Co-evolution of host and endosymbiont genomes during Eukaryogenesis: the role of sexual reproduction

Major research project

Alkmini Zania (1659499)

Supervisors: Prof. Paulien Hogeweg and Dr. Samuel von der Dunk

Examiners: Prof. Paulien Hogeweg and Prof. Berend Snel

Theoretical Biology and Bioinformatics, Department of Biology, Science Faculty, Utrecht University

Abstract

Sexual reproduction is a trait shared in all eukaryotes, and was present in the last eukaryotic common ancestor. Current eukaryotes mostly inherit mitochondria from one parent, but the mechanisms that ensure this vary hugely in different species, and the ancestral trait is unknown. It is unknown when sex appeared in relation to the acquirement of mitochondria and its role in the evolution of endosymbiosis. In this project, a multilevel, individual based model of endosymbiosis (von der Dunk et al., 2023) is used to study the role sex in the co-evolution of host and symbiont genomes. Evolutionary simulations were performed on populations of cells that live on a 2D grid, each cell consisting of one host and one or more symbionts, the latter inherited by one parent (asexual reproduction) or two parents (sexual reproduction). Our results indicate that sex poses threats, due to the emergence of selfish symbionts that can drive populations to extinction. However, resolution of this conflict is possible, and can drive the evolution of signalling, allowing the host to control the symbiont's cell cycle. In fact, in many cases the populations that survive sexual reproduction outperform the asexual populations that evolve under the same conditions.

Introduction

The last eukaryotic common ancestor (LECA) appeared ~1.7-2.4 billion years ago [1, 2], and was the product of an endosymbiotic event between an archaeal host and an alphaproteobacterial endosymbiont [3]. While LECA had the characteristics of modern eukaryotes, including active cytoskeleton, endomembrane system, phagocytosis, nucleus, mitochondria and meiosis, it is unknown when and in what order these arose during eukaryogenesis [4]. Especially the timing of the endosymbiont acquirement and the initial relationship between host and symbiont (phagotrophic or syntrophic) are heavily debated [4, 5].

This first endosymbiotic event was a key transition during eukaryogenesis. Endosymbiotic gene transfer from endosymbiont to host contributed to the expansion and complexification of the eukaryotic genome [6], while the energetic advantage that mitochondria provide could be the force that allowed the evolution of eukaryote complexity and multicellularity [7]. The nucleus and meiotic sex have also been suggested to be adaptations to the influx of introns and the ROS stress generated by the proto-mitochondrion respectively [8].

During the transition from endosymbiont to organelle, the mitochondrion underwent massive changes in structure and function. Its genome size decreased dramatically, which is often observed in symbionts [26]. Many of its genes migrated to the nucleus or were replaced by host proteins, allowing the host to take it under full control [27]. In order for this to happen, signal peptides and the mitochondrial protein import machinery were required. The latter probably evolved from the symbionts protein export and membrane protein insertion systems [28]. These processes were fundamental in the transition of individuality that happened during eukaryogenesis, where host and symbiont were combined to form a higher level of selection. They could be viewed as conflict mediation mechanisms, that smoothed competing interests between the different levels of selection [24].

Mitochondria of modern eukaryotes are usually inherited by one parent, most commonly the mother in animals and plants. However, uniparental inheritance (UPI), and therefore lack of mixing and recombination between the symbiont population, makes organelles prone to Muller's ratchet; accumulation of deleterious mutations and loss of viability. Various, non-mutually exclusive hypothesis have been proposed to explain the prevalence of UPI, but the question has not been resolved. The oldest hypothesis is that it prevents the spread of selfish symbionts [10]. Additionally, UPI could increase genetic variance among hosts, enabling selection against deleterious mutations [9]. Alternative explanations include direct selection against heteroplasmy [12, 13] and promotion of mito-nuclear co-adaptation [11].

Despite various strong arguments for the omni-presence of uniparental inheritance, there are increasing evidence that UPI is not as strictly followed as it was previously thought. Paternal leakage and mtDNA recombination have been reported in plants, animals and fungi [15, 16]. The vast variety of mechanisms that ensure UPI in eukaryotes [17, 18] suggests frequent changes in selection pressures and turnover of transmission mechanisms. Greiner et al. (2015) suggest that this turnover is driven by selfish elements and Muller's ratchet; selfish elements promote the evolution of UPI, while muller's ratchet promotes BPI/ paternal leakage, leading to a continuous shift between inheritance patterns [18].

While progress has been made in understanding the prevalence of UPI in modern eukaryotes, what happened between FECA and LECA is still a mystery. LECA possessed mitochondria, was sexual and capable of cell fusion, but we do not know when sex appeared, in what form, and how endosymbionts were inherited in primordial eukaryotes. Given the vast variety of sex determination and mating type systems across eukaryotes, Heitman (2015)

proposes that LECA was a unisexual organism [19], and mating types appeared later. The same can be said for mitochondrial inheritance mechanisms; modern eukaryotes possess various tools to ensure UPI, but all of them seem to have evolved later, leaving no traces from LECA. While it has been proposed that early eukaryotes should have evolved UPI, in order to avoid the spread of selfish symbionts [20, 21], we could also assume that primordial cell fusion didn't include endosymbiont segregation/ destruction mechanisms [22]. Theoretical models have examined whether the evolution of sexual cell fusion could have been driven by the bi-parental inheritance of symbionts; either as a means of promoting proto-mitochondrial complementation [23] or because of the short-term advantage of homogenizing the cytoplasm, when there are negative epistatic interactions among mitochondria and weak mito-nuclear associations [14]. However, Radzvilavicius & Blackstone [24] suggest that in order for cell – cell fusion to evolve, the new holobiont should have already evolved conflict mediation mechanisms, like signaling and transfer of mitochondrial genes to the nucleus, to prevent the spread of selfish symbionts. Therefore they propose that sex appeared in the later stages of eukaryogenesis.

Here, we study the impact of bi-parental symbiont inheritance in a multilevel model of eukaryogenesis [25], where hosts and symbionts are individual entities with their own genome that, through a non-linear genotype-phenotype map, results in their growth dynamics. This framework allows properties like selfishness, co-operation and cell cycle co-ordination to evolve. In our simulations there is no pre-imposed fitness criterion, but selection happens implicitly, based on the ability of holobionts to regulate their cell cycle, grow and outcompete each other. We assume that early eukaryotes do not have mechanisms that allow symbiont segregation, so cell fusion always leads to BPI. Our focus is on host-symbiont dynamics and co-evolution, so we implement sexual reproduction as BPI, without the additional factor of genome recombination (neither host nor symbiont). First, we study how strict bi-parental inheritance alters the outcomes of evolution. Then, we allow hosts to mutate from asexual to sexual and study which tactic is favored and under which conditions.

Materials and methods

Genome codes a Boolean regulatory network

The modeling framework we use is based on the cell cycle regulation model by von der Dunk et al., 2022 [29]. It is an individual-based model, where cells live on a two-dimensional grid. Each cell has a genome made up by regulatory genes and binding sites, forming a Boolean regulatory network. The genome is linear and each binding site or gene is a discrete bead (beads-on-a-string genome). There is an origin and a terminus of replication at the opposite ends of the genome. Apart from sites and regulatory genes, a bead can also be a household gene (inactive in the regulatory network).

The state of the cell is defined by the expression pattern of the 5 core genes (g1-5). Cells need to correctly regulate the expression of these core genes in order to go from G1 to the S phase, stay in the S phase long enough to fully replicate their genome, go to stage G2, and finally to stage M to divide. If a cell does not manage to spend enough time in the S-phase to complete replication, but tries to divide by entering the M stage, it dies (fig. 1).

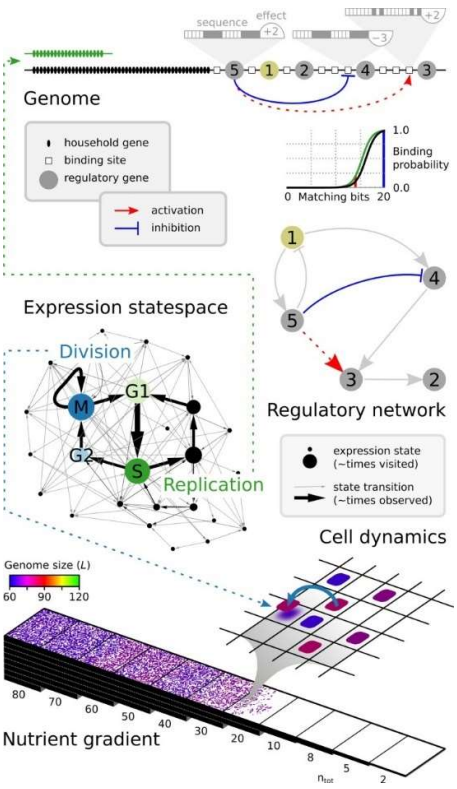


***Figure 1 Overview of the modeling framework**. A linear genome consists of discrete beads that can be household genes, regulatory genes or binding sites. Regulatory interactions give rise to gene expression dynamics shown in the state space, where four states (G1, S, G2,andM) define the cell cycle. In stage S cells duplicate their genome and at stage M they divide, possibly killing a neighbor in the process. Figure from von der Dunk et al., 2022.*

Genome replication is explicitly modeled in this framework; cells use externally provided nutrients to replicate each bead of their genome. During 1 timestep spent in the S-phase, if *n* nutrients are available, the cell replicates *n* beads. Therefore genome size is a constraint for growth speed. Upon successful completion of the cell cycle, the cell divides; the daughter cell inherits one copy of the genome and inhabits a random adjacent grid point. If the grid point is not free, the cell that until then occupied that grid point is killed and the new one takes its place. The grid is divided in 11 sectors, with decreasing influx of nutrients, forming a nutrient gradient (fig. 1).

At every timestep, expressed gene products bind stochastically to binding sites, with a probability based on bitstring similarity between gene and binding site. Even genes with dissimilar bitstring have a small probability of binding to a site. Only one gene product can bind to each site per timestep (eq. 1). A gene is expressed when activating products have bound to its upstream sites, if their effect surpasses its activation threshold (eq. 2). The effect of a binding depends on the regulatory weight of the binding site and the bound gene (eq. 3).

$$p_{x\to b} = \frac{\epsilon_x k_0 \, e^{S(x,b)}}{1 + \sum_y^{n_g} \epsilon_y k_0 \, e^{S(y,b)}}$$

Equation 1 *S(x, b) is the number of matching bits between the binding sequences of gene x and site b (bitstrings of length l = 20); n_g is the total number of genes in the genome, and e_x is the expression status of gene x (0 or 1).*

$$\epsilon_i = \begin{cases} 1 & \text{if } \sum_b w_{x\to b} > \theta_i, \\ 0 & \text{otherwise} \end{cases}$$

Equation 2 *A gene is expressed if the sum of regulatory effects at its upstream binding sites reaches its specific activation threshold u.*

$$w_{x\to b} = w_x \cdot w_b$$

Equation 3 *The regulatory effect $w_{x->b}$ results from the binding of expressed gene x to binding site b.*

Beads that have been replicated also participate in the regulatory dynamics. This allows cells to evolve low affinity interactions that create checkpoints during the cell cycle; dosage-sensitive genes located at the end of the genome have a high probability of being expressed (and thus promote progression of the cell cycle) after they have been replicated (von der Dunk et al., 2022)

Different types of mutations can take place on the genome (table 1). Individual beads can be duplicated (μ_{dup}), relocated (μ_{rel}) or deleted (μ_{del}). The new location of a relocated or duplicated bead is random. In case of genes, the mutation affects its upstream binding sites as well. The properties of binding sites and genes; bitstring, regulatory weight w and the activation threshold θ, can also mutate with rates μ_B, μ_w, μ_θ respectively. Finally, there is innovation of genes ($\mu_{g,in}$) and binding sites ($\mu_{b,in}$), that can create a bead with random properties at a random location.

Obligate endosymbiosis

An extension of the above model was used to study endosymbiosis [30]. Here, every cell ('holobiont') consists of one host and one or more symbionts, each with its own genome that encodes a regulatory network. Host and symbiont(s) regulate their cell-cycle autonomously. The holobiont divides when the host does, while symbionts divide and die inside their host. The holobiont dies when the host dies, or when the last symbiont dies. Upon division, each symbiont has a 50% chance of being passed to the daughter cell; this means that a holobiont with few symbionts has a high chance of producing an offspring without any symbionts that dies immediately. Nutrients are now equally distributed between host and symbionts in the 3-by-3 neighborhood of the grid. If too many symbionts are present, they can slow down the host's growth or even kill it.

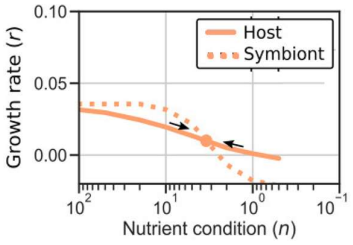


Figure 2 Example of implicit host–symbiont cell-cycle coordination. *Host and symbiont growth rates are assessed independently. They form a stable equilibrium at the nutrient concentration for which their growth rates are equal. Figure from . Figure from von der Dunk et al., 2023.*

Each holobiont needs 100 household (HH) genes. The HH genes of the host and the average of all the symbionts are summed, allowing them to be transferred from host to symbiont or the reverse.

In this model, symbionts and hosts evolved co-ordinate their cell cycle implicitly, through nutrient availability. When there are few symbionts (thus more nutrients), they replicate faster than the host, increasing their numbers. When there are too many symbionts (fewer nutrients), the host

divides faster. Thus they form a stable equilibrium at a specific nutrient concentration, where their growth rates are equal (fig. 2, [30]).

Host – symbiont communication

In the extended version of the model, described in [25], hosts and symbionts can communicate directly. This allows explicit cell cycle co-ordination to evolve, and integration of the host-symbiont regulatory networks. Communication happens either through passive leakage of regulatory products, or through active signaling.

Products can leak (rate *l=0.01*) from host to symbiont and reverse, or be actively transported according to the signal peptide that defines their localization; host, symbiont, dual or no relocation (product stays in the compartment where it is expressed). Foreign products can bind to the genome and interfere with its regulation, as well as define the cell cycle stage if they are identical to the native gene. The effect of leakage from symbionts to host is proportional to the number of symbionts. In addition to the mutations described above, transfer of genes (copy- and cut-paste with rate μ_t) between host and symbionts is also implemented, as well as bitstring mutations (μ_s) of signal peptides (table 1).

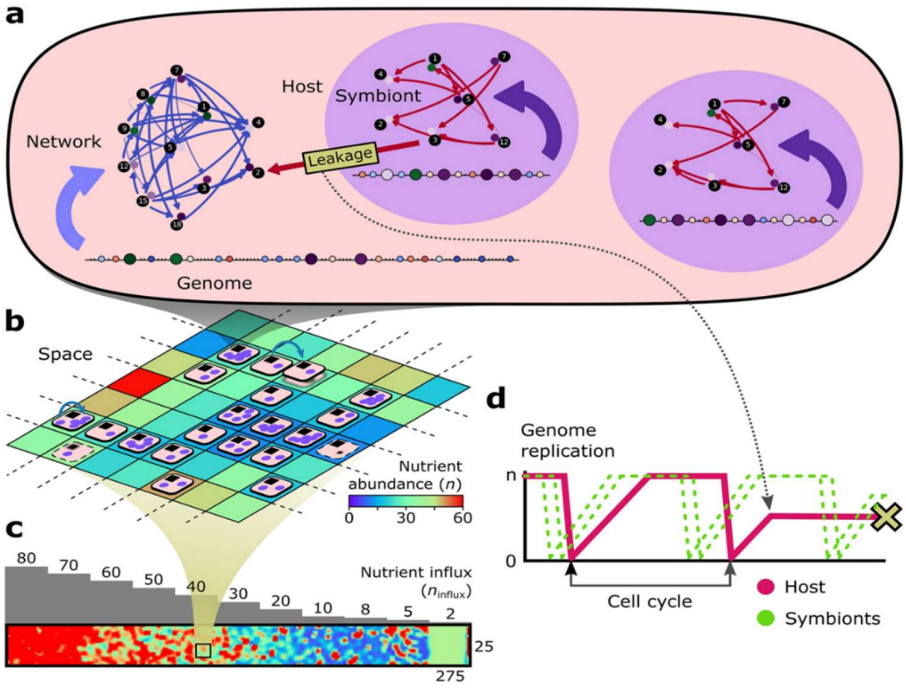


Figure 3 Summary of the model *(a) The holobiont consists of 1 host and 1 or more symbionts, each with its own genome. A genome consist of discrete beads that can be regulatory genes, binding sites or inactive household genes. Interactions of regulatory gene products and binding sites create a regulatory network. Products leak from host to symbionts and from symbionts to host or are actively transported through signaling. (b, c) Holobionts live on a two-dimensional grid, divided in sectors with decreasing nutrient influx (nutrient gradient). They compete for space and nutrients. (d) Example of a cell cycle. Hosts and symbionts replicate their genomes, enter the G2 and then the M phase and divide. A host does not manage to completely replicate its genome, tries to divide and dies. Figure from von der Dunk et al., 2024.*

Leakage drives the evolution of signaling and explicit cell cycle co-ordination between host and symbionts [25]. Low affinity interactions between host and symbiont genomes can create checkpoints for symbiont numbers, so that the holobiont only divides when there is a safe amount of symbionts. Another strategy that emerged was cell-cycle synchronization, were host and symbionts replicate their genomes and divide at the same time.

Sexual reproduction

The model is extended to include sexual reproduction. As we want to examine the effects of bi-parental inheritance of symbionts (BPI) in host symbiont co-evolution, sexual reproduction here is restricted to BPI. No genome recombination between hosts or symbionts takes place.

When a host is in the M-stage, it must find a mate; a randomly picked host that is also in the M-stage and inhabits the same sector of the gradient. If no mate is found in a timestep, the holobiont enters a dormant state and pauses its cell cycle. This happens when there is an odd number of dividing holobionts in one sector of the gradient. If a mate is found, the hosts divide normally: they pass their genome to the offspring, which grows on an adjacent square, possibly killing its previous inhabitant. No genome recombination takes place between hosts. However, now each of their symbionts has a 50% chance of being passed to a symbiont pool and from there to either offspring. In some experiments, holobionts are allowed to mutate from asexual to sexual reproduction and vice versa (μ_r=*0.001*). In this case, mates can only be other sexually reproducing holobionts.

Table 1 Mutation rates

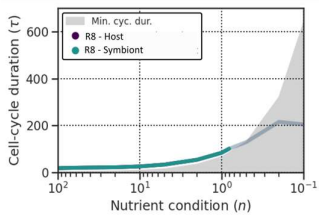
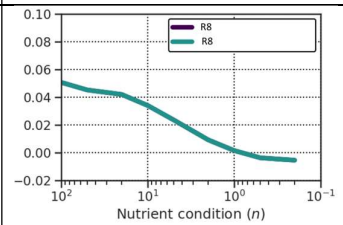
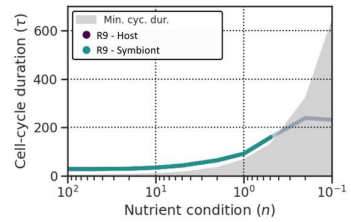
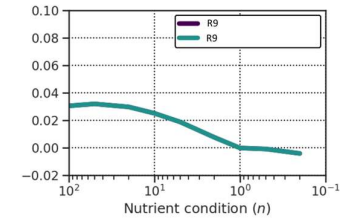
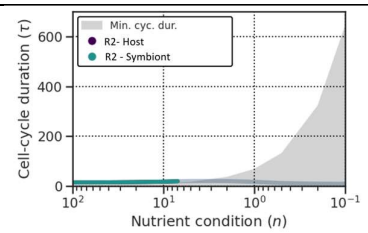
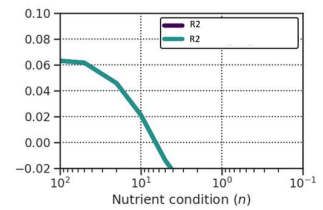
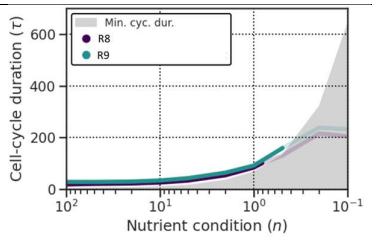
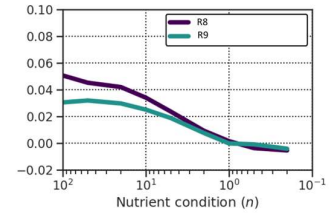
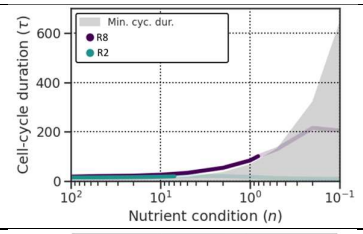
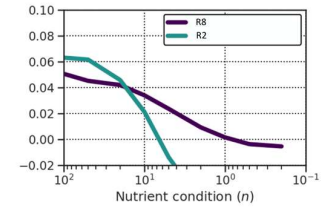
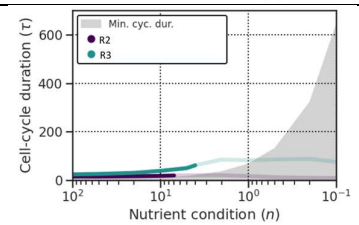
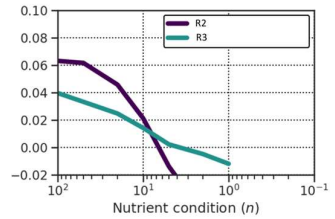
Parameter	Value	Description
μ_B	$3.3 \cdot 10^{-5}$	Per-bit mutation rate of binding sequence
μ_S	$1.0 \cdot 10^{-5}$	Per-bit mutation rate of signal peptide
μ_w	$1.67 \cdot 10^{-4}$	Regulatory weight mutation rate
μ_θ	$1.67 \cdot 10^{-4}$	Activation threshold mutation rate
μ_{dup}	$1.67 \cdot 10^{-4}$	Per-bead duplication rate
μ_{del}	$1.67 \cdot 10^{-4}$	Per-bead deletion rate
μ_{rel}	$1.67 \cdot 10^{-4}$	Per-bead relocation rate
μ_t	$2.0 \cdot 10^{-5}$	Per-bead transfer rate
$\mu_{b,in}$	$1.67 \cdot 10^{-3}$	Per-genome innovation rate for binding sites
$\mu_{r,in}$	$1.67 \cdot 10^{-4}$	Per-genome innovation rate for regulatory genes
μ_r	$1 \cdot 10^{-3}$	Reproductive mode mutation

Strains used in experiments

Experiments were performed with primitive holobionts, where the host and symbiont have identical genomes that can perform a basic cell cycle. They can only survive in the nutrient rich sectors of the gradient in the beginning, but adapt to harsher conditions as they evolve better cell cycle regulation and co-ordination. Populations evolved with and without allowing host – symbiont communication (leakage and signaling), with obligate BPI and mutable reproduction (table 3). Here we investigate the evolution of selfish symbionts and weather communication can allow the host to control the symbionts' cell-cycle, preventing their spread.

