Farming the mitochondrial ancestor as a model to endosymbiotic establishment by natural selection
Supplementary Material

István Zachar[1,2,3], András Szilágyi[1,2,3], Szabolcs Számadó[4,2], Eörs Szathmáry[1,2,3]

[1] Department of Plant Systematics, Ecology and Theoretical Biology Eötvös Loránd University (ELTE), Pázmány Péter sétány 1/C, Budapest 1117, Hungary

[2] Evolutionary Systems Research Group, MTA, Centre for Ecological Research, Hungarian Academy of Sciences, Klebelsberg Kunó str. 3., Tihany 8237, Hungary

[3] Center for the Conceptual Foundations of Science, Parmenides Foundation, Kirchplatz 1, 82049 Pullach/Munich, Germany

[4] MTA TK "Lendület" Research Center for Educational and Network Studies (RECENS), Tóth Kálmán u. 4., Budapest 1097, Hungary

This file includes:

Materials and Methods

Ecological minimum model

Benefit during decline (D)

Two predator species are considered: X is a conventional phagocytotic cell (*non-farmer*), Y is a phagocytotic cell that can set aside and farm engulfed prey bacteria (*farmer*). In the decline period D, when resources are scarce, the dynamics for X is a simple decline with death rate d, but the initial death rate of Y, d_0, is much less than that of X, as an indirect benefit of having provisions (*farm*). Starvation of farmers thus starts with a slower decay rate d_0 ($d_0 < d$), and with time, d_0 converges to d in a sigmoid curve, corresponding to the exhaustion of the farm. The parameter r affects the shape of the sigmoid: the higher the value of r is the faster d_0 tends to d. The decline curves at different r values and the pure exponential decay are compared in Figure S1. The corresponding differential equations for X and Y are as follow:

$$\dot{x}(t) = -dx(t) \qquad \text{(Eq. 1)}$$

$$\dot{y}(t) = -\frac{d\,d_0\,\mathrm{e}^{rt}}{d+d_0(\mathrm{e}^{rt}-1)}y(t) \qquad \text{(Eq. 2)}$$

Eqs. 1 and 2 have analytic solutions:

$$x(t) = x_0\mathrm{e}^{-dt} \qquad \text{(Eq. 3)}$$

$$y(t) = y_0\left[1+\frac{d_0}{d}(\mathrm{e}^{rt}-1)\right]^{-\frac{d}{r}}. \qquad \text{(Eq. 4)}$$

Thus, the population density of Y after time t in period D can be expressed by the above formula. Note that farming could give some further benefit to Y due to some growth, but this is neglected by a worst-case argument.
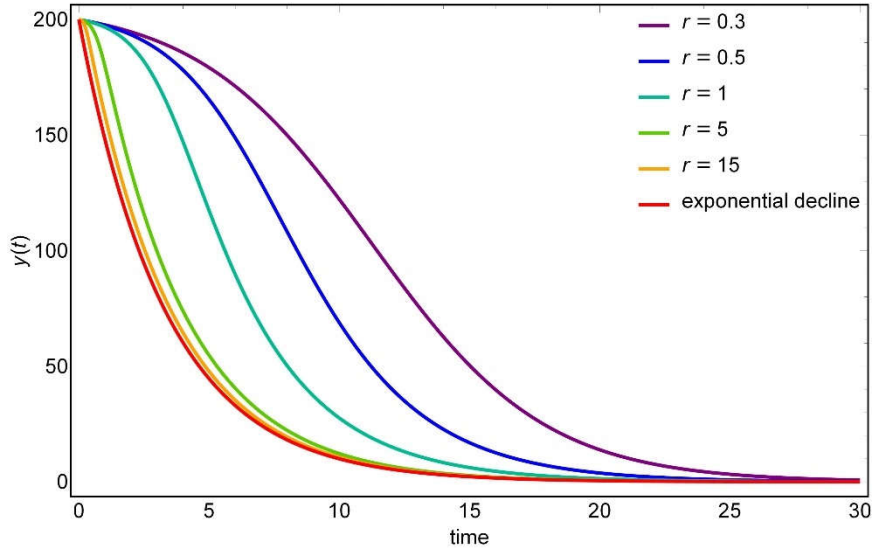
Figure S1. Comparison of decline of farmers and non-farmers in the ecological minimum model. Pure exponential decay of non-farmers and delayed decay of farmers at different r values ($r = 15, 5, 1, 0.5, 0.3$). Parameters: $\{d = 0.3, d_0 = 0.01, y_0 = 200\}$.

Cost during growth (G)

In the growth period (G), resource is abundant and populations of X and Y grow in an environment with carrying capacity k. The cost of Y is due to two factors: (*i*) accumulation of non-digested prey, and (*ii*) maintenance and internal growth of the symbiont population to match cell division of the farmer. Mechanistically these component processes may result in complicated dynamics (see the individual-based model later). We simply assume that the Malthusian growth rate of Y is below that of X through the whole period. If nothing else happened, this would guarantee extinction of Y by competition.

In order to be analytic, we construct an Eigen-type equation-system (see (1)):

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = ax(t) - \frac{F(t)}{k}x(t), \tag{Eq. 5}$$

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = by(t) - \frac{F(t)}{k}y(t), \tag{Eq. 6}$$

where F is the total production of the system:

$$F(t) = ax(t) + by(t).$$

Note that here we do not demand the condition that $x(t) + y(t) = k$, although the sum of x and y converges to k in the long run. Combining the above two equations we obtain:

$$\frac{1}{x(t)}\frac{\mathrm{d}x(t)}{\mathrm{d}t} - a = \frac{1}{y(t)}\frac{\mathrm{d}y(t)}{\mathrm{d}t} - b,$$

from which we obtain:

$$\frac{\mathrm{d}\ln x(t)}{\mathrm{d}t} - \frac{\mathrm{d}\ln y(t)}{\mathrm{d}t} = a - b = \frac{\mathrm{d}\ln z(t)}{\mathrm{d}t},$$

from which we have:

$$\frac{x(t)}{y(t)} = \frac{x(0)}{y(0)}e^{(a-b)t}, \tag{Eq. 7}$$

which is always true in period G. Using this relation, we can solve the original differential equations:

$$x(t) = \frac{e^{at}kx_0}{k+(e^{at}-1)x_0+(e^{bt}-1)y_0}$$

$$y(t) = \frac{e^{bt}ky_0}{k+(e^{at}-1)x_0+(e^{bt}-1)y_0}.$$

Thus, we again have closed formulae for the population values at any time t, in period G. The full dynamics of the system can be obtained by iterating GDGDGD... for indefinite time with constant period lengths for G and D (see dynamics in Figure 1B in main text). To test the robustness of the system, we have also modelled random fluctuations of the environment, i.e. random lengths of successive G and D periods (see Figure S2 and Figure S3). Farmers can invade at certain parameter combinations (see Figure 1 in main text and Figure S3).
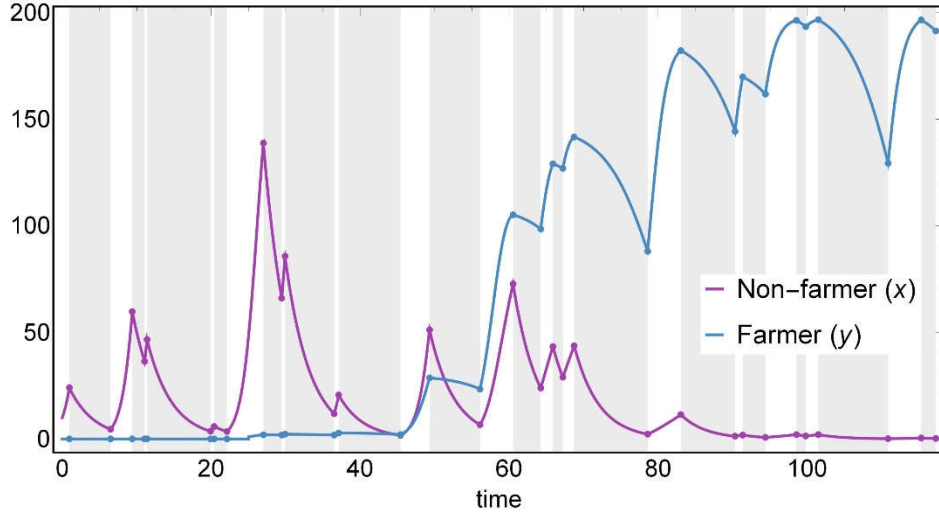

Figure S2. Small scale dynamics of the ecological minimum model with fluctuations of random period lengths. Purple curve indicates non-farmer abundance x, blue curve indicates farmer abundance y. Parameters: $\{x_0 = 10, y_0 = 0, t_G = 5, t_D = 10, d_0 = 0.01, d = 0.3, r = 0.3, a = 1, b = 0.8, k = 200\}$; actual period lengths $\hat{t}_G$ and $\hat{t}_D$ are randomly drawn from uniform distributions over the intervals $\{0, t_G\}$ and $\{0, t_D\}$, respectively. 16 cycles of G-D period pairs are shown, shaded zones indicate poor periods (D). Mutant is introduced at $t_{inv} = 25$ with concentration $y(t_{inv}) = 1$.

