaderns de Filosofia* 6(1): 59-76.

[144] The term 'adaptation', or 'adaptive trait' is ambiguous in the scientific and philosophical literature, as it can express at least two different concepts. Sometimes, a trait is considered an adaptation of an organism if it serves a purpose, that is, if it makes that the organism looks as if it were 'designed' by an engineer to fit its environment *now* ('engineering concept' of adaptation). However, sometimes the term is used to refer to traits that have been selected *for*—in Sober's terminology—because they were useful to perform certain activity *in the past*. In the latter sense, a trait is an adaptation for performing a concrete activity if it has the right type of causal history, which is the reason why the trait is there (Williams 1966; Sober 1984; Sober

research implicitly assumed—even though it did not explicitly discuss—a MLS1 approach to hologenomic evolution. The purpose of the studies was to discover which traits were undergoing (or had undergone) positive selection *in the particles* that integrated the holobiont.[145] The explanation of why these traits were undergoing positive selection was attributed to the existence of a common selection pressure affecting the group, so that, in each case, trait-frequency evolves towards an optimal state *for the group*. In the case of Osmanovic et al. (2018), though, the research had a profound limitation, as it assumed that the microbiome species were vertically transmitted, and thus hologenome evolution was somehow reduced to host-species and microbiome-species cospeciation, in which the holobiont has the capacity to form parent-offspring species-lineages.[146] Notice that this way of accounting for holobiont evolution has a serious limitation since, as I argued in **section 2**, assuming that SoS is necessary for holobiont selection is a wrong characterization of HCE and conflates MLS1 and MLS2. Despite this limitation, Osmanovic et al.'s result shows an important consequence. From a MLS1 point of view, particle evolution can be traced by considering the holobiont as the 'collective' unit of selection in which particle demography is studied.

On the other hand, the case of Mendoza et al. (2018) does not offer the same limitations as Osmanovic et al.'s model, since their analysis sticks not only to the taxonomic level, but also—and *especially*—to the functional level.[147] Given an organism with a highly specialized diet (obligate blood-sucking),

---

and Wilson 2011). This notion of adaptation has been called 'selection product' (Lloyd 1992, 2001, 2017a). This second meaning is the one I want to explore here because my argument will be the following: each taxa in the microbiome might bear a trait that is being positively selected because it is useful for the taxa *now*, in that concrete environment (**chapter 4)**; but the reason why the trait is there, in the microbiome, and thus why the trait has been acquired by the taxa now is because the trait has been useful *in the past* for the holobiont. In other words, I will argue that, historically, the reason why those traits are there and are adaptations is because they have been selected at the level of the holobiont.

[145] I use 'particle' intentionally, to avoid any commitment to the nature (or the way of individuating) the components of the microbiome. Indeed, as my claim is that the components are usually (but not always) the traits of the microbiome, rather than the taxa that bear the traits, using 'particle' introduces a necessary degree of ambiguity.

[146] Nonetheless, their reflection about what would happen in the case in which transmission were horizontal is glossed in MLS1 language. But, importantly, it is only a reflection, and not something that the authors have been able to mathematically prove yet, probably for the complexity of computing horizontal transmission.

[147] Other approaches to holobiont evolution have suggested doing the same. The most noticeable defence of this approach is probably Lemenceau et al. (2017), although a previous informal version of it appeared in Catania et al. (2016).

Mendoza et al. seek to understand which are the functional traits that make that behaviour possible, as well as whether they are undergoing positive selection. To discover the traits that underlie the evolution of the specialized diet, Mendoza et al. study not only the genomic level, but also the hologenomic level. That is to say, they study the traits that are undergoing or have undergone positive selection both in the host genome and in the microbiome. Why is this so? It is now necessary to introduce a brief digression about the notion of scientific explanation.[148] First, what is the *explanandum* in Mendoza et al.'s research? The evolution of a specialized diet (sanguivory). Their *explanandum* is simply:

(*Explanandum*) Why does sanguivory evolve in the vampire bat?

Second, what is their *explanans*? As sanguivory has necessarily have evolved by cumulative selection, i.e. it is clearly a lifestyle that can only evolve as a consequence of the accumulation of adaptations—in the historical sense—their *explanans* will necessarily be natural selection. That is to say:

(*Explanans*) Because some traits have been evolving by natural selection in a cumulative way such that sanguivory is possible in the vampire bat.

Third, in virtue of what do the scientists think that the *explanans* accounts for the *explanandum*? In virtue of the existence of a pattern or a regularity that allows natural selection *and only natural selection*, to *causally* produce that type of outcome in some individuals.[149] Assuming that natural selection has

[148] This digression does not come out of the blue. It is a reconstruction of what leds Mendoza et al. to postulate the existence of hologenomic selection which is partially based in José Díez's theory of scientific explanation (Díez 2014), as Roger Deulofeu and I have understood it and applied it in our papers (see Deulofeu and Suárez 2018, Deulofeu et al. 2019, Suárez and Deulofeu 2019; see also Moreno and Suárez (submitted)). To avoid extending myselef too much, I will not discuss the issue of scientific explanation here, nor why we rely on Díez's account. Such discussion can be seen, however, in Roger Deulofeu's dissertation.

[149] 'Only natural selection' because natural selection is the unique factor that stabilizes causally trait distributions in biological populations. The other possibility—drift—would be a form of sorting (**chapter II, section 5**).

produced the outcome, a second question arises. What is the material basis on which natural selection has operated to produce the outcome? At this point, scientists need to find the structures that have produced that complex lifestyle, that is, the traits that have been selected for, as well as the mechanisms that have guaranteed the transgenerational continuity of these traits in the population. In their research, the scientists will depart with an advantage, as they will know *a priori* which traits need to exist so that sanguivory evolves (namely, those that allow the vampire bat to cope with the nutritional and non-nutritional challenges posed by its lifestyle).

The hologenomic approach is thus indispensable because the genome of the vampire bat encodes only *part* of the traits that allow sanguivory. However, the bat microbiome offers itself as a possibility, as it encodes part of the traits that allow sanguivory. The vampire bat microbiome seems to have adapted (be adapting) to cope with sanguivory. Therefore, the evolution of sanguivory has only been possible because the bat genome and its microbiome are evolving together. Thus, the bat genome *and* the bat microbiome constitute a unit of selection. But, to recall the message I conveyed in **chapter IV**, this does not mean that the host and the taxa that compose its microbiome have necessarily cospeciated. They might have cospeciated but notice that this is irrelevant. All that is needed to argue that the holobiont is the unit of selection is to show that the host evolves as a holobiont, that is, that the host evolves together with its microbiome. And, importantly, this coevolution might leave a mark exclusively *at the functional level*, which is what in fact Mendoza et al.'s results suggest.

The last claim explicitly connects with the fact that HCE defenders claim to embrace a MLS perspective. To summarize what I already showed in **chapter I**, for HCE defenders, coevolution, cospeciation, or cocladogenesis are highly evolved states that require of a selective explanation. The hypothesis that the holobiont is a unit of selection, when it is conceived from the perspective of MLS, does not *a priori* require the existence of any of these states, at least when it comes to the MLS1 approach. The only requirement is that some collectives (i.e. holobionts) survive longer than others in virtue of the traits of the particle types that compose the collectives. Because of this, these particles

types—the entities whose fitness is tracked in MLS1—spread more, and thus their probability of eventually forming new collectives increases. To recall, a collective is a unit of selection in MLS1 in virtue of the parent/offspring lineages that the particle types form, and how some particle types increase in the population because they belong to the collective (**chapter III, section 3.4**).

Notice that so far, I have used the notion 'particle type' to remain as neutral as possible about the biological commitments of MLS in the context of HCE, but it is now time to unpack its meaning. In HCE, 'particle type' might either refer to the *host*, or to its *microbiome*, though not necessarily to the concrete *species* or taxa that compose the microbiome, but rather to the functions that are encoded in the microbiome (Catania et al. 2016; Lemenceau 2017). This should not sound strage to the reader, as this is the logical consequence of the processual-based analysis of biological individuality I presented in **chapter III** and **chapter IV**, as this is applied to the evolutionary holobiont. HCE is a framework to tell the story of how hosts evolve, *in virtue of the functions that their microbiome encodes*. It is a story of how the host genome evolves—or does not evolve—in virtue of the functions that can be 'externalized' in its microbiome. It is, thus, a story of how the collection of traits that is instantiated in a host-microbiome system gets 'filtered' by natural selection and is shaped by other forces, such as neutral evolution.[150] But it is not a story of how the bacterial species that mereologically compose the host's microbiome speciate. Therefore, the types that evolve are the host genome, or traits that are encoded in the host's microbiome for, in principle, host and microbiome are exposed to the same environmental pressures in virtue of mereologically forming the holobiont.

With this notion in mind, there is now a way of formalizing the role of the holobiont as a unit of selection from a MLS1 perspective. My claim is that holobionts can be regarded as units of selection provided that the notion of SoS is substituted by the notion of 'stability of traits' (SoT, hereafter), or simply trait-

[150] It is important to bear in ind that HCE does not assume that everything that occurs in the holobiont is necessarily a product of natural selection (i.e. it is not committed to a radical adaptationism). Thanks to Seth Bordenstein for emphasizing this point to me in several email exchanges.

recurrence *at the particle level*. The idea might be explained as follows. To detect possible cases of holobiont selection, one needs to regard the different traits that can be recognized in the microbiome of the holobiont. If there is an increase in the distribution of particle traits in virtue of their capacity of making some hosts survive longer than others in the host population, then holobionts are units of selection for the trait from a MLS1 perspective. To put it more formally:

> *Stability of traits (SoT)*: A holobiont will be a unit of selection if and only if for at least one trait *Tx* that appears in the microbiome of a host *n*, its appearance increases the probability that *Tx* increases its frequency in the global population of particles, provided *Tx* increased the survival and reproduction of the host *n*.

The main notions that support my account are that, first, SoT is possible without SoS. Or, in other words, that it is logically possible that the frequency of a trait increases in the global population of particles due to its effects on the survival of the host, even though the frequency of the microbial species does not increase, as the trait increases its frequence in different bacterial lineages. Second, that SoT is the way in which holobionts evolve. Let me now develop the two ideas.[151]

First, our current evidence suggests that SoT and SoS can be logically disconnected. On the one hand, the functional traits that are encoded in the microbiome are usually redundant (Boon et al. 2014; Vieira-Silva et al. 2016), which indicates that different microbial species in the microbiome can—and *do*—carry the same traits, despite belonging to different lineages. On the other hand, bacterial microorganisms are known to constantly engage in horizontal gene transfer (HGT, hereafter) with each other, which had led some philosophers and biologists to argue that microorganism evolution follows a web-pattern, rather than tree-pattern (Bapteste et al. 2009). Importantly, our

[151] My argument will be conceptual. It is about logical possibility, rather than about biological contingency. Nonetheless, it is based on some of our current biological evidence.

current evidence suggests that the frequency of HGT is substantially increased in the host's microbiome, which had led some researchers to argue that the holobiont is a 'hot spot' for HGT (Liu et al. 2012; Lerner et al. 2017; Blow et al. 2019; Schnorr, personal communication). The conjunction of these two pieces of evidence suggests that the functional traits encoded in the microbiome are not as strongly linked to the 16S rRNA as it might seem for the way in which bacteria reproduce (by binary fussion), and thus some of these traits can be evolving *simultaneously* in different bacterial lineages. Moreover, this suggests that the longer the host survives, the higher the probability of HGT that leads to the increase in the frequency of these traits in the microbiome that increase the fitness of the host, as they will be subject to natural selection due to the environmental pressures that the host experiences, as Osmanovic et al. (2018) have demonstrated. And, inversely, the shorter the host survives, the lower the probability for HGT, and thus the lower the probability that the microbiome traits that increase the fitness of the host spread, first, in the microbiome, and second, in the global population of particles. Therefore, this shows that the host and its functional microbiome engage with each other in fitness-affecting interactions, such that the higher the increase in host fitness, the higher the increase in the fitness of the traits encoded by its microbiome. And this is conceptually possible because the traits can be encoded by different microbial species in the microbiome of the host, which suggests that the species composition of the microbiome is to a certain extent irrelevant to conceive the role of the holobiont as a unit of selection.[152] [153]

Second, viewing holobiont evolution from a MLS1 perspective framed in terms of SoT entails a new view on holobiont evolution that has not been analysed in the literature yet (cf. Rosenberg and Zilber-Rosenberg 2008, 2014; Roughgarden et al. 2017; Doolittle and Booth 2017), but that corresponds to the type of reasoning that underlies the study by Mendoza et al. (2018). Remember that from the perspective of MLS1, whether holobionts form parent-offspring lineages is irrelevant to decide whether they are units of selection (**chapter III,**

[152] It is obviously important, though, in the case of tight genotype-genotype connections, such as vertically transmitted endosymbionts.