Eukaryogenesis was a merger between two distantly related, complex prokaryotes. To better model this process, we also evolved holobionts formed from pre-evolved strains (table 2), that adapted to the gradient as free-living "prokaryotes". These strains are R2, R3, R8 and R9, described in [29]. Strain R2 is a specialist while R3, R8 &R9 are generalists. The holobionts are

evolved under intermediate nutrient conditions (n_{influx}=30, 50x50 grid, table 3) with sexual reproduction, in order to compare the results with those of asexual reproduction. They are also evolved on the nutrient gradient, with asexual, sexual and mutable reproduction, in order to investigate how BPI affects adaptation to different conditions and if/when it can be the favored mode of reproduction.

Holobiont	Host	Symbiont	Cell cycle timing	Growth rate
1	R8	R8		
2	R9	R9		
3	R2	R2		
5	R8	R9		
6*	R9	R8		
7*	R8	R2		
8	R2	R8		
9*	R3	R2		
10	R2	R3		
11	R8	R3		

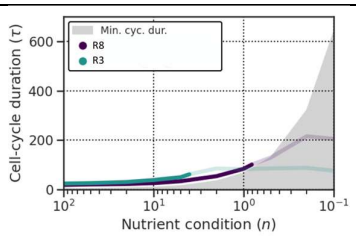
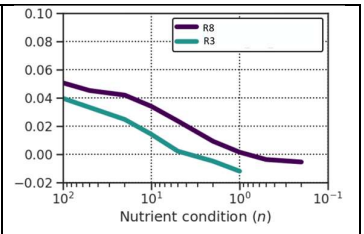
12	R3	R8		

Table 2 Pre-evolved strains as host-symbiont pairs. *The growth curves and cell cycle duration were calculated independently for every strain, using the protocols described in [29]. Holobiont 4 was not viable and excluded from the experiments.*
**stable equilibrium*

Experiment	Strains	Leakage	Signaling	Reproduction	Environment
K	Primitive	No	No	Sexual	Gradient (275x25)
H	Primitive	No	No	Mutable	Gradient
L	Primitive	Yes	Yes	Sexual	Gradient
E	Primitive	Yes	Yes	Mutable	Gradient
ESI	Pre-evolved	No	No	Sexual	n_{influx}=30 (50x50)
PPIA, PPIB, PPIC	Pre-evolved	No	Yes	Sexual	n_{influx}=30
PPIIA, PPIIB, PPIIC	Pre-evolved	Yes	Yes	Sexual	n_{influx}=30
GAP	Pre-evolved	Yes	Yes	Asexual	Gradient
GP	Pre-evolved	Yes	Yes	Sexual	Gradient
GCP	Pre-evolved	Yes	Yes	Mutable	Gradient

Table 3 Overview of experimental conditions. *Experiments K, H, L and E take place with primitive holobionts on the gradient. Experiments ESI, PPIA-C, PPIIA-C, GP, GAP and GCP take place with pre-evolved strains, on the gradient and on intermediate nutrient influx.*

Results 1: Primitive holobionts

Direct advantage of BPI

We observed that was sex selected against in the primitive populations, but persistent in adapted ones. To find out why, we need to take a closer look into how reproduction works. When a cell divides asexually, each symbiont of the parent has 50% chance of being passed to the offspring. It is possible that if a cell has few symbionts, the parent or the offspring will die, if all or none of the symbionts are passed down. When cells reproduce sexually, the dynamics of parents are the same; they pass down symbionts with 50% chance each. However now each offspring has a 50% chance to inherit symbionts from either parent. In populations with very few symbionts per cell, this can hinder reproductive success; for example, if two parents with 1 symbiont each pass down their symbionts, there is a chance that both of them will end up to the same child, leaving the population with 1 instead of 2 viable holobionts (table 4). When the population is better adapted, however, symbiont rich holobionts can 'rescue' the children of holobionts with few symbionts.

Sex/ div		#symbionts parent 1									
		1	2	3	4	5	6	7	8	9	10
#symbionts parent 2	1	-0.123 / 0.0	0.40474 / 0.49813	0.73913 / 0.75093	0.96223 / 0.8761	1.11282 / 0.93633	1.21496 / 0.96854	1.29137 / 0.98445	1.34642 / 0.99223	1.38521 / 0.99613	1.41393 / 0.99796
	2		0.86278 / 0.9948	1.1503 / 1.25177	1.33139 / 1.37479	1.45162 / 1.43677	1.53458 / 1.47048	1.59212 / 1.4835	1.63316 / 1.49278	1.6624 / 1.49515	1.6877 / 1.496
	3			1.39778 / 1.49954	1.54596 / 1.62441	1.64702 / 1.68777	1.70879 / 1.71799	1.75463 / 1.73146	1.7848 / 1.74333	1.80967 / 1.7441	1.82662 / 1.74732
	4				1.67171 / 1.75087	1.75432 / 1.81054	1.81164 / 1.8439	1.84525 / 1.85767	1.8708 / 1.86718	1.88813 / 1.87031	1.90159 / 1.87298
	5					1.82617 / 1.87488	1.86949 / 1.90554	1.89674 / 1.92135	1.9172 / 1.92944	1.93069 / 1.93321	1.94159 / 1.93468
	6						11.90584 / 1.93782	1.93126 / 1.95404	1.94529 / 1.96145	1.95511 / 1.96462	1.96378 / 1.96699
	7							1.94794 / 1.96855	1.96207 / 1.9766	1.97013 / 1.98094	1.97639 / 1.98244
	8								1.97287 / 1.98419	1.97915 / 1.9885	1.98364 / 1.99
	9									1.9839 / 1.9918	1.98875 / 1.99457
	10										1.99168 / 1.99631

Table 4 Average increase in number of holobionts in one reproductive step*, depending on how many symbionts each parent has. Red indicates advantage of asexual division, green advantage of BPI. When each parent has 1 symbiont, BPI can lead to negative growth, if both parents and one child die. On the other hand, if one parent has 1-3 symbionts and the other 5 or more, BPI has a clear advantage.*

Holobionts evolving without leakage and transfer are susceptible to selfish symbionts

These experiments were performed in order to determine the evolutionary stability of implicit co-ordination under BPI. We started with primitive holobionts; host and symbionts have identical genomes and can only survive in nutrient-rich sectors, performing a basic cell-cycle. Since there is no leakage of regulatory products nor signaling, host and symbiont can only

communicate indirectly, through nutrient availability, and adapt by forming a stable equilibrium at a specific nutrient concentration (fig. 2).

However, when we introduce bi-parental inheritance of symbionts, this adaptation mechanism becomes unstable. This is due to the advantage of fast-growing symbionts, who outcompete more co-operative ones when found in the same host: BPI creates selection for very fast, small symbionts that can replicate even with very few nutrients. They increase in numbers to the point that their own host is starved. Most populations died out before hosts could adapt to the nutrient gradient. The populations that did adapt where quickly overtaken by fast growing ('selfish') symbionts that deplete the host of nutrients (fig. 4), evolving to extinction. Symbionts with fewer HH genes are faster and spread in the population, causing hosts with more HH genes survive more. This results in the evolution of large host and small symbiont genome size.

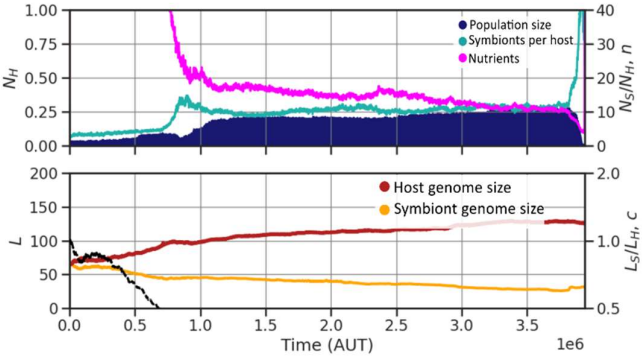


Figure 4 Timeline of a population without host - symbiont communication, obligate BPI *(K2). Top panel: This primitive population starts with a small size, few symbionts per cell and a lot of nutrients. Later holobionts adapt somewhat to the gradient, increasing the population size. However symbionts start growing faster and increase their numbers, eventually starving the host to extinction. Bottom panel: evolution of genome sizes. As symbionts compete with each other, they evolve a small genome size, which allows them to complete their cell cycle in fewer timesteps. HH genes are indirectly transferred to the host, who increases its genome size.*

The same experiment was repeated, starting with asexual populations that can mutate to BPI (μ_r=*0.001*). These populations do not die out, as they can restrict the spreading of selfish symbionts by switching to asexual reproduction. The dynamics of populations can be described in 3 main phases: pre-adaptation, stable sexual reproduction and host-symbiont conflict. In the pre-adaptation period, when holobionts have few symbionts per host, sexual reproduction is mainly avoided (see direct benefits of BPI). Increase in population size and symbiont numbers is accompanied by switch towards sexual reproduction. In the phase of stable sexual reproduction host – symbiont dynamics are stable, i.e. the host has a cell cycle with speed comparable to the symbiont's and can survive with fewer nutrients than the symbiont (fig. 5c). The host-symbiont conflict phase appears if symbionts are faster and start depleting nutrients: when the fraction of sexual holobionts fluctuates, along with symbiont numbers (fig. 5b).

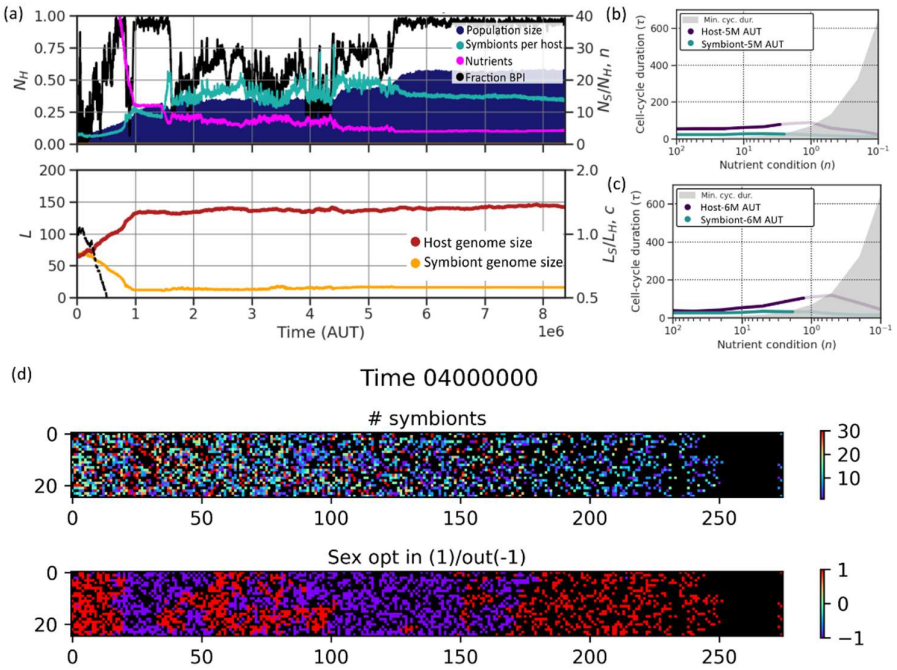


Figure 5 Evolution of population with no H-S communication, mutable reproduction *(a) Timeline of population H7, with mutable reproduction, and no host – symbiont communication. When the population is small and with few symbionts per host, cells mostly reproduce asexually. While cells adapt, BPI is dominant. Later, cells alternate between division and sex, to control symbiont numbers. BPI is finally re-established and stable. For more timelines, see table S1 (b) Cell cycle duration of host and symbiont MRCA at 5M timesteps, when symbionts fluctuate; they are faster that the host, who reproduces asexually to keep them in check. (c) Cell cycle duration of host and symbiont MRCA at 6M timesteps, when hosts reproduce sexually; they are more efficient than their symbionts. (d) Snapshot of the population at 4M AUT. Sexual and asexual reproduction are present in different sectors, while symbiont numbers are high.*

Obligate sexual populations remain stable by evolving communication

We now investigate if the ability to signal can rescue populations from selfish symbionts. Experiments started with the same conditions, but now host and symbionts can communicate. There is always passive leakage (l=0.01) of regulatory products from host to symbiont and vice versa. We also introduce signal peptides that can mutate (μ_s=10^{-5}), changing the localization of a regulatory product.

These settings allow hosts and symbionts to evolve explicit communication and coordination of their cell cycle. In the results described in [25] where holobionts evolved asexually, different control strategies emerged. Shortly, host to symbiont, bi-directional and symbiont to host signaling evolved and created check points for symbiont and host division. The most successful strategy, which outcompeted all other populations, evolved host-to-symbiont signaling (host control) and a synchronized cell cycle: hosts and symbionts replicate their genome and divide at the same timesteps. In all these cases, signaling evolved when there was product leakage, which initially causes problems in holobionts due to interference between host and symbiont cell cycles. When holobionts evolved without leakage, they only evolved implicit coordination.

We evolved 10 populations with obligate sexual reproduction, leakage and the possibility of signaling and transfer of genes. Out of the 10, 3 when extinct early, which is similar to the asexual model [25]. One more went extinct later, due to selfish symbionts taking over the population (fig. 6, population 3). Of the remaining 6 populations, 3 evolved a synchronized cell-cycle, with bi-directional control. In these populations the host – symbiont conflict seems to be

resolved; they maintain dense populations and stable symbiont numbers (fig. 6, populations 1,2, 7). This is the most noticeable difference from the asexual model, where synchronization only evolved once, and under host control. The remaining populations evolved various forms of signaling without synchronization, allowing hosts to keep their symbionts under control. However, these populations have relatively large symbiont numbers that fluctuate over time, indicating an on-going battle between hosts and selfish symbionts (fig. 6, populations 5, 8).

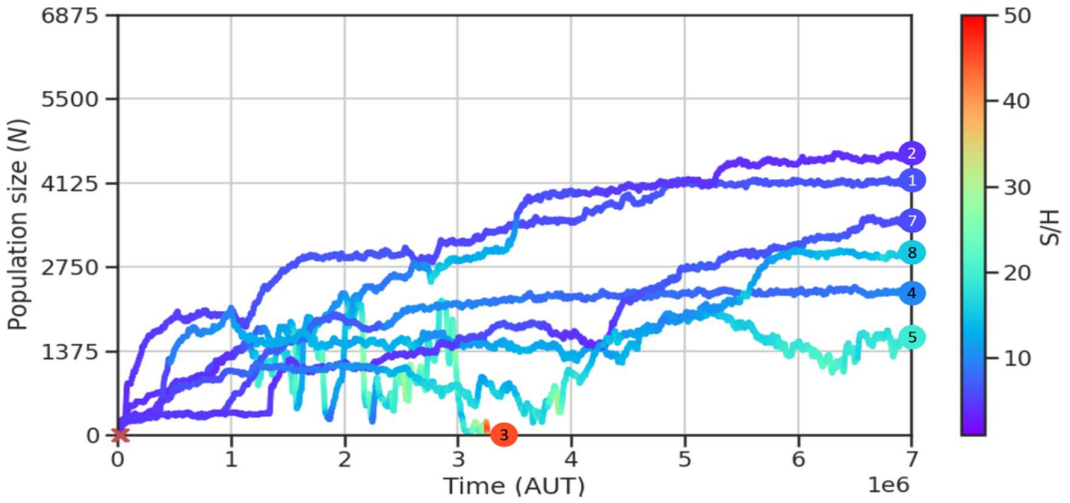


Figure 6 Timelines of populations that evolved with H-S communication and obligate BPI. *Tree types of evolutionary outcomes: a) extinction (3) b) ongoing host – symbiont conflict (5, 8) c) conflict resolution, often with cell-cycle synchronization (1, 2, 4, 7). White numbers indicate synchronized cell cycle; these populations are the best adapted to the gradient.*

When we allow populations to mutate from asexual to BPI, most present a clear pattern; they remain asexual while they are still small and primitive and switch to sexual reproduction while adapting. Five out of 7 well adapted populations evolve cell-cycle synchronization and remain sexual (fig. 7). Synchronization is achieved either with host or bi-directional control.

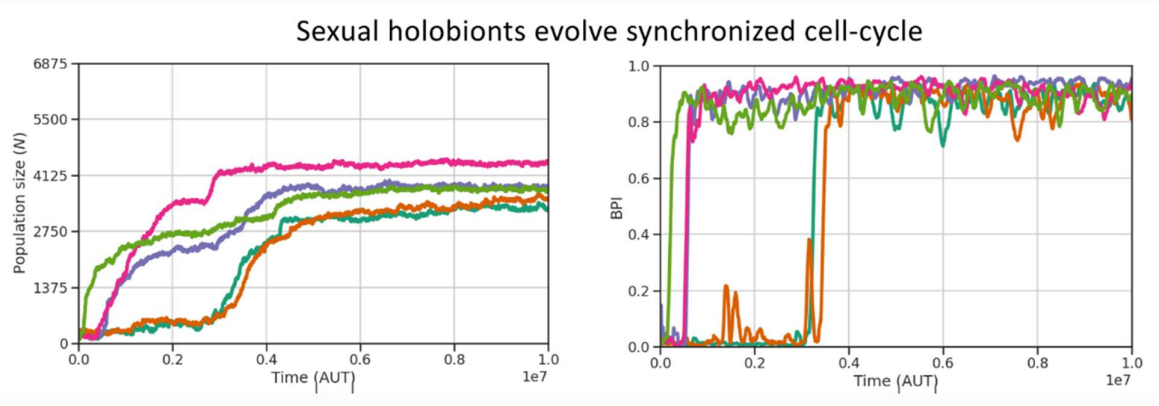


Figure 7 Populations that evolved with H-S communication and mutable reproduction. *Timelines (left) and evolution of reproductive mode (right). These populations evolved synchronized cell cycles, have stable size and stable symbiont numbers. Increase in population size and symbiont numbers is accompanied by switch to BPI. Populations E3,4,9,13&16.*

Two populations did not evolve synchronization and go through the host-symbiont conflict phase that appeared in the populations without communication. Population E14 evolved signaling, but not synchronization and the pattern of fluctuating symbiont numbers and reproductive modes appears (fig. 8a). Population E20 did not evolve signaling, and resolved the host – symbiont conflict by switching to an r- strategy, with very few, large symbionts (fig. 8b).