The ratios $y(t)/x(t)$, at the end of the respective periods D and G:

$$z_D(t_D) = \frac{y(t_D)}{x(t_D)} = z_0 e^{dt_D}\left[1 + \frac{d_0}{d}(e^{rt_D}-1)\right]^{-\frac{d}{r}}$$

$$z_G(t_G) = \frac{y(t_G)}{x(t_G)} = z_0 e^{(b-a)t_G}.$$

The ratio $y(N)/x(N)$, after N full cycles ($N \times \{G, D\}$) can be expressed as follows:

$$z(N) = z_0[z_G(t_G)z_D(t_D)]^N = z_0 e^{N[(b-a)t_G+dt_D]}\left[1 + \frac{d_0}{d}(e^{rt_D}-1)\right]^{-N\frac{d}{r}},$$

$\lim_{N\to\infty}(z(N)) = \infty$ indicates that farmer wins over non-farmer. In all simulations the parameters are $\{r = 0.3, d = 0.3, d_0 = 0.01, k = 200, t_G = 5, t_D = 10\}$.
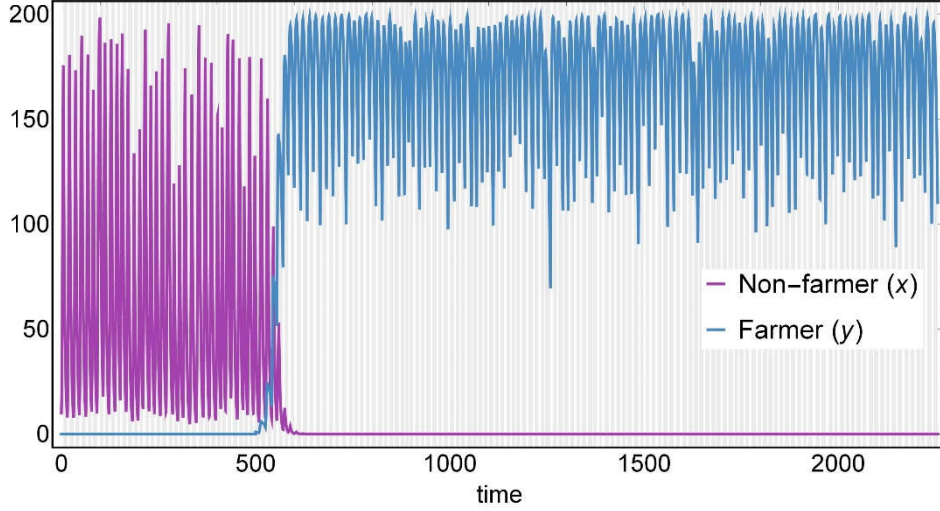
3

Figure S3. Successful invasion of farmers within the ecological minimum model. Similar to Figure S2, with longer temporal span and later invasion time. Purple curve indicates non-farmer abundance x, blue curve indicates farmer abundance y. Parameters: $\{x_0 = 10, y_0 = 0, t_G = 5, t_D = 10, d_0 = 0.01, d = 0.3, r = 0.3, a = 1, b = 0.8, k = 200\}$; actual period lengths $\hat{t}_D$ and $\hat{t}_G$ are randomly drawn from normal distributions $N(t_G, 1)$ and $N(t_D, 1)$, respectively. 150 cycles of G-D period pairs are shown, shaded zones indicate poor periods (D). Farmer is added to the system at $t_{inv} = 500$, with concentration $y(t_{inv}) = 1$.

Adaptive evolution

We have simulated possible evolutionary routes of the emergence of farming in a non-farmer population. According to Eq. 2 or Eq. 4, farmers and non-farmers are identical in the limit of $r \to \infty$ only. In these simulations we treat farmers with $r \geq 20$ and $b = 1$ as non-farmers.

Simulations start with an established resident population of non-farmers ($r = 20, b = 1$) with density of 10. $r = 20$ effectively mimics the behaviour of non-farmers, see Figure S1. After two complete cycles of growth-decline periods (G-D) of random lengths, at a random time during the third G-D period, a mutant is introduced. The initial density of the invader is 1% of the actual resident population. The mutated traits b' and r' of the invader are calculated as follows:

$$b' = b \cdot \left(1 + N(0, \sigma_e)\right)$$
$$r' = r \cdot \left(1 + N(0, \sigma_e)\right),$$

where $N(0, \sigma_e)$ denotes a Gaussian distributed random number with zero mean and σ_e variance (actual value used is $\sigma_e = 0.05$).

To comply with realistic assumptions, we introduced a trade-off between the traits r and b. As any decrease in r results in better survival in poor times, it must have a fitness cost realized in simultaneously reducing b, the growth rate in rich periods. If $b' > b$ and $r' > r$, we accept the mutant as invader only if $0.5 < \frac{r'-r}{b'-b}\frac{b}{r} < 2$. Similarly, if $b' < b$ and $r' < r$, we accept the mutant only if $0.5 < \frac{r-r'}{b-b'}\frac{b}{r} < 2$. In any other case, the mutant is ignored and a new mutant species is generated.

The invasion is considered successful if the density of the resident population is reduced below the extinction limit 10^{-3} or unsuccessful if the invader density goes below this limit. The system is then reinitialized with the successful species as the new resident and after two more G-D cycles, a new mutant invades the population.

The lengths of the rich (growth) and poor (decline) periods, t_G and t_D, respectively, have small fluctuations due to environmental variability, defined as follows:

4

$$t_{\mathrm{G,D}} = T_{\mathrm{G,D}} \cdot (1 + \mathrm{N}(0, \sigma_T)),$$

where $T_{\mathrm{G}} = 5$ and $T_{\mathrm{D}} = 15$ are the means of the respective period lengths and σ_T is the variance (actual value used is $\sigma_T = 0.35$).

We run 100 simulations each with 40 000 successful invasions and monitored the evolution of traits b and r (see Figure S4). Parameters not specified are the same as in the previous section. About 60% of the simulations are converged during the last 8 000 invasions to the region bounded by $0 < r < 4$ and $0 < b < 0.2$, 30% did not yet converged but tended towards small (r, b) pairs and about 10% converged then left the target (r, b) region at the end of the simulation. Continuation of these runs ultimately results in convergence.
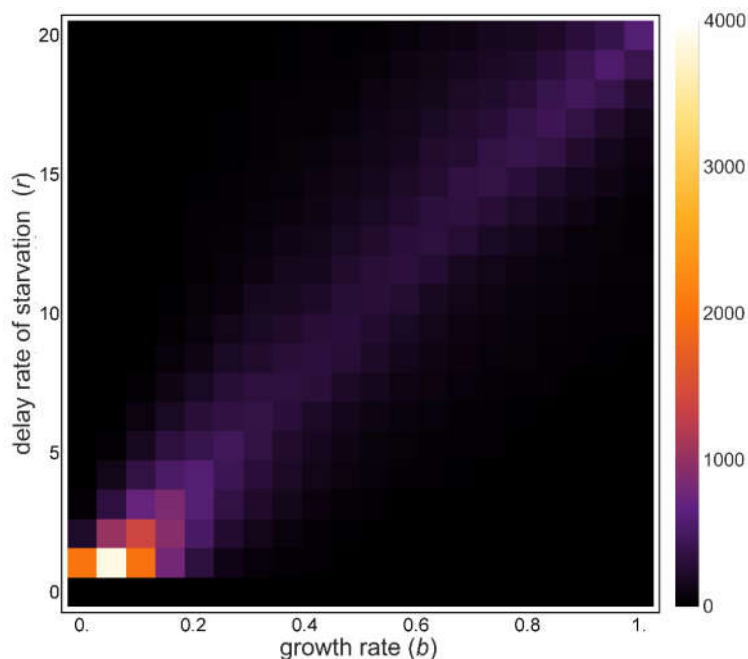

Figure S4. Density plot of the number of failed invasion steps in the adaptive evolution of the ecological minimum model. The lower left corner with small growth rate b (disadvantage in rich periods) and small r (better delay of starvation in poor periods) is a strong attractor. Average of 200 independent simulations, all starting from the top right corner ($b = 1, r = 20$). The b-r plane was partitioned into 20 by 20 rectangular bins for discretization.

Fusion, fission and stealing

Here we expand the minimal model with the possibility of fusion-fission in poor periods. During fusion, cells exchange cytoplasmic components (and farmed prey), and then split. A rudimentary cytoskeleton (which some Archaea might possess (2), and possibly the eukaryotic cytoskeleton originates from Archaea (3)) might have been able to assist the non-farmers to pinch off a few phagosomes or farmed cells from the farmer. Also, a cytoskeleton helps to withhold the bulk of the stock preventing too large donations to an empty partner. We assume that fusion happens in the poor period, as is the case with many bacteria, where sexual phase is triggered by poor conditions in their habitat. It is easy to see why having sex at the onset of resource scarcity is advantageous. Fusion has a cost both in energy and time (it is slow, and predators cannot feed during fusion) hence it poses an immediate disadvantage in the growth period against those who invest all their time and energy into grow (at least in the ecological timescale). In contrast, fusion in the poor period might provide much needed food for a farmless cell if it receives a piece of its partner's stock. In the ecological timescale during which these considerations apply, we can safely ignore the exchange of genes due to the much lower probability of gene insertion than sharing stock.

We assume that both types (farmer and non-farmer) can possess an internal food stock, either by having allocated prey to the farm or by having received prey from a fusion partner. Farmer and non-farmer types have densities $x(t)$ and $y(t)$ and average internal stock sizes $f(t)$ and $p(t)$, respectively, at time t. Fusion is triggered by starvation, hence it only happens in the poor period D, when there is neither growth, nor further farming happens. Farmers start the poor period with an average stock size $f(0) = f_0 > 0$ that they have built up in the good period, non-farmers with $p(0) = p_0 = 0$, and with time, the stock is eliminated with rate q. Both types can possibly receive stock from their fusion partner: non-farmers take a fraction s, farmers take a fraction u of the partner' stock. By this way, we can model asymmetric fusion situations where one of the types has a type-specific advantage over the other. We assume that mating has a cost c, and happens with probability m in unit time. The change in average internal food levels in the poor period is as follows:

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = -q\,f(t) - m\,x(t)\,y(t)\,s\,f(t) + m\,x(t)\,y(t)\,u\,p(t) - m\,c, \qquad \text{(Eq. 8)}$$

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t} = -q\,p(t) + m\,x(t)\,y(t)\,s\,f(t) - m\,x(t)\,y(t)\,u\,p(t) - m\,c. \qquad \text{(Eq. 9)}$$

Note, that we ignore the cases where same-typed cells fuse (with probability $m\,x(t)^2$ for non-farmers and $m\,y(t)^2$ for farmers), as the average stock size never changes in those fusions. The minimum condition for farming to be useful is:

$$f_0 \geq \frac{m\,c}{q}(e^{t_D\,q} - 1),$$

otherwise everyone starves to death in the poor period, even if $u = s = 0$. The worst for farmers is when mating always happens ($m = 1$), as they are most exposed to thieves; $m < 1$ increases the chance of farmers against stealing. Since both types must pay the cost for mating, the term $-m\,c$ can be eliminated from the equations. Also, for the moment, we consider $x(t)$ and $y(t)$ to be constant during the poor period, and we introduce $\pi = m\,x(t)\,y(t) = m\,x_0\,y_0$ to denote the probability of farmers and non-farmers mating. Eqs. 8 and 9 become:

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t} = -q\,f(t) - \pi\,s\,f(t) + \pi\,u\,p(t), \qquad \text{(Eq. 8b)}$$