[153] These two observations also relate to the interest that HCE defenders have in developing concepts such as 'phylosymbiosis' and 'community heritability'. See **section 3.2.**

section 3.4). It is enough to determine, first, that the particles that compose the holobiont engage in fitness-affecting interactions and, second, that the differential contribution to the global population of particles is a consequence of the mutual interaction between the particles that compose the holobiont. In our case, it is enough to show that the differential replication of the traits in the microbiome, or the differential reproduction of the hosts that form the holobiont is a consequence of the fitness-affecting interactions between the host and its functional microbiome. In that vein, it is possible to determine how the traits in the microbiome *must* evolve to consider that the holobiont is the unit of selection. I will now offer an argument to support that the results by Mendoza et al. (2018) suggest that the holobiont is a unit of selection from a MLS1 perspective.[154] Notice, though, that the argument I will offer is conceptual. To test its validity for every biological system it is necessary to empirically validate whether the microbiome has evolved how I suggest here that the microbiome must evolve if the holobiont is a unit of selection.[155]

The key elements of my argument for holobiont selection in terms of SoT are sketched in **Figure 28**. According to my account, a holobiont will be a unit of selection for a trait T_a of the microbiome if and only if SoT is true. On the contrary, if SoT is not satisfied, then the holobiont will not be a unit of selection for T_a. MLS1 holobiont selection is possible in the following circumstances. Suppose we have a population formed by two variants of a host species (*red*, *blue*, in the figure), and that each variant interacts with a maximum of three microbial species, each taxon represented by a different colour (*green*, *yellow*, *grey*). Following the account of the holobiont that I presented in **chapter IV**,

[154] I explore the consequences of this argument in a paper with Vanessa Triviño (under review) 'Holobionts as emergent manifestors of adaptation'. There, we argue that the case of the vampire bat suggests that the traits are dispositional adaptations of the bacterial lineages, and etiological adaptations of the holobiont. We need to develop the approach further and explore the possible consequences, as well as the set of biological mechanisms that could support the validity of our approach, but notice that if our hypothesis were true, it would dramatically change the extension of the concept of 'lineage', thus having serious consequences for MLS2.

[155] Notice, however, that by arguing this way I am not neglecting that an account of holobiont evolution in terms of MLS2 (i.e. in terms of collectives that stand in parent-offspring relations) could be provided. Indeed, in **section 4** I will suggest what that type of account would look like. Rather, what I am arguing is that (1) from a conceptual point of view, that type of account seems unnecessary to argue that holobionts are units of selection (**chapter IV, section 3.4**); (2) the account I am presenting squares with part of the current scientific practice without the necessity of discussing how to determine the existence of inheritance (i.e. parent-offspring lineages) at the holobiont level (**section 3.1**); (3) it agrees with the claims made by HCE defenders that they embrace a MLS framework for the study of holobionts (**chapter I**).

assume that at *t1*, the host interacts only with one of the species of the microbiome in virtue of a trait that admits two variants, represented by the *white* and the *black* circles within the bacterial taxa, respectively. In this scenario, the holobiont (represented by the dashed red circle) would consist in the host plus the traits in the bacterial taxa. Now assume, by hypothesis, that the host that bears the black trait survives longer than the host that bears the white trait, so that at *t2* it is still alive, while the red host (and thus, its holobiont) is not.[156] Assume that the black trait can participate in events of HGT, such that at *t2*, it is present in all the taxa that compose the microbiome of the blue host. In this scenario, the blue host would have reproduced in *t1* and in *t2*, and so would have done the microbes of its microbiome, whereas the red host would have only reproduced at *t1*. Importantly, this has the following consequences in the global population of particles:

- There will be two blue hosts, and only one red host;
- Four microorganisms will bear the black trait, and one individual will bear the white trait;
- There will be five members of the green taxon, three members of the yellow taxon, and one member of the grey taxon.

There three effects are consequence of the effects of HGT, and the effects of the trait in the survival of the host and are disconnected of the taxa that used to bear the trait at *t1*, as the growth in the members of the yellow taxon suggests. My hypothesis is that this model runs for several generations the black trait will become stable in every microbiome of the host population, or it will reach a situation of equilibrium with the white trait, in case that the latter were advantageous in the case that two taxa with the white and the black trait were competing in the same host. SoT, thus, can be used to model holobiont evolution from a MLS1 perspective, without assuming the necessity of species transmission. Everything that is required for holobiont evolution is that the *traits* encoded in the microbiome can be preserved outside the holobiont, and that

[156] Importantly, in this scenario, *t1* and *t2* correspond to reproductive events, i.e. moments during the ontogeny of the host when it can reproduce.

they increase their presence in the global population of particles *as a consequence of their effects on the fitness of the host*.

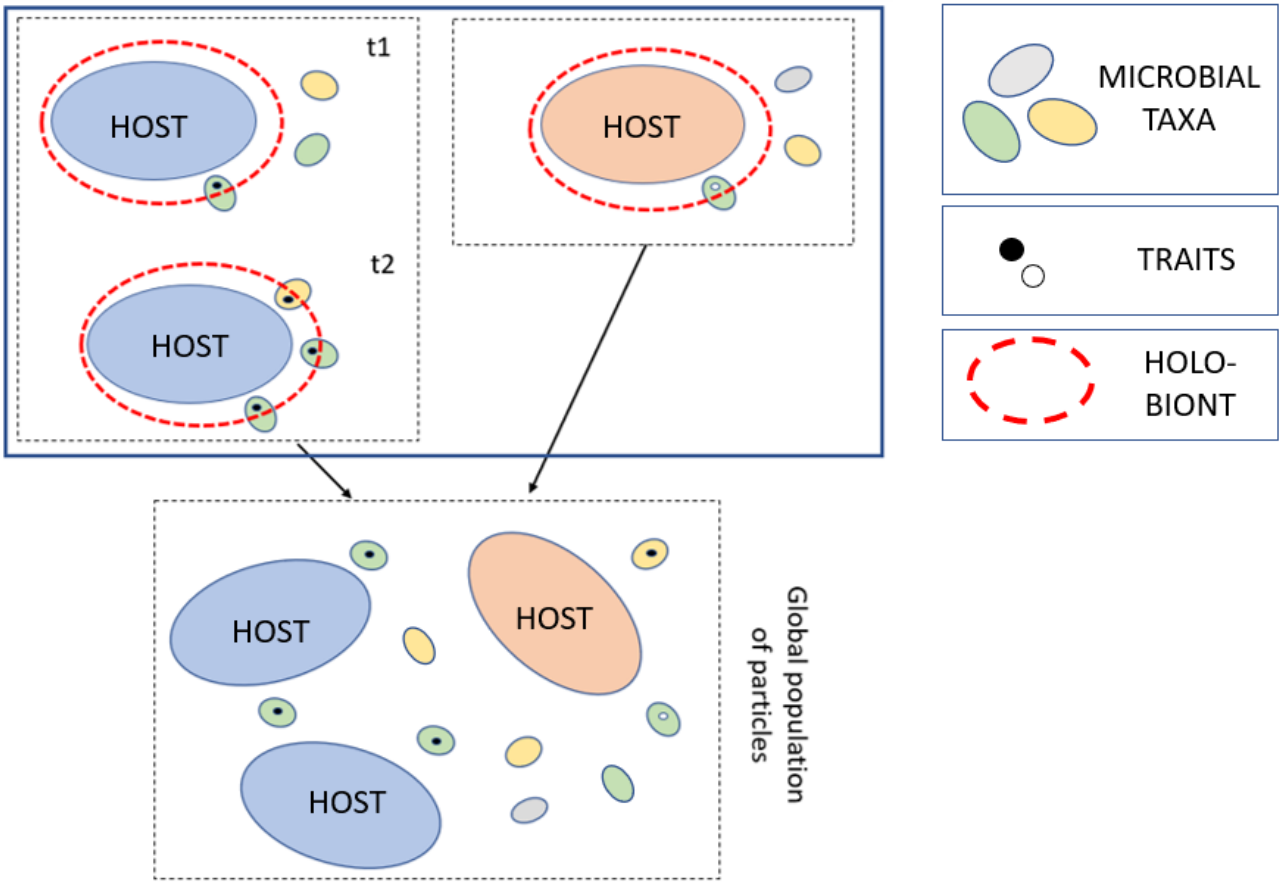


Figure 28. The holobiont as a unit of selection from a MLS1 perspective, according to the SoT account.

3.3. The relevance of phylosymbiosis and community heritability from a MLS1 perspective. A brief proposal

As I discussed extensively in **chapter III, section 3.4**, what is relevant to determine whether there is MLS1 is that the collective has an influence on the fitness of the particle types that form the collective since the fitness of the collective equals the averge fitness of the particles types that compose it (**Figure 23**). To repeat, whether collectives stand (or do not stand) in parent/offpring relations is irrelevant for MLS1. As, in my account, the particle types are the different variants of the host and the traits encoded in the microbiome, and what is counted is how the traits that the lattest encodes

evolve in different host generations, what is traced in MLS1 is how the traits encoded by the microbiome increase, and how their increase affects the increase in the variant of the host that harbours them. I said before that a consequence of the non-nested structure of the biological hierarchy is that the host is the focal unit of HCE, as HCE only aims to explain host evolution (**chapter IV**). It is now time to connect my MLS1 through SoT account to some of the concepts and research that biologists working under the framework of HCE have developed. In this section, I will briefly relate the research on phylosymbiosis and community heritability to the MLS1 approach to holobiont selection.

Phylosymbiosis refers to the 'eco-evolutionary pattern in which evolutionary changes in the host associate with ecological changes in the microbiota.' (Brooks et al. 2016: 3). In other words, a phylosymbiotic pattern between two (or more) species obtains when the evolutionary patterns of the species mirror each other, such that their changes are concordantly related in a statistically relevant manner. For the case of the holobiont, a phylosymbiotic pattern exists when the sets of species that compose the microbiome of the host mirrors the phylogeny of the latter (**chapter I, section 3.4.2**; **Figure 9**). So stated, phylosymbiosis seems to have little connection with the claim that holobionts are units of selection according to MLS1, at least as I have interpreted it in terms of the notion of SoT. But this interpretation would be different if phylosymbiosis research were reframed *functionally*, that is, in terms of the OTUs, rather than in terms of the taxa. If they did so, then defenders of HCE could discover how some microbiomes reproduce their types more, as well as how some host species reproduce their variants more than other host species, and thus show how holobionts are units of selection from a MLS1 perspective. Why from a MLS1 perspective and not from a MLS2 perspective? Because these functional groups can increase independently of causing the formations of holobiont parent/offspring lineages. Each host might produce offspring that are not required to acquire the same functional microbiome as their parents. The key element for MLS1 evolution is that certain variants of the host species increase over time, and that certain variants in the functional

microbiome increase over time, *due to the existence of holobionts*[157] (**Figure 28**). If phylosymbiosis were studied this way, then it could serve as a proxy for MLS1 holobiont selection.

What about community heritability, H^2_c? Remember that H^2_c measures how the degree of differences in the genetic makeup of hosts in a population determines the structure of their microbiomes, or whole-community phenotype. This is sharp contrast with traditional heritability measures h^2, which measure how hosts with different genetic backgrounds might bear different microbial species in their microbiome (**chapter I, section 3.4.3**; **Figure 10**). In this vein, H^2_c can represent three possible states: host control of microbiome assembly; host susceptibility to microbiome 'infection'; or a combination of host control and host susceptibility. Why and how is this relevant for MLS1? Because it could be used to measure the probability that the SoT in the microbiome and/or in the hosts would be preserved, even in the abscense of parent/offspring lienages of holobionts. Suppose that we have a population of hosts H_1, H_2, H_3, with an associated microbiome composition m_1, m_2, m_3, and such that H_3 is the genetic variant with more control over microbiome composition, and with a higher susceptibility to infection, and such that m_2, is the variant that increases more the survival of the host it associates with. Initially, by assumption, m_2, appears in H_2, but then it can spread in the population of hosts. Assuming that H_3 is also the variant with higher rate of survival (assuming that the host do not interaction with the microbiome), then what is expected from a MLS1 perspective is that H_3 will increase its presence in the population of hosts. But, additionally, H_3 will tend to create an association with the functional microbiome m_2 that, because of that, would increase its presence in the population. Notice that the increase of both types is not necessarily a consequence of H_3-m_2 individuals forming parent-offspring lineages. On the contrary, what is expected to happen is that the benefit of each 'type' will spread across the population, and the two fitter types will fuse with each other in the long term. As, *ex hypothesi*, H_3 is the fitter type and the one with higher degree of H^2_c, natural selection will tend to favour the formation of H_3-m_2 individuals in the long term, other things being equal. If this happens, this would be a consequence of the action of MLS1 on holobionts.

[157] It is thus a question of causality, and thus selection, rather than mere sorting.

If, on the contrary, this does not happen, then holobiont selection is not occurring or it is too weak in relation to selection at lower levels.[158]

In conclusion, both phylosymbiosis and H^2_c can be conceptualized as proxies for detecting selection on holobionts of MLS1 type, and thus they become two relevant concepts to argue that holobionts are units of selection.