During the lifetime of E20 there are periods when BPI is dominant (e.g. 4-6M AUT) and periods with fluctuations that correspond to symbiont number fluctuations (e.g. 3M AUT). However, after 7M AUT, no clear pattern exists and the fraction of sexual cells seems to fluctuate somewhat randomly. It is not clear how this strain evolved and r-strategy, in contrast to all other experiments, nor the exact forces that lead to the changes in reproductive mode in this period.

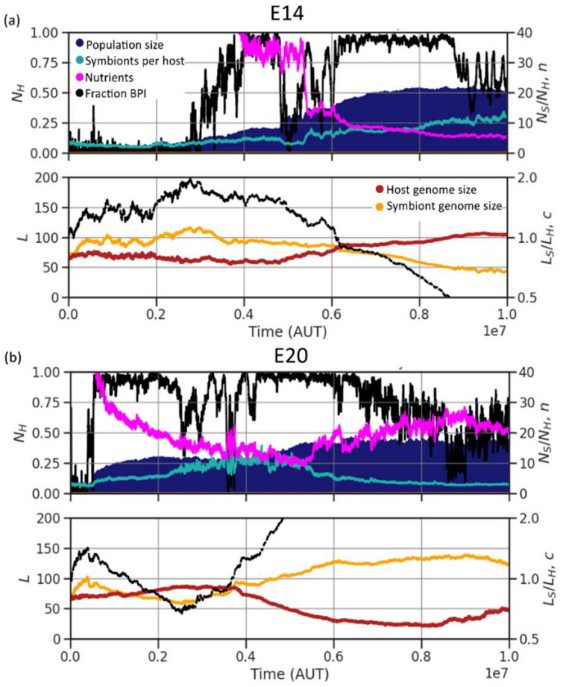


Figure 8 Populations with non – synchronized cell cycles *(a) The only population (E14) that did not resolve the host – symbiont conflict, though it evolved signaling. Symbiont numbers are controlled by periodically switching to asexuality. (b) The only population (E20) that evolved an r-strategy and no signaling. Towards the end of the simulation there is no clear pattern in the reproductive mode.*

Genome size evolution

In experiments with obligate BPI, symbionts usually evolve smaller genomes and thus faster cell cycles (table S2, fig. S1). In the populations that did not evolve synchronization, symbionts hold no or few HH genes, and have a very streamlined genome. The same happens if there are periods of host-symbiont conflicts, before synchronization appears. Since genome replication takes place explicitly in this model, symbionts with fewer HH genes are able to complete replication and divide more quickly.

However, once the cell cycle is synchronized, the holobiont if most efficient if genome replication has the same duration in host and symbionts. Otherwise, one compartment will have to wait for the other to finish replication before the holobiont divides. Consequently, these hosts and symbionts tend to evolve equal genome sizes, that share HH genes. Symbionts might increase their genome size again, if synchronization appears after a long period of conflict that results in small symbionts (e.g. table S2.2). Notably, some of these holobionts have smaller symbionts, but this is compensated by slower replication speed (e.g. table S2.7), so that the host and symbiont complete their genome duplication simultaneously.

Results 2: pre-evolved prokaryotes

BPI promotes the evolution of signaling in pre-evolved prokaryotes

So far, in the asexual model, signaling has only evolved in primitive populations and only in the presence of leakage. When holobionts reproduced asexually, pre-evolved strains do not evolve signaling, regardless of the presence of leakage; implicit co-ordination was the only outcome [25]. Here we test whether BPI can lead to the evolution of signaling between complex or genetically distant host – symbiont pairs. We also test if signaling can evolve even in the absence of leakage. The experiments discussed below took place on a 50x50 grid, with intermediate nutrient influx ($n_{influx}=30$).

We establish that bi-parental inheritance of symbionts is dangerous and implicit co-ordination unstable. When host – symbiont communication is not allowed (neither leakage nor signaling), almost all populations die out or maintain a very large amount of symbionts which deplete hosts from nutrients (fig. 9a). Symbiont genome sizes are streamlined here too, as they transfer (through duplication and deletion) most of their household genes to the host (fig. S2), allowing them to complete replication in a few timesteps.

Strain 9 is capable of stable implicit co-ordination. In contrast to all other populations, the genome of this symbiont increased and many household genes transferred there (fig. S2). Strains 7 and 11 also survive and have relatively stable populations. However, these populations have an extreme amount symbionts (40-80 per host) and are very slow growing as the environment has practically no nutrients (fig. S3).

The next experiments took place in the same conditions, but now signal peptides of genes can mutate ($\mu_s=10^{-5}$), allowing signaling from one compartment to the other. Endosymbiotic gene transfer can also take place, in either direction ($\mu_t=2*10^{-5}$). In these settings, many populations still go extinct. However, some manage to survive and even stabilize symbiont numbers through signaling (fig. 9b, table S3&S4). Strains that started with identical host and symbiont evolved a synchronized cell-cycle with very large, stable populations and stable symbiont numbers. This is the first experiment were signaling and even synchronization evolve without leakage and the first time signaling and synchronization appear in pre-evolved strains. These results resemble the outcome of eukaryogenesis and only appeared in out model when we introduced conflict through BPI.

Finally, populations evolved with leakage ($l=0.01$) in addition to signaling. Now, even more populations manage to survive and evolve signaling (fig. 9c, table S1). Some populations evolve synchronization too; mostly from identical host-symbiont strains (1-3), but also from different in the case of strain 10. In addition, these populations maintain high density and small symbionts (when they are not synchronized). In the case of synchronized cell cycles, host and genome size tend to be equal, because, as mentioned before, it is beneficial for the holobiont if genome duplication ends simultaneously in all compartments.

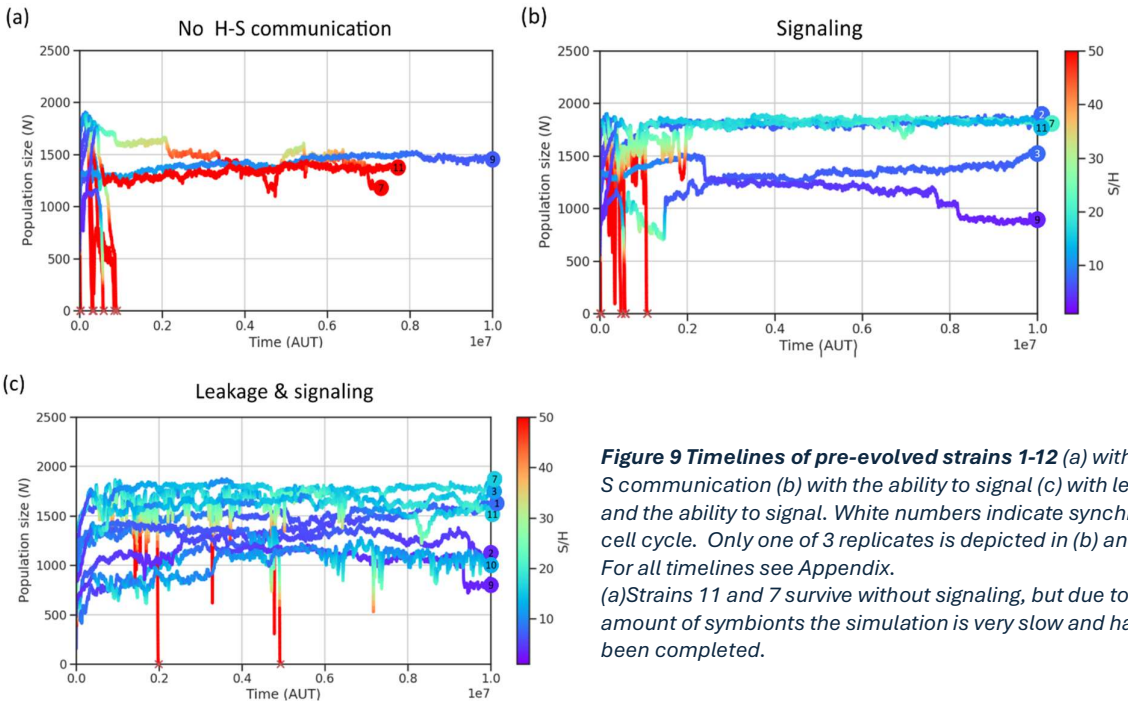


Figure 9 Timelines of pre-evolved strains 1-12 *(a) with no H-S communication (b) with the ability to signal (c) with leakage and the ability to signal. White numbers indicate synchronized cell cycle. Only one of 3 replicates is depicted in (b) and (c). For all timelines see Appendix.*
(a)Strains 11 and 7 survive without signaling, but due to the amount of symbionts the simulation is very slow and has not been completed.

The ability of signaling and the leakage seem to allow hosts to survive and restrict the spread of selfish symbionts. Indeed, the more communication we allow, the more strains survive: strains with leakage and signaling survive more compared to those with just signaling. This is in contrast to populations that reproduce asexually, where leakage and gene transfer were deleterious [25]. Another difference is the evolution of symbiont genome size, which here tends to be small and streamlined, in contrast to the large symbionts of asexually reproducing strains that suffer from leakage (fig. 10). However, not all cases of signaling are the same. In some populations signaling is weak or only appears for a short period during the evolutionary history of the population (table S4). The host – symbiont conflict is still present there, as reflected by large fluctuations in symbiont numbers and population size (e.g. table S3.11).

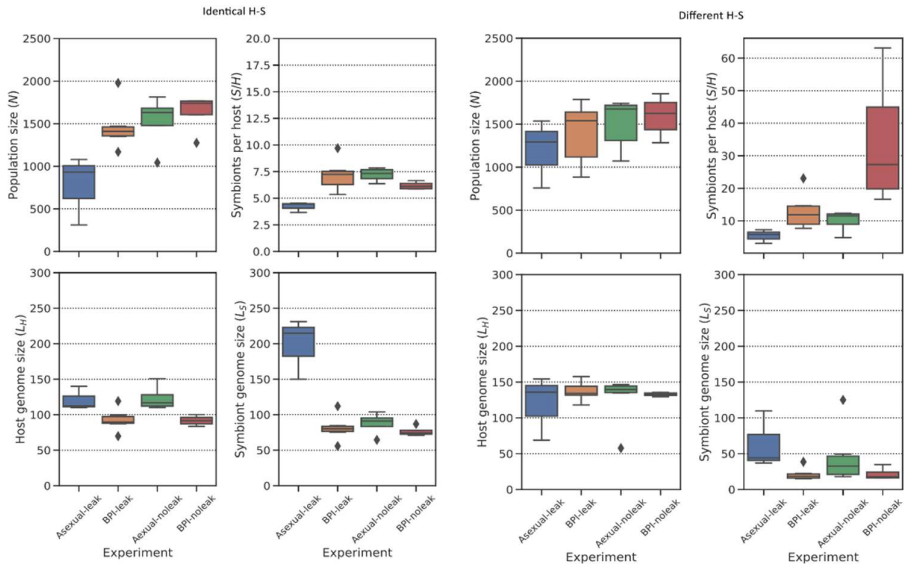


Figure 10 Overview of at 3*10^6 AUT*. Only populations that survived (and their asexual counterparts) were taken into account. Sexual strains maintain larger populations, especially in the presence of leakage.*

In general, BPI makes the outcome of endosymbiosis unpredictable between different replicates of the same experiment. Asexual populations resulted in relatively similar population sizes in independent experiments (starting with the same strains), and did not evolve signaling. With BPI, the same strain can die or evolve a very dense population with stable symbiont numbers in different simulations. The co-ordination of the cell cycle is also unpredictable; cell cycle synchronization, strong signaling, weak signaling and (rarely) implicit co-ordination can appear.

However, across different strains the results are largely the same: populations that manage to survive maintain large populations with many symbionts and asymmetric genome sizes (large host, small symbiont). Alternatively, they synchronize their cell cycle and maintain equally sized genomes.

Adaptation to the nutrient gradient: holobionts choose BPI

We now introduce holobionts 1-12 to the more diverse and challenging environment of the nutrient gradient, in order to investigate how BPI affects their adaptation. We evolve them asexually, sexually and with mutable (μ_r=*0.001)* reproduction. These experiments took place with full communication (leakage & signaling). Similar patterns as above were observed in the sexual populations: identical host – symbiont pairs evolved a synchronized cell cycle, 2 strains evolved signaling and relatively stable populations while the rest fluctuate due to the host – symbiont battle. Strain 9 evolved a stable population without any signaling in this setting as well.

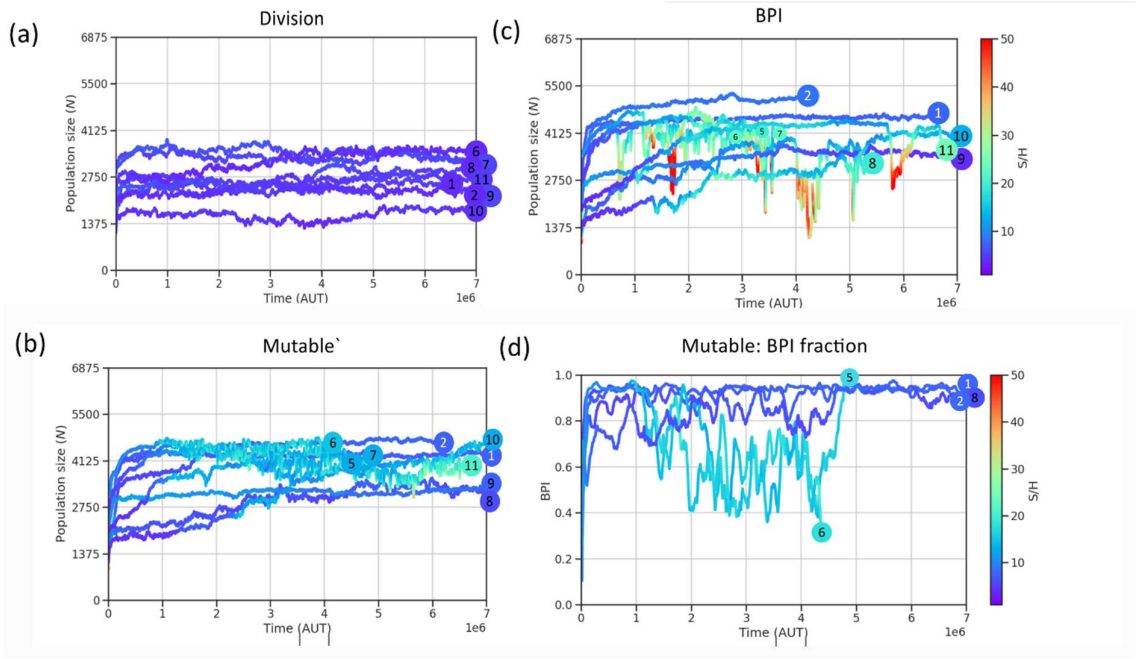


Figure 11 Pre-evolved strains on the nutrient gradient, with H-S communication *(a) Timelines of strains 1-12 on the gradient, reproducing by division. All populations are stable, with few symbionts. (b) with obligate BPI. Many population fluctuate, but populations 1,2,9&10 have resolved the conflict and are stable. White numbers indicate synchronized cell cycle. (c-d) Mutable reproduction. Populations 1,2&8 are stable and consistently opt BPI. In populations (e.g. 5&6) the fraction of holobionts reproducing sexually fluctuates. For all timelines see Appendix 2.*

Holobionts that can mutate from one reproductive mode to the other, quickly change from asexual to BPI (fig. 11d, Appendix 2). BPI provides an immediate advantage due to increased success in division: hosts with many symbionts rescue the offspring that did not

inherit any (table 3). Mutable and sexual populations are able to adapt better to the gradient and evolve higher population size (fig. 12). In strictly sexual populations, signaling between different host and symbiont appears and allows hosts to control symbionts. However, holobionts that can mutate to asexual reproduction, can restrict the spread of selfish symbionts by periodically opting out of asexual BPI (fig. 11d, Appendix 2). This solution is more accessible than signaling, though signaling did appear in population 11. The exception to this are identical host – symbiont pairs, that evolve synchronization very quickly after switching to sexual reproduction. These populations remain stable and sexual.

When prokaryotes evolve on the gradient, populations don't die out. There are two factors that contribute to that. Firstly, the gradient (25x275) is by default bigger than the 50x50 grid, allowing for larger populations with increased diversity, as strains can specialize in different nutrient concentrations. Secondly, the different sectors are somewhat reproductively isolated, as mating is only allowed between cells in the same sector. This way there can be local extinction and re-colonization from other sectors, by strains with different properties.

Remarkably, one asexual population (GAP8) also evolved signalling on the gradient (fig. S4). In contrast to all sexual populations however, here the symbiont is self- sufficient and the host does not survive without S->H signalling (fig. S5).

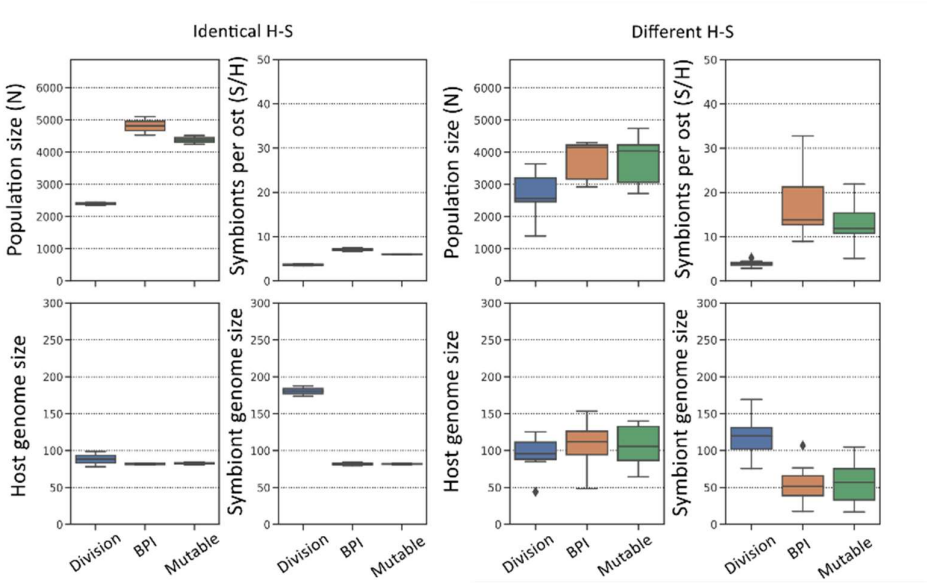


Figure 12 Overview of populations that evolved with leakage and signaling on the gradient at 2.5*10⁶ AUT. *Identical H-S strains evolve a synchronized cell cycle, and thus stable symbiont numbers end equal H-S genome sized. Non-identical strains maintain many, small symbionts. Asexual strains tend to have larger, fewer symbionts and smaller population size in all cases.*

Types and function of signaling

We have seen that signaling can restrict the spread of selfish symbionts and even completely stabilize their relationship with the host. Now, we take a closer look at the regulatory networks to understand how cell cycle synchronization is achieved, how other types of control mechanisms function and to what extent hosts and symbionts retain an autonomous regulatory network.

(a) Cell cycle synchronization resolves host-symbiont conflicts

The most successful strategy so far is cell cycle synchronization, which leads to large populations with stable symbiont numbers. There are different ways this strategy is implemented; with host or bi-directional signaling and different levels of host and symbiont autonomy. Cell cycle synchronization mostly appears with identical host-symbiont pairs. However, population PPIIC10 (strain 10), which started with a specialist host and generalist symbiont also evolved a host controlled synchronized cell cycle. This is so far the only population to do so when starting with non-identical strains, resembling the outcome of eukaryogenesis. All populations with a synchronized cell cycle maintain stable symbiont numbers, at a lower level compared to other sexual populations (fig. 12).

In the case of PPIB2 (fig. 13a, f – strain 2, no leakage) synchronization appears after a short period of intense host symbiont conflict (fig. S6), before 10^6 AUT. By the end of the simulation, signaling is bi-directional, with core genes in host and symbiont being activated by the other compartment. Here the host controls cell cycle, however, some core host genes are activated by the symbiont, upon the host's signal. Indeed, this strain cannot survive without signaling, as both compartments lost their autonomy (fig 13a).

Population PPIIC10 (fig. 13b, g – strain 10, leakage) evolves synchronization at around $8x10^6$ AUT, after a long period of host – symbiont conflict (fig. S7), as full genetic integration is difficult between unrelated genomes. In this case, the host remains fully autonomous and takes over full regulatory control of the symbiont, who cannot survive without host signaling.

Population PPIIC1 (fig. 13c-e, h – strain 1, leakage) also evolved bi-directional synchronization. In this population both host and symbiont maintain their autonomy, can perform their cell cycle independently (fig 13d) and have their own checkpoints for genome replication. However, the symbiont's checkpoint for replication is highly degenerate (fig. S8); the interaction at the end of the genome is no longer of low affinity, causing the symbiont to exit the S-stage prematurely. Thus, the symbiont cannot effectively regulate its growth in poor nutrient conditions (fig. 13d). The host however keeps the symbiont in the S-phase until it has replicated its own genome, and then gives the signal for synchronized division. This strain can survive with host signaling, though maintains a bigger population when bi-directional signaling is allowed (fig. 13b).

