SUPPLEMENTARY INFORMATION

Contents

Supplementary text S1: Hamilton's rule and Price's covariance formalism

We recall here how Hamilton's rule can be derived using Price's covariance formalism [1,2].

A. Price equation

Consider a population composed of groups, indexed by g, themselves composed of individuals, indexed by i. An individual i can either be a producer, $p_i = 1$, or a non-producer, $p_i = 0$. Initially, a group g contains a fraction x_g of the total population and has a proportion p_g of producers; after a given amount of time, the group size is multiplied by a factor w_g, which is assumed to depend only on p_g (and not on the absolute size of the groups as it could more generally do).

The so-called Price equation can be written at two levels. At the population level, it gives the overall change in the proportion of producers $\Delta\bar{p} = \bar{p}' - \bar{p}$, where $\bar{p} = \sum_g x_g p_g$ represents the initial proportion, and $\bar{p}' = \sum x'_g p'_g$, with $p'_g = p_g + \Delta p_g$ and $x'_g = x_g w_g / \sum_h x_h w_h$, the final proportion (an implicit assumption here is that we are only interested in a mixture of all the groups after a fixed period of time defining a "final" time; otherwise quantities other than the global mean could also be of interest). From these definitions, the following identity, known as the Price equation, follows:

$$\langle w_g \rangle \Delta\bar{p} = \mathrm{Cov}(w_g, p_g) + \langle w_g \Delta p_g \rangle. \tag{1}$$

Here, averages and covariances are taken with weights depending on the initial relative sizes of the groups: for any quantities a_g and b_g defined at the group level, $\langle a_g \rangle = \sum_g x_g a_g$ and $\mathrm{Cov}(a_g, b_g) = \langle a_g b_g \rangle - \langle a_g \rangle \langle b_g \rangle$. The second term in Eq. (1) involves Δp_g, the change in proportion of producers within group g, which can be expressed in terms of a Price equation at the group level:

$$w_g \Delta p_g = \mathrm{Cov}_g(w_i, p_i). \tag{2}$$

Here, w_i corresponds to the multiplicative factor by which producers (if $p_i = 1$) or non-producers (if $p_i = 0$) are multiplied (at this level, $x_i = 1$ and $\Delta p_i = 0$ since the unit is an individual and no conversion between producer and non-producer is assumed). $\mathrm{Cov}_g(w_i, p_i) = \langle w_i p_i \rangle_g - \langle w_i \rangle_g \langle p_i \rangle_g$ where the subscript g indicates that averages are taken for individuals i belonging to the group g; in particular, we have by definition $\langle w_i \rangle_g = w_g$ and $\langle p_i \rangle_g = p_g$.

B. Covariance and regression

Given a set of pairs (p_g, w_g) with weights x_g we can always write

$$\mathrm{Cov}(w_g, p_g) = \beta(w_g, p_g)\mathrm{Var}(p_g) \tag{3}$$

where $\mathrm{Var}(p_g) = \mathrm{Cov}(p_g, p_g) = \langle p_g^2 \rangle - \langle p_g \rangle^2$ is the variance of p_g and $\beta(w_g, p_g)$ corresponds to the regression coefficient of w_g against p_g. $\beta(p_g, w_g)$ can also be interpreted as the value of β which minimizes, together with the other regression coefficient α, the residual sum

$$R^2 = \langle (w_g - (\alpha + \beta p_g))^2 \rangle = \sum_g x_g (w_g - \alpha - \beta p_g)^2. \tag{4}$$

Graphically, $\beta(p_g, w_g)$ is therefore the slope of the best linear interpolation, in the mean-square sense (using weights x_g) of the data points (p_g, w_g) (see Fig. 3 of main text).

Similarly, we can write

$$\mathrm{Cov}_g(w_i, p_i) = \beta_g(w_i, p_i)\mathrm{Var}_g(p_i). \tag{5}$$

With these subtitutions, the condition $\Delta\bar{p} > 0$ is equivalent to

$$\beta(w_g, p_g)\mathrm{Var}(p_g) + \langle \beta_g(w_i, p_i)\mathrm{Var}_g(p_i) \rangle > 0. \tag{6}$$

C. From Price to Hamilton

If $\beta_g(w_i, p_i)$ is independent of g, corresponding to an intrinsic individual cost independent of the nature of the group g to which an individual belongs, Eq. (6) can be rewritten

$$\beta(w_g, p_g) r + \beta_g(w_i, p_i) > 0. \tag{7}$$

where

$$r = \frac{\mathrm{Var}(p_g)}{\langle \mathrm{Var}_g(p_i) \rangle} \tag{8}$$

is a purely "geometrical" parameter, that depends only on the initial composition of the groups.