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t} = -q\,p(t) + \pi\,s\,f(t) - \pi\,u\,p(t). \qquad \text{(Eq. 9b)}$$

Later, we will examine the case where cell densities also change due to starvation (see Eqs. 1b, 2b and Figure S5). We note however, that in case of variable densities, non-farmers are clearly in disadvantage as they start the poor period with no stock, hence their density will drop faster than that of farmers.
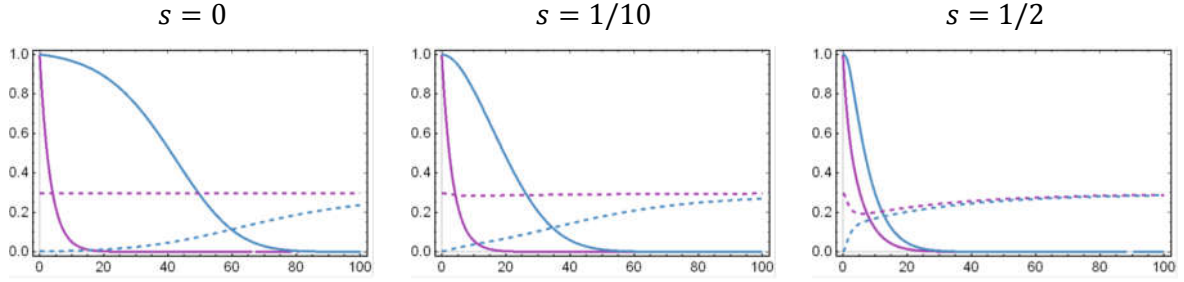
Figure S5. Decay dynamics in the poor period D, with different fusion rates (integrating Eqs. 1b, 2b). The farmer ($y(t)$, blue curve) and non-farmer ($x(t)$, purple curve) densities drop as the different types slowly starve using up their resources. Dashed curves denote time-dependent deat rates $d_x(t)$ (dashed purple) and $d_y(t)$ (dashed blue). As thieving rate s is increased for non-farmers, the advantage of delayed death for farmers reduces until non-farmers steal so much that they get the advantage. Case $s = 0$ (no stealing, no sharing) approximates the decay dynamics of the minimal model without sex (Eqs. 1, 2). Parameters are $\{f_0 = 1, p_0 = 0, x_0 = 1, y_0 = 1, q = 0.01, d = 0.3, m = 1, u = s\}$.

The basic asymmetry emerges from the fact that only farmers can allocate directly to their farms (and only in the good period), hence they start the poor period with nonempty stocks. Furthermore, as a worst case, we assume that no leftover stock from a poor period is carried over to the following good period for either type. Otherwise, both types are capable of digesting and stealing stock. The two types follow asymmetric strategies only if $s \neq u$, which assumes that farming and stealing are exclusive traits and cannot mix. While this is unlikely, we again refer to this as a worst-case scenario. Furthermore, we assume, that there is no preferential fusion, because types cannot recognize each other (which could be a later evolutionary invention).

No fusion. If $u = s = 0$, there is no fusion at all. Our basic models discussed above (Eqs. 1, 2) apply to this case.

Diffusion. If $0 < u = s \leq 1$, fusion equalizes stock levels. While maximum efficiency is at $u = s = 1$, this is unrealistic as fusion cannot manifest in full exchange. Hence, we assume $u = s \leq 1/2$. The longer the time allocated to fusion, the more material is exchanged and higher s is. The solutions for Eqs. 8, 9 are:

$$f(t) = \frac{1}{2}e^{-qt}(f_0(1 + e^{-2\pi st}) + p_0(1 - e^{-2\pi st})),$$

$$p(t) = -\frac{1}{2}e^{-qt}(f_0(e^{-2\pi st} - 1) - p_0(1 + e^{-2\pi st})$$

It is clear, that $f(t) > p(t)$ is always true whenever $f_0 > p_0$ (assuming $0 < q, m$), and only in infinite time (after infinite fusions) will the two stock levels perfectly equalize (in the deterministic case). Our game theoretical model applies to the case of finite mating times and a small number of matings (see later). The critical time when farmer and non-farmer stock levels converge (below an arbitrary small difference ε):

$$f(t) - p(t) < \varepsilon,$$

$$\hat{t} = \frac{\ln\left(\frac{f_0 - p_0}{\varepsilon}\right)}{q + 2\pi s},$$

and if we assume $p_0 = 0$ and $s = \frac{1}{2}$:

$$\hat{t} = \frac{\ln\left(\frac{f_0}{\varepsilon}\right)}{q + \pi}.$$

Internal stock levels define the time the cells can survive in the poor period: if one type has more stock than the other, it can survive for a longer time. Thus, depending on the length of the poor period, better provisions might imply survival in the poor period, or better reproductive chances in the next good period. With time, however, the average stock level changes due to feeding (q) and death (as starving predator cells are removed). Due to this, fusion probability π reduces along a hyperbolic curve, and mating becomes less and less frequent (an Allee effect), which further favors farmers.

We assume that early fusion was a form of simple diffusion ($u = s$). In case of diffusion (or if the farmer takes more ($u > s$)), non-farmers cannot acquire more farm than farmers. In the limit of infinite fusions, stock levels settle at identical values and different types have the same chance of survival. Only if non-farmers steal more than farmers ($s > u$), can non-farmers have more stock than farmers. While this is an interesting aspect, we emphasize, that for early fusion, we cannot expect more evolved mechanism than fusion that can steal proportionally more from the partner. Furthermore, thieving in this sense should be type-specific, and only available to non-farmers, which is not a parsimonious assumption.

The critical time $\hat{t}$ defines the time threshold within the poor period until farmers possess a larger stock on average than non-farmers. But this is not enough for ecological stability. When fusion is entirely based on diffusion, and no asymmetric stealing is possible, stock levels equalize in the limit, but there is a chance that 1) the poor period ends well before stochastic events would remove farmers, perhaps even before the critical time; and/or 2) fusion becomes so rare due to loss of partners to starvation, that equalization practically stops. Ultimately, in any case **farmers must maintain a farm large enough to survive the poor period with an end period density that prohibits non-farmers to outgrow them in the following good period.**

Game of fusion, fission and farming

We follow the assumptions of the ecological minimal model. There are two predatory types, a farmer F and a non-farmer P, which can fuse in the poor period to exchange their farms, and the split. When P cells fuse, they do not receive any payoff. When F cells fuse, they, in average, retain their stock levels at an average size. This stock provides f benefit in unit time (as digested food), between two probabilistic fusions. When farmers fuse with non-farmers (and split), the non-farmer receives a fraction $0 < s < 1$ of the stock, leaving the fraction $(1 - s)$. The cost of mating is c, and its probability is m. The average payoffs after the first fusion-fission are given in the following matrix.

	P (donor)	F (donor)
P (recipient)	$-mc$	$mfs - mc$
F (recipient)	$(1-m)f + mf(1-s) - mc$	$f - mc$

As the expected value of the mating cost ($-mc$) is the same for all cases, the matrix simplifies to:

$$E = \quad \begin{array}{c|c|c} & P \text{ (donor)} & F \text{ (donor)} \\ \hline P \text{ (recipient)} & 0 & fms \\ \hline F \text{ (recipient)} & f(1 - ms) & f \end{array}$$

We assume that the farm lasts longer than the time period between two fusions (plus the cost of fusion):

$$f > q + mc,$$

where q is the amount of farm digested until the next fusion event. F is an evolutionarily stable strategy (ESS) if $E(FF) > E(PF)$. While $s < 1$, that is, the non-farmer cannot steal the whole farm, this inequality holds and F is an ESS. If $s = 1$, the result is neutral, but $s > \frac{1}{2}$ is and unrealistic assumption (see above). If $s \leq \frac{1}{2}$, it is easy to see, that there could be no non-farmer that ever has more stock than any farmer, and ultimately, in the limit of infinite fusions, stock levels are equalized.

Individual-based evolutionary-ecological model

Fully continuous dynamics in continuous time do not allow individual differences among farming hosts, eliminating thus any stochastic effect that could work for the benefit of farming. Hence the above differential equations cannot be used directly to model internal differences of farmers and non-farmers, to provide a more realistic model. One can design a class-based continuous-time model (extending the above defined dynamics) where each class consists of the farmers having the exact same farm size and the dynamics of each class is calculated in time (as in (4)). This, however, would unnecessarily complicate the model introducing a potentially huge number of differential equations (in the thousands, considering the maximally allowed farm size in the model discussed below), while it still would not provide the necessary stochastic dynamics we assume to be important during the emergence of farmers.

Consequently, to provide a more realistic model that accounts for cellular interactions and internal states, we have designed an individual, agent-based model where stochastic effects are inherently present. Farmer and non-farmer cells are represented individually, have internal states (cell growth- and farm size-states), perform different actions according to their states, and possess traits that are subject to mutation and evolution.