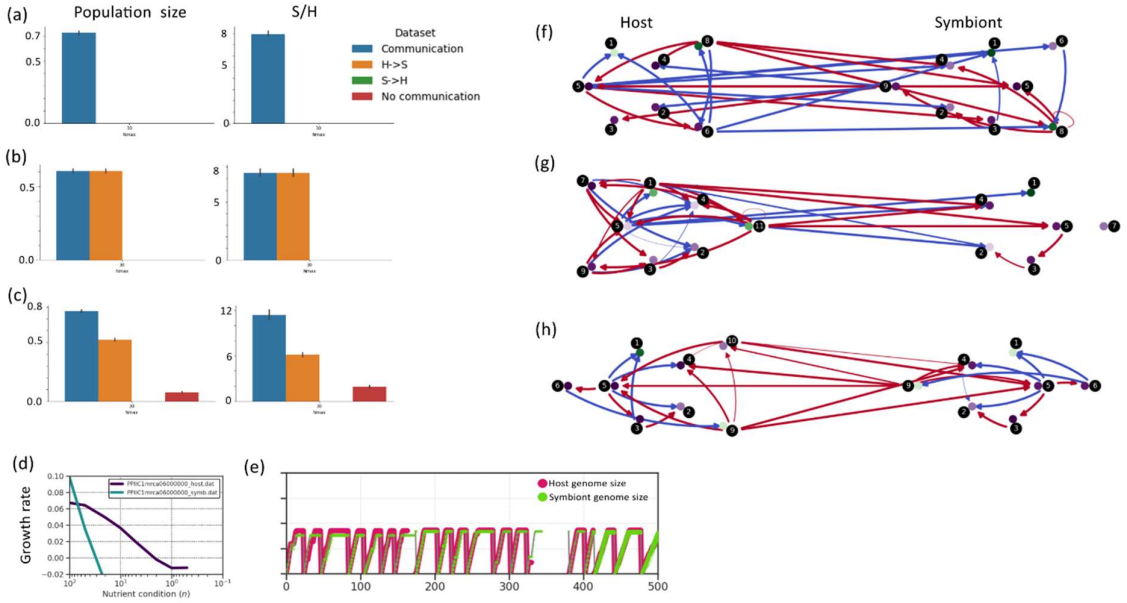


Figure 13 Strains with synchronized cell cycle. *(a-c) Survival of populations under different conditions: full signaling, no signaling, H->S signaling and S->H signaling. Leakage background is the same as in original populations: off for PPIB2 (a) on for PPIIC10 (b) & PPIIC1 (c). (f-h) Regulatory networks that give rise to synchronized cell cycle. Interactions show the aggregate regulatory effect (binding probability times regulatory weight) of gene types (black disks) on loci (colored disks adjacent to the gene types). Arrow thickness shows probability of interaction and color type of interaction (red: activation, blue: deactivation). Genes are colored based on their activation threshold (green < 0, purple > 0). (f) PPIB2: bi-directional control, low autonomy. Hosts and symbionts do not survive without bi-directional signaling. (g) PPIIC10: host controlled. Symbionts require signaling from the host to survive. (h) PPIIC1: bidirectional signaling, high autonomy. Host to symbiont signaling is most important, but cells are fitter with bi-directional signaling. (d) Growth curve of PPIIC1: host is an efficient regulator, symbiont only survives in nutrient rich conditions. (e) Cell cycle progression of PPIIC1 in host and symbiont: measured by how much of the genome has been replicated (genome size can be read off from the maxima). Genome duplication and division happen simultaneously for host and symbionts.*

(b) Symbiont waits for host genome duplication

Population PPIC7 has largely resolved the conflict, resembling synchronized strains, though it maintains relatively many and small symbionts (fig. 14). It evolved from strain 7 (different host and symbiont) without product leakage. Both host and symbiont remain autonomous (fig. 14b) and can regulate their own cell cycle. However, host g9, which regulates its own S-phase, activates symbiont g7, which regulates the symbiont's S-phase. This way, the symbiont remains in S-phase together with the host. g9 only turns off after a low affinity interaction at the end of the genome. In order for the host to progress to G2 and M, g9 needs to be off for at least 4 consecutive timesteps. However, 2 timesteps are enough for the symbiont, thus their cell cycle is not perfectly synchronized, though symbionts do stall their division for a long time (fig 14c).

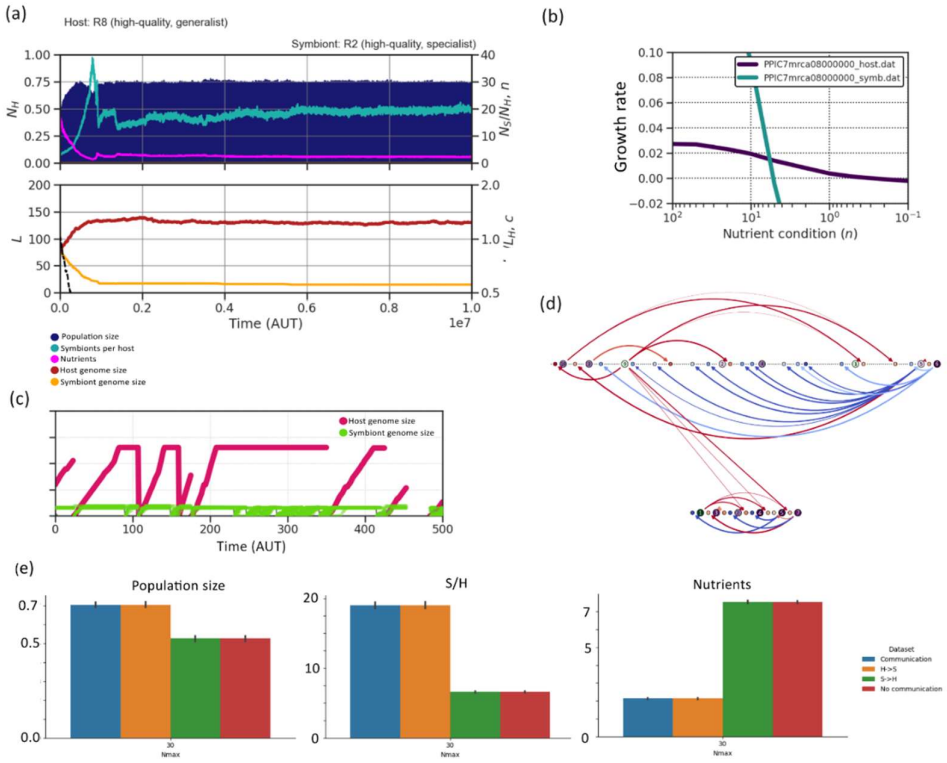


Figure 14 Timeline (a), growth curves of MRCAs (b), cell cycle co-ordination (c) and genome (d) of strain PPIC7. *(a) The population maintains stable, though high, numbers of symbionts. (b) growth curves of host and symbiont MRCA (t=8*10⁶ AUT). Both host and symbiont remain autonomous. The host is a slow generalist and the symbiont a fast specialist that only survives in rich environments. (c) Cell cycle progression of host and symbiont: measured by how much of the genome has been replicated (genome size can be read off from the maxima). Both compartments replicate their genome at the same time and the symbiont stalls it division, but often divides before the host (see text). (d) Genome of host (top) and symbiont (bottom) MRCA at 8*10⁶ AUT. While the host remains in the S-phase, it activates the symbiont's S-phase genes too.*
HH genes are denoted as dots. Big circles: regulatory genes, colored based on their activation threshold (green < 0, purple > 0). Small circles: binding sites, colored based on their activity (red: activation, blue: deactivation). Arrow thickness shows probability of interaction and color type of interaction (red: activation, blue: deactivation)
(e) Population size, symbionts per host and nutrients under different communication conditions. Host signaling allows symbionts to survive in low nutrient, by keeping them in the S-phase long enough to duplicate their genome. Without host signaling, the strain survives with implicit co-ordination, close to the nutrient concentration of their equilibrium.

(c) Host delays the symbiont's cell cycle in distantly related pairs

In distantly related host and symbionts, genetic integration is not easy, with the regulatory networks and binding site sequences having diverged significantly. In some cases, the most accessible solution to the host is to interfere with the symbionts' cell cycle and delay its progression. This signaling is often weak and not stable over time, as the symbionts that escape the host's influence can spread faster. Both host and symbiont remain completely self-sufficient and can survive without any signaling (fig. 15e-f). In these cases, the symbiont's genome remains small with no or very few HH genes, to allow fast growth.

In population PPIC11 (fig. 15a-b – strain 11, no leakage) the host's weak signals activate g4 and g5, preventing the symbiont from completing its cell cycle while the host is in the S-phase. Weak signaling from host g9 appeared at 2.2M AUT and became stronger over time.

Indeed, the conflict de-escalates from that point onwards, though still present. Host signaling lowers symbiont numbers in this strain (fig. 15e).

In population PPIIB10 (fig. 15c-d – strain 10, leakage) the host has genes that are always on and target symbiont genes, possibly to interfere with its cell cycle. When signaling is not allowed, this holobiont maintains more symbionts, however signaling from the symbionts is also required to keep symbiont their numbers low (fig. 15f).

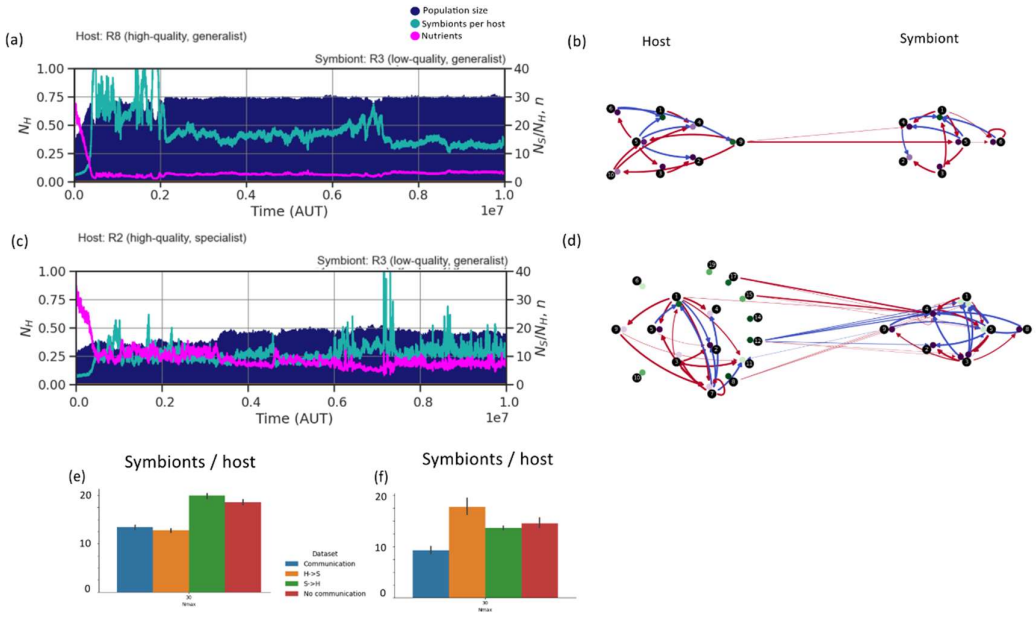


Figure 15 Strains with weak signaling. *Timeline of PPIC11(a) and PPIIB10 (c) Gene regulatory network of PPIC11 (b) and PPIIB10 (d). Host – symbiont conflict is present, as symbiont numbers fluctuate sometimes decreasing host population size. Symbionts per host in PPIC11(e) and PPIIB10 (f) with full signaling, no signaling, H->S signaling and S->H signaling. Leakage background is the same as in original populations: off for PPIC11 on for PPIIB10. Signaling decreases symbiont numbers in these strains.*

<u>Host's dependance on symbiont genes prevents invasion from foreign symbionts</u>

When populations with fast growing symbionts are allowed to mate with healthy ones, they 'infect' them with their selfish symbionts and kill them (table 5). Populations that evolved bi-directional synchronization and the host lost its autonomy (e.g. PPIB2 and PPIIA2) are resistant to invasion of foreign symbionts. These hosts depend on their native symbionts to activate 1 or more the cell cycle's core genes. Thus they are incompatible with foreign symbionts die if their own symbionts are outcompeted, preventing foreign ones from invading the population.

	PPIIC1 (bi-dir sync)	PPIIA3 (bi-dir sync)	PPIIB7 (host signaling)	PPIIA9 (implicit)	PPIIC10 (host sync)	PPIIA11 (host signaling)	PPIB1 (host sync)	PPIB2 (bi-dir sync)	PPIC2 (bi-dir sync)
PPIIA2 Invasive symbionts			Symbionts invade and kill native host population						
PPIIB6 Invasive symbionts									

Table 5 Result of symbiont invasion experiments. *Strains with host signaling are prone to invasion from foreign symbionts. Synchronized stains, and especially bi-directional ones, are resistant to invasion. HH genes were not taken into account for these experiments, so that small symbionts don't kill their host right away.*

Speciation and phylogeny

The environment of the gradient allows diversification and even speciation to take place in the population. While in asexual populations symbionts are bound to their hosts and follow their phylogeny (fig. 16b), symbionts of sexual populations do now necessarily follow the hosts. Horizontal symbiont transfer introduced by BPI allows one strain of symbiont to take over the whole population (fig. 16c). When cells that are located on the border between sectors divide, their offspring can inhabit a grid point of the neighboring sector. These new cells mate with cells of the neighboring sector, allowing symbionts to jump from sector to sector and possibly spread over the whole gradient.

However, this is not a strict rule, as there are populations where symbionts have a phylogeny that reflects their hosts'. In these populations some host – symbiont incompatibility might exist, either explicit (through signaling) or implicit (incompatible growth curves). Primitive populations also follow similar patterns (fig. S9). Especially bi-directionally synchronized strains (L1 & L2) present deep branching in host and symbiont clades. In these population explicit incompatibility is possible to evolve, as network compatibility between host and symbiont is needed for a viable holobiont.

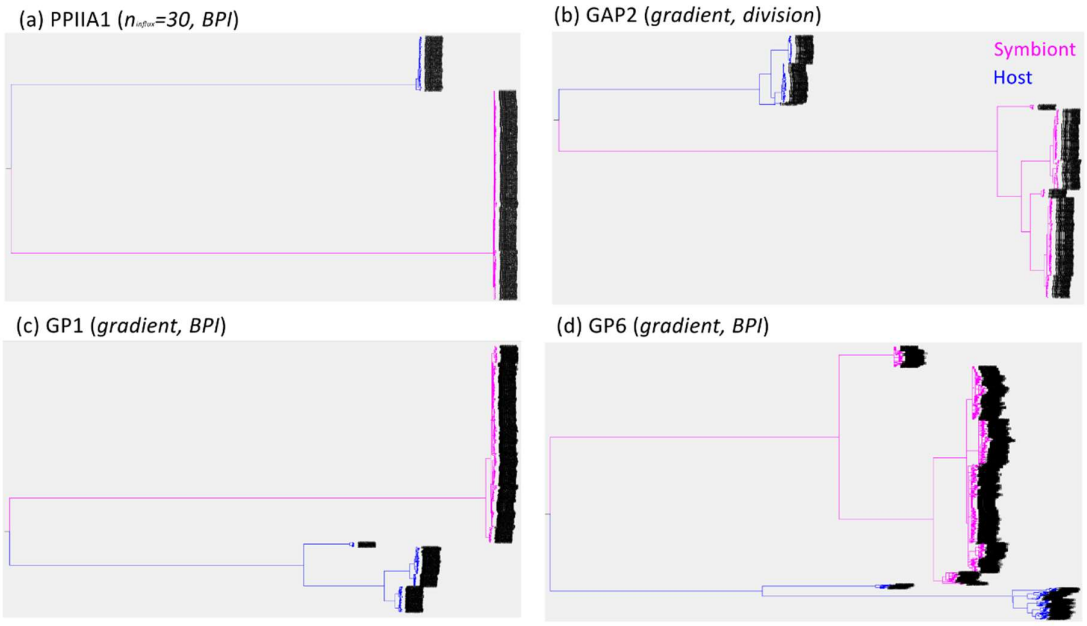


Figure 16 Examples of phylogenetic trees of pre-evolved strains. *(a) Populations that evolve in constant nutrient conditions do no diversify. One strain of host and symbiont is dominant (b) In asexual populations, the host and symbiont follow the same branching, as symbionts are bound inside their hosts. (c) In sexual populations, where symbionts mix, it is possible that one symbiont strain is present in diverging hosts. (d) In sexual populations speciation is also possible, with different host-symbiont strains specialized in different nutrient concentrations.*

Discussion

In this study we have attempted to understand the implications of bi-parental inheritance of symbionts, using an evolvable system with a complex, non-linear genotype-phenotype map. We modeled hosts and symbionts individually, allowing different levels of selection (symbiont, host, holobiont) to emerge from the modeling framework. This approach can give some insight into the evolution of selfishness, co-operation and regulatory integration of initially autonomous replicating entities.

Periodic asexuality rescues the population from selfish symbionts

A prominent result of our simulations is that obligate BPI can lead to the accumulation of selfish symbionts and drive a population to extinction. BPI destabilizes the formation of host – symbiont growth rate equilibrium, by providing an advantage to faster and better regulating symbionts. Surprisingly, when reproduction is mutable, populations often choose BPI due to the immediate advantage it provides: holobionts with few symbionts can rescue their offspring when mating with hosts with more symbionts. This feature is in line with the assumption that primordial eukaryotes would not have mechanisms that ensure fair symbiont distribution and cell fusion might have been a source of symbiont acquisition [24]. Because of this advantage, holobionts evolve sexual reproduction early on, and later revert to asexuality to keep symbionts in check. As LECA was capable of both sexual and asexual reproduction, it is plausible that occasional cell fusion and BPI could have taken place without immediate danger for the population. Moreover, we observe switching of the reproductive mode created by opposing selective pressures, analogous to the model proposed by [18]. The benefit of BPI in our model (efficient reproduction) is different from proposed benefits of BPI in modern eukaryotes (escape from Muller's ratchet). However, in both models UPI/asexuality is selected in order to restrict selfish symbionts and results in a continuous turnover of reproductive mode.

Synchronization resolves host – symbiont conflict

The potential to evolve signaling between hosts and symbionts changes the evolutionary outcomes profoundly, especially when starting with primitive hosts and symbionts. Primitive holobionts have an inefficient, malleable regulatory network, identical in host and symbiont. With identical initial networks, most strains evolve cell-cycle synchronization, which stabilizes symbiont numbers. This can be achieved with a dominant host that subjugates the symbiont or with a bi-directionally connected regulatory network where both partners are needed for a viable holobiont. Synchronized strains, and especially the bi-directional ones, are not invadable by foreign selfish symbionts, due to host-symbiont incompatibility. While the conditions of experiments with primitive holobionts are not analogous to eukaryogenesis, where host and symbionts where complex and distantly related, they can provide insights to the possible solutions to this system. Conflict resolution and transition of individuality from autonomous host and symbiont to holobiont with non-autonomous compartments can take place, even if there is selection for selfishness. Indeed, this strategy is favored precisely because of, and as the means to escape, the conflict bi-parental inheritance introduces.

Cell cycle synchronization in asexual strains might have a disadvantage compared to other co-ordination strategies: holobionts that, due to stochastic distribution, carry too many or too few symbionts cannot regulate their number. Symbionts always divide with the host and keep their number stable. However, with bi-parental inheritance symbiont numbers of the offspring can be corrected, as symbionts from 2 holobionts are mixed and redistributed.

Conflict introduced by BPI promotes adaptation and regulatory integration in pre-evolved strains

To investigate the dynamics of complex and divergent host – symbiont pairs, that reflect eukaryogenesis more closely, pre-evolved strains were used to form holobionts. The evolutionary outcomes of these simulations vary widely, and the initial properties of the host and the symbiont define to a large degree the fate of the holobiont. Pairs with symbionts more efficient than the host are often doomed to die under BPI, as the host is unable to keep up with the symbiont. Very few of these pairs are able to survive with BPI and without host – symbiont communication, even if the initial pair forms a stable equilibrium.

Asexual populations of pre-evolved strains have not evolved explicit cell cycle co-ordination through signaling (with one exception). However, the host – symbiont conflict that sex introduces results in the evolution of signaling that allows the population to survive and maintain low symbiont numbers. The added stress due to product leakage increases the chance that holobionts survive and evolve signaling, while it was deleterious for asexual holobionts.