4. The holobiont as a unit of selection from a MLS2 perspective: some suggestions

So far, I have argued that holobiont evolution can be explained in terms of MLS1 by appealing to the concept of SoT as a way of developing the concept of the holobiont that I introduced in **chapter IV**. However, a question now arises: is it also possible to apply MLS2 by appealing to the functional definition of holobionts that I put forward? In other words, is there a way of finding a consistent pattern of parent/offspring *regression* among functional holobionts to argue that holobionts can be units of selection from a MLS2 perspective? And, if so, how would this pattern look like? In this section, I will suggest that there are some mechanisms that could account for the existence of parent/offspring regression among holobionts *conceived as functional units*, i.e. conceived as units formed by a host plus the traits endoded in its microbiome, rather than by a collection of genomes. In that vein, it is coherent to think that at least some holobionts can be units of selection from a MLS2 perspective. Or, in other words, I will argue that there are reasons to think that it is plausible to find some biological mechanisms that guarantee that parent-holobiont/offspring-holobiont functional similarity is higher than the functional similarity between random holobionts in the same population (**Figure 29**).

[158] Of course, my account here assumes that host-microbiome relations are linear, and different effects might arise if, in fact, there is some non-linearity going on. As my case was only to illustrate a conceptual possibility, rather than to develop a mathematical model that takes these complexities into account, I will not consider these cases here. Thanks to Caglar Karaça for making this point to me.

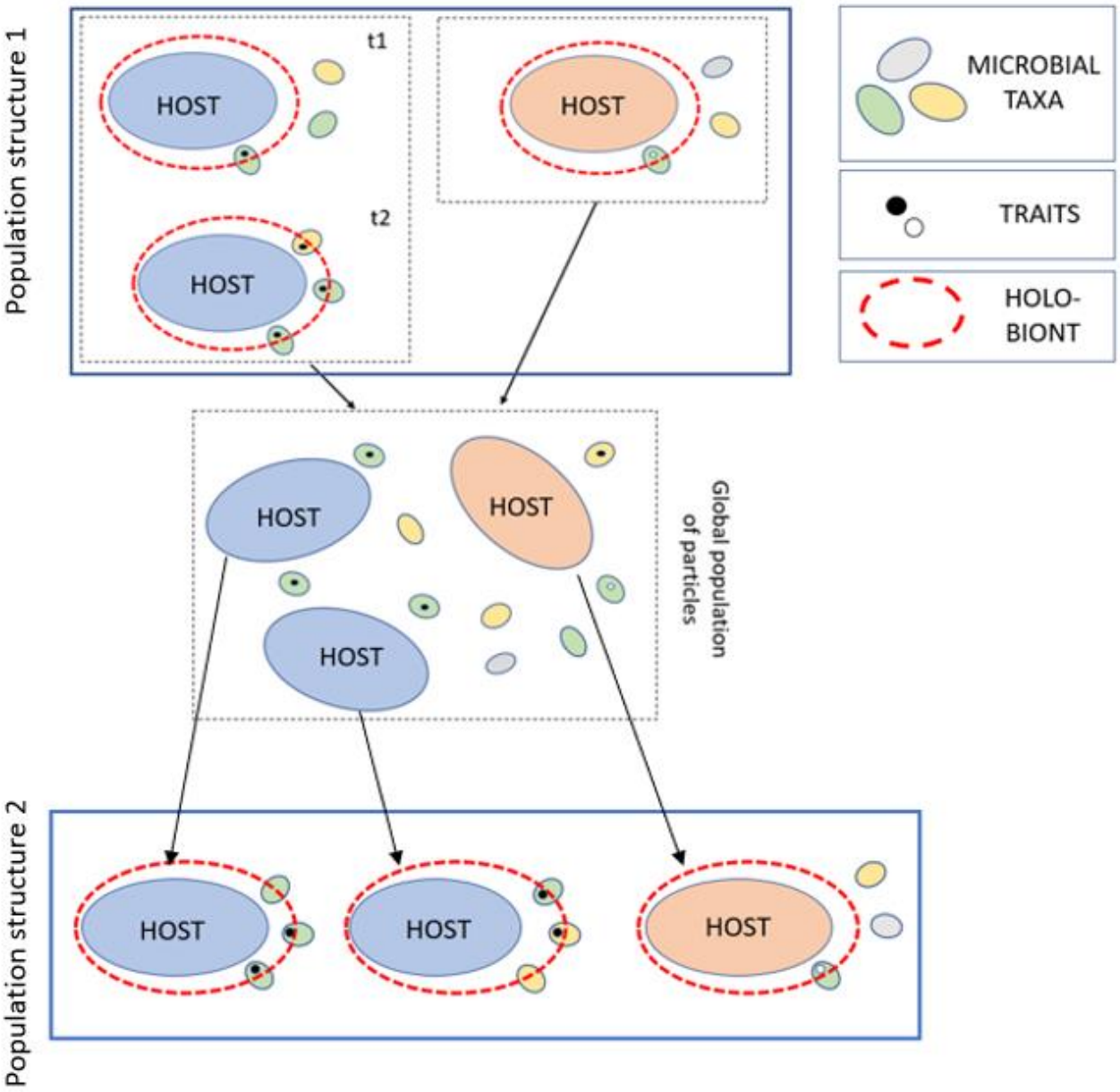


Figure 29. The holobiont as a unit of selection from a MLS2 perspective, according to the SoT account.

These mechanisms constitute a form of extended inheritance, i.e. they are non-Mendelian mechanisms through which inheritance can be channelled leading to parent/offspring regression (Jablonka and Lamb 2005, for the concept of 'extended inheritance'). It goes without saying that these mechanisms do not necessarily guarantee the same degree of parent/offspring similarity as conventional Mendelian mechanisms. But notice that this is not problematic, for as I argued in **chapter III**, the first inheritance mechanisms *must have been* very unrealiable. Furthermore, epistemologically, the process view (plus the

non-nested nature of the hierarchy) suggests that the discovery of inheritance mechanisms is always posterior to the discovery of the patterns of similarity. It is now time to see how the patterns of similarity that are observed in natural populations, and that I reviewed in **chapter IV**, could be transgenerationally realized through parent/offspring lineages. To do so, I will first review the evidence that I consider relevant to make my case, and I will later introduce a conceptual argument that makes use of that evidence to support the consideration of holobionts as units of selection from a MLS2 perspective.

4.1. Hosts as 'niche constructors', the role of the host in the acquisition of the microbiome, and the role of the microbiome as 'hot spots' for horizontal gene transfer

The type of evidence I review in this section has not been so far related in any consistent manner to the possibility of MLS2 selection among holobionts. I will only introduce it here as a basis to support the conceptual framework for MLS2 selection among holobionts that I will introduce in the next section. Notice that the argument I will introduce will be about the *conceptual possibility* of holobionts being units of selection in virtue of the existence of some biological mechanisms that could guarantee a certain degree of transgenerational preservation, i.e. in virtue of the existence of some mechanisms that support the existence of host-functional microbiome parent/offspring regression. It will therefore be conceptual, and it should be taken as a suggestion about what biologists should look for to investigate the degree of parent/offspring similarity among holobionts.

The first mechanism I will mention is the role of hosts as niche constructors (Oddling-Smee et al. 2003; Laland et al. 2016). Concretely, about the role that hosts play in spreading their microbiome in their environment. The holobiont is known to be an 'open system', where microorganisms can enter and leave. In a very interesting paper entitled 'Transmission of the gut microbiome: spreading of health' (see also the references in there), Browne et al. (2017) suggest the existence of several mechanisms that guarantee the survival of different elements of the micobiome outside their hosts. These microorganisms can use different reservoirs, and can be in a sort of 'latent'

state until they can find a new host to infect. Interestingly, hosts that live longer will have more opportunities of spreading their 'health' to the environment, and thus the density of the microbes that increase their fitness could increase. Notice that the increase in the microbial density does not need to be an increase in the density of the microbial species, but an increase in the microbes that bear the appropriate types of traits, i.e. those that have been selected due to their effects in the fitness of the host (MLS1).

Second, and not less important for the way in which this phenomenon can be related to the former as a mechanism that supports the existence of MLS2 among holobionts, some current evidence suggests that the host is partially in control of the acquisition of its microbiome (again, H^2_c). For instance, Goodrich et al. (2014, 2016a) and Turpin et al (2016) have recently discovered that the acquisition of the microbiota might be sometimes regulated by the influence of the genetics of the host, so that part of the variation in microbiome composition might be dictated by genetic factors. Second, the immunological system of the host has been hypothesized, and empirically proven, as being one of the main sources that shapes the microbiome of the holobiont, i.e. the immune system plays a basic role in regulating the microorganisms that will be part of the host's microbiota, as well as those that will be excluded (Rakoff-Nahoum et al. 2004; Cullen et al. 2015).

Finally, as I already explained in **section 3.2**, some recent empirical data demonstrates the existence of a high level of functional redundancy (or 'degeneracy') in the microbiome (Boon et al. 2014; Vieira-Silva et al 2016), which suggests that a holobiont with the same traits might be biologically possible without any single species of the original microbiome reappearing in the next generation. In addition, our current evidence also indicates that the level of HGT among microorganisms is higher between those microorganisms that compose a host's microbiome, than between free-living microorganisms (Liu et al. 2012; Lerner et al. 2017). HGT is well known for being the main mechanism of genetic exchange among bacteria, being one of the main mechanisms involved in bacterial adaptation and acquisition of new *functional* capabilities. Thus, if one species spends a longer time in a host's microbiota,

the likelihood that it will acquire the functional factors responsible for the reconstruction of the traits of the holobiont will in principle increase. This would reinforce the existence of a functional redundancy in the microbiome as well as decrease the necessity of the host association to concrete bacterial species to acquire the traits that are being selected for, and this could even increase exponientally, if the effects of HGT are noticeable.

4.2. On the possibility of holobiont lineages. A conceptual suggestion for MLS2 selection among holobionts

Based on the previous evidence, I will now argue that there are modes of 'extended inheritance' that guarantee that the probability that the offspring of a holobiont that bears the traits that have been selected for in its population *P* also bears those traits is higher than the probability that a random member of *P* bears those traits, other things being equal.[159] The argument I will present here does not aim to be exhaustive. I will only suggest there are reasons to think that these mechanisms can lead to a situation in which holobionts are units of selection from a MLS2 perspective, and that thus this possibility should not be discarded *a priori*. In addition to that, the argument only aims to show how SoT would account for the role of holobionts as units of selection from a MLS2 perspective, and how it is different from other conventional approaches to holobiont evolution that require genotype-to-genotype transmission, including the work by Doolittle and Booth (2017), and by Roughgarden et al. (2017).[160]

---

[159] Notice that, in agreement with the holobiont account presented in **chapter IV**, holobiont populations are identified as host populations since, as I argued there, the proper trajectory of evolution of hosts is *qua* holobionts. Furthermore, the problem will be how to discover that the transmission of the functional microbiome acroos host generations is reliable enough, so that the same 'functional holobionts' are transgenerationally formed. Thus, the discussion will be about the mechanisms of microbiome transmission, and I will take for granted that the mechanisms of host transmission are well known (Mendelian inheritance).

[160] Roughgarden et al. (2017), Lloyd (2017b), and Lloyd and Wade (2019)'s models of holobiont evolution are, in reality, MLS1 models, as they are grounded on Drawn et al. (2013) and Drawn and Wade (2014). But they assume that SoS holds (or can hold), and do not make the conceptual distinction between *traits* and *species* (and, consequently, between *cospeciation* and *coevolution*) that I advocate here, and thus they have still failed to make the step forward to characterize holobiont evolution in terms of MLS2 (see also Suárez 2018c). In addition to that, their models aim to explain, among other things, how different types of microbiome transmission could evolve, so that the transition from horizontal to vertical transmission becomes expectable, for example. Notice that none of this is required in my model, as I presuppose that traits will travel faster than the species that bear them, and thus the evolution of species-transmission modes might be highly irrelevant for holobiont evolution, and might be indeed not expected in many cases (namely, those where the trait can be disconnected from the bacterial genome more easily than vertical modes of transmission can evolve) (check later in this section). Thanks to the fantastic feedback I have received from Lisa Lloyd, Seth Bordenstein, Ehud Lamm, Ford

The argument will thus be grounded on a functional conception of the holobiont, i.e. it will be grounded on a view of the holobiont according to which the microbiome composition is functionally determined, and thus what becomes relevant is how SoT is transgenerationally preserved from parent-holobiont to offspring-holobiont.

The recruitment of the microbiome across different generations of the holobiont to guarantee SoT might take several forms. On the one hand, it can occur through direct transmission, as in cases of vertical transmission or transmission during birth (in mammals); on the other hand, the recruitment might occur through environmental acquisition, especially through parental feeding, daily contact with the family members, social interaction with other members of the group, or diet (**Figure 30**).

Doolittle, and Pierrick Bourrat, who very recently helped me see the key difference between our models, and to illuminate how the SoT model, properly characterized, was an original model that went somehow beyond theirs.

		Increases the probability of transgenerational SoT	Increases the probability of transgenerational SoS
DIRECT TRANSMISSION	Vertical parent-offspring transmission	YES	YES
	Transmission at birth (birth channel)	YES	SOMETIMES
ENVIRONMENTAL ACQUISITION	Parental feeding (coprophagy, breast milk)	YES	NO
	Daily contact	YES	NO
	Social interaction	YES, UNDER CERTAIN CONDITIONS*	NO
	Diet		NO

Figure 30. Possible scenarios under which SoT would lead to a higher parent-offspring similarity with respect to a trait T_a than the similarity of the offspring to an average member of the population for the same trait, and comparison with SoS. See text for the discussion of the cases of social interaction and diet.