Note, that the internal farm must have a reasonable size comparable to (perceived) external prey density, to make an effective difference in poor periods. If farm is too small, it does not matter much for the host during starvation, as it will have a low relative frequency compared to the low external abundance and host might not rely on it. Furthermore, a full farm should provide enough food for the host to survive poor periods. However, due to the implementation of the model, simple survival is not enough. For tractability, farmers and non-farmers have a closed population size and compete in a Moran process for food and space (prey is modelled as a separate population, as it is not in competition with predators). To avoid ending up with empty population slots and to focus on competitive dynamics, cells do not lose their growth state when starving but simply stagnate. Consequently, farmer can only outcompete non-farmers, if it can grow and reproduce in the poor period (as it always has a disadvantage in the rich period). Thus, if the farm size is not large enough for the host to reach split density in poor periods, farmers won't increase in number, see Figure S6. The Moran process also means that when dynamics stop due to lack of resources, it effectively represents extinction for both species.
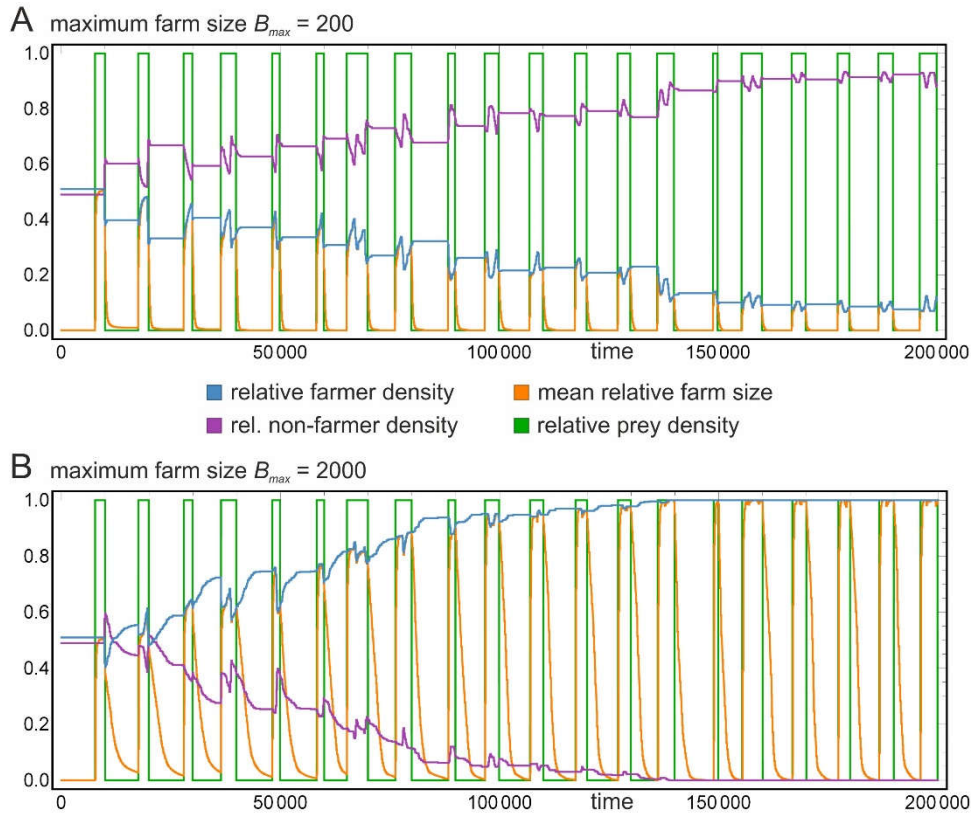
Figure S6. Effect of maximal farm size on farmer invasion in the individual-based model. A: Starting from a mixed population of farmers and non-farmers, farmers cannot invade, if farm size is not sufficiently large to last till the end of the poor period. **B**: Farmers can invade successfully, if their maximal farm size B_{max} has roughly the same magnitude as the external prey density ($A_{max} = B_{max} = S_{max} = 2000$). For further parameters, see Table S3.

The following traits are heritable and are subject to mutation during evolution:

- **Farm allocation rate** f specifies the probability that an engulfed prey is allocated to the internal farm instead of being directly consumed.

- **Farm culling rate** h (*harvesting*), parameterizing the switching function that governs whether the host consumes a free prey or its farm, based on prey-availability.

- **Farming cost** c, specifying the explicit cost (if negative) or benefit (if positive) of maintaining a farm. The actual cost (or benefit) is dependent on c and the actual farm size. Fitness cost due to maintaining the farm is simulated as the host not growing when a prey (free or farmed) is consumed. Explicit benefit (e.g. the farm leaking some metabolite for the host) is simulated by the host spontaneously growing. Since cost and benefit is not bounded to unit regions (-1, 0) and (0, 1) respectively, but to regions (-β, 0) and (0, +β) (β being possibly larger than 1), they cannot be directly converted to probabilities of losing or gaining a unit in cell growth. Hence, only the fractional part of c is ever used to simulate probabilistic starvation/growth of the host, while the integer part is deterministically subtracted/added from the host's growth state.

Model parameters, functions and evolutionary traits are as follow:

P	population size; fixed throughout a simulation
S_{max}	maximal cell size, at this size the cell inevitably splits
A_{max}	maximal external prey amount
B_{max}	maximal farm size, farm cannot grow above this
t_{max}	maximum number of time steps; the population is updated for each timestep
g	inherent farm growth rate
μ_f	mutation rate of farm allocation rate
μ_h	mutation rate of farm culling rate
μ_c	mutation rate of farming cost rate
sd	mutation standard deviation, uniform for each trait
C_b	baseline cost of farming: the cost of maintaining the apparatus
β	maximum benefit of the farm ($-\beta$ is the minimum cost)
$\rho(t)$	temporal distribution of the abiotic resource R, returning a value within (0, 1)
$\sigma(r, h)$	actual farm-culling probability based on relative farm density r (relative to external prey density) and culling property h
f	farm allocation rate: probability of a captured prey being farmed; evolutionary trait
h	farm culling ("harvesting") rate: governs the digestion of the farm; evolutionary trait
c	factor of farm size dependent cost (or benefit); evolutionary trait

Pseudocode of the model:

```
S := [S₁, S₂, …, Sₚ];              set up individual cell size vector for the population of size P
B := [B₁, B₂, …, Bₚ];              set up individual farm sizes
f := [f₁, f₂, …, fₚ];              set up individual farm allocation rates
h := [h₁, h₂, …, hₚ];              set up individual farm culling ("harvesting") rates
c := [c₁, c₂, …, cₚ];              set up individual farm cost rates

for T := (0 → Tₘₐₓ) {              iterate time variable from 0 to Tₘₐₓ
   R := ρ(T);                       set external abiotic resource state
   A := round(R·Aₘₐₓ);             reset external prey density is by resource state

   for (1 → P) {                    update population
      i := randI(1, P);             randomly select an individual from the population
      Bᵢ := max(min(Bᵢ + round(R·Bᵢ·(g - 1))·sgn(fᵢ), Bₘₐₓ), 0);  inherent farm
                                                                        growth

      a := A/Aₘₐₓ;                  relative external prey density
      b := Bᵢ/Bₘₐₓ;                 relative internal prey density
      H := σ(b/(a + b), hᵢ);        actual probability of culling farm

      if (fᵢ > 0) and (randR() < abs(C_b)) {
         Sᵢ := max(min(Sᵢ - 1, Sₘₐₓ), 0);        baseline cost of farming apparatus
      }
      if (fᵢ > 0) {
         x := 0;                    use fractional part of cost cᵢ as probability
         if(randR() < abs(frac(cᵢ)·b)) x := sgn(cᵢ);
         Sᵢ := max(min(Sᵢ + int(cᵢ) + x, Sₘₐₓ), 0);   cost/benefit of farm
      }
      if (rand() < a) {             capture external prey
         A := max(A - 1, 0);
         if (randR() < fᵢ) and (randR() < (1 - b)) {
            Bᵢ := min(Bᵢ + 1, Bₘₐₓ);    add captured prey to farm
         } else {
            Sᵢ := min(Sᵢ +1, Sₘₐₓ);     eat captured prey and grow
         }
      } else {                      if no external prey is captured, cull farm
         if (Bᵢ > 0) and (randR() < H) {
            Bᵢ := max(Bᵢ − 1, 0);       cull farm and grow
            Sᵢ := min(Sᵢ + 1, Sₘₐₓ);
         }
      }

      if (Sᵢ == Sₘₐₓ) {             when maximum size is reached, cell splits
                  select another cell than i to overwrite with daughter cell
         while (j == i) {j := randI(1, P);}
         Sᵢ := 0; reset cell sizes
         Sⱼ := 0;
         K := randB(Bᵢ);           randomly distribute farm among daughter cells
         Bᵢ := Bᵢ - k;
         Bⱼ := k;
         hⱼ := hᵢ;                 inherit evolutionary traits
         cⱼ := cᵢ;
         fⱼ := fᵢ;
                                    mutate traits
         if (randR() < μ_f) {fⱼ := max(min(fⱼ + randN(0, sd), 1), 0);}
         if (randR() < μ_h) {hⱼ := max(min(hⱼ + randN(0, sd), 1), 0);}
         if (randR() < μ_c) {cⱼ := max(min(cⱼ + randN(0, sd), β), -β);}
      }
   }
}
```

Variables T, A, B and S (and T_{max}, A_{max}, B_{max} and S_{max}) have integer values. The function $sgn(x)$ is the sign function returning -1 for negative x, 1 for positive x, and 0 for $x = 0$; $frac(x)$ and $int(x)$ return the signed fractional and integer parts of x, respectively; $randR()$ returns a uniformly distributed random real from the interval [0, 1]; $randI(min, max)$ returns a uniformly distributed random integer from the interval [min, max]; $randN(\mu, \sigma)$ returns a Gaussian random real with mean μ and standard deviation σ; $randB(max)$ returns a binomially distributed random integer from the interval [0, max]. At any time, farmer amount F is the number of hosts with nonzero farm allocation rate f, non-farmer amount is the rest of the population (P - F), independent of farm culling rate h.

The resource function is a temporally regularly periodical function, setting the resource variable R, with transitions between poor and rich environments (roughly corresponding to decline and growth periods of the minimal model). The richness is following the fact that A depends directly on R and A is only replenished once every timestep, i.e. during a timestep the population is updated randomly and hence all individuals in the population compete for the same limited pool of preys (A). The following box wave function is used to define R:

$$\rho(t) = R_{\min} + (R_{\max} - R_{\min}) \begin{cases} 0, & \mod(t, v) < t_D \\ 1, & \text{else} \end{cases}$$

where v is the period length, t_D is the relative length of the resource-poor decline period within a period, $R_{\min}$ and $R_{\max}$ scale the function and mod is modulo division. If for each period, a random Gaussian factor is applied to t_D (and also to t_G in case of other resource functions), the result is a periodically fluctuating function with slight irregularities in period lengths, see Figure S7.
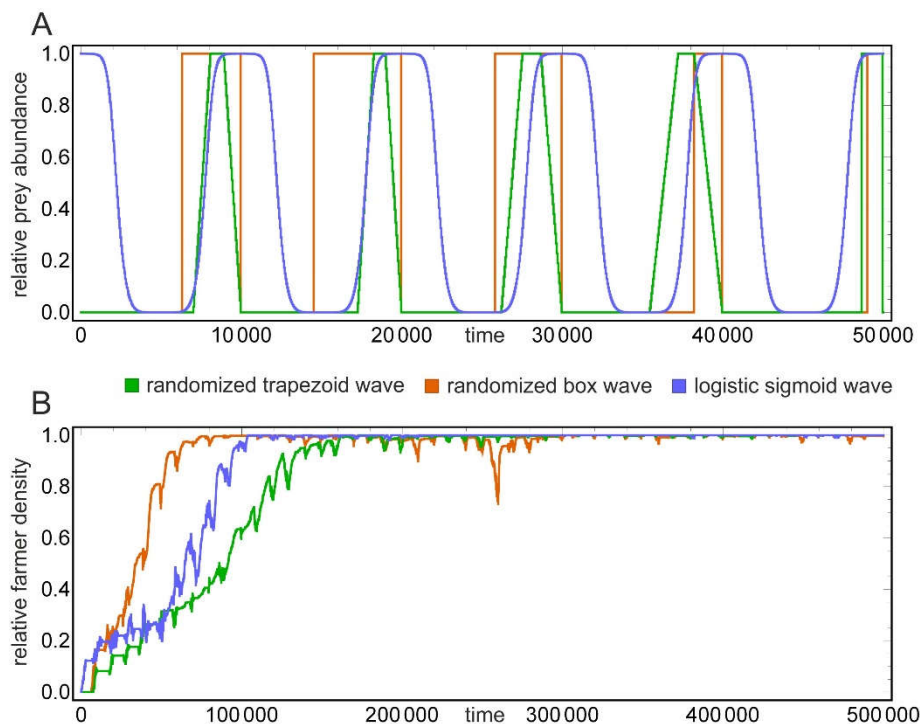