Despite the ability to signal, a lot of strains still die out. The most popular survivors are strains with identical host and symbiont networks (that die out without signaling), as regulatory integration and cell cycle synchronization is easily accessible. Strains that survive BPI without communication are also more likely to keep symbionts in check, as they stay alive long enough to evolve signaling. Remarkably, host controlled synchronization evolved in a strain with divergent host and symbiont, resembling the outcome of eukaryogenesis.

In general, the conflict introduced by BPI creates pressure for genetic integration of host and symbiont and gives access to more regulatory solutions; asexual strains usually have stationary populations and don't evolve host – symbiont signaling. On the other hand, sexual populations that survive, even if they suffer from selfish symbionts, have larger populations than their asexual counterparts that suffer from the effects of leakage.

Our results suggest that cell fusion that leads to bi-parental inheritance of endosymbionts could have appeared before the complete domestication of mitochondria. Firstly, the ability to periodically opt out of sexual reproduction seems to be enough to rescue a population from death due to selfish symbionts. Secondly, BPI promotes regulatory integration between distant strains, providing that the foundations for the evolution of signaling are present.

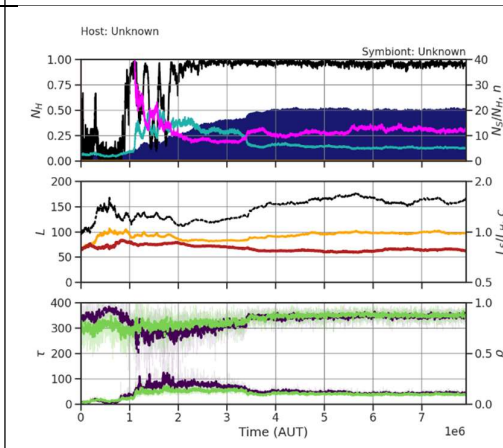
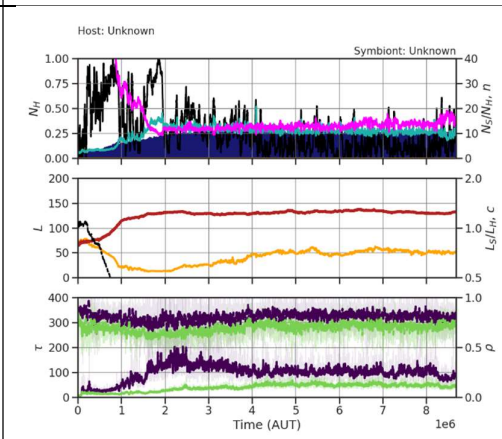
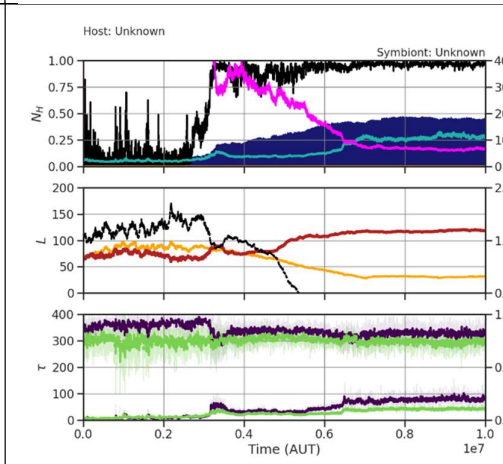
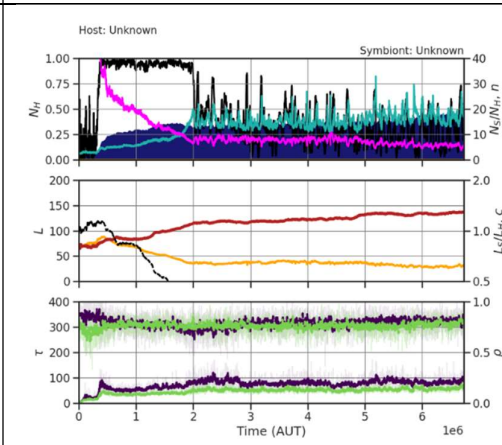
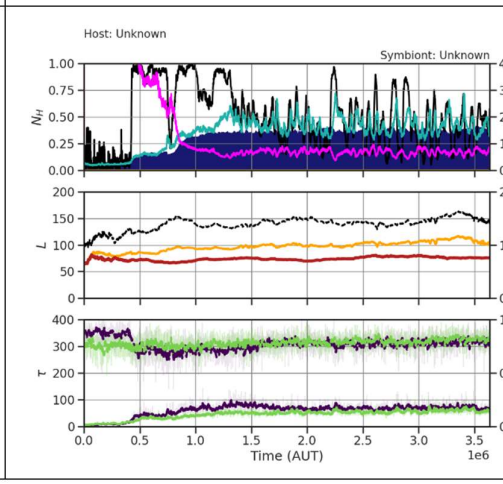
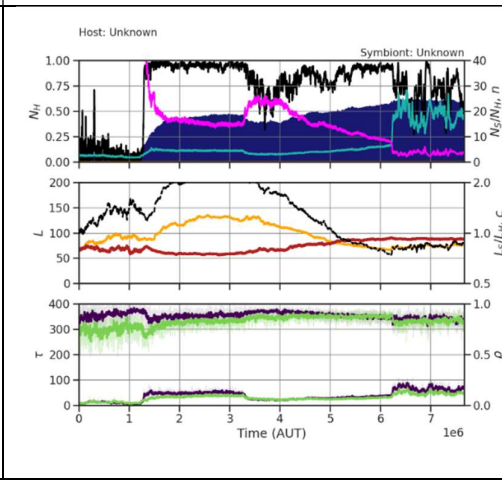
References

1. Betts, H.C., Puttick, M.N., Clark, J.W. *et al*. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat Ecol Evol* **2**, 1556–1562 (2018). https://doi.org/10.1038/s41559-018-0644-x
2. Parfrey LW, Lahr DJ, Knoll AH, Katz LA. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci U S A*. 2011;108(33):13624-13629. doi:10.1073/pnas.1110633108
3. Spang, A., Saw, J., Jørgensen, S. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521, 173–179 (2015). https://doi.org/10.1038/nature14447
4. Koumandou VL, Wickstead B, Ginger ML, van der Giezen M, Dacks JB, Field MC. Molecular paleontology and complexity in the last eukaryotic common ancestor. *Crit Rev Biochem Mol Biol*. 2013;48(4):373-396. doi:10.3109/10409238.2013.821444
5. Poole AM, Gribaldo S. Eukaryotic origins: How and when was the mitochondrion acquired?. Cold Spring Harb Perspect Biol. 2014;6(12):a015990. Published 2014 Jul 18. doi:10.1101/cshperspect.a015990
6. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet*. 2004;5(2):123-135. doi:10.1038/nrg1271
7. Lane, N., Martin, W. The energetics of genome complexity. Nature 467, 929–934 (2010). https://doi.org/10.1038/nature09486
8. Speijer D, Lukeš J, Eliáš M. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc Natl Acad Sci U S A*. 2015;112(29):8827-8834. doi:10.1073/pnas.1501725112
9. Bergstrom CT, Pritchard J. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. Genetics. 1998;149(4):2135-2146. doi:10.1093/genetics/149.4.2135
10. Hoekstra RF. Nucleo-cytoplasmic conflict and the evolution of gamete dimorphism. In: Togashi T, Cox PA, eds. The Evolution of Anisogamy: A Fundamental Phenomenon Underlying Sexual Selection. Cambridge University Press; 2011:111-130.
11. Hadjivasiliou Z, Pomiankowski A, Seymour RM, Lane N. Selection for mitonuclear co-adaptation could favour the evolution of two sexes. Proc R Soc Lond B Biol Sci. 2012;279(1734):1865–72. pmid:22158961
12. Christie JR, Schaerf TM, Beekman M. Selection against heteroplasmy explains the evolution of uniparental inheritance of mitochondria. PLoS Genet. 2015;11(4):e1005112. Published 2015 Apr 16. doi:10.1371/journal.pgen.1005112
13. Sharpley MS, Marciniak C, Eckel-Mahan K, et al. Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell. 2012;151(2):333-343. doi:10.1016/j.cell.2012.09.004
14. Radzvilavicius AL. Evolutionary dynamics of cytoplasmic segregation and fusion: Mitochondrial mixing facilitated the evolution of sex at the origin of eukaryotes. *J Theor Biol*. 2016;404:160-168. doi:10.1016/j.jtbi.2016.05.037
15. Barr CM, Neiman M, Taylor DR. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol. 2005;168(1):39-50. doi:10.1111/j.1469-8137.2005.01492.x
16. Breton S, Stewart DT. Atypical mitochondrial inheritance patterns in eukaryotes. Genome. 2015;58(10):423-431. doi:10.1139/gen-2015-0090

17. Sato M, Sato K. Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *Biochim Biophys Acta*. 2013;1833(8):1979-1984. doi:10.1016/j.bbamcr.2013.03.010
18. Greiner S, Sobanski J, Bock R. Why are most organelle genomes transmitted maternally?. *Bioessays*. 2015;37(1):80-94. doi:10.1002/bies.201400110
19. Heitman J. Evolution of sexual reproduction: a view from the Fungal Kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biol Rev*. 2015;29(3-4):108-117. doi:10.1016/j.fbr.2015.08.002
20. Speijer D. What can we infer about the origin of sex in early eukaryotes?. *Philos Trans R Soc Lond B Biol Sci*. 2016;371(1706):20150530. doi:10.1098/rstb.2015.0530
21. Havird JC, Hall MD, Dowling DK. The evolution of sex: A new hypothesis based on mitochondrial mutational erosion: Mitochondrial mutational erosion in ancestral eukaryotes would favor the evolution of sex, harnessing nuclear recombination to optimize compensatory nuclear coadaptation. *Bioessays*. 2015;37(9):951-958. doi:10.1002/bies.201500057
22. Radzvilavicius AL. Mitochondrial genome erosion and the evolution of sex. *Bioessays*. 2016;38(10):941-942. doi:10.1002/bies.201600148
23. Tilquin A, Christie JR, Kokko H. Mitochondrial complementation: a possible neglected factor behind early eukaryotic sex. *J Evol Biol*. 2018;31(8):1152-1164. doi:10.1111/jeb.13293
24. Radzvilavicius AL, Blackstone NW. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. *J R Soc Interface*. 2015;12(111):20150584. doi:10.1098/rsif.2015.0584
25. von der Dunk SHA, Hogeweg P, Snel B. Intracellular signaling in proto-eukaryotes evolves to alleviate regulatory conflicts of endosymbiosis. *PLoS Comput Biol*. 2024;20(2):e1011860. Published 2024 Feb 9. doi:10.1371/journal.pcbi.1011860
26. McCutcheon, J., Moran, N. Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**, 13–26 (2012). https://doi.org/10.1038/nrmicro2670
27. Gabaldón T, Huynen MA. From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PLoS Comput Biol*. 2007;3(11):e219. doi:10.1371/journal.pcbi.0030219
28. Roger AJ, Muñoz-Gómez SA, Kamikawa R. The Origin and Diversification of Mitochondria. *Curr Biol*. 2017;27(21):R1177-R1192. doi:10.1016/j.cub.2017.09.015
29. von der Dunk SHA, Snel B, Hogeweg P. Evolution of Complex Regulation for Cell-Cycle Control. Genome Biol Evol. 2022;14(5):evac056. doi:10.1093/gbe/evac056
30. von der Dunk SHA, Hogeweg P, Snel B. Obligate endosymbiosis enables genome expansion during eukaryogenesis. Commun Biol. 2023;6(1):777. Published 2023 Jul 25. doi:10.1038/s42003-023-05153-x

Appendix 1

Table S1 Timelines of primitive holobionts, no communication, mutable reproduction (H)

N		N	
1		5	
2		9	
4		10	

Population size
Symbionts per host
Nutrients
Host genome size
Symbiont genome size

Table S2.1 Timelines of primitive holobionts, communication (L)

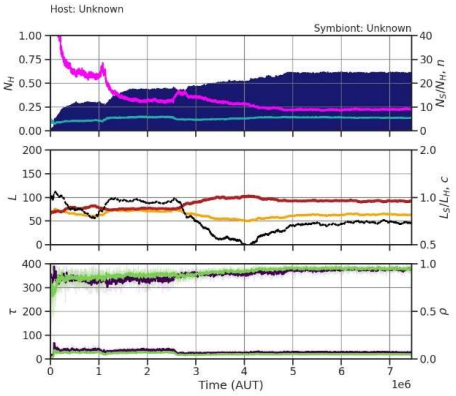
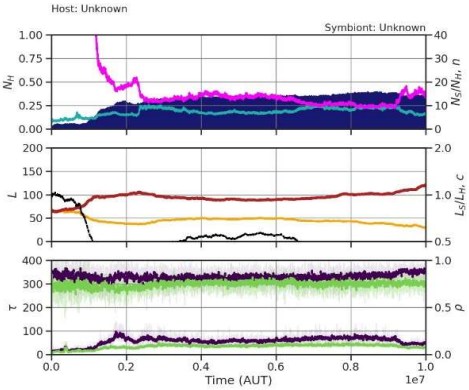
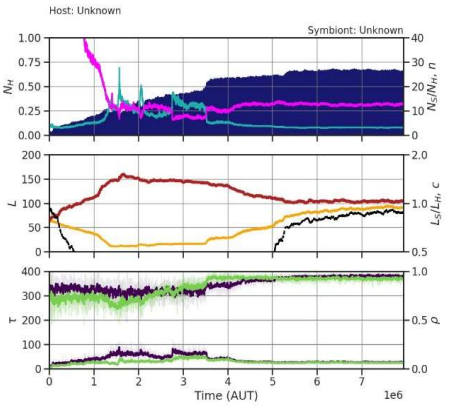
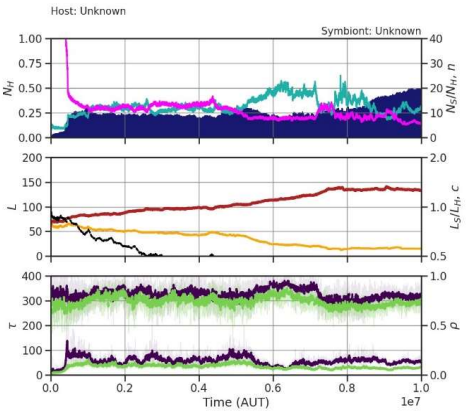
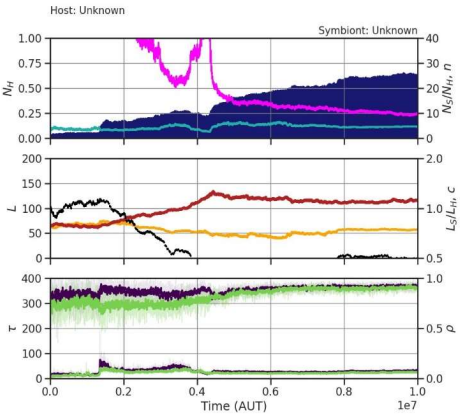
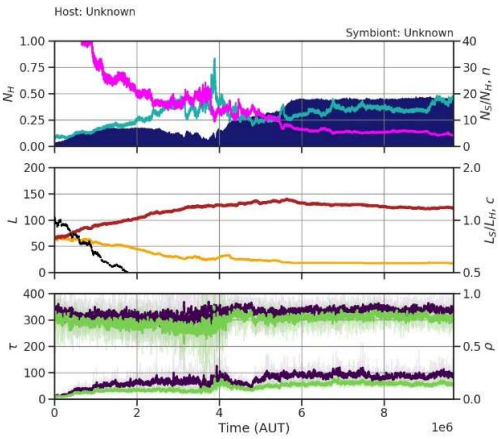
N	Synchronized	N	Non - synchronized
1		4	
2		5	
7		8	

Table S2.2 Primitive holobionts, communication (L) – genomes of most recent common ancestors

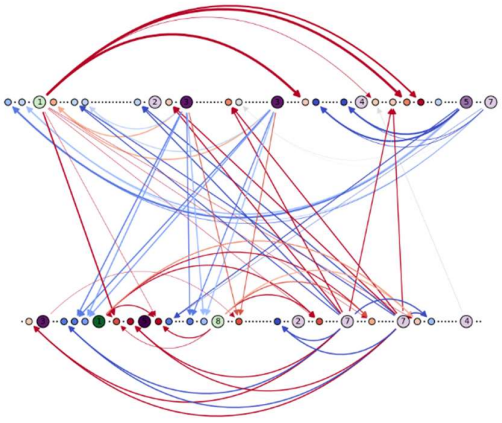
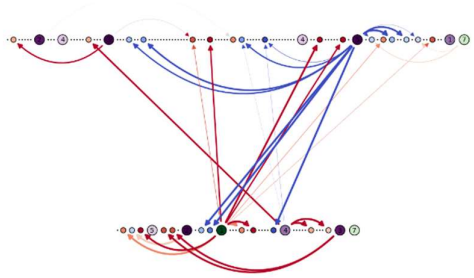
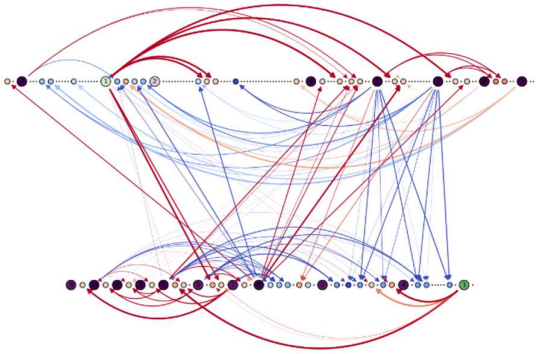
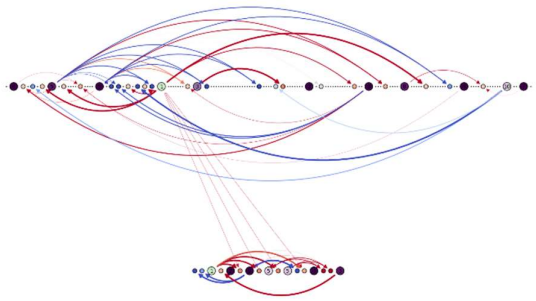
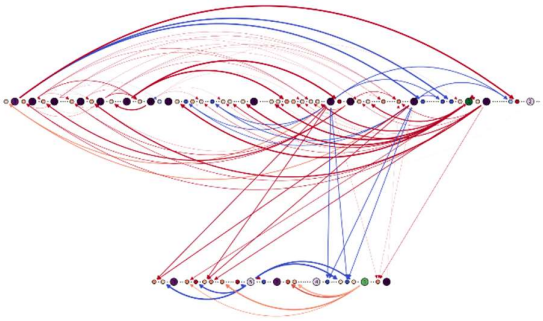
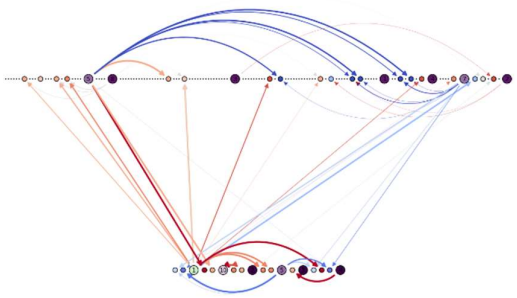
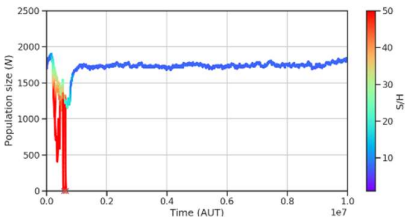
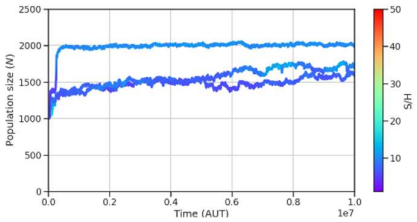
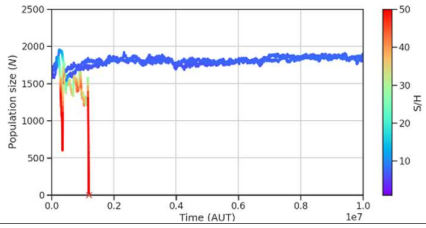
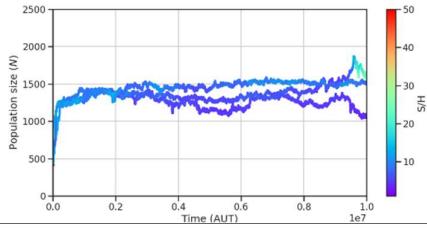
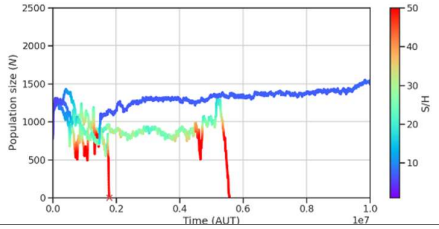
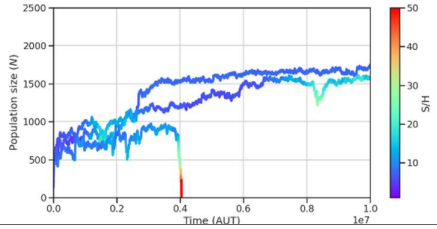
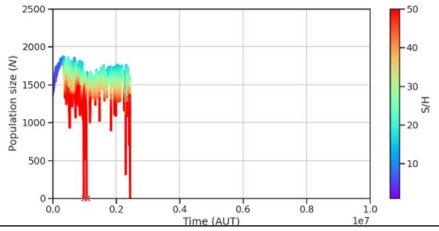
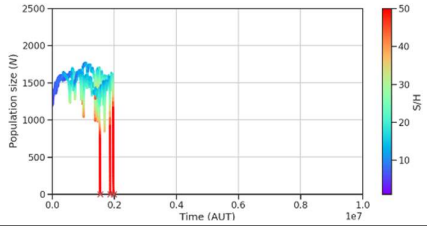
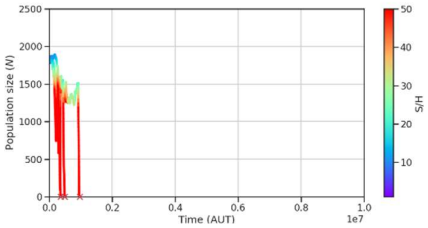
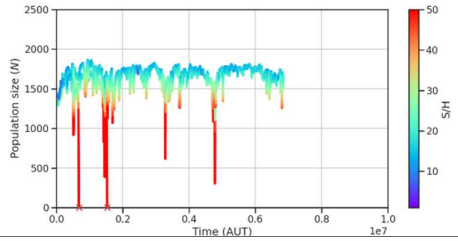
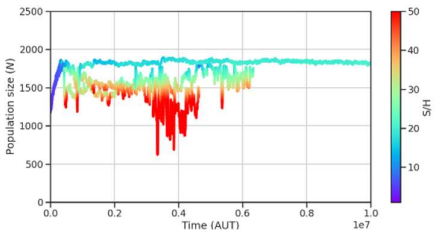
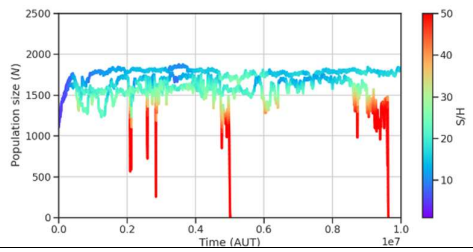
N	Synchronized	N	Non - synchronized
1		4	
2		5	
7		8	

Table S2.2 Genomes of MRCAs. HH genes are denoted as dots. Big circles: regulatory genes, coloured based on their activation threshold (green < 0, purple > 0). Small circles: binding sites, colored based on their activity (red: activation, blue: deactivation). Arrow thickness shows probability of interaction and color type of interaction (red: activation, blue: deactivation).