The cases of direct transmission pose no problem to explain the routes how traits might get transgenerationally transmitted, as in most cases that transmission would occur simply by species transmission (thus, also satisfying SoS). For instance, in the species where there is a certain form of vertical transmission across host generations, like the cases of the organelles in eukaryotic cells, *B. aphidicola* among aphids, *Wolbachia* among flies, etc., vertical transmission constitutes the main source of recruitment of the components of the microbiome that guarantee trait-recurrence. Second, many of those elements will be transmitted either during birth, or through the vaginal channel (Funkhouser and Bordenstein 2013; Bäckhed et al. 2015; Chu et al. 2017). In these two cases, especially in the case of vertical transmission, it is likely that the same species will be transgenerationally acquired, thus both SoS

and SoT would be satisfied. However, notice that even if in any of these cases SoS were not satisfied, the possibility that SoT is satisfied would be higher, and thus the probability of parent/offspring regression would be higher. How? Because the high degree of functional redundancy in the microbiome, together with the abundance of HGT, would increase the probability that the same functional microbiome will be passed on, despite not passing on the same taxonomic microbiome. As what needs to be reconstructed transgenerationally in MLS2 is the functional holobiont, functional redundancy and HGT will be enough to guarantee that, through direct contact, the same functional microbiome will be passed on to the offspring with enough fidelity.

The cases of environmental acquisition present their own problems, though, because our current evidence suggests that microbiome species composition is transgenerationally disrupted (thus, SoS is not satisfied). However, as in the case of direct transmission, our current evidence also suggests that the traits might be transgenerational-ly preserved. The risk of horizontal acquisition, however, is that it might disrupt parent/offspring regression, and thus it could disrupt MLS2 selection. That is to say, that the 'benefits' of a 'fitter' functional microbiome could be shared among all the members of *P*, and not only between the offpring of the fitter holobiont.[161] Even while this is true, I will argue that our current evidence can be interpreted suggesting that the similarity of traits is in most cases higher between holobiont parent/offspring generations than it is with any other random member of the population, thus allowing a MLS2 interpretation. Let us see how.

Environmental acquisition comes in a spectrum that involves many different forms of microbe acquisition: firstly, direct acquisition of the microbiota from the mother through feeding, as it happens in case of coprophagy, a standard phenomenon of microbiome acquisition among many animals (koala-bears, rabbits, elephants, hippos, iguanas, etc.; see e.g. Osawa et al. (1993); Kovacks et al. (2006)), or in the case of brest feeding (Funkhouser and Bordenstein 2013). Secondly, by direct daily contact (e.g. acquisition of the

[161] Remember that MLS2 selection requires the existence of collective parent/offspring lineages, that is, the existence of a form of inheritance.

microbiome as a consequence of the relationship with close members of the family; see e.g. Song et al. (2013); Nayfach et al. (2016)). Thirdly, through social interaction with other members of the population (see e.g. Tung et al. (2015); Moeller et al. (2016)). Fourthly, through diet (see e.g. Leff and Fierer (2013); Lang et al. (2014)). Drawing on the mehcanisms I mentioned in **section 4.1**, I will now suggest that MLS2 selection is possible, and plausible, if what is required is transgenerational host-functional microbiome matching through SoT.[162] The argument I introduce here is more powerful for the situations of parental feeding and daily contacts with the family members, where the likelihood that the hosts acquire microbial components that lead to the acquisition of the same traits as their parents is increased for the duration of the contact. However, it also works in the situations where the microbial composition is driven by diet or by social interaction with other members of the population.

Let us imagine a situation where, by random acquisition of certain microbial members of the environment (M_1, M_2), a holobiont *X* in a population *P* acquired a trait *Ta* that increases its fitness with respect to the fitness of other members of *P*. As *Ta* increases the fitness of *X*, it will be expected that: (1) *X* survives longer than an average member of *P*; and (2) *X* produces more offspring than an average member of *P*.[163] As a consequence of (1), it is expected that the relative population of M_1 and M_2 in the bacterial pool will increase, due to the possibility of reproducing within *X* and getting transferred to the environment through *X*'s depositions; second, it is expected that M_1 and M_2 will engage in more relations of HGT with other members of *X*'s microbiome than the average member of M_1 and M_2 living non-symbiotically. Therefore, it is expected that the relative numbers of M_1 and M_2 in the environmental bacterial pool increases and that new members of the bacterial pool (M_3, M_4) will acquire,

[162] Notice that, in contrast with e.g. Rosenberg and Zilber-Rosenberg (2016, 2018), the account I will present here depends on a functional (SoT), and not a structural (SoS) conception of the holobiont.

[163] I am conscious that an immediate objection would be that natural selection does not predict what could happen to *a particular individual*, but what would happen to *the trait distributions in a population*, provided that there's heritable variance in fitness (e.g. Sober and Lewontin 1982; Sober 1984). However, the problem with holobionts is that we would need to prove, first, that there's *inheritance*. In other words, we need to prove first that the mechanisms aforementioned lead to parent/offspring regression. So, the argument is just an idealization to prove that this could be so, regarding our current evidence.

through HGT, the trait *Ta*, and thus the capacity of transmitting *Ta* to other holobionts in *P* as an effect of niche construction. If this is so, then the probability of *Ta* reappearing in future generations of holobionts in *P*, provided that new hosts can acquire the microorganisms that bear *Ta*, will increase and, what's more important, will increase as a consequence of *Ta* being naturally selected for in the holobiont, rather than in any of the M_1-M_4 microorganisms that bear the trait.

However, this situation will increase the probability of *Ta* reappearing *for any member of* P, not necessarily for *X*'s offspring and, therefore, as I argued in **section 3.2**, it would be a case of MLS1. Is there a possibility of making this a case for MLS2? Or, in other words, is there some mechanism to guarantee that the benefits of *Ta* are not shared among all the members *P*, but restricted to the host/holobiont lineage? I think that there are reasons to believe that this might be so. Let me mention three mechanisms and how these could determine the existence of MLS2, despite the environmental acquisition of the microbiome.

1. *Influence of the host genetics in the acquisition of the functional microbiome*. If this is the case, it will be expected that the offspring of *X*—as it is genetically related to *X*—will probably acquire M_1 and M_2, or any of its functional equivalent variants, more easily than any random member of *P*.
2. *Influence of the host immunological properties in the acquisition of the functional microbiome*. If this is the case, it will be expected that the offspring of *X*, as it will probably have similar immunological features to *X*—due to the genetic relationship between the hosts—will probably acquire M_1 and M_2 more easily than any random member of *P*.
3. *Host contact with an environment where* M_1 *and* M_2*, or any of their functional equivalent variants are present*. If this is the case, the expectation for the offspring of *X* to be more likely to acquire M_1 and M_2 will depend on the breeding conditions of the host that composes *X*.

3.a. If the host nurtures and raises its offspring for a long time, the probability of its offspring acquiring M_1, M_2 or any other element of its microbiome that causally brings about *Ta* will be greater than the probability of any random member of *P*.

3.b. Otherwise, the probability of *X*'s offspring acquiring *Ta* to be similar to the probability of any random member of *P*, and thus MLS2 selection will be hampered.

The argument presented here is purely conceptual and merely illustrative of how some currently known biological mechanisms could lead to MLS2 among holobionts. However, these types of conceptual framework have some clear limitations, as under a closer mathematical scrutiny they might prove mistaken. And, furthermore, even these mathematical models would need a close empirical scrutiny, because what might seem caonceptually and/or mathematically plausible, might be biologically impossible.

For now, the only population biologists who, to my knowledge, have tried to model a similar scenario have been Devin M. Drawn and Michael J. Wade (Drawn et al. 2013; Drawn and Wade 2014).[164] In Drawn et al. (2013), the authors analysed how a case of an obligate symbiont that is horizontally acquired can evolve into a case of vertical transmission and concluded that epistasis for fitness between host and fitness genes was a precondition for this to happen. A limitation of their model, though, was that it did not include the possibility of HGT guaranteeing that the epistasis could be transgenerationally maintained despite the lack of vertical transmission. Or, in other words, they did not consider a possible scenario in which the existence of an epistasis between symbiont genes and host genes could lead to the HGT of the symbiont genes to other symbionts of different taxa, so that the epistatic relationship would be

[164] Since I wrote this section, a new paper has appeared that modelled the same scenario, Lloyd and Wade (2019). Interestingly, their paper put the emphasis on what I have been putting the emphasis all this chapter, namely, that what matters to decide whether holobionts are units of selection is not species coevolution, but coevolution of traits. They do not use, however the evidence that I use here to argue that coevolution of traits can be disconnected from coevolution of species for the taxa that compose the host microbiome. In fact, their examples presuppose that the species are transgenerationally preserved, which I think is arguably sufficient for holobiont evolution, but it is definitely not necessary.

transgenerationally maintained despite the lack of the vertical transmission of the different taxa—that is to say, their research depended on the assumption of SoS.[165] In Drawn and Wade (2014), they proved that the genes adapting to a biotic heritable environment (such as the case of epistasis) can make the selection process self-accelerating. In other words, they prove that selection on one gene that stays in epistatic relationships with a set of genes in its environment—i.e. genes to which the first one is not physically bounded in the same genome—leads to an increase in the genetic environment where the gene is selected for. If their model were applicable the case of hosts and their functional microbiomes, it could show that, indeed, the possible generation of extra-genomic regulatory networks can lead to MLS2 selection among holobionts, as selection for some hosts would lead to selection for the functional microbiomes where these hosts are favoured.[166]

On the empirical side, experiments do not yet exist, since, to my knowledge, no experiment has been able to demonstrate the existence of MLS2 for certain functional microbiomes associated to a host. Nonetheless, as it happens in biology in general, if some cases of this were discovered, their value would be discretional. That is to say, the discoveries would be valid for the focus organisms that biologists are studying, but nothing would guarantee that they will necessarily apply to any other organism. In any case, the evidence of 'extended inheritance' mechanisms I have reviewed here should serve to suggest that holobiont selection could be glossed in MLS2 terms for some organisms, and that this possibility should not be excluded simpy *a priori*.

[165] I have been asked a few times, when I introduce HGT, how HGT can guarantee that the symbiont is transferred to the right host, so that it is possible to detect parent-offspring lineages, and thus MLS2. The key, of course, is that if 'the right symbiont' means the right taxa, then HGT does not guarantee it. But notice that the SoT model requires that *any of the taxa that bears the trait is transferred to the right host*. In this case, it becomes quite evident why HGT plays a key role in guaranteeing that this is so.

[166] Notice that their model, if extended to host-microbiome interactions, could be used to conceptualize Burcker and Bordenstein´s experiment on the *Nasonia* wasps. See **chapter 1, section 3.4.4**.

5. Differences between the stability of traits account and Doolittle and Booth's 'song not singer' model

To date, the only account of the role of the holobiont as a unit of selection in functional terms has been presented by Ford Doolittle and his collaborators, in three different papers: Doolittle (2017), Doolittle and Booth (2017), and Doolittle and Inkpen (2018). Doolittle's account is presented under the explicit mission of not 'throwing out the baby with the bathwater' (Doolittle and Booth 2017: 10), i.e. to save what is conceptually valuable in HCE, while getting rid of the main postulates of the hypothesis. In their view, the holobiont is the biological entity that *performs* a set of biological processes, including metabolic, immunologically, or developmental processes, among others, in virtue of the existence of *networks* of functional genes whose interactions cause these processes. These processes have a high degree of transgenerational stability and can result from the interaction of different lineages (or taxa) of bacterial species, provided that the lineages that interact transgenerationally can carry out the same biological function as their predecessors—or, in other words, provided that the interacting taxa bring the same functional genes, thus giving rise to the creation of the same networks (**chapter I, section 5**; **Figure 13**). I think Doolittle's account of the holobiont captures part of the intuition that is behind the proposal of HCE. However, I suspect his proposal does, in fact, through out the baby with the bathwater. Let me explain why.