Figure S7. Simulations with various resource functions in the individual-based model. A: Relative prey abundance. The resource function directly defines the relative prey abundance $a = A/A_{max}$. A full G-D period cycle has length $v = 10\,000$. The actual length of poor period $\hat{t}_D$ is drawn from a normal distribution $N(t_D, \sigma_{t_D})$ with $t_D = 7000$ and $\sigma_{t_D} = 0.2$ for the box and trapezoid waves; constant 5000 for the sigmoid wave. The actual rich period length is defined, for each full G-D cycle, as $\hat{t}_G = v - \hat{t}_D$. **B**: Relative farmer density. More realistic resource functions (trapezoid-wave, spike-wave or logistic sigmoid-wave functions) produce qualitatively similar farmer behaviour in time (note the time scale difference between panels A and B). For parameters, see Table S3.

13

The culling function is defined to be dependent on the relative amount of farmed prey r (relative to the amount of external prey) and on a single parameter, h, that is subject to mutation within the interval (0, 1):

$$\sigma(r,h) = \begin{cases} 2rh, & \text{if } h \leq \frac{1}{2} \\ 1, & h \geq 1 - \frac{r}{2}, \\ \frac{r}{2(1-h)}, & \text{else} \end{cases} \qquad \text{(Eq. 10)}$$

where $r = b/(a+b)$ and $a = A/A_{max}, b = B/B_{max}$. If $h = 0$, host never culls its farm, hence internalized prey effectively parasitize the host unless they provide an explicit benefit ($c > 0$, which must have been a late evolutionary adaptation). If $h = 1$, host always culls farm, regardless of farm size and prey density (but only if it couldn't capture a free prey in the given timestep); see Figure S8.
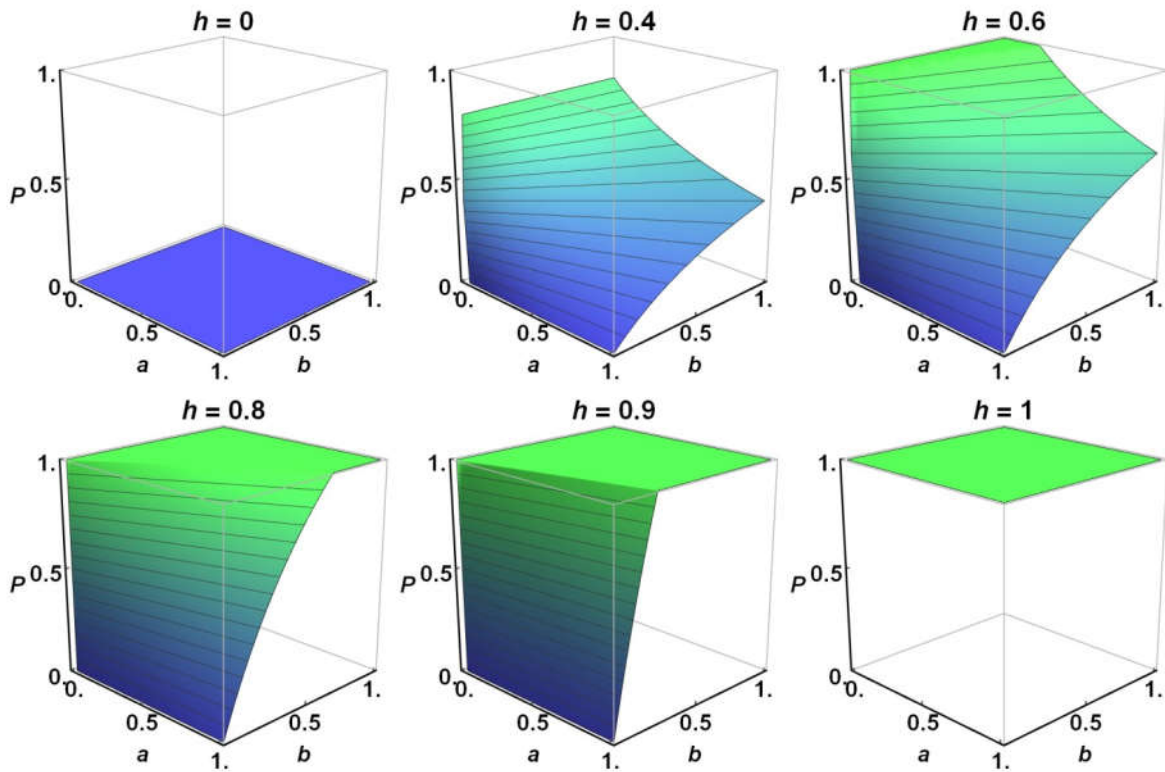

Figure S8. Culling probability function of the individual-based model at various values of culling rate h. If $h = 0$, host never culls its farm, hence internalized prey population (if exists) effectively parasitizes the host unless it provides an explicit benefit. If $h = 1$, host always culls farm, regardless of farm size (but only if it couldn't capture a free prey in the given timestep). Axes are: relative external prey density $a = A/A_{max}$; relative farm density $b = B/B_{max}$; actual probability of culling farm $P = (\sigma(b/(a+b), h))$, where σ is the culling function as defined by Eq. 10.

Additional figures for the individual-based model
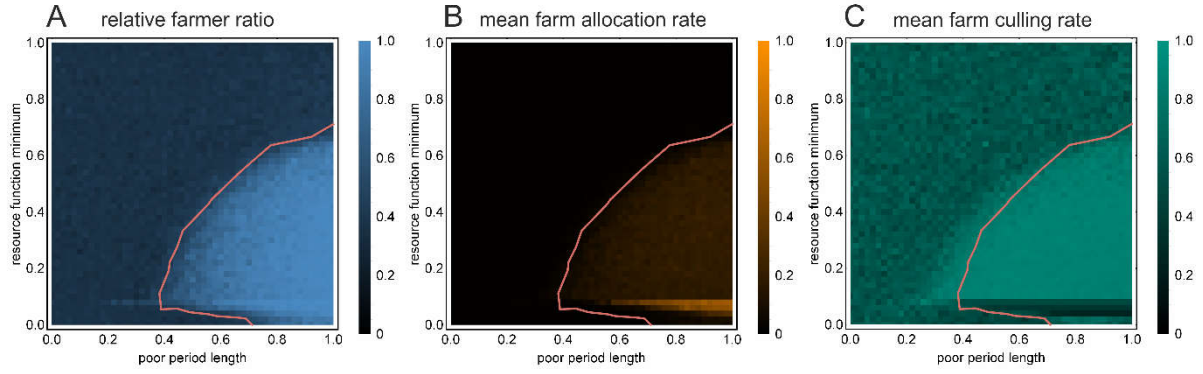These are referenced in the main text.

Figure S9. Evolutionary outcome of farming and culling. Panels show the various properties of the same equilibrium populations after 2×10^6 timesteps, depending on the poor period length t_D and resource function minimum R_min. Each pixel is an average of 30 independent simulations. The light red curve on each plot indicates the region where equilibrium farmer ratio is above ½. Farm allocation and farm culling rates were allowed to evolve while the explicit cost of farming was fixed at $c = $ -0.4. **A**: **Relative farmer frequency**. Farmers can stably invade a resident non-farming population only if the environment is sufficiently harsh (poor period is long). Non-farmers (constantly generated by mutations) cannot invade an established farming population. Figure is qualitatively the same as the *y-z* slice of Figure 3A at $c = $ -0.4 (main text). **B**: **Mean farm allocation rate** f. Non-farmers have zero farming rate, but even the farmer population has low values, as farm is rather maintained by its own growth instead of the host storing more. **C**: **Mean farm culling rate** h. The culling rate is maximized toward $h = 1$ in farmers, while it is in neutral drift for non-farmers. For parameters and explanation of the behavior at low R_min values, see Figure S10.
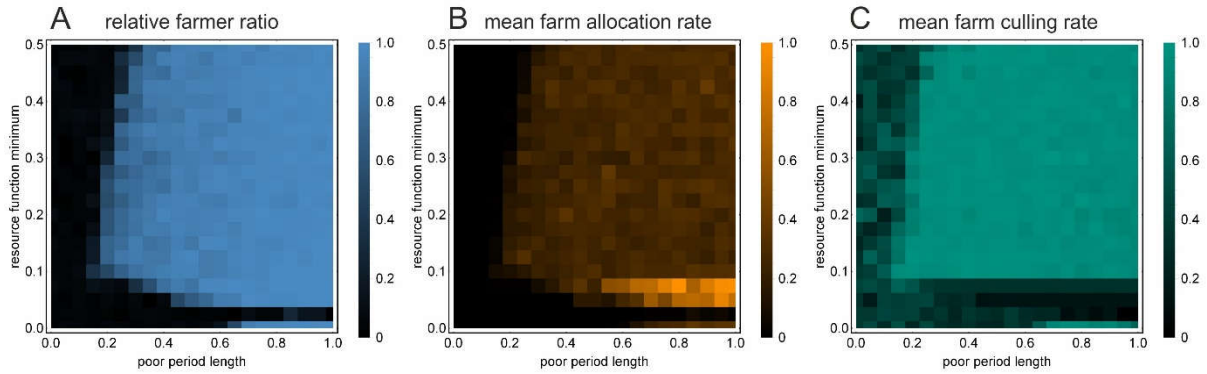