Table S3 Timelines of pre-evolved strains

Strain	Signaling (PPI)	Leakage & Signaling (PPII)
1		
2		
3		
5		
6		
7		

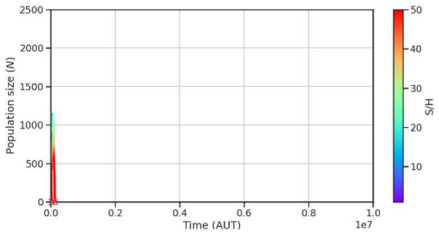
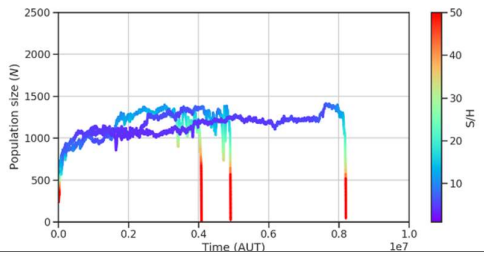
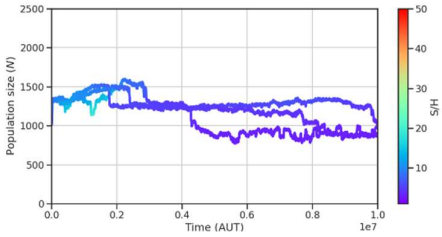
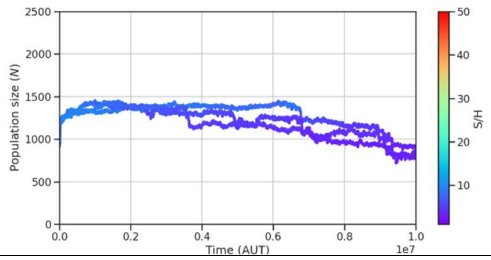
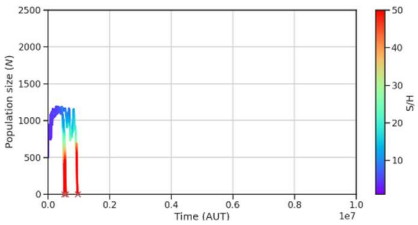
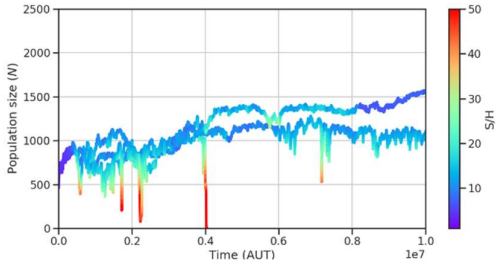
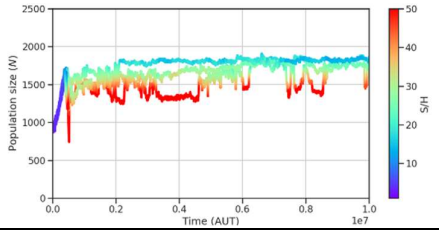
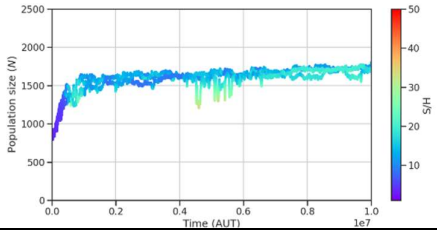
8		
9		
10		
11		
12	Dies immediately	Dies immediately

Table S4 Evolutionary outcomes of pre-evolved host symbiont pairs

	Death & unstable population	Implicit coordination (stable)	Weak signaling / disappears (unstable)	Strong signaling or synchronization
Intermediate nutrients – no communication	10/11 (ESI7&ESI11 alive)	ESI9	(blocked)	(blocked)
Intermediate nutrients – signaling – no leakage	20/33	PPIABC9	PPIA&B7 *PPIA&B11 (?)*	PPIB1* PPIB2 PPIC2* PPIC3* PPIC7 PPIC11
Intermediate nutrients – signaling – leakage	14/30	PPIIB2 PPIIABC9	PPIIA&B1 PPIIB10 PPIIB11 *PPIIC11 (?)*	PPIIC1* PPIIC2 PPIIA3* PPIIB3 PPIIC10* PPIIB7 PPIIC7 PPIIA11
Gradient signaling – leakage	4 (no deaths, fluctuating symbionts)	GP9		GP1* GP2* GP10 GP11

Table S4 Evolutionary outcomes of pre-evolved host symbiont pairs. The first column contains populations that died out or are highly unstable due increasing symbiont numbers. The second column contains populations that evolved stable implicit co-ordination. The only populations that systematically opt this solution come from strain 9 (R3 host, R2 symbiont). The third column contains populations that evolved signaling but it is either weak, disappears and reappears during their evolution or are highly unstable despite signaling. In the case of 11 (?), relatively weak signaling only appears for a short period. The last column contains populations with a synchronized cell cycle (*) or strong signaling. These populations are mostly stable. The names of the experiments are thus:
***ESI**: intermediate nutrients, no communication, **PPI(A-C)**: intermediate nutrients, signaling, **PPII(A-C)**: intermediate nutrients, signaling, leakage, **GP**: nutrient gradient, signaling, leakage. Numbers refer to initial host – symbiont pairs (see methods).*

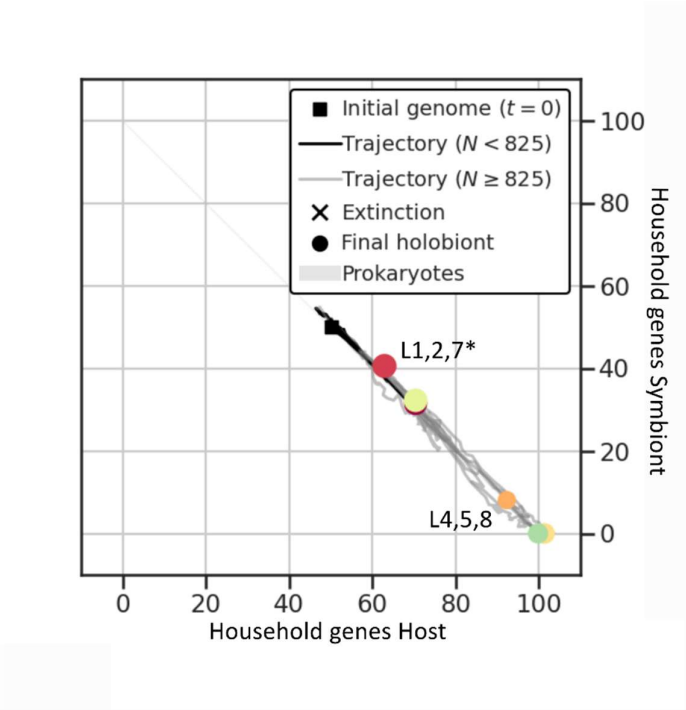


Figure S1 Household gene (HH) evolution in primitive eukaryotes. Populations L1, L2 and L7, that have synchronized cell cycle, share HH genes between host and symbiont. In non-synchronized populations the host carries most of the HH genes.

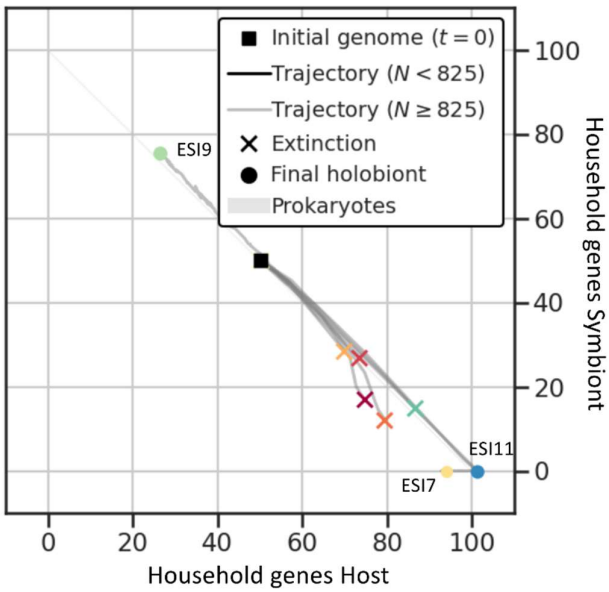


Figure S2 Household genes evolution in pre-evolved strains (no communication). With the exception of ESI9, symbionts evolve smaller genomes, with fewer HH genes. Hosts need to carry more HH genes, for a viable holobiont.

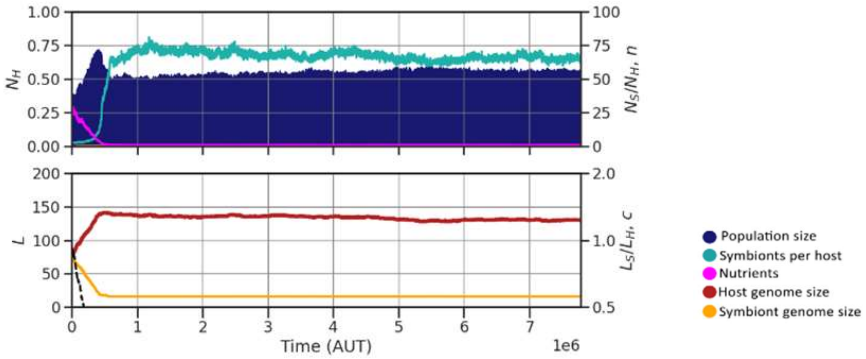


Figure S3 Timeline of population ESI11 (pre-evolved strain 11, no communication). This population carries around 70 symbionts per host, and survives with the minimum amount of nutrients.

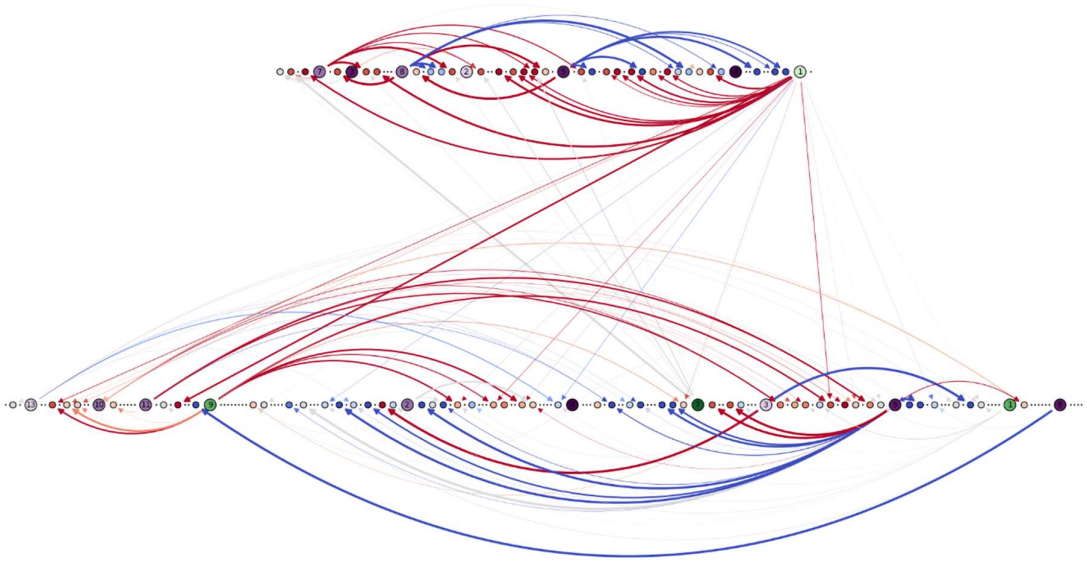


Figure S4 Genome of population GAP8, common ancestor at 8M AUT. For notation see table S2.2. Here host signaling improves population survival, but symbiont signaling is necessary for the hosts survival. This is the only pre-evolved population that evolved signaling while reproducing asexually.

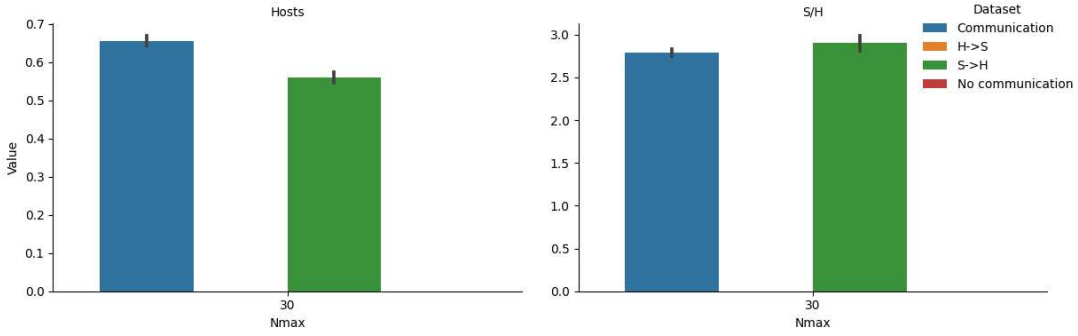


Figure S5 Survival of population GAP8 under different conditions: full signaling, no signaling, H->S signaling and S->H signaling. Host signaling improves population (host) size. The host cannot survive without its symbiont.

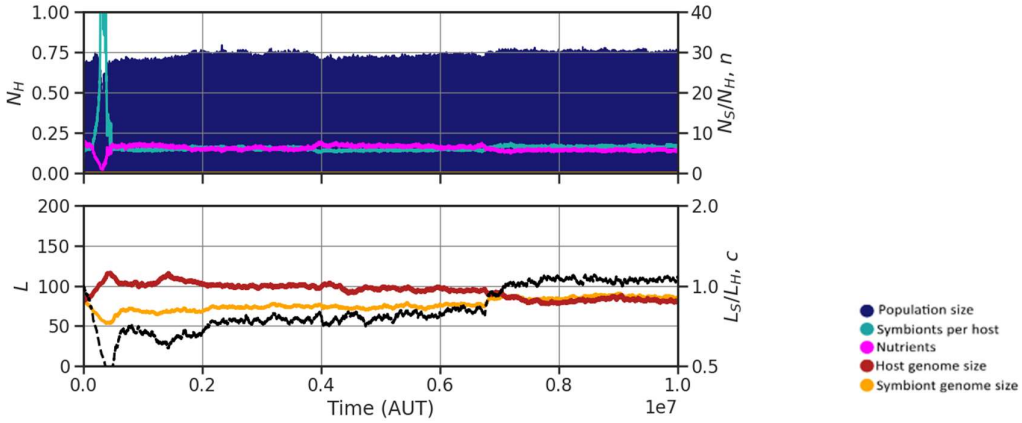


Figure S6 Timeline of population PPIB2. Synchronization appeared after a short period of host – symbiont conflict.

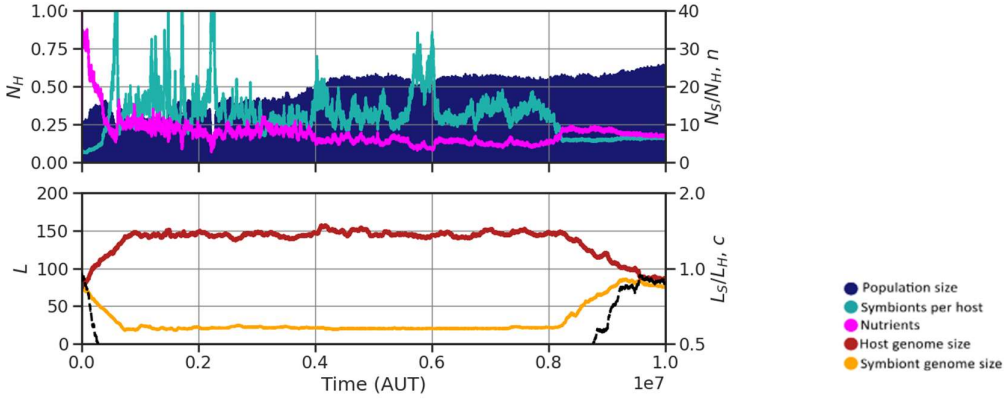


*Figure S7 Timeline of population PPIIC10. This strain evolved a synchronized cell cycle around 8.2*10^6 AUT, after a long period of host – symbiont conflict.*

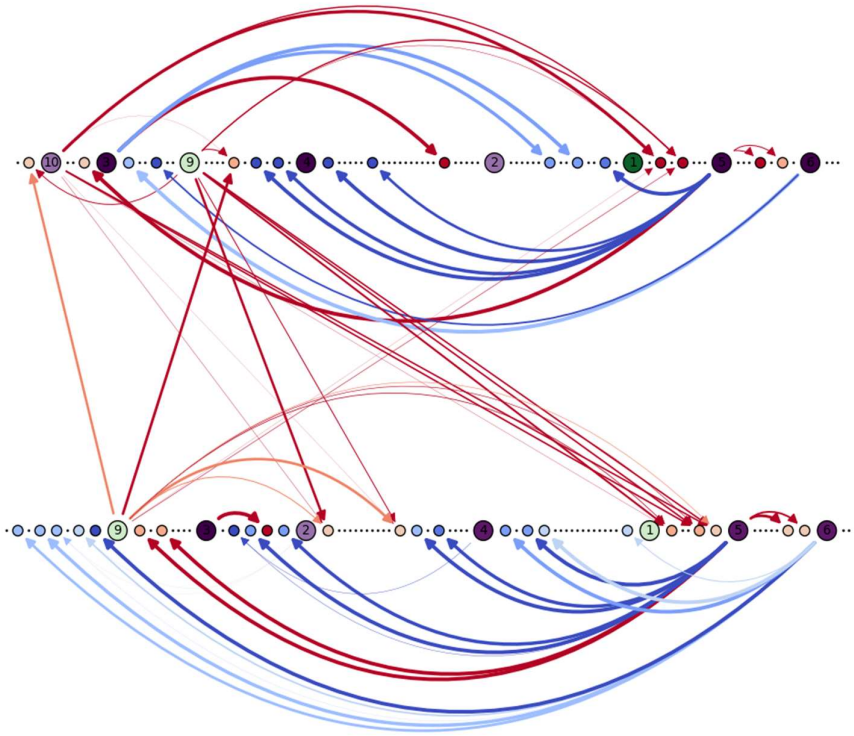


Figure S8 Genome of PPIIC1 host (top) and symbiont (bottom). For notation see table S2.2. There are checkpoints for genome replication at the end of the genomes: low affinity interactions that turn on G2-phase and M-phase. However, the symbiont's checkpoint for replication is highly degenerate; the interaction at the end of its genome is no longer of low affinity, causing the symbiont to exit the S-stage prematurely in nutrient poor conditions.