As I argued several times, HCE was a proposed as a hypothesis to explain why hosts, or macrobes, bear certain phenotypic traits whose maintainance was unexplainable if it had to be accounted for solely in terms of the genetics of the host. This evidence included the ability of corals to get rid of the infection caused by *V. shiloi*, the mating preferences in *D. melanogaster*, the hybrid incompatibility in *Nasonia* wasps, etc. Each of these phenomena constitutes an evolved host phenotype that results, allegedly, from the interaction between the host and its functional microbiome. I introduced the hypothesis of the non-nested nature of the biological hierarchy with the aim of explaining each of these cases. Doolittle's song/singer account, on the contrary, loses track of the significance of each of these evolved states *for the host*,

because in his view, what matters are the ‘songs’, or patterns of interaction that somehow ‘emerge’, rather than their ‘singers’. That is to say, what matters is how certain genetic networks are formed so that some patterns of interaction (metabolic, structural, developmental) obtain, *no matter the lineages where they obtain*. But, obviously, if the lineages do not matter *at all*, then the host lineage does not matter *at all* either. The only thing that is relevant and should be studied is, say, the metabolic pattern, the developmental pattern, the structural pattern, etc. And each of these patterns will obviously be multiply realizable in many different types of hosts. In other words, the host where the pattern of interaction obtains is as irrelevant as each of the bacterial species is.[167]

I recognise that, seen in this in this way, Doolittle’s account has some merit. Of course, there is a scientifically legitimate question about how certain patterns of interaction evolve. However, the account is also problematic. First, it is not clear what these patterns might be, or how they could be individualized. He mentions general patterns such as homeostatic processes and biogeochemical cycles, and even goes on to extend his idea to Gaia (Doolittle 2018). I think in doing so Dollittle is substituting an obscure concept (HCE, when the hypothesis is conceptualized as asserting that the whole set of species of the microbiome gets transmitted) for one that is even more obscure. Interaction patterns, as an abtract concept that is not linked to a particular recurrent structure, are impossible to detect, since we lack every kind of diagnostic criteria for identifying them.[168] Second, but related, I cannot see in any way how Doolittle’s account could avoid ‘throwing out the baby’, to use his expression. What would be the exact role of the host species in his story? If the aim is to explain the evolution of some patterns regardless of the lineages that create them, then no feature of the host lineage could be explained. It seems to me that this is really throwing out the baby, as this is, in my opinion, the key discovery of HCE, namely, that the microorganisms that interact with a host

[167] A very recent and long discussion both by email and in ISHPSSB in Oslo with Ehud Lamm made me realize that the key distinction between SoT and Doolittle’s account can be fleshed out by saying that my account of the holobiont assumes that the holobiont includes the host singer plus the songs that are sung in the microbiome. I think that analogy is really good to capture my intituion, and I would really like to thank Ehud for suggesting it to me. Furthermore, I would also like to thank Ford Doolittle for his generosity, as he has also dedicated some time to discuss my ideas and the differences between our accounts in ISHPSSB.

[168] As I explained before, a diagnostic criterion is a basic identification method of the form ‘the entity whch is in <*x, y*> at time *t*’.

have more influence on its evolution than a random non-biotic environmental factor. Discovering how this influence actually occurs is the key to HCE, and it would be the right way of 'elucidating' the meaning of the hypothesis. The song/singer account, unfortunately, is far from doing so.

Finally, there is another substantial problem with Doolittle's account, which I anticipated in **chapter I, section 5.1**. The song/singer model seems to me completely based on Dawkins/Hull's interactor/replicator model of the units of selection. In fact, Doolittle explicitly recognizes that his is an account of how holobionts could be interactors in virtue of certain sets of regulatory genetic networks being their replicators. At this point, it must be clear to the reader why the interactor/replicator framework is inadequate to capture the essence of the units of selection debate, so I will not repeat the arguments here (**chapter III, section 3.3**). In any case, his account has one virtue. If, indeed, if it's true that these regulatory networks exist and work as replicators, then it would be a strong proof to support the claim that holobionts *have been* units of selection. Nonetheless, as I argued, a MLS framework would be preferable, since it would allow us to disentangle the properties that the units that we call holobiont have, both as partially 'cohesive units' (MLS1), and as units that can get their properties 'reproduced' transgenerationally (MLS2). Assuming, on the contrary, that we should only focus on highly evolved states commits the mistake of forgetting about the diachronic dimension of the units of selection debate.

6. Brief summary of chapter V: The holobiont is a unit of selection

In this chapter, I have articulated my account of the role of the holobiont as a unit of selection, which was my way of 'uploading' HCE, so that it scores the empirical evidence. To do so, I have departed from the non-nested nature of the biological hierarchy and argued that it could be applied to holobionts via a MLS approach. After that, I have argued why the criticism to HCE based on the notion of SoS does not work. First, I have argued that it reduces discussions about holobiont evolution to discussion about the replicator. Second, I have argued that reducing the debate about the units of selection that way is

illegitimate, as it puts the cart before the horses. Therefore, I have argued, SoS must not be the condition that holobionts need to satify to be considered units of selection from a MLS approach. Drawing upon my criticism of SoS, I have elaborated a MLS1 account of holobiont evolution where, I argued, the ‘particle types’ that get multiplied and that increase their presence in future generations are either host variants, or the traits that are encoded by the microbiome. By contrast with SoS, I called my approach SoT, for in my account what needs to be counted is how traits increase their presence, rather than how microbial species increase their frequency. Secondly, I have presented some notes to suggest how a MLS2 approach would look, and I have contrasted MLS1 and MLS2, to show what the difference between the twe approaches is. Finally, I have distinguished between my account and Doolittle and Booth’s ‘it’s the song, not the singer’ account.

Global Summary

‘The holobiont as a unit of selection. A non-nested metaphysics for the biological hierarchy’

1. General reflection and concluding remarks

This thesis was about the hologenome concept of evolution. I started making clear that HCE was the project of understanding the features that made holobionts biological individuals, *and* units of selection. Was there a possible argument to support the claim that the symbiotic assemblages composed by a macrobe (host) and a collection of microbes (symbionts) are biological individuals in a significant sense? That is, was there a way of fleshing out the biological properties of that specific type of communities, apart from the very basic ‘diagnostic criteria’? And, if so, would this entail a change in some of the basic metaphysical commitments generally assumed by evolutionary biologists and philosophers of biology? The response to these questions has been driven by a sort of ‘processual’ view of the biological world—widely inspired by the recent work of John Dupré and the rest of the members of his ERC project: ‘A process ontology for contemporary biology’—which has been combined with a particular view of the topic of scientific explanation—as this has been developed in my research group in Barcelona, under the financial support of the MINECO project: ‘Laws, explanation, and realism in the physical and biomedical sicences’. More importantly, these two views have been embedded within the wider framework of John Dupré’s ‘promiscuous realism’, the view that scientific classification is always done by scientists (*human beings*) in a goal-oriented

fashion, and those classifications that are chosen are the ones that serve the purposes of the scientists—in Spanish, promiscuous realism could be presented by summarizing it under the (in)famous motto: '*vale quien [lo que] sirve*'.[169]

With this framework in mind, the project has achieved a double goal. First, it has diagnosed the ontological committments that underlay the rejection of HCE. At this stage I want to make a point clear. I have never said—and would never dare to do so—that those scientists that rejected HCE were not doing science, or were not doing it properly, or rigurously. I indeed think they were making fair points and criticisms to some ways of understanding HCE that were empirically questionable. They were behaving as scientists behave, and how scientists deal with scientific controversies. Given a scientific controversy, what does the empirical evidence say? But, importantly, I applied the same charity principle to those scientists defending HCE. What is there of value in the scientific evidence that they have been collecting? How is it best to represent it? However, the thesis is *a philosophy thesis*; that is, it has a philosophical nature. Thus, my purpose was neither limited to certify that there was a disagreement—that is rather obvious—nor to scour the evidence to see which of the sides in the controversy was, so to speak, 'winning'. That would have been a valuable goal but, in my view, not a philosophical goal. My purpose was rather to investigate which were the metaphysical (conceptual) assumptions that underlay the disagreement. It is now a good moment to restate Samir Okasha's advice, that I used as a guide to open my doctoral project.

> 'Obviously, empirical data is crucial for resolving the levels-of-selection question, as for all scientific questions; but conceptual clarity is a prerequisite too. Unless we can agree on what it means for there to be selection at a given hierarchical level, on what the criteria for individuating "levels" are, on whether selection at one level can ever be "reduced" to selection at another, on how multi-level selection should be modelled, and

[169] 'Purposes' should not be understood in a phychological or sociological way. According to Dupré, the classifications are choosen if they serve the *epistemological* purposes of the scientists, that is, if they achieve the epistemic goals that the scientist aims to achive (tracing phylogenies, discriminating flavours, etc.).

> on whether there is always “one true fact” about the level(s) at which selection is acting, then there is little prospect of empirical resolution, however much data we collect. Focusing on conceptual questions such as these is not meant to downplay the significance of empirical data, but rather to help provide the clarification needed for addressing the issues empirically.’ (2006: 2)

The whole project has been inspired by that dictum plus a key idea about the nature of philosophical inquiry that I learnt during my B.A. degree at University of Oviedo, namely, that every view on a concrete scientific matter will necessarily be embedded in a metaphysical framework, in a particular way of seeing the ‘deep structure’ of the world, even when this is (most times) not explicit. In this thesis, I claimed to have found one of these commitments in the current debate about the status of the holobiont as a unit of selection. The guiding commitment that both sides followed was a commitment to the nested nature of the biological hierarchy, followed by a commitment to a view of the process of natural selection that depended on nestedness. That commitment grounded why both defenders and detractors of HCE were assuming that a very strict condition about species co-transmission needed to be established to prove HCE. The thesis has shown that this is a metaphysical commitment, and that precisely some of the evidence gathered by HCE defenders supported a different version of HCE than the one they assumed. Namely, it supported a version of HCE in which the host and its functional microbiome coevolve, rather than one in which the host and its microbiome cospeciate. In other words, what coevolves is the family of *functional traits* that the holobiont expresses, and that could have their origin either in the host genome, or in its microbiome. Recognizing that this is so, while at the same time accepting the existence of a biological hierarchy, plus the requirement of studying each level of the hierarchy according to the processes that entities at that level engage in, lead me to propose the non-nested nature of the biological hierarchy.

Secondly, I have explored the consequences of a non-nested hierarchy to the debate about the units of selection and have provided a multilevel selection account of the role of the holobiont as a unit of selection. I have

argued that the empirical evidence strongly supports the claim that holobionts are units of selection from a multilevel selection 1 perspective, provided the microbiome is adequately individuated (i.e. functionally, not taxonomically). Also, I have suggested a possible framework to conceive holobionts as units of selection from a multilevel selection 2 perspective. The latter framework, however, still needs more development, so I have only been able to sketch the main ideas, to argue what biologists should look for if they want to prove the existence of a form of extended inheritance among holobionts.

To conclude, my doctoral project has served four key purposes. First, to uncover a 'hidden' metaphysical assumption that could mask what could be worth developing in HCE. Second, to propose an alternative framework to save what was valuable from HCE without renouncing to what is also valuable from alternative perspectives. Third, to propose an alternative framework to concive the role of the holobiont as a unit of selection. Fourth, to prove the usefulness of philosophical reflection in enriching and making clearer some current scientific debates.

Bibliography

- Adair KL & AE Douglas (2017) Making a microbiome: the many determinants of host-associated microbial community composition. *Curr opin microbiol* 35: 23-29.
- Ariew A & RC Lewontin (2004) The confusions of fitness. *Br J Philos Sci* 55: 347-363.
- Bäckhed F, J Roswall, et al. (2015) Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host Microbe* 17: 690–703.
- Bapteste E & J Dupré (2013) Towards a processual microbial ontology. *Bio Philos* 28 (2): 379-404.
- Bapteste E, MA O'Malley, et al. (2009) Prokaryotic evolution and the tree of life are two different things. *Biol Direct* 4: 34.
- Bennett GM & NA Moran (2015) Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *PNAS USA* 112: 10169-10176.
- Birch J (2012) The negative view of natural selection. *Stud Hist Philos Sci C* 43(2): 569-573.
- Blow F, A Gioti, et al. (2019) Functional genomics of a symbiotic community: shared traits in the olive fruit fly gut microbiota. Cold Spring Harbor Laboratory. https://doi.org/10.1101/590489.
- Boon E, CJ Meehan, et al. (2014) Interactions in the microbiome: communities of organisms and communities of genes. *FEMS Microbiol Rev* 38: 90–118.
- Booth A (2014) Symbiosis, selection and individuality. *Bio Philos* 29: 657-673.
- Bordenstein SR & KR Theis (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoSBiol* 13.8: e1002226.

- Bouchard F (2018) Symbiosis, Transient Biological Individuality, and Evolutionary. In *Everything Flows: Towards a Processual Philosophy of Biology*. D Nicholson & J Dupré (eds.). Oxford University Press.
- Bourrat P & PE Griffiths (2018) Multispecies individuals. *Hist Phil life sci* 40 (2): 33.
- Brandon R (1982) The levels of selection. *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association* 1(1): 315-323.
- Brandon R (1985) Adaptation Explanations: Are Adaptations for the Good of Replicators or Interactors? In *Evolution at a Crossroads: The New Biology and the New Philosophy of Science*. David J. Depew & Bruce H. Weber (eds.). The MIT Press.
- Brandon R (1988) The levels of selection: A hierarchy of interactors. In *The Role of Behaviour in Evolution*. HC Plotkin (ed.). The MIT Press.
- Brandon R & G Ramsey (2006). What's wrong with the emergentist statistical interpretation of natural selection and random drift. In *The Cambridge companion to philosophy of biology.* M Ruse & D Hull (eds.). Cambridge University Press.
- Bright M & S Bulgheresi (2010) A complex journey: transmission of microbial symbionts. *Nat Rev Microbiol* 8: 218–230.
- Brooks AW, Kohl KD, et al. (2016) Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoSBiol* 14 (11): e2000225.
- Browne HP, BA Neville, et al. (2017) Transmission of the gut microbiota: Spreading of health. *Nat Rev Microbiol* 15: 531–543.
- Brucker RM & SR Bordenstein (2012a) Speciation by Symbiosis. *Trends Ecol Evol* 27 (8): 443–451
- Brucker RM & SR Bordenstein (2012b) The roles of host evolutionary relationships (genus: Nasonia) and development in structuring microbial communities. *Evolution* 66 (2): 349-362.
- Brucker RM & SR Bordenstein (2013a). The capacious hologenome. *Zoology* 116 (5): 260-261.
- Brucker RM & SR Bordenstein (2013b) The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia. *Science*: 1240659.