Figure S10. Invasion of farming when cost of farming is diminished in the individual-based model depending on the poor period length t_D and the resource function minimum in poor periods R_min. Same as Figure S9, with farming cost $c = -0.2$ instead of $c = -0.4$. Panels show different mean properties of the same equilibrium populations, each pixel is an average of 10 independent simulations. Non-farmers (panel **A**) can only survive poor periods if those are sufficiently short (left on each panel). The high farming (panel **B**) but low culling rates (panel **C**) at low R_min values (bottom of each panel; also in Figure S9) is due to the sensitive balance between the culling rate and the farm's own growth rate (with rate constant $g = 1.04$). When R_min is 0, host can only cull its farm right after the rich period, then farm is quickly depleted as it cannot grow on its own (as there is no resource at all). Culling rate is nevertheless kept high to squeeze every drop from the temporary presence of the farm. In the range $0 < R_\mathrm{min} < 0.1$, there is almost no reproduction in the population as the farm is already large enough to consume resources of the host (cost of farming has to be paid). Host cells cannot reach split density and are stagnating. As R_min increases, culling is balanced solely by allocating more prey into the farm but not by the farm's own growth. This is because the farm's growth exactly balances its own maintenance cost, and only a minimum amount can be culled. Therefore, culling has to be finetuned to a minimum tolerable rate. Above the critical level $R_\mathrm{min} \approx 0.1$, the farm's own growth practically tolerates any culling and the farm is maintained in a stable density against host feeding on it, thus culling can be maximized toward 1. For parameters, see Table S3.
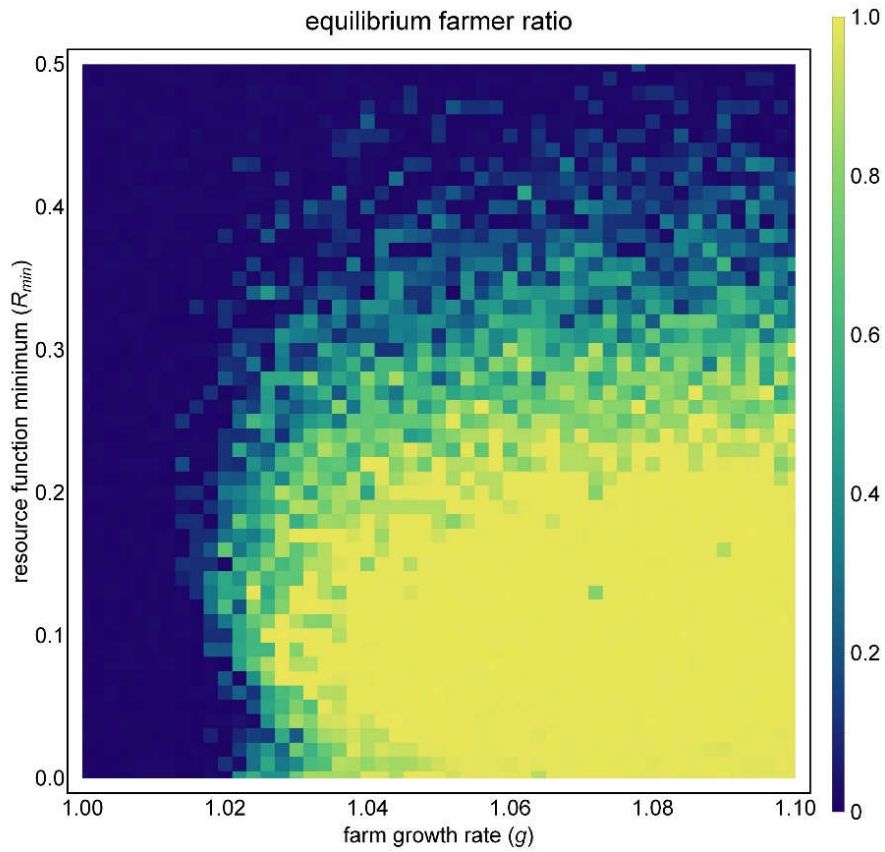
Figure S11. Invasion of farming depending on farm's own growth rate g and resource minimum in poor periods $R_{\mathbf{min}}$ **in the individual-based model**. Each pixel is an average of 10 independent simulations, recording the equilibrium farmer/non-farmer ratio after 2×10^5 timesteps. All 51 x 51 x 10 simulations have identical parameters (see Table S3) and were initiated from quantitatively identical initial conditions.
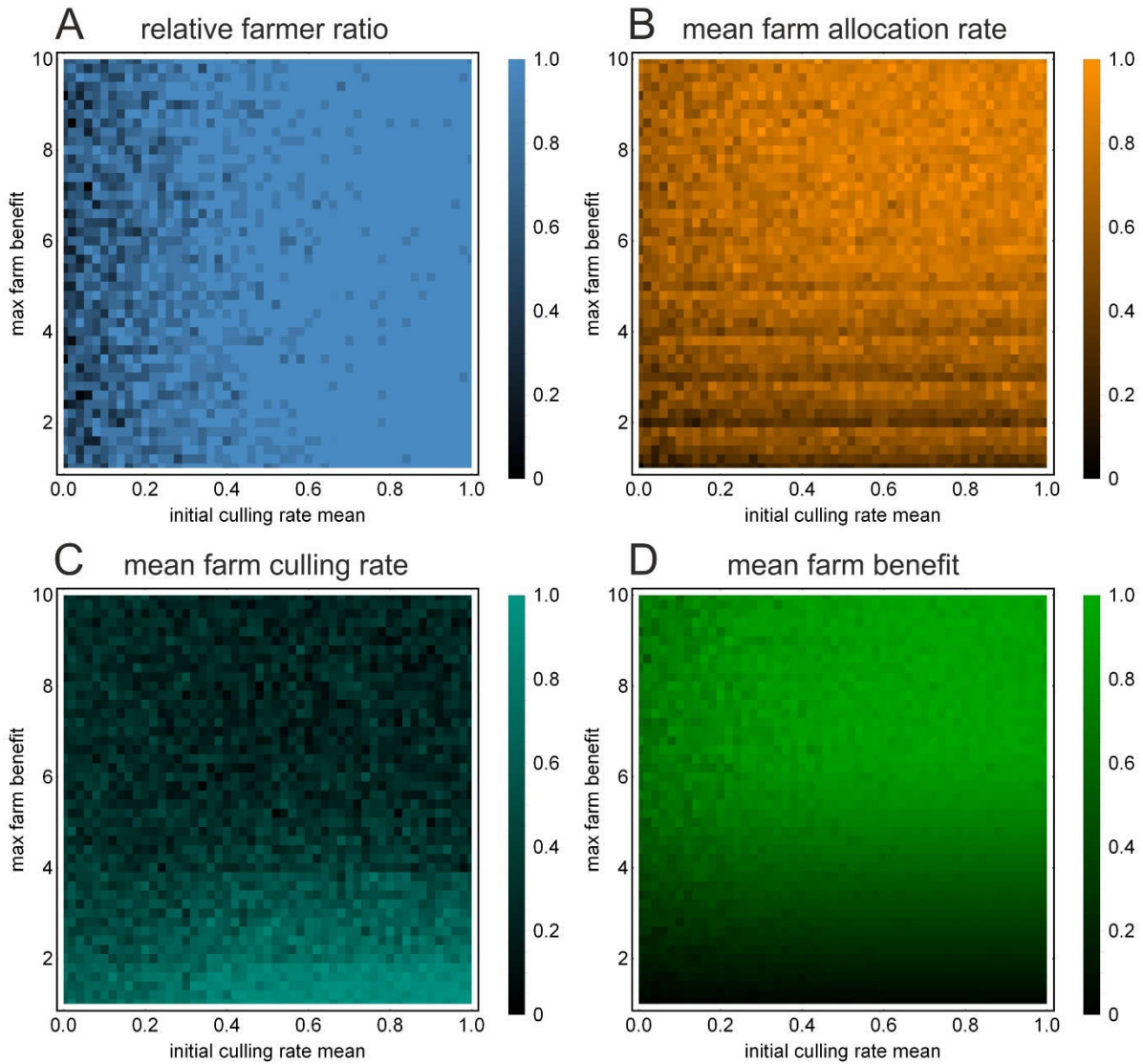
Figure S12. Evolution of a beneficial symbiont in the individual-based model. Panels show the various equilibrium properties in the same simulations. Simulations start from already farming populations with low farm allocation rate and various culling rates (*x* axis, assuming prior evolution has settled at these culling rates). The *y* axis shows the maximum possible explicit benefit the farm can provide. Non-farmers can only invade (or farmers cannot invade non-farmers at all), if the initial culling rate is small (left side of panel **A**). There are two distinct evolutionary outcomes. Farmers can evolve to always feed on the farm with high farming rate (**B**) and culling rates (**C**), so whatever is skimmed from the farm by digestion is quickly restored by allocating to the farm (bottom half in each panel); the actual explicit benefit gained from the farm remains low in these cases (**D**). If the maximum possible explicit benefit is high enough, farmers can evolve toward minimal culling and maximally enjoying the explicit metabolic help of the farm (top half of each panel). At the start of each simulation, farming is costly ($c = -0.4$), but cost can evolve toward the maximum benefit (*y* axis), however, slower than the evolution of the culling rate (also see Figure S13 for more details). The culling rate is only reduced to zero if a farmed cell provides more explicit benefit than eating it (always yielding one unit). Each pixel is an average of 30 independent simulations. Horizontal stripes (pronounced in panel **B**) are results of the benefit-discretization method of the model (when the farm benefit crosses an integer value during evolution, an extra unit of food is granted for the host). For more details and parameters, see Figure S13.