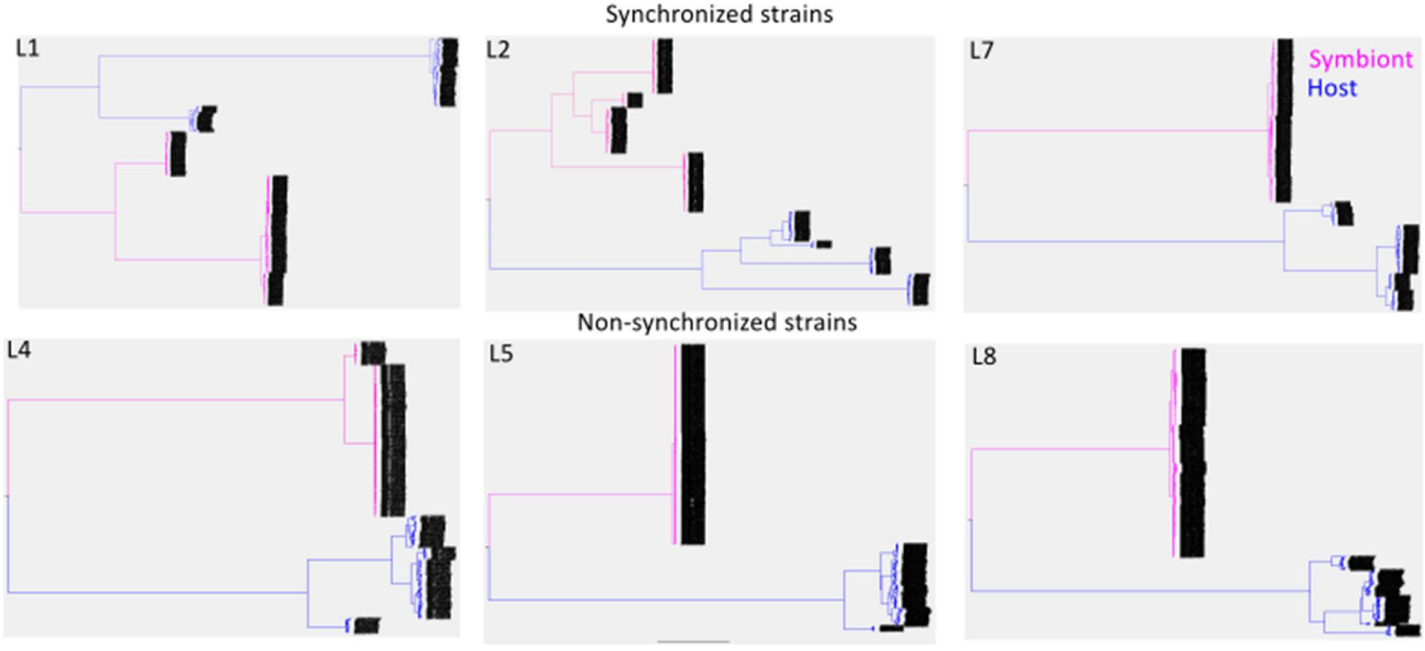
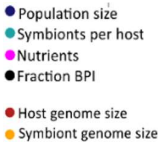
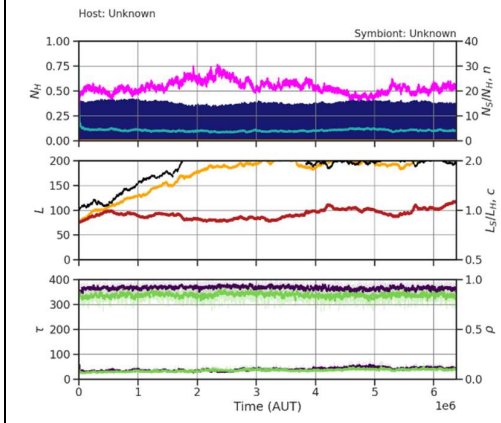
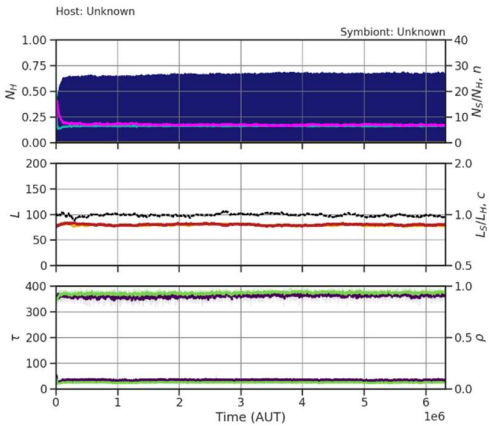
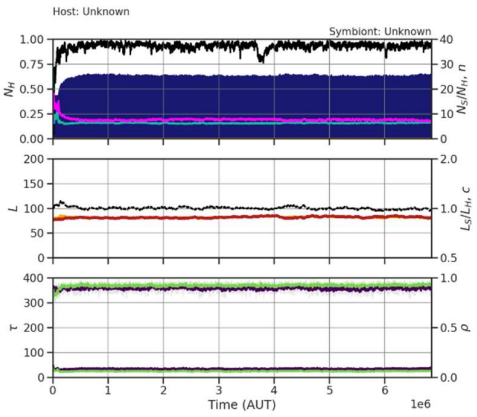


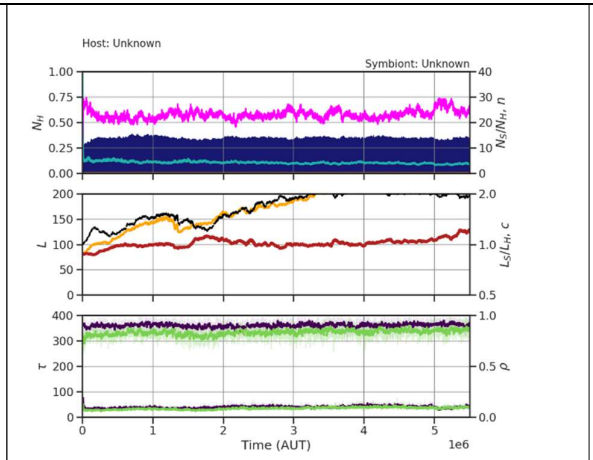
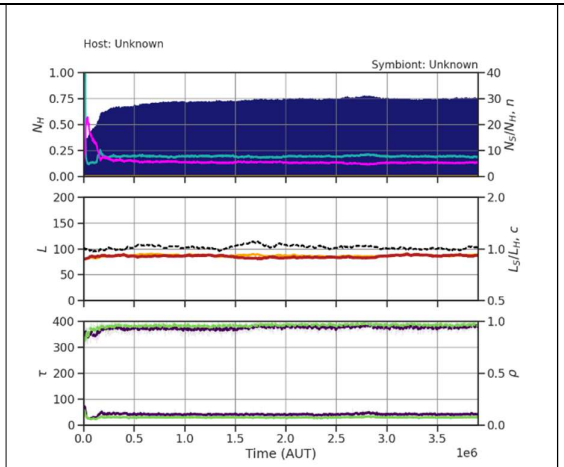
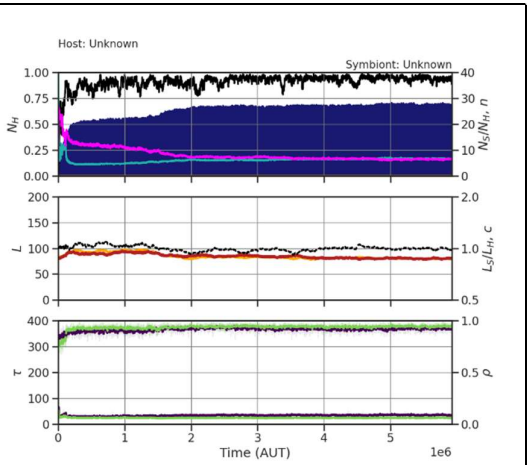
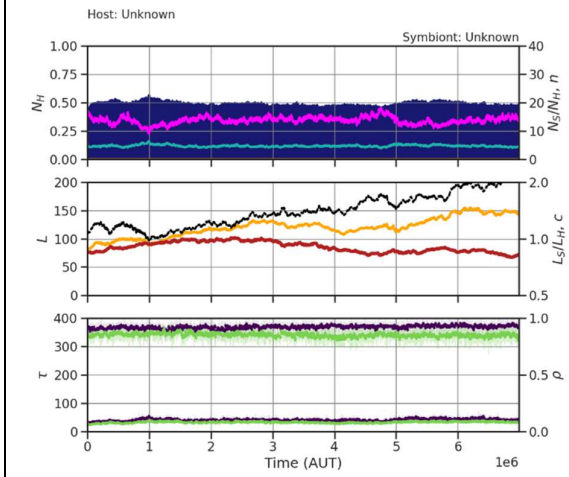
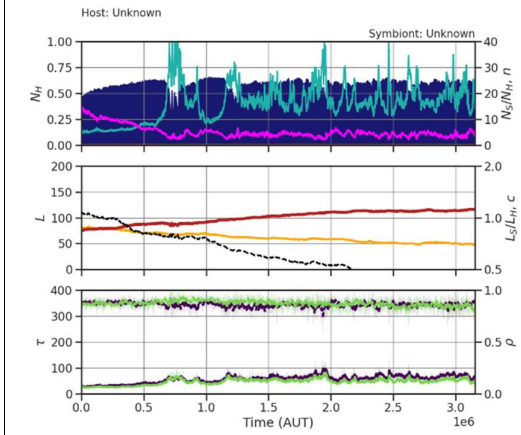
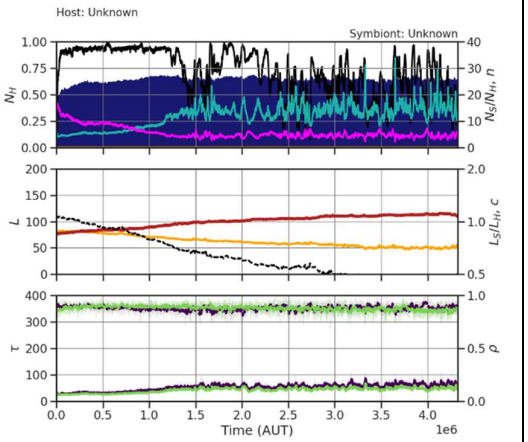
Figure S9 Phylogenetic trees of populations L. Symbiont trees are a lot flatter than host trees, as mixing allows strains to takeover the population regularly. However, speciation is possible, especially in synchronized strains.

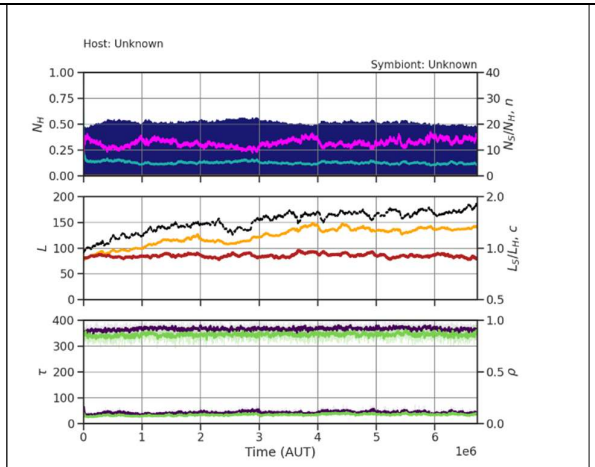
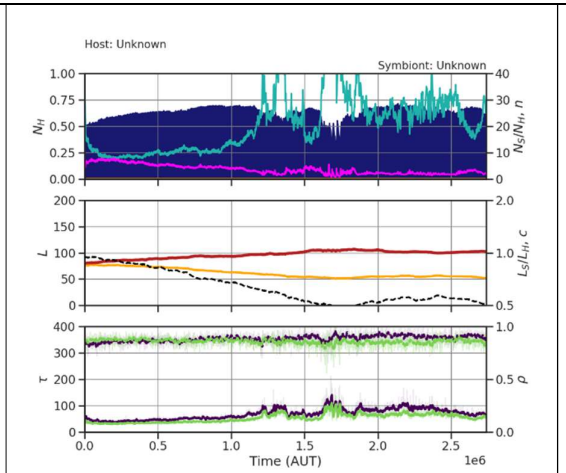
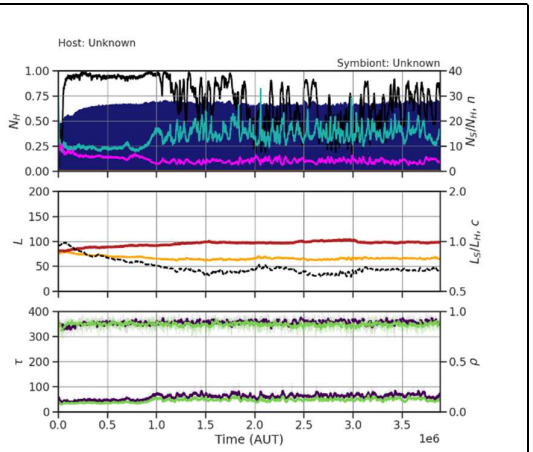
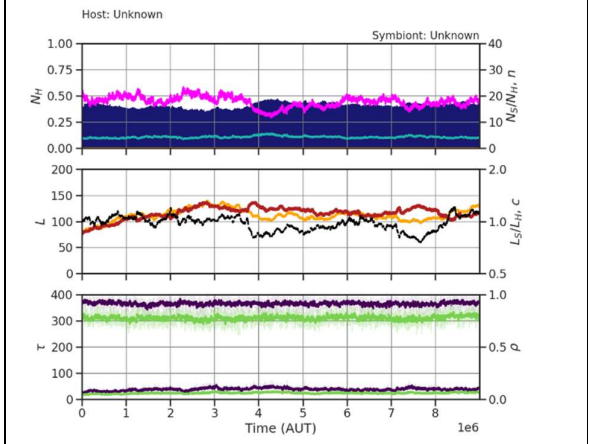
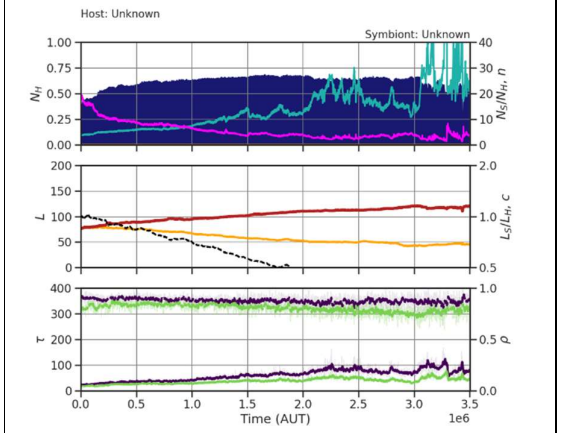
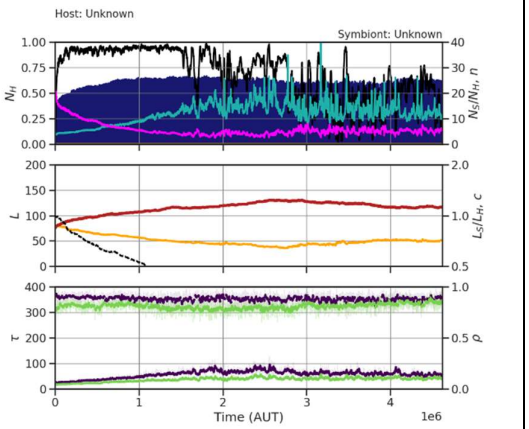
Appendix 2.1

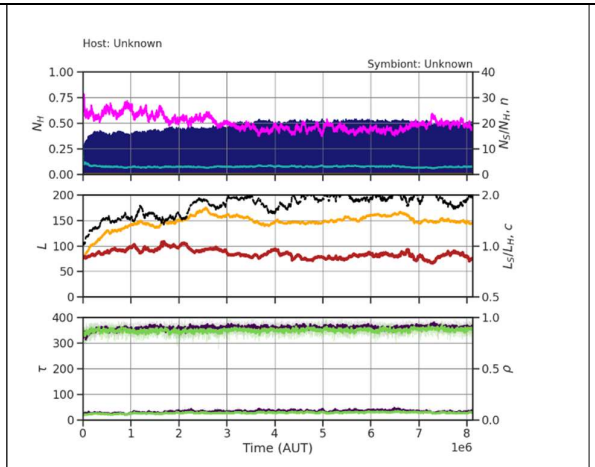
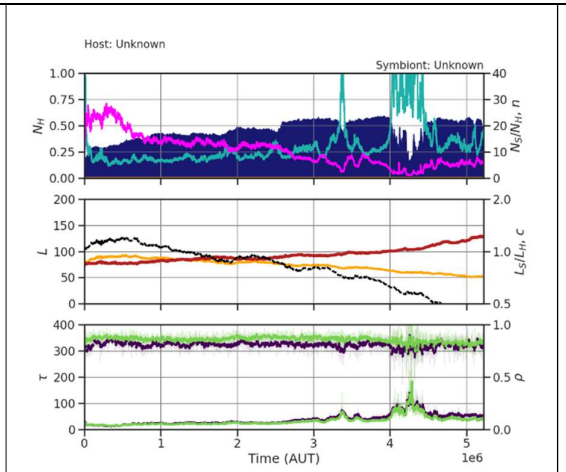
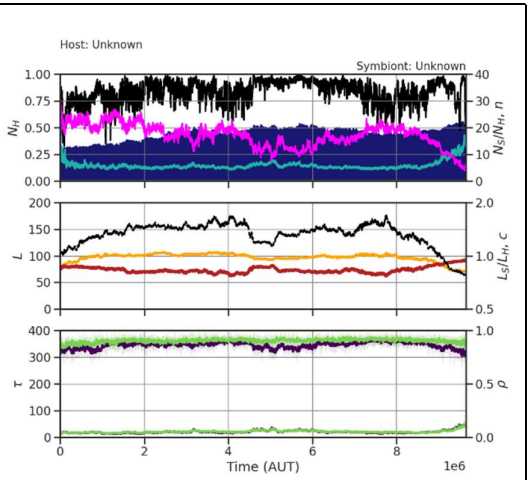
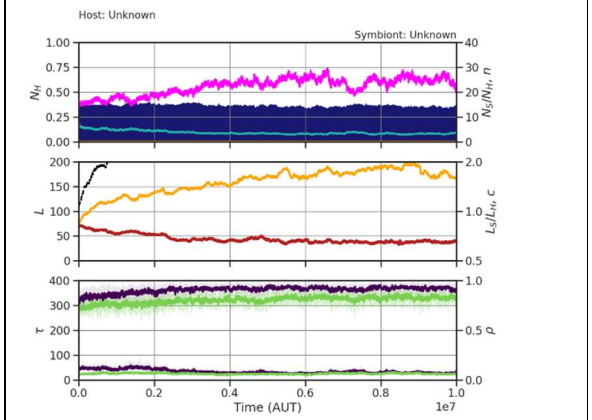
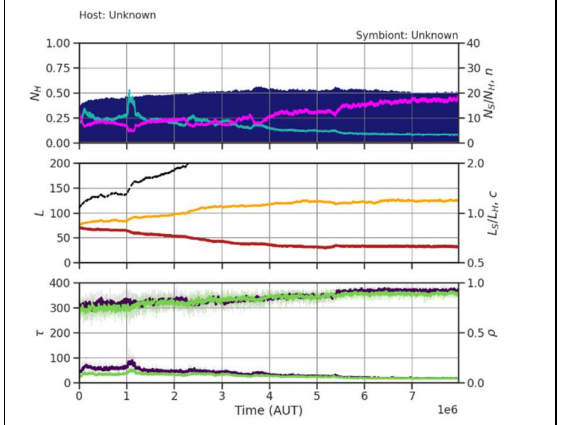
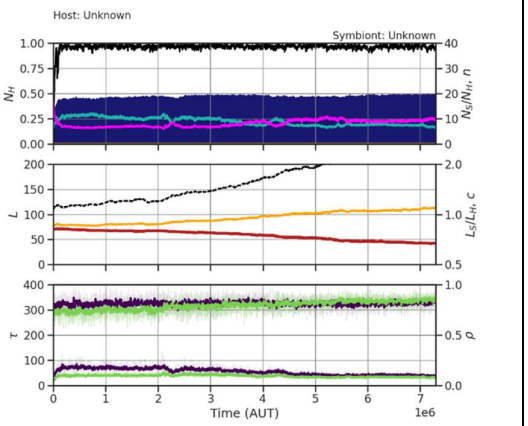
Timelines of pre-evolved strains on the nutrient gradient.