- Brucker R M & SR Bordenstein (2014). Response to Comment on “The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia”. *Science* 345(6200): 1011-1011.
- Bulgarelli M, R Garrido-Oter, et al. (2015) Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host Microbe* 17(3): 392–403.
- Burke C, P Steinberg, et al. (2011) Bacterial community assembly based on functional genes rather than species. *PNAS USA* 108 (34): 14288–14293.
- Buss LW (1987) *The Evolution of Individuality*. Princeton University Press.
- Cáceres-Vázquez E & C Saborido (2018) ¿Realmente mató la bacteria al coronel? Perspectiva sistémica, causación internivélica e intervalos de cuasi-descomponibilidad en las explicaciones mecanísticas. *Theoria* 33 (1): 129-148.
- Calcott B & K Sterelny (2011) *Major transitions in evolution revisited.* The MIT Press.
- Catania F, U Krohs, et al. (2016) The hologenome concept: we need to incorporate function. *Theory Biosci* 136 (3-4): 89-98.
- Chandler J & M Turelli (2014) Comment on “The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus Nasonia”. *Science* 345(6200): 1011.
- Charbonneau M (2014) Populations without reproduction. *Philos Scie* 81 (5): 727–740.
- Chiu L & G Eberl (2016) Microorganisms as scaffolds of host individuality: an eco-immunity account of the holobiont. *Bio Philos* 31(6): 819-837.
- Chu DM, J Ma, et al. (2017) Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nat Med* 23(3): 314–326.
- Clark KA, DK Howe, et al. (2012) Selfish little circles: Transmission bias and evolution of large deletion-bearing mitochondrial DNA in *Caenorhabditis briggsae* nematodes. *PLoS ONE* 7: e41433.
- Clarke E (2010) The problem of biological individuality. *Biol Theory* 5(4):312–325.
- Clarke E (2013) The multiple realizability of biological individuals*. J Philos* CX (8): 413–435.
- Clarke E (2014) Origins of evolutionary transitions. *J Biosci* 39(2):307–317.

- Clarke E (2016) Levels of selection in biofilms: multispecies biofilms are not evolutionary individuals. *Biol Philos* 31: 191-212.
- Cullen TW, WB Schofield, et al. (2015) Gut microbiota. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* 347 (6218):170–175.
- David LA, CF Maurice, et al. (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505:559–563.
- Dawkins R (1976/2006) *The Selfish Gene*. Oxford, Oxford University Press.
- Dawkins R (1982a) *The extended phenotype*. Oxford, Oxford University Press.
- Dawkins R (1982b) Replicators and vehicles. In King's College Sociobiology Group 1982: 45-64. http://www.stephenjaygould.org/library/dawkins_replicators.html
- De Queiroz, K. (2005a) A unified concept of species and its consequences for the future of taxonomy. *Proc Cal Acad Sci* 56: 196-215.
- De Queiroz, K. (2005b) Different species concepts and their resolution. *BioEssays* 27 (12): 1263-1269.
- De Queiroz, K (2007) Species concepts and species delimitation. *Syst Biol* 56: 879-86.
- De Sousa R (2005) Biological individuality. *Croat J Philos* 14: 195-218.
- DeMonte S & PB Rainey (2014) Nascent multicellular life and the emergence of individuality. *J Biosci* 39(2):237–248.
- Derrien M & JET van Hylckama Vlieg (2015) Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends Microbiol* 23(6):354–366.
- Dethlefsen L, M McFall-Ngai, et al. (2007) An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* 449:811–818.
- Deulofeu R & J Suárez (2018) When mechanisms are not enough. The origin of eukaryotes and scientific explanation. In *Philosophy of Science*. A Christian, D Hommen, N Retzlaff & G Schurz (eds.). European Studies in Philosophy of Science, vol 9. Springer. https://doi.org/10.1007/978-3-319-72577-2_6.

- Deulofeu R, J Suárez & A Pérez-Cervera (2019) Explaining the behaviour of random ecological networks. The stability of the microbiome as a case of integrative pluralism. *Synthese*. https://doi.org/10.1007/s11229-019-02187-9.
- Díaz JS (2015) El mecanismo evolutivo de Margulis y los niveles de selección. *Contrastes* XX(1): 7–24. http://dx.doi.org/10.24310/Contrastescontrastes.v20i1.2298.
- Díez J (2014) Scientific w-explanation as ampliative, specialized embedding: A neo-Hempelian account. *Erkenntnis* 79: 1413–1443.
- Díez J & P Lorenzano (2013) Who got what wrong? Fodor and Piattelli on Darwin: guiding principles and explanatory models in natural selection. *Erkenntnis* 78(5): 1143-1175.
- DiFrisco J (2017) Kinds of biological individuals: Sortals, projectability, and selection. *Br J Philos Sci*.
- Dodd DMB (1989) Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution* 43.6: 1308-1311.
- Dominguez-Bello MG, EK Costello, et al. (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *PNAS USA* 107: 11971-11975.
- Doolittle WF (2013a) Is junk DNA bunk? A critique of ENCODE. *PNAS USA* 110(14): 5294-5300.
- Doolittle WF (2013b) Natural selection through survival alone. *Biol Philos* 29(3):415–423.
- Doolittle WF (2017) Darwinizing Gaia. *J Theor Biol* 434:11–19.
- Doolittle WF & A Booth (2017) It's the song not the singer: an exploration of holobiosis and evolutionary theory. *Biol Philos* 32:5–24.
- Doolittle, WF & SA Inkpen (2018). Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking. *PNAS USA* 115(16), 4006-4014.
- Doolittle, WF & C Sapienza (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284: 601-603.
- Douglas AE (2010) *The symbiotic habit.* Princeton University Press.
- Douglas AE & JH Werren (2016) Holes in the hologenome: why host-microbe symbioses are not holobionts. *mBio* 7(2): e02099-15.
- Drown DM & MJ Wade (2014) Runaway coevolution: adaptation to heritable and nonheritable environments. *Evolution* 68(10): 3039-3046.

- Drown DM, PC Zee, et al. (2013) Evolution of transmission mode in obligate symbionts. *Evol Ecol Res* 15:43-59.
- Dunbar HE, Wilson ACC, Ferguson NR, Moran NA (2007) Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol* 5(5): e96.
- Dupré J (1993) *The Disorder of Things*. Harvard University Press.
- Dupré J (2010) The polygenomic organism. *Sociol Rev* 58 (s1):19-30.
- Dupré J (2012) *Processes of life: Essays in the Philosophy of Biology*. Oxford, Oxford University Press.
- Dupré J (2017) The metaphysics of evolution. *Interface focus* 7:20160148. http://dx.doi.org/10.1098/rsfs.2016.0148
- Dupré, J and MA O'Malley (2007) Metagenomics and biological ontology. *Stud Hist Philos Sci C* 38: 834-846.
- Dupré J & MA O'Malley (2009) Varieties of living things: life at the intersection of lineage and metabolism. *PTPBio* 1 (December).
- Dupressoir A, C Lavialle, & T Heidemann (2012) From Ancestral Infectious Retroviruses to Bona Fide Cellular Genes: The Role of Captured Syncytins in Placentation. *Placenta* 33(9): 663–671.
- Duron O, D Bouchon, et al. (2008) The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol* 6:27.
- Eberl G (2016) Immunity by equilibrium. *Nat Rev Immunol* 16(8): 524.
- Elwick J (2017) Distrust that particular intuition: Resilient essentialism and empirical challenges in the history of biological individuality. In *Biological Individuality. Integrating Scientific, Philosophical and Historical Perspectives*. Scott Lidgard, and Lynn K. Nyhart (eds.). The University of Chicago Press.
- Ereshefsky M & M Pedroso (2013) Biological individuality: the case of biofilms. *Biol Philos* 28(2): 331-349.
- Ereshefsky M & M Pedroso (2015) Rethinking evolutionary individuality. *PNAS USA* 112(33): 10126-10132.
- Ereshefsky M & M Pedroso (2016) What biofilms can teach us about individuality. In A Guay & T Pradeu (eds.) *Individuals across the sciences*. Oxford, Oxford University Press: 103-122.
- Fraune S & TC Bosch (2007) Long-term maintenance of species-specific bacterial microbiota in the basal metazoan Hydra. *PNAS USA* 104(32): 13146-13151.

- Fernández L, S Langa, et al. (2013) The human gut microbiota: origin and potential roles in health and disease. *Pharmacol Res* 69:1-10.
- Fitzpatrick BM (2014) Symbiote transmission and maintenance of extra-genomic associations. *Front Microbiol* 5: 46.
- Folse HJ & J Roughgarden (2010) What is an individual organism? A multilevel selection perspective. *Q Rev Biol* 85:447-472.
- Fontana W & LW Buss (1994) "The arrival of the fittest": Toward a theory of biological organization. *Bull Math Biol* 56 (1): 1-64.
- Frank SA (2014) Somatic mosaicism and disease. *Curr Biol* 24:R577-R581.
- Funkhauser LJ & SR Bordenstein (2013) Mom knows best. The universality of maternal microbial transmission. *PLOS Biol* 11(8):e1001631.
- Gage DJ (2004) Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev* 68: 280-300.
- Garcia-Suárez A (2011) *Modos de significar. Una introducción temática a la Filosofía del lenguaje.* Tecnos.
- Gilbert SF (2017) Biological individuality. A relational reading. In *Biological Individuality. Integrating Scientific, Philosophical and Historical Perspectives*. Scott Lidgard, and Lynn K. Nyhart (eds.). The University of Chicago Press.
- Gilbert SF & L Chiu (2015) The birth of the holobiont. Multispecies birthing through mutual scaffolding and niche construction. *Biosemiotics* 8 (2): 191-210.
- Gilbert SF, E Rosenberg, et al. (2017) The holobiont with its hologenome is a level of selection in evolution. In SB Gissis, E Lamm & A Shavit (eds.) *Landscapes of collectivity in the life sciences*. London, The MIT Press: 305-324.
- Gilbert SF, J Sapp, et al. (2012) A symbiotic view of life: We have never been individuals. *Q Rev Biol* 87(4): 325-341.
- Gilbert SF & AI Tauber (2016) Rethinking Individuality: The Dialectics of the Holobiont, *Biol Philos* 31(6): 839-853.
- Godfrey-Smith P (2000) The replicator in retrospect. *Biol Philos* 15: 403-23.
- Godfrey-Smith P (2009) *Darwinian populations and natural selection*. Oxford NY, Oxford University Press.

- Godfrey-Smith P (2013) Darwinian Individuals. In *From group to individuals: evolution and emerging individuality*. F Bouchard & P Huneman (eds.). The MIT Press.
- Godfrey-Smith P (2015) Reproduction, symbiosis, and the eukaryotic cell. *PNAS USA* 112 (33):10120-10125.
- Gómez A, JM Rothman, et al. (2016) Temporal variation selects for diet-microbe co-metabolic traits in the gut of Gorilla spp. *ISME Journal* 10:514-526.
- Gontier N (2010) Evolutionary epistemology as a scientific method: a new look upon the units and levels of evolution debate. *Theory Biosci* 129.2-3: 167-182.
- Goodrich JK, Davenport ER, et al. (2017) The relationship between the human genome and microbiome comes into view. *Annu Rev Genet* 51: 413-433.
- Goodrich JK, ER Davenport, et al. (2016a) Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* 19:731-743.
- Goodrich JK, ER Davenport, et al. (2016b) Cross-species comparisons of host genetic associations with the microbiome. *Science* 352(6285): 532-535.
- Goodrich JK, JL Waters, et al (2014) Human genetics shapes the gut microbiome. *Cell* 159:789-799.
- Gould SJ (1977) Caring groups and selfish genes. *Nat Hist* 86(10): 20-24.
- Gould SJ (2002) *The Structure of Evolutionary Theory*. Harvard University Press.
- Gould SJ & E Lloyd (1999) Individuality and adaptation across levels of selection. How shall we name and generalize the unit of Darwinism? *PNAS USA* 96 (21): 11904-11909.
- Griesemer J (2000a) The units of evolutionary transition. *Selection* 1 (3): 67-80.
- Griesemer J (2000b) Development, culture, and the units of inheritance. *Philos Sci* 67:2348-2368.
- Guerrero R, Margulis L & M Berlanga (2013) Symbiogenesis: the holobiont as a unit of evolution. *Int Microbiol* 16:133–143
- Hallen-Adams HE & MJ Suhr (2017) Fungi in the healthy human gastrointestinal tract. *Virulence* 8(3): 352-358.