A high initial culling, fast mutation of culling
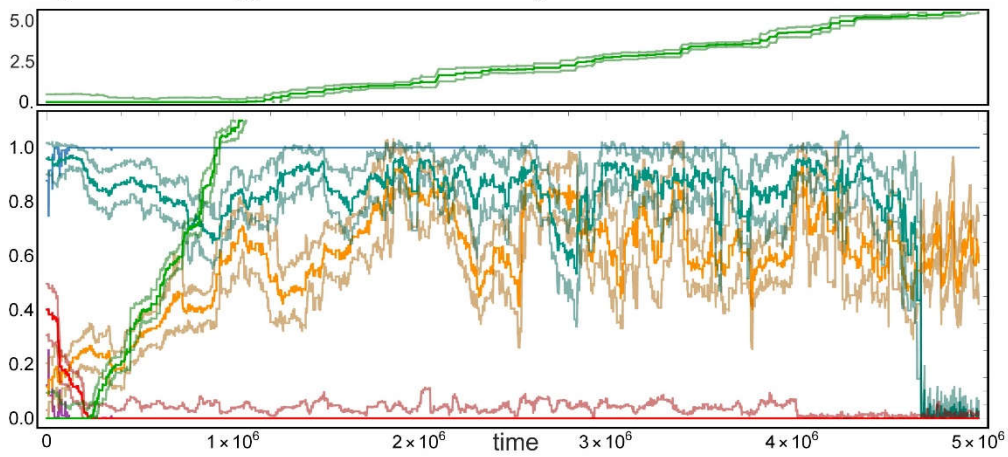
B low initial culling, slow mutation of culling
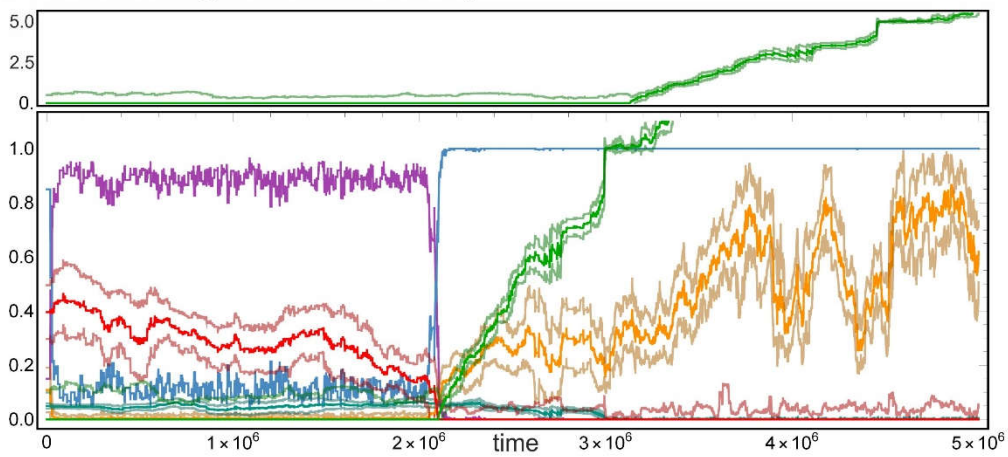
C high initial culling, slow mutation of culling
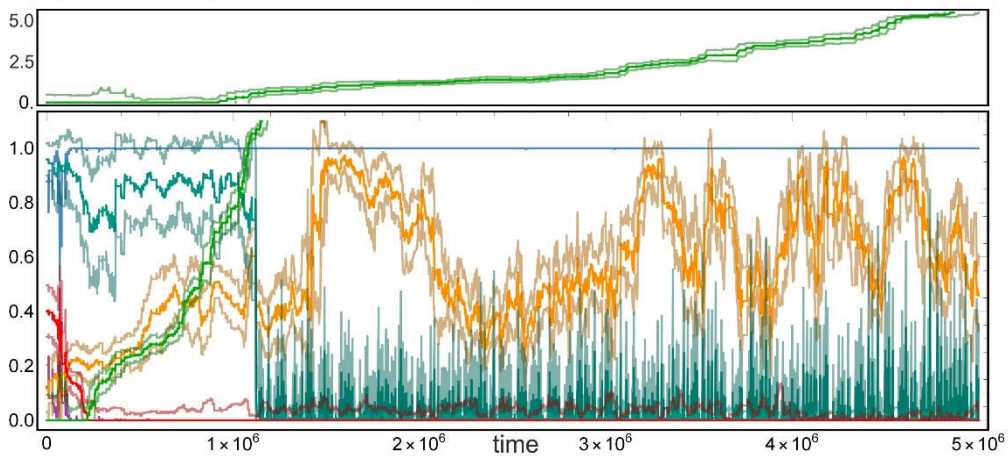
- relative farmer density
- rel. non-farmer density
- mean farm allocation rate
- mean farm culling rate
- mean farm benefit
- mean farm cost (times -1)

Figure S13. The two evolutionary stable outcomes of farming in the individual-based model. The farmed bacterium can be turned to a "powerhouse" only if keeping a farmed cell for its metabolic benefit provides more advantage than the implicit benefit of eating it. Actual mean farm cost (red curve, mirrored from the negative to the positive halfplane) is can evolve toward explicit mean benefit (green curve). Mean culling rate (turquoise curve) drops to zero only if the explicit benefit grows larger than the implicit benefit of eating a farmed bacterium (which is 1 unit). Lighter curves indicate plus and minus standard deviations of the trait in the population. Since

18

maximum farm benefit can go above 1, it is displayed separately above each plot (note the different scaling of the y axis). The mean farm allocation rate (orange curve) is mostly a neutral trait, drifting around $f = 0.5$, independent of the culling rate. **A**: The population starts from high initial culling rate. A sudden transition around $t \approx 4.6 \cdot 10^6$ turns full culling (and digestion) to almost zero culling. This happens with increasing probability as the maximum explicit benefit is increased, as is apparent from Figure S12C. Parameters: $\{m_h = 1.0, s_h = 0.1, \mu_h = 0.05, \sigma_h = 0.05\}$. **B**: If the population starts from low culling rates and culling rate mutations are smaller than the mutation steps of the explicit benefit, it is possible for the population to increase the benefit *before* culling rate tops at 1. In this case, when farmers invade (at $t \approx 2 \cdot 10^6$), they immediately start with high benefit and low culling rate, which leads to the elimination of culling. Parameters: $\{m_h = 0.05, s_h = 0.01, \mu_h = 0.05, \sigma_h = 0.005\}$. **C**: When initial culling rate is high in the population (as in panel A), larger mutation steps make it easier to reduce the culling rate (and the culling ability can in fact be lost), which immediately pays off due to the explicit benefit and is fixed in the population. Parameters: $\{m_h = 1.0, s_h = 0.1, \mu_h = 0.05, \sigma_h = 0.2\}$. For further parameters, see Table S3.

Parameterization

We have carefully designed our models to rely on as few explicit parameters as necessary. In case of the ecological minimum model, the relative growth rates of non-farmers and farmers (a and b, respectively) are important instead of absolute values. We prove that, depending on r, the value of b can be as small as $a/10$, or even smaller, and *vice versa* (see Figure 1C and Figure S4). The benefit of farming, a reduced death rate $d_0 = 0.01$ converges to the default death rate $d = 0.3$. Any value chosen for d fixes the time scale, of which a and b can be interpreted. The relevant parameters are d_0/d and r, for which we have found that no fine tuning is necessary as the behavior of the system is qualitatively the same with differently chosen values (results not included).

In the individual-based model, most of our parameters are defined by prey influx per unit time (A_{max}), which ultimately determines dynamics. Maximum cell size S_{max} is defined to allow competition. For the basic values (e.g. Figure 3 in the main text, also see Table S3), in one timestep, A_{max} = 2000 prey cells enter the system, which are distributed among p = 500 predators; in average, each predator receives 4 prey cells per timestep. In this rate, 50 timesteps are required to reach division size S_{max} = 200 for a non-farmer predator, more for a farmer. In the $v\,t_D$ = 2 000 timesteps of the rich period, a cell will divide 40 times, which provides ample time to overgrow inferior competitors. All evolvable parameters (farming rate f, culling rate h and explicit cost/benefit c) are defined in the same units of growth (i.e. a single prey cell) directly comparable to cell growth. Furthermore, storing and culling are exclusive, only one can happen per timestep, to prevent indirect advantage of farmers. Similarly, the choice between eating free-living or farmed prey is exclusive, only one could happen per timestep (the implicit cost of farming). This is comparable to the case reported by Brock et al. that "nonfarmers eat all the bacteria whereas farmers leave many bacteria unconsumed, roughly half the number present as compared with bacteria grown alone" ((5) p. 395).

Apart of these considerations, lack of data for the assumed time period of the Proterozoic and the lack of modern examples of prokaryotic farming prevented us to use data from the literature, and we provide reasonable assumptions instead.

- **Endosymbiont abundance** (B_{max}): Farm size is B_{max} = 2000. Considering modern examples of unicellular endosymbioses (6-10), a value in the order of 10^2-10^3 seems to be entirely reasonable. This of course means that the archaeal host had to be large enough to house hundreds of bacteria, but this is a general assumption of the phagocytic lifestyle and not specific to farming.

- **Cell size** (S_{max}): The size where cells divide must be in roughly the same order of magnitude as farm size B_{max}, as farmers have to grow/survive solely by consuming their farms in poor times (like consuming free-living cells in good times), as initially there is no other metabolic benefit

provided by the farm. We used 10 $S_{max} = B_{max}$ due to the fact that most of the time the farm is far from being full.

- **Baseline farming cost** (c_b), independent of farm size: Farming is costly (11), and in our models, cost values are interpreted as negative growth. Laboratory measurements can only measure hindered growth, but cannot distinguish between farm dependent and independent cost (like maintaining farming mechanism, proteins, genes, etc.). Farmers have to pay the farm independent baseline cost even if they do not have a single farmed bacterium. According to Brock et al., *Dictyostelium* "farmers do as poorly as non-farmers, if they are previously made bacteria free using antibiotics" ((5) p. 394), hence the baseline cost must be marginal and can be safely omitted. We nevertheless assumed a baseline cost of -5 % of normal growth (c_b = 0.05), but it is easy to see that our results still apply in case of c_b = 0.

- **Farm dependent cost** ($c < 0$, initial farming cost mean m_c): Most of the reported cases for endosymbioses indicate increased growth for the host in presence of the symbiont (12), as the partnership depends on direct metabolic help. We do not know about any such case where costs were distinguished from benefits. Even harder is to distinguish the explicit cost of farming (dependent of farm size) from the implicit cost of farming (as discussed above). Brock et al. reported, that "farmers produce fewer spores than non-farmers when provided a fixed amount of live bacteria" ((5), p. 394), roughly half of those of farmers. This value, of course, is a time average of growth and spore formation, so we do not know how the host's growth constant is affected. We used the extreme value of m_c = -40 % (or -20 %) cost on growth, but we also tested the range of growth reduction up to -100 % (Figure 3 in the main text), which means that to maintain each cell in the farm, the host has to eat two.