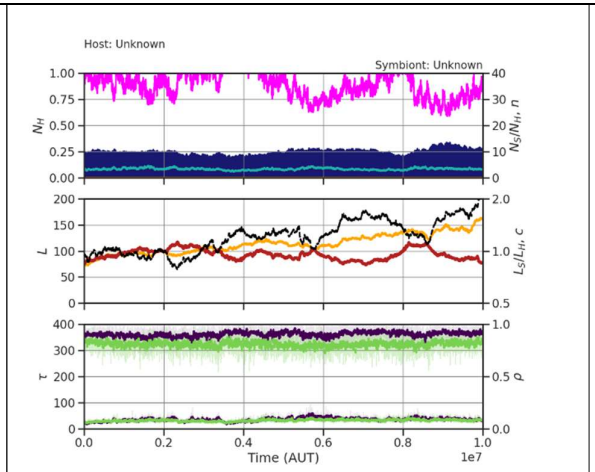
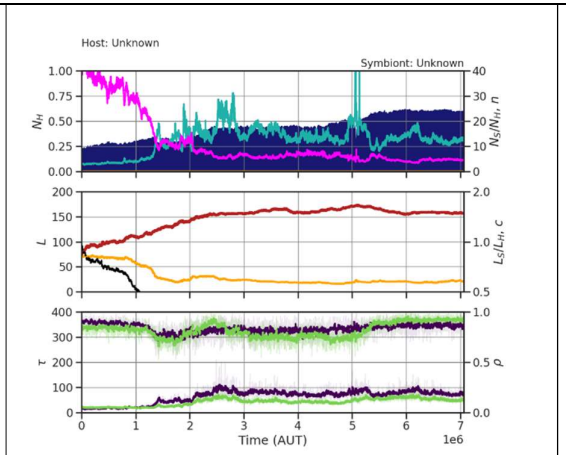
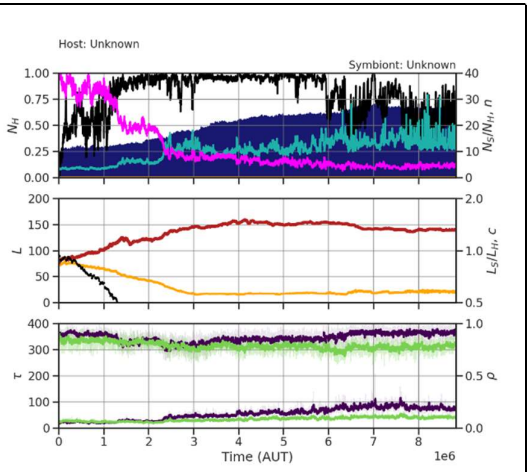
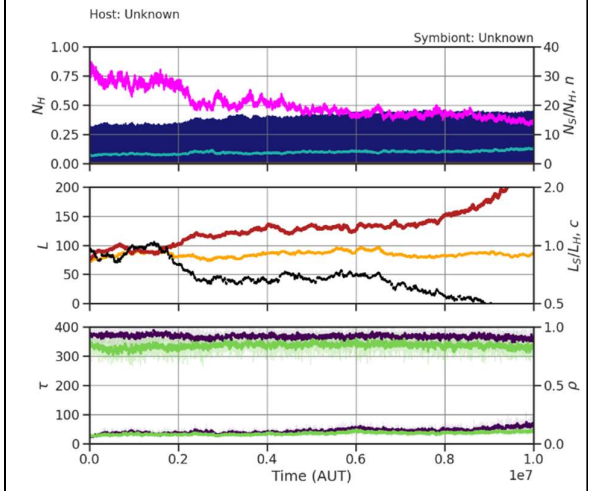
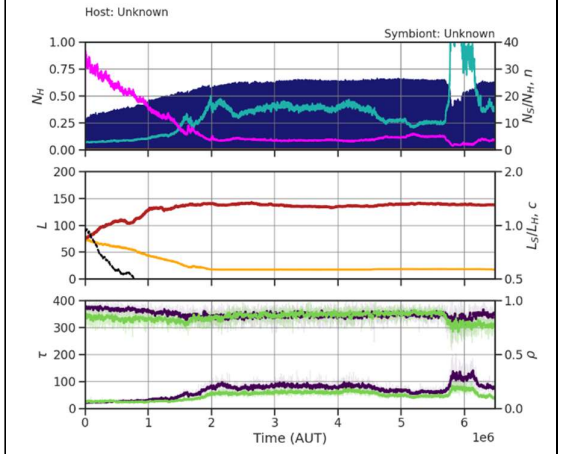
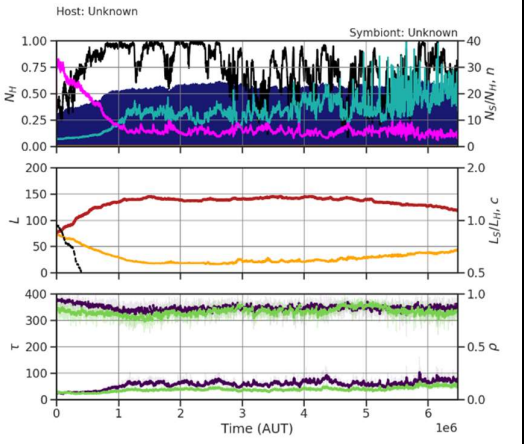


	Asexual (GAP)	Sexual (GP)	Mutable (GCP)
1			

2	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6
5	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6

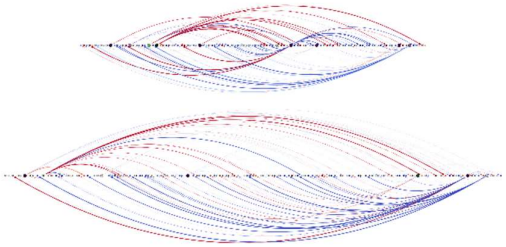
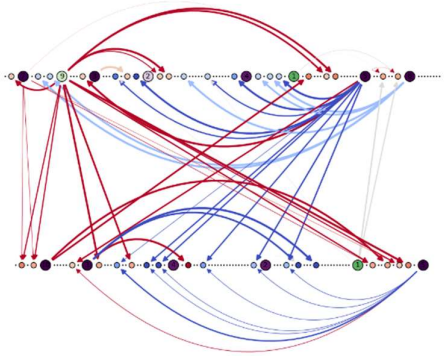
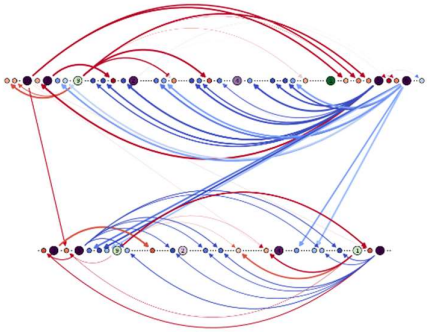
6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6
7	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6

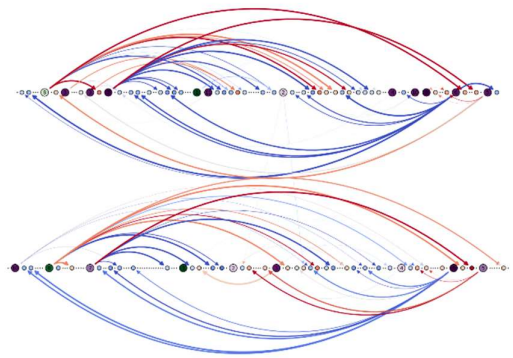
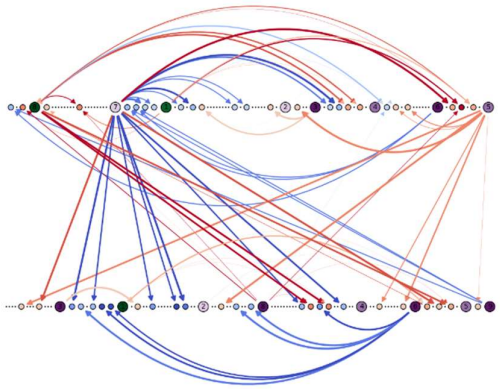
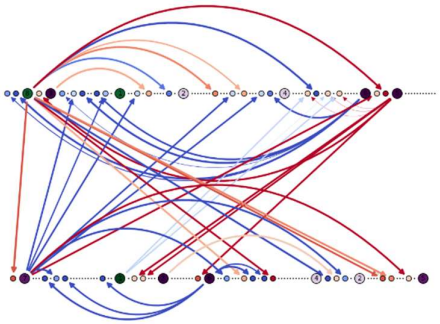
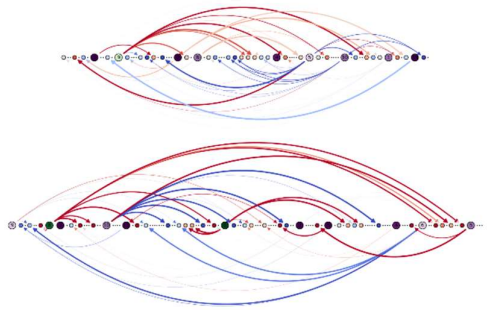
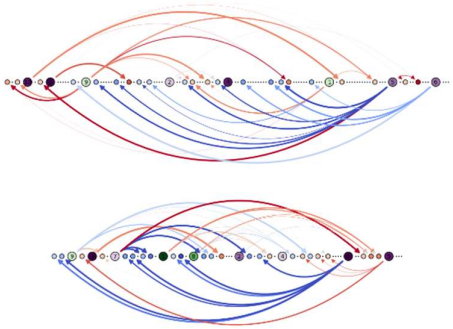
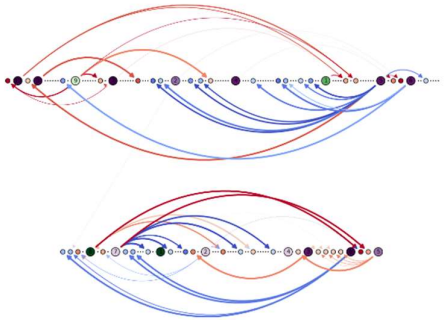
8	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6
9	 Host: Unknown Symbiont: Unknown Time (AUT) 1e7	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6

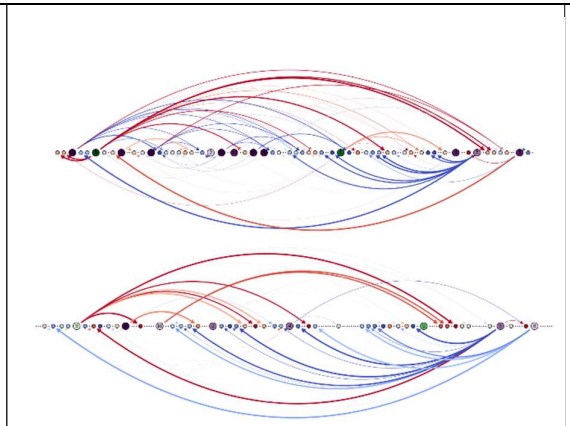
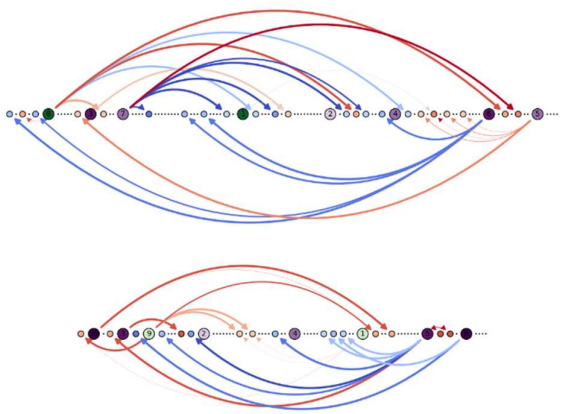
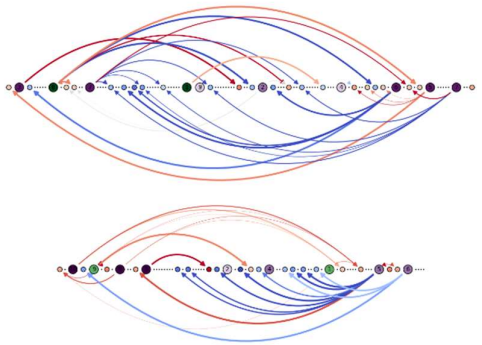
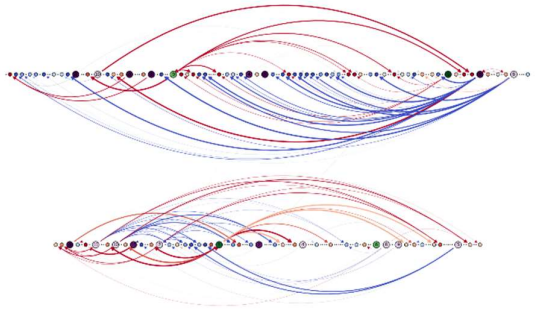
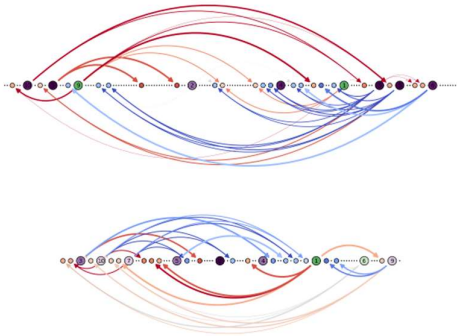
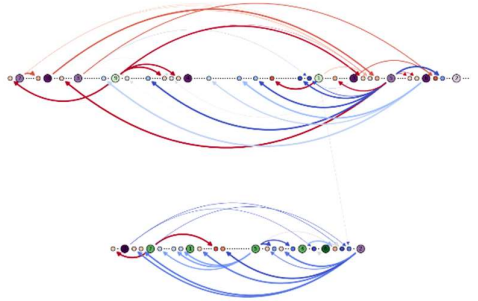
10	 Host: Unknown Symbiont: Unknown Time (AUT) 1e7	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6
11	 Host: Unknown Symbiont: Unknown Time (AUT) 1e7	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6	 Host: Unknown Symbiont: Unknown Time (AUT) 1e6

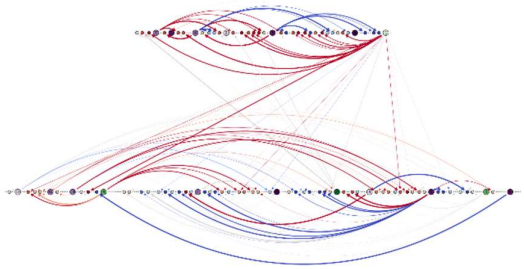
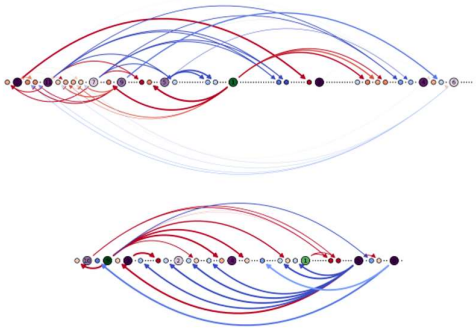
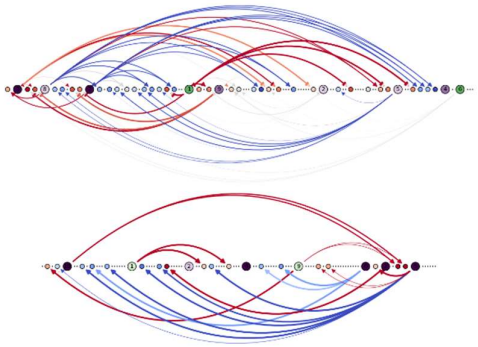
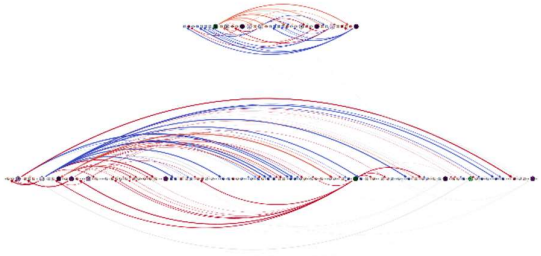
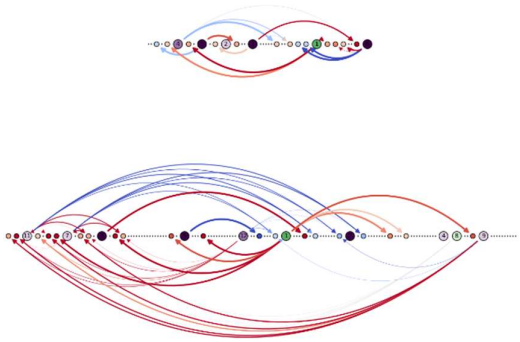
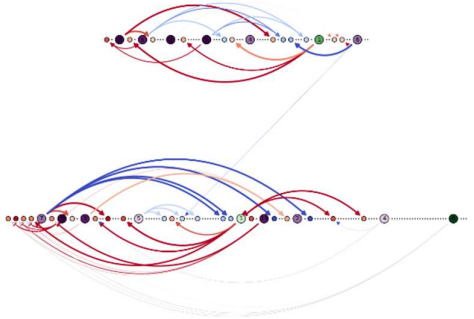
Appendix 2.2

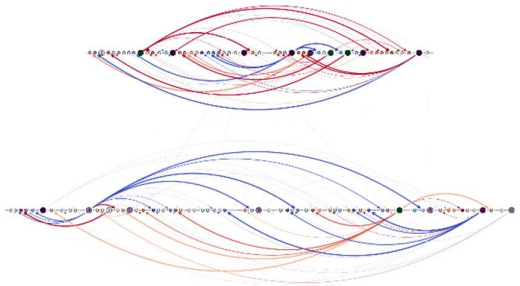
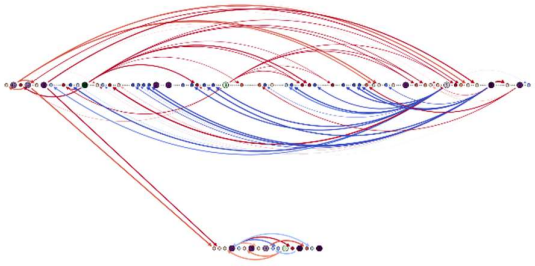
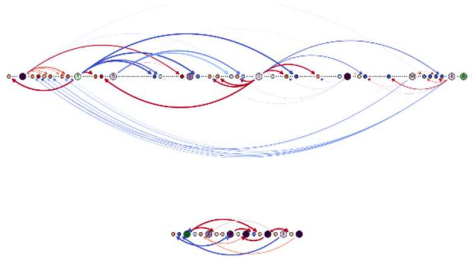
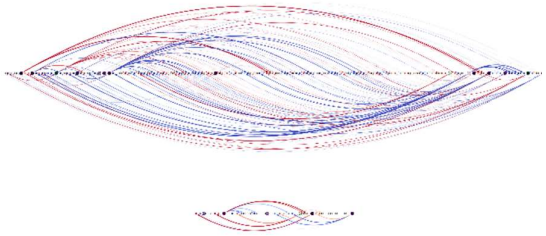
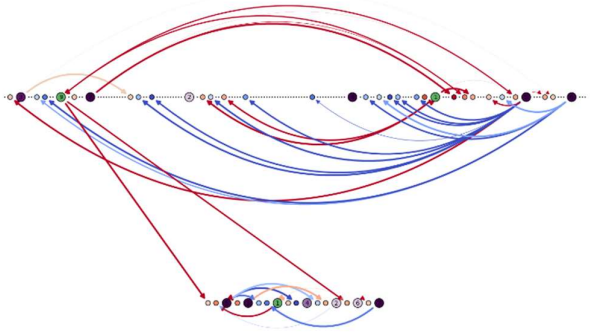
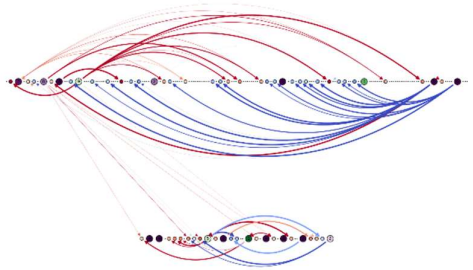
Most recent common ancestors (MRCAs) of pre-evolved strains on the nutrient gradient. For notation see table S2.2.

	Asexual (GAP)	Sexual (GP)	Mutable (GCP)
1			

2			
5			

6			
7			

8			
9			

10			
11			

Appendix 2.3 Phylogenetic trees of strain 9