- Hester ER, KL Barott, et al. (2015) Stable and sporadic symbiotic communities of coral and algal holobionts. *The ISME Journal* 10:1157-1169.
- Hey J (2001) The mind of the species problem. *Trends Ecol Evol* 16(7): 326-329.
- Hull DL (1980) Individuality and Selection. *Annu Rev Ecol Evol Syst* 11:311–332.
- Hull DL (1981) Units of evolution: A metaphysical essay. In Jensen UJ, Harré R (eds.) *The Philosophy of Evolution*. Harvester Press, pp. 23-44.
- Hull DL (1988) *Science as a Process*. University of Chicago Press.
- Huneman P (2012) Natural Selection: A Case for the Counterfactual Approach. *Erkenntnis* 76(2): 171-194.
- Hurst GDD (2017) Extended genomes: symbiosis and evolution. *Interface Focus* 7: 20170001.
- Jablonka E & M Lamb (2005) *Evolution in four dimensions*. The MIT Press.
- Jaenike J, R Unckless, et al. (2010) Adaptation via symbiosis: recent spread of a Drosophila defensive symbiont. *Science* 329:212-215.
- Jakobsson HE, AM Rodríguez-Piñero, et al. (2015) The composition of the gut microbiota shapes the colon mucus barrier. *EMBO Reports* 16 (2): 164-177.
- Jantzen BZ (2017) Kinds of process and the levels of selection. *Synthese*. https://doi.org/10.1007/s11229-017-1546-1.
- Janzen DH (1977) What are dandelions and aphids? *Amer Naturalist* 111 (979): 586-589.
- Karaça C (2019) Relational basis of the organism's self-organization. PhD Dissertation. University of Exeter. https://ore.exeter.ac.uk/repository/handle/10871/37126. Accesed 30 June 2019.
- Kettlewell HD (1958) A survey of the frequencies of *Biston betularia* (L.)(Lep.) and its melanic forms in Great Britain. *Heredity* 12(1): 51.
- Klein J (1982) *Immunology: The Science of Self-Nonself Discrimination*. Wiley.
- Knowlton N & F Rohwer (2003) Multispecies microbial mutualisms on coral reefs: the host as a habitat. *Amer Naturalist* 162(S4): S51-S62.
- Koga R, T Tsuchida, et al. (2003) Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the

essential endosymbiont *Buchnera* in an aphid*. Proc R Soc Lond B* 270: 2543–2550.

- Kovacs M, Z Szendro, et al. (2006) Effect of nursing methods and faeces consumption on the development of *Bacteroides*, *Lactobacillus* and *Coliform* flora in the caecum of the new-born rabbits. *Reprod Nutr Dev* 46: 205-210.
- Kushmaro A, Rosenberg E, Fine M & Y Loya (1997) Bleaching of the coral Oculina patagonica by Vibrio AK-1. *Mar Ecol Progr Ser* 147: 159-165.
- Laland K, B Matthews & MW Feldman (2016) An introduction to niche construction theory. *Evol Ecol* 30: 191-202.
- Lamm E (2018) Big Dreams for Small Creatures: Ilana and Eugene Rosenberg's path to the Hologenome Theory. In Oren Harman & Michael Dietrich (eds.) *Dreamers, Visionaries, and Revolutionaries in the Life Sciences*. Chicago University Press.
- Lang JM, JA Eisen, et al. (2014) The microbes we eat: Abundance of taxonomic microbes consumed in a day's worth of meals for three diet types. *Peer J* 2: e659.
- Lederberg J (2000) Infectious history. *Science* 288: 287-293.
- Lederberg J & AT McCray (2001) "Ome sweet omics": A genealogical treasure of words. *Scientists* 15: 8.
- Leff JW & N Fierer (2014) Bacterial communities associated with the surfaces of fresh fruits and vegetables. *PLoS ONE* 8:e59310.
- Leggat W, T Ainsworth, et al. (2007) The hologenome theory disregards the coral holobiont. *Nat Rev Microbiol* 5(10).
- Lemanceau P, M Blouin, et al. (2017) Let the core microbiota be functional. *Trends Plant Sci* 22(7): 583-595.
- Lerner A, T Matthias, et al. (2017) Potential Effects of Horizontal Gene Exchange in the Human Gut. *Front Immunol* 8: 1630.
- Lewens T (2009). The natures of selection. *Br J Philos Sci* 61(2): 313-333.
- Lewontin RC (1970) The units of selection. *Annu Rev Ecol Evol Syst* 1: 1-18.
- Lewontin RC (1985) Adaptation. In *The Dialectical Biologist*. R Levins & RC Lewontin. Harvard University Press.
- Ley RE, M Hamady, et al. (2008) Evolution of mammals and their gut microbes. *Science* 320 (5883): 1647-1651.
- Lidgard S & LK Nyhart (2017a) *Biological Individuality. Integrating biological, historical, and philosophical perspectives*. The University of Chicago Press.

- Lidgard S & LK Nyhart (2017b) The work of biological individuality. Concepts and contexts. In *Biological Individuality. Integrating biological, historical, and philosophical perspectives*. S. Lidgard & LK Nyhart (eds.). The University of Chicago Press.
- Lipnicki LL (2015) The role of symbiosis in the transmission of some eukaryotes from aquatic to terrestrial environments. *Symbiosis* 65: 39-53.
- Liu L, X Chen, et al. (2012) The human microbiome: a hot spot of microbial horizontal gene transfer. *Genomics* 100 (5): 265-270.
- Lloyd E (1988) *The Structure and Confirmation of Evolutionary Theory*. Princeton University Press.
- Lloyd E (1992) Units of selection. In *Keywords in Evolutionary Biology*. EF Keller & E Lloyd. Harvard University Press.
- Lloyd E (2001) Units and Levels of Selection: An Anatomy of the Units of Selection Debates. In *Thinking about evolution*. SS Rama, CB Krimbas, et al. (eds.). Cambridge University Press.
- Lloyd E (2017a) Units and Levels of selection. In EN Zalta (ed.) *Stanford Encyclopaedia of Philosophy*. https://plato.stanford.edu/entries/selection-units/. Accessed 15th May 2019.
- Lloyd E (2017b) Holobionts as units of selection: Holobionts as interactors, reproducers, and manifestors of adaptation. In SB Gissis, E Lamm & A Shavit (eds.) *Landscapes of collectivity in the life sciences*. London, The MIT Press: 351-367.
- Lloyd E (2017c) A glimpse of philosophy of biology and collectivities today. In: Gissis SB, Lamm E, Shavit A (eds) *Landscapes of collectivity in the life sciences*. The MIT Press, London, pp 291–301.
- Lloyd E & MJ Wade (2019) Criteria for holobionts from community genetics. *Biol Theory*. https://doi.org/10.1007/s13752-019-00322-w.
- Louca S, SMS Jacques, et al. (2016) High taxonomic variability despite stable functional structure across microbial communities. *Nat Ecol Evol* 1:15.
- Love AC & I Brigandt (2017) Philosophical dimensions of individuality. In S Lidgard & LK Nyhart (eds.) *Biological individuality: Integrating scientific, philosophical and historical perspectives*. Chicago, The University of Chicago Press.

- Lynch KE, Parke EC, O'Malley MA (2019, forthcoming). How causal are microbiomes? A comparison with the Helicobacter pylori explanation of ulcers. *Bio Philos*.
- Marchesi JR & J Ravel (2015) The vocabulary of microbiome research: A proposal. *Microbiome* 3(31).
- Margolis E & S Laurence (2014) Concepts. In EN Zalta (ed.) *Stanford Encyclopaedia of Philosophy.* https://plato.stanford.edu/entries/concepts/. Accessed 15th May 2019.
- Margulis L (1990) Words as battle cries: Symbiogenesis and the new field of endocytobiology. *Biosci* 40(9): 673-677.
- Margulis L (1991) Symbiogenesis and symbioticism. In L Margulis & R Fester (eds.) *Symbiosis as a source of evolutionary innovation*. London, The MIT Press: 1-14.
- Margulis L (1998) *Symbiotic planet. A new look at evolution.* Basic books. New York.
- Margulis L (2010) Symbiogenesis. A new principle in evolution. Paleontol J 44(12): 1525-1539.
- Margulis L & D Sagan (2001) The beast with five genomes. *Nat hist* 110 (5): 38-41.
- Margulis L & D Sagan (2002) *Acquiring genomes. A theory of the origin of species*. Basic Books, New York.
- Margulis L & R Fester (eds) (1991) Symbiosis as a source of evolutionary innovation. The MIT Press, Cambridge.
- Martin BD & E. Schwab (2012) Current usage of symbiosis and associated terminology. *Int J Biol* 5(1): 32.
- Matthen M & A Ariew (2002) Two Ways of Thinking About Fitness and Natural Selection. *J Philos* 99(2):55-83.
- Mazel F, Davis KM, et al. (2018) Is Host Filtering the Main Driver of Phylosymbiosis across the Tree of Life? *mSystems* 3(5): e00097-18.
- Maynard-Smith J (1991) A Darwinian view of symbiosis. In L Margulis & R Fester (eds.) *Symbiosis as a source of evolutionary innovation*. London, The MIT Press: 26-39.
- Maynard-Smith J (1987) Evolutionary progress and levels of selection. In: Dupré J (ed) *The latest on the best: essays on evolution and optimality*. MIT Press, Cambridge, pp 119–131

- Maynard-Smith J & E Szathmary (1995) *The Major Transitions in Evolution*. Oxford. Oxford University Press.
- Mazmanian SK, JL Round, & DL Kasper (2008) A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 453: 620-625
- McFall-Ngai M (2015) Giving microbes their due – animal life in a microbially dominant world. *J Exp Biol* 218:1968-1973.
- McFall-Ngai M (2002) Unseen forces: the influence of bacteria on animal development. *Dev biol* 242(1): 1-14.
- McFall-Ngai M, MG Hadfield, et al. (2013) Animals in the bacterial world, a new imperative for the life sciences. *PNAS USA* 110(9):3229-3236.
- McShea DW & C Simpson (2011) The miscellaneous transitions in evolution. In *The Major Transitions in Evolution Revisited*. B Calcott & K Sterelny (eds.). The MIT Press.
- Mendoza MLZ, Z Xiong, et al. (2018) Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat Ecol Evol* 2: 659-668.
- Michod RE (1999) *Darwinian Dynamics. Evolutionary Transitions in Fitness and Individuality*. Princeton University Press.
- Michod RE & D Roze (2001) Cooperation and Conflict in the Evolution of Multicellularity. *Heredity* LXXXVI(1):1-7.
- Mitchell SD (1987) Competing units of selection? A case of symbiosis. *Phil Sci*54(3): 351–367.
- Miura T, C Braendle, et al. (2003) A comparison of parthenogenetic and sexual embryogenesis of the pea aphid *Acyrthosiphon pisum* (Hemiptera: Aphidoidea). *J Exp Zool* 295B:59-81.
- Moeller AH, S Foerster, et al. (2016) Social behavior shapes the chimpanzee pan-microbiome. *Sci Adv* 2:e1500997.
- Mogensen AL (2016) Do evolutionary debunking arguments rest on a mistake about evolutionary explanations? *Phil Studies* 173(7): 1799-1817.
- Moran N & DB Sloan (2015) The Hologenome Concept: Helpful or Hollow? *PLoS Biol* 13(12): e1002311.
- Moreno A & M Mossio (2015) *Biological Autonomy: A Philosophical and Theoretical Inquiry*. Dordrecht: Springer.
- Morris JJ (2018) What is the hologenome concept of evolution? *F1000Research* 7.

- Nyholm SV & MJ McFall-Ngai (2004) The winnowing: establishing the squid–*Vibrio* symbiosis. *Nat Rev Microbiol* 2: 632-642.
- Ochman H, Worobey M, et al. (2010) Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *Plos Biol* 8(11):e10000546.
- Odling-Smee FJ, KN Laland, MW Feldman (2003) *Niche construction: the neglected process in evolution.* In *Monographs in population biology*, vol 37. Princeton University Press
- Okasha S (2001) Why won't the group selection controversy go away? *Br J Philos Sci* 52(1): 25-50.
- Okasha S (2006) *Evolution and the levels of selection*. Oxford. Oxford University Press.
- Okahsa S (2011) Biological ontology and hierarchical organization: A defence of rank freedom. In *The Major Transitions in Evolution Revisited*. B Calcott & K Sterelny (eds.). The MIT Press.
- Okahsa S (2018) The strategy of endogeneization in evolutionary biology. *Synthese*. https://doi.org/10.1007/s11229-018-1832-6.
- Oliver KM, PH Degnan, et al. (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* 325:992-994.
- Olszak T, Di An, et al. (2012) Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* 336: 489–493.
- O'Malley MA (2014) *Philosophy of microbiology*. Cambridge University Press.
- O'Malley MA (2016) Reproduction expanded: Multigenerational and multilineal units of evolution. *Phil sci* 83:835-847.
- O'Malley MA (2017) From endosymbiosis to holobionts: evaluating a conceptual legacy. *J Theor Biol* 434:34–41. https://doi.org/10.1016/j.jtbi.2017.03.008
- O'Malley MA & J Dupré (2007). Size doesn't matter: Towards a more inclusive philosophy of biology. *Biol Philos* 22:155–191.
- Osawa R, WH Blanshard, et al. (1993) Microbiological studies of the intestinal microflora of the koala, Phascolarctos cinereus. II. Pap, a special maternal faeces consumed by juvenile koalas. *Aust J Zool* 41:611–620.
- Osborne SE, YS Leong, et al. (2009) Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. *PLoS Pathog* 5: e1000656. (doi:10.1371/journal.ppat.1000656).