- **Inherent symbiont growth rate constant** (g): Most of the studied endosymbioses rely on evolved partnerships, where symbiont and host already synchronized their cell cycles (e.g. (8)), hence $g = 1$ in general. This of course would not be enough for early farming where division constantly dilutes farms. We used g = 1.04 for most cases as a reasonably small value, but we also explored the behavior of the model in the range g = 1.0 … 1.1 (Figure S11).

- **Farm dependent benefit** ($c > 0$, maximum benefit β): The metabolic benefit, when evolves in later stages (Figure S12, Figure S13), is also interpreted in the same unit as costs: cell growth per timestep. Moreover, the actual benefit can grow above unity, indicating that an evolved metabolic partnership allows "more free energy per unit of food consumed" ((12), p. 20).

Supplementary tables

Table S1. Comparison of the minimal and the individual-based models.

components	minimal model	individual-based model
resource dynamics	fluctuating between rich and poor periods, depending on the environment	
prey dynamics	prey is implicitly represented, as the rich or poor period of the environment; has no dynamics	prey density is explicitly represented, dynamics directly depend on the resource and the consumption by predators
cell properties	farmer and non-farmer populations have uniform properties per population: individuals share the same growth rate a in the non-farmer and b and delayed decay rate r in the farmer population	cell size, farm size, farm allocation rate, farm culling rate and explicit farm cost are defined for each cell; populations are naturally polymorphic
farm properties	implicitly modelled through the delayed decay rate (r) of farmers	farm is explicitly modelled, farm size is managed for each individual independently
cell dynamics	growth and decay are implicitly continuously modelled	capture of prey, consumption of prey (external or internal), cell growth and decay dynamics and reproduction are explicitly discretely modelled for each cell individually
farm dynamics	none	allocation to farm, digestion of farm, internal growth of farm are explicitly discretely represented;
mutation	only in adaptive evolution experiment: growth rate (b) and delayed decay rate in poor period (r) are subjects to evolution	farm allocation rate (f), farm culling rate (h) and farm cost (c) are subjects to evolution
farming cost (trade-off between fitness and farming)	explicitly modelled between b and r in the adaptive evolution experiment	implicitly modelled, the trade-off naturally emerges as consuming the farm conflicts with maintaining the farm (for its explicit benefit, if any)
reseeding environment with prey from the farm	cannot be modelled	implemented, but omitted

Table S2. List of the different assumptions tested in the various experiments. MM = minimum model, IBM = individual-based model.

main text	Figure 1	MM	A: Equilibrium distribution of farmers and non-farmers depending on relative poor period length t_D/v, relative farmer growth rate b/a, and relative farmer decay rate d_0/d; B: Time evolution of a single characteristic simulation; C: Evolutionary trajectory of the population when both r and b are evolvable.
	Figure 2		Predator-prey interactions in the farming archaeon.
	Figure 3	IBM	A: Equilibrium distribution of invading farmers depending on relative poor period length t_D/v, explicit farming cost c, and resource minimum in poor periods R_min. B: Characteristic run of a winning farmer population. C: Simulation of B in longer time span.
Supplementary Information	Figure S1	MM	Comparison of decline of farmers and non-farmers in case of various starvation delay values (r).
	Figure S2		Competitive dynamics when period lengths are randomly drawn from a uniform distribution.
	Figure S3		Competitive dynamics when period lengths are randomly drawn from normal distribution (longer time than in Figure S2).
	Figure S4		Adaptive evolution, number of failed invasion steps depending on the $\{b, r\}$ values of mutants invading.
	Figure S5		Decay dynamics in the poor period, with different fusion rates (s).
	Figure S7	IBM	Simulations with various resource functions, random period lengths from uniform distribution.
	Figure S8		Culling probability function of the individual-based model at various values of culling rate (h).
	Figure S6		Effect of maximal farm size on farmer invasion, starting from mixed populations, random period lengths.
	Figure S9		A: Invasion of farming when explicit cost of farming is constant $c = -0.4$ (similar to Figure S10), depending on poor period length t_D/v and resource minimum in poor periods R_min. B: Time evolution of a single characteristic simulation.
	Figure S10		Invasion of farming when cost of farming is diminished compared to Figure S9: $c = -0.2$.
	Figure S11		Invasion of farming, depending on farm's own growth rate and resource scarcity.
	Figure S12		Evolution of a beneficial symbiont: explicit cost of farming evolves toward benefit.
	Figure S13		The effect of the mutation rate of culling to the evolution of explicit benefit.

Table S3. Parameter values of the individual-based model for the various figures in the main text and supplement. Parameters specified in figure captions take precedence over parameters listed in the table. Abbreviations: SD = standard deviation, B = box wave, T = trapezoid wave, S = logistic sigmoid wave, U(*min*, *max*) = uniformly distributed in the interval (*min*, *max*).

	description	Figure 3	Figure S9, Figure S10	Figure S12, Figure S13	Figure S7	Figure S6	Figure S11
t_{max}	max time step	10 000 000	2 000 000	5 000 000	500 000	100 000	200 000
p	population size	500	500	500	500	500	500
A_{max}	max prey amount	2000	2000	2000	2000	2000	2000
S_{max}	max cell size	200	200	200	200	2000	200
B_{max}	max farm size	2000	2000	2000	2000	200 or 2000	2000
m_S	initial cell size mean	100	U(0, 200)	100	100	100	U(0, 200)
s_S	initial cell size SD	10	N/A	10	10	10	N/A
m_B	initial farm size mean	0	0	100	0	0	0
s_B	initial farm size SD	0.0	0.0	10	0.0	0	0.0
m_h	initial culling rate mean	0.0	0.0	0.0	0.0	0.2	0.3
s_h	initial culling rate SD	0.0	0.0	0.1	0.0	0.1	0.0
m_f	initial farming rate mean	0.0	0.0	0.0 … 1.0	0.0	0.1	0.0
s_f	initial farming rate SD	0.0	0.0	0.1	0.0	1.0	0.0
m_c	initial farming cost mean	-1 … 0	-0.4 or -0.2	-0.4	-0.1	-0.2	-0.2
s_c	initial farming cost SD	0.0	0.0	0.1	0.0	0.1	0.0
c_b	baseline farming apparatus cost	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
β	max benefit	1	1	1 … 5	1	1	1
g	inherent farm growth rate	1.04	1.04	1.04	1.04	1.04	1.0 … 1.1
μ_h	farm allocation mutation rate	0.05	0.05	0.05	0.05	0.0	0.005
sd_h	farm allocation mutation step SD	0.05	0.05	0.05	0.05	0.0	0.05
μ_f	culling mutation rate	0.05	0.05	0.05	0.05	0.0	0.005
sd_f	culling mutation step SD	0.05	0.05	0.05	0.05	0.0	0.05
μ_c	farm cost mutation rate	0.0	0.0	0.05	0.0	0.0	0.0
sd_c	farm cost mutation step SD	0.0	0.0	0.05	0.05	0.0	0.05
ρ	resource function	B	B	B	B, T, S	B	T
v	resource function period length	10 000	10 000	10 000	10 000	10 000	5000
	resource function period length SD	0.2	0.0	0.8	0.2	0.2	0
R_{min}	resource function min	0.0 … 1.0	0.0	0.0	0.0	0.0	0.0 … 0.5
R_{max}	resource function max	1.0	1.0	1.0	1.0	1.0	1.0
t_D	mean length of poor period (D)*	0.0 … 1.0	0.8	0.2	0.7	0.7	0.5
t_G	mean length of rich period (G)**	(1.0 - t_D)	N/A	N/A	0.1	N/A	0.3
	sigmoid wave steepness***	N/A	N/A	N/A	30	N/A	N/A
	sigmoid wave sharpness***	N/A	N/A	N/A	4	N/A	N/A

* Given as a fraction of the resource function period length.
** Given as a fraction of resource function period length. Only used for the random trapezoid-wave function.
*** Only used for the logistic wave function.

Supplementary references

[1] Szathmáry E (1991) Simple growth laws and selection consequences. *Trends in Ecology and Evolution* 6(11):366–370.

[2] Dey G, Thattai M, Baum B (2016) On the archaeal origins of eukaryotes and the challenges of inferring phenotype from genotype. *Trends in Cell Biology* 26(7):476–485. Special Issue: 25 Years of Trends in Cell Biology.

[3] Ettema TJG, Lindås AC, Bernander R (2011) An actin-based cytoskeleton in archaea. *Molecular Microbiology* 80(4):1052–1061.

[4] Dean AD et al. (2016) Host control and nutrient trading in a photosynthetic symbiosis. *Journal of Theoretical Biology* 405(Supplement C):82–93. Advances in Modelling Biological Evolution: Linking Mathematical Theories with Empirical Realities.

[5] Brock DA, Douglas TE, Queller DC, Strassmann JE (2011) Primitive agriculture in a social amoeba. *Nature* 469(7330):393–396.

[6] Fenchel T, Bernard C (1993) A purple protist. *Nature* 362(6418):300–300.

[7] Jeon KW (1995) The large, free-living amoebae: wonderful cells for biological studies. *Journal of Eukaryotic Microbiology* 42(1):1–7.

[8] Fenchel T, Finlay BJ (1991) Synchronous division of an endosymbiotic methanogenic bacterium in the anaerobic ciliate *Plagiopyla frontata* Kahl. *The Journal of Protozoology* 38(1):22–28.

[9] Horn M et al. (2001) Members of the Cytophaga-Flavobacterium-Bacteroides phylum as intracellular bacteria of acanthamoebae: proposal of `Candidatus *Amoebophilus asiaticus*`. *Environmental Microbiology* 3(7):440–449.

[10] Horn M, Wagner M (2004) Bacterial endosymbionts of free-living amoebae. *Journal of Eukaryotic Microbiology* 51(5):509–514.

[11] DiSalvo S et al. (2015) *Burkholderia* bacteria infectiously induce the proto-farming symbiosis of *Dictyostelium* amoeba and food bacteria. Proceedings of the National Academy of Sciences 112(36):E5029–E5037.

[12] Fenchel T, Finlay BJ (1991) Endosymbiotic methanogenic bacteria in anaerobic ciliates: Significance for the growth efficiency of the host. *The Journal of Protozoology* 38(1):18–22.