- Osmanovic D, Kessler DA, Rabin Y & Y Soen (2018) Darwinian selection of host and bacteria supports emergence of Lamarckian-like adaptation of the system as a whole. *Biol direct* 13(1), 24.
- Paul, SS., & A Dey (2015) Domesticated Rare Animals (Yak, Mithun and Camel): Rumen Microbial Diversity. In *Rumen Microbiology: From Evolution to Revolution*. A Puniya, R Singh, & Devki Kamra (eds). Springer.
- Pedroso M (2017) Inheritance by recruitment. A reply to Clarke's "Levels of selection in biofilms". *Biol Philos* 32:127-131.
- Pepper JW & MD Herron (2008) Does biology need an organism concept? *Biol Rev Camb Philos Soc* 83 (4): 621-627.
- Phillips CD, Phelan G, et al. (2012) Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. *Mol ecol* 21(11): 2617-2627.
- Pollock FJ, McMinds R, et al. (2018) Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat commun* 9(1): 4921.
- Pound R (1893) Symbiosis and mutualism. *Am Nat* 27(318): 509–520.
- Pontarotti G (2016) Extended inheritance as reconstruction of extended organization: the paradigmatic case of symbiosis. *Lato Sensu* 3(1): 93-102.
- Pradeu T (2010) What is an organism? An immunological answer. *Hist Phil Life Sci* 32 (2-3): 247-267.
- Pradeu T (2012) *The Limits of the Self: Immunology and Biological Identit.* Oxford University Press.
- Pradeu T (2016a) The many faces of biological individuality. *Biol Philos* 31: 761-773.
- Pradeu T (2016b) Organisms or biological individuals? Combining physiological and evolutionary individuality. *Biol Philos* 31:797-817.
- Pradeu T & E Vivier (2016) The discontinuity theory of immunity. *Science immunology* 1(1).
- Queller DC & JE Strassmann (2009) Beyond society: The evolution of organismality. *Philos Trans Royal Soc B* 364: 3143-3155.
- Queller DC & JE Strassmann (2016) Problems of multispecies organisms: endosymbionts to holobionts. *Biol Philos* 31:855-873.
- Rainey PB & B Kerr (2011) Conflicts among levels of selection as fuel for the evolution of individuality. In *The Major Transitions in Evolution Revisited*. B Calcott & K Sterelny (eds.). The MIT Press.

- Rakoff-Nahoum S, J Paglino, et al. (2004) Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118(2):229-241.
- Raychoudhury, R., C. A. Desjardins, et al. (2010) Behavioral and genetic characteristics of a new species of Nasonia. *Heredity* 104: 278-288.
- Reeve HK & Hölldobler B (2007) The emergence of a superorganism through intergroup competition. *PNAS USA* 104(23):9736-9740.
- Reshef L, Koren O, Loya Y, Zilber-Rosenberg I & E Rosenberg (2006) The coral probiotic hypothesis. *Environ Microbiol* 8(12): 2068-2073.
- Richards RA (2010) *The Species Problem. A Philosophical Analysis*. Cambridge University Press.
- Rohwer F, Seguritan V, Azam F & N Knowlton (2002) Diversity and distribution of coral-associated bacteria. *Mar Ecol Progr Ser* 243: 1-10.
- Rosenberg E, Sharon G & I Zilber-Rosenberg (2009) The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environmental microbiology* 11.12: 2959-2962.
- Rosenberg E & I Zilber-Rosenberg (2014) *The Hologenome Concept*. London, Springer.
- Rosenberg E & I Zilber-Rosenberg (2016) Microbes drive evolution of animals and plants: the hologenome concept. *mBio* 7(2): e01395-15.
- Rosenberg E, O Koren, et al. (2007a) The role of microorganisms in coral health, disease and evolution. *Nat Rev Microbiol* 5(5): 355.
- Rosenberg E, O Koren, et al. (2007b) The hologenome theory disregards the coral holobiont: reply from Rosenberg et al *Nat Rev Microbiol* 5(10).
- Roughgarden J, Gilbert SF, et al. (2017). Holobionts as Units of Selection and a Model of their Population Dynamics and Evolution. *Biol Theory*.
- Round JL, SM Lee, et al. (2011) The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota.” *Science* 332: 974–977.
- Round JL & SK Mazmanian (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 9: 313–323
- Ruiz-Mirazo K, et al. (2000) Organisms and their place in biology. *Theory Biosci* 119: 209.

- Sanos SL, VL Bui et al. (2009) RORγt and commensal microflora are required for the differentiation of mucosal interleukin 22–producing NKp46+ cells. *Nat Immunol* 10 (1): 83-91.
- Santelices B (1999) How many kinds of individuals are there? *Trends Ecol Evol* 14: 152-155.
- Sapp J (2003) *Genesis: the evolution of biology*. Oxford University Press.
- Sapp J (2004) The dynamics of symbiosis: an historical overview. *Can J Bot* 82(8): 1046-1056.
- Shapira M (2016) Gut microbiotas and host evolution: Scaling up symbiosis. *Trends Ecol Evol* 31(7):539-549.
- Sender R, Fuchs S & R. Milo (2016) Revised estimates for the number of human and bacteria cells in the body. *PLoS biology*, *14*(8), e1002533.
- Sharon G, Segal D, Ringo J, et al. (2010). Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *PNAS USA* 107(46): 20051-20056.
- Shigenobu S & ACC Wilson (2011) Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. *Cell MolLife Sci* 68(8): 1297-1309.
- Shigenobu S, H Watanabe, et al. (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. *Nature* 407:81-86.
- Shropshire JD & SR Bordenstein (2016) Speciation by symbiosis: the microbiome and behavior. *mBio* 7(2): e01785-15.
- Skillings D (2016) Holobionts and the ecology of organisms: Multi-species communities or integrated individuals? *Bio Philos* 31:875-892.
- Smith PM & WS Garrett (2011) The gut microbiota and mucosal T cells. *Front Microbiol* 2: 111.
- Sober E (1984) *The nature of selection. Evolutionary theory in philosophical focus*. University of Chicago Press.
- Sober E & Lewontin RC (1982). Artifact, cause and genic selection. *Phil Sci* 49(2): 157-180.
- Sober E & Wilson DS (1994) A critical review of philosophical work on the units of selection problem. *Phil Sci* 61(4): 534-555.
- Sober E & Wilson DS (1998) *Unto Others: The Evolution and Psychology of Unselfish Behaviour*. Cambridge MA, Harvard University Press.

- Sober E & DS Wilson (2011) *Adaptation and Natural Selection* revisited. Journal of Evolutionary Biology 24 (2): 462-468.
- Sommer F, M Stahlman, et al. (2016) The gut microbiota modulates energy metabolism in the hibernating brown bear *Ursus arctos*. *Cell Reports* 15 (7): 1655-1661.
- Souza DT, Genuário DB, et al. (2017) Analysis of bacterial composition in marine sponges reveals the influence of host phylogeny and environment. *FEMS microbiol ecol* 93 (1).
- Stappenbeck TS, LV Hooper, et al. (2002) Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *PNAS* (USA) 99:15451–15455.
- Stegmann U (2010) What can natural seletion explain? *Stud Hist Philos Sci C* 41(1): 61-66.
- Stencel A (2016) The relativity of Darwinian populations. *Biol Philos* 31(5): 619-637.
- Stencel A & AM Proszewska (2017) How Research on Microbiomes is Changing Biology: A Discussion on the Concept of the Organism. *Found Sci* 23(4): 603-620.
- Stencel A & DM Wloch-Salamon (2018) Some theoretical insights into the hologenome theory of evolution and the role of microbes in speciation. *Theory Biosci* 137(2): 197-206.
- Sterner B (2015) Pathways to pluralism in biological individuality. *Biol Philos* 30 (5): 609-628.
- Sterner B (2017) Individuating population lineages: a new genealogical criterion. *Biol Philos* 32: 683-703.
- Suárez J (2016) Bacterial species pluralism in the light of medicine and endosymbiosis. *Theoria* 31 (1): 91-105. http://dx.doi.org/10.1387/theoria.13242.
- Suárez J (2018a) The importance of symbiosis in philosophy of biology: an analysis of the current debate on biological individuality and its historical roots. *Symbiosis* 76: 77-96. https://doi.org/10.1007/s13199-018-0556-1.
- Suárez J (2018b) Review of Biological Individuality. Integrating scientific, philosophical, and historical perspectives, by Scott Lidgard and Lynn K. Nyhart. *Hist Phil Life Sci* 40 (67). https://doi.org/10.1007/s40656-018-0233-y.

- Suárez J (2018c) Review of Landscapes of Collectivity in the Life Sciences, by SB Gissis, E Lamm and A Shavit. *Hist Phil Life Sci* 40 (37). https://doi.org/10.1007/s40656-018-0202-5.
- Suárez J & V Triviño (2019) A metaphysical approach to holobiont individuality. Holobionts as emergent individuals. *Quaderns de Filosofia* VI(1): 59-76. https://doi.org/10.7203/qfia.6.1.14825.
- Tauber AI (2016) Immunity in Context: Science and Society in Dialogue. *Theoria.* 31(2): 207–24.
- Taxis TM, S Wolff, et al. (2015) The players may change but the game remains: network analyses of ruminal microbiomes suggest taxonomic differences mask functional similarity. *Nucleic Acids Res* 43(20): 9600-9612.
- Theis KR, Dheilly NM, et al. (2016) Getting the hologenome concept right: an ecoevolutionary framework for hosts and their microbiomes. *mSystems* 1(2):e00028-16.
- Triviño V & L Nuño de la Rosa (2016) A causal-dispositional account of fitness. *Hist Phil Life Sci* 38 (6).
- Tsuchida T, R Koga, et al. (2010) Symbiotic bacterium modifies aphid body color. *Science* 330:1102-1104.
- Tung J, LB Barreiro, et al. (2015) Social networks predict gut microbiome composition in wild baboons. *eLife* 4:e05224.
- Turpin W, O Espín-García, et al. (2016) Association of host genome with intestinal microbial composition in a large healthy cohort. *Nat Genet* 48 (11):1413-1417.
- Valdés-Villanueva LM (2005) *La búsqueda del significado: Lecturas de filosofía del lenguaje*. Tecnos.
- van Opstal EJ & SR Bordenstein (2015) Rethinking heritability of the microbiome. *Science* 349:1172-1173.
- Veigl SJ (2017) Use/disuse paradigms are ubiquitous concepts in characterizing the process of inheritance. *RNA Biology* 14(12): 1700-1704.
- Veigl SJ (2019) Seeing “Lamarckian” More Positively: The Use/Disuse Paradigm Increases Understanding. *BioEssays* https://doi.org/10.1002/bies.201900054.
- Vieira-Silva S, G Falony, et al. (2016) Species–function relationships shape ecological properties of the human gut microbiome. *Nat Microbiol* 1:1-8.

- Visscher PM, WG Hill & NR Wray (2008). Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 9 (4): 255.
- Vrba ES & SJ Gould (1986) The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology* 12 (2): 217-228.
- Walsh D, Lewens T, & A Ariew (2002). Trials of life: Natural selection and random drift. *Phil Sci* 69: 452-473.
- Werren JH, L Baldo, et al. (2008) *Wolbachia*: master manipulator of invertebrate biology. *Nat Rev Microbiol* 6:741-51.
- Westra U, RR Rivera, et al. (2010). Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J Comp Neurol* 518(19): 3981–4000.
- White JA (2011) Caught in the act: Rapid, symbiont-driven evolution. *BioEssays* 33:823829.
- Wilkins JS & P Bourrat (2018) Replication and reproduction. In *The Stanford Encyclopedia of Philosophy*. EN Zalta (ed.). https://plato.stanford.edu/entries/replication/. Accessed 29th May 2019.
- Williams GC (1966) *Adaptation and Natural Selection*. Princeton University Press.
- Wilson DA & E Sober (1989) Reviving the superorganism. *J Theor Biol* 136(6): 337-356.
- Wilson J (2000) Ontological Butchery: Organism Concepts and Biological Generalizations. *Phil Sci* 67: S301-S311.
- Wilson RA & M Barker (2013) The Biological Notion of Individual. In *The Stanford Encyclopedia of Philosophy*. EN Zalta (ed.). https://plato.stanford.edu/entries/biology-individual/. Accessed 29th May 2019.
- Wolfe CT (2010) Do organisms have an ontological status? *Hist Phil Life Sci* 32 (2-3): 195-231.
- Zilber-Rosenberg I & E Rosenberg (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Ecol* 32.5: 723-735.
- Zheng Y, PA Valdez, et al. (2008) Interleukin-22 mediates early host defence against attaching and effacing bacterial pathogens. *Nat med* 14 (3): 282-289.