If we consider for instance groups of equal size with a proportion p_g of producers in each group, we have $p_i = 1$ with probability p_g and $p_i = 0$ with probability $1 - p_g$, so that $\langle p_i^m \rangle_g = p_g$ for all m, and in particular $\mathrm{Var}_g(p_i) = p_g - p_g^2$. In such a case,

$$r = \frac{\langle p_g^2 \rangle - \langle p_g \rangle^2}{\langle p_g \rangle - \langle p_g^2 \rangle}. \tag{9}$$

D. Linear models

Eq. (7), known as Hamilton's rule, is most easily interpretable when the regression coefficients $\beta(w_g, p_g)$ and $\beta_g(w_i, p_i)$ are independent of the distribution of the p_g's. In the model introduced in Box 1, it is thus assumed that there is a fixed production cost per individual that is independent from group properties. For this model, we can write the linear relation

$$w_i = a + kp_g - cp_i, \tag{10}$$

where $a + kp_g$ is the multiplicative factor for non-producers ($p_i = 0$) and $a + kp_g - c$ for producers ($p_i = 1$) when they are in a group with a proportion p_g of producers. We have then $\beta_g(w_i, p_i) \simeq -c$ (see next paragraph) and, since averaging within a group leads to $w_g = a + (k - c)p_g$, we have $\beta(w_g, p_g) = k - c$. By introducing $b \equiv k - c$, we thus obtain Hamilton's rule under the form $br - c > 0$, with r given by Eq. (8).

When deriving these formulas, p_i and p_g should not be treated as independent variables (for instance, if $p_g = 0$, then necessarily $p_i = 0$). Introducing p_{g-i}, the fraction of producers in the subgroup of size $n_g - 1$ where i is excluded (n_g representing the total size of group g), we have $p_g = p_{g-i}(n_g - 1)/n_g + p_i/n_g$. Therefore, $w_i = 1 + k(n_g - 1)/n_g p_{g-i} + (k/n_g - c)p_i$ where now, *conditionally on* p_g, the variables p_{g-i} and p_i are uncorrelated. We thus get $\beta_g(w_i, p_i) = -c + k/n_g$, which simplifies to $\beta_g(w_i, p_i) \simeq -c$ when the size of the group n_g is large.

E. Non-linearities and interpretation of b

In general, the relation between w_g and p_g is non linear and the "benefit" $b = \beta(w_g, p_g)$ depends on the distribution into groups (see Fig. 3 for an illustration). Formally, Eq. (7) still holds but since both b and r change when the composition of the group changes, and since the change of $b = \beta(w_g, p_g)$ cannot be from the current values of b and c only, the relation cannot indicate how the direction of selection is affected when the system is perturbed. Only when operating in a regime where w_g varies linearly with p_g can a single number, b, provide a sufficient characterization.

Note also that, even in the linear case, the notion of benefit that enters in Hamilton's rule is that of a "differential benefit" that addresses only changes in the relative proportion of producers and non-producers: if the relation $w_g = a + bp_g$ is changed to $w_g = a' + bp_g$ with $a' > a$, there is an (absolute) "benefit" in the sense that the population globally improves its growth, but no (differential) "benefit", in the sense of Hamilton's rule, since the ratio between producers and non-producers is not affected. This situation is illustrated with our system in Fig. 3.

Supplementary text S2: A model for the dynamics of producer and nonproducer strains

We introduce here a simple model for the dynamics of the synthetic producer and nonproducer strains. The population dynamics depend on the concentration of autoinducer, AI, which is assumed to modulate the growth rate between $s_{\min}$ and $s_{\max}$ following

$$s(AI) = s_{\min} + (s_{\max} - s_{\min}) \left(\frac{AI}{AI + K_M} \right), \tag{11}$$

where K_M is the concentration of autoinducer at which half-maximal response occurs. Assuming a logistic growth with a common carrying capacity K and an autoinducer dependent growth rate, the dynamics of producers is described by

$$\frac{d[P]}{dt} = s(AI)\ P \left(1 - \frac{P + NP}{K} \right), \tag{12}$$

and the dynamics of producers by

$$\frac{d[NP]}{dt} = \kappa\ s(AI)\ NP\ \left(1 - \frac{P + NP}{K} \right), \tag{13}$$

where $\kappa > 1$ represents the relative advantage of nonproducers. Finally, the autoinducer production rate is taken to be proportional to number of producers

$$\frac{d[AI]}{dt} = \alpha P. \tag{14}$$

This model is simply focused on the growth dynamics of producers and nonproducers and notably does not explicitly incorporate on/off rates, transport rates, transcription/translation/degradation rates. When this model was simulated using parameter values estimated from experimental data and initial conditions corresponding to our experimental conditions, the simulation qualitatively reproduced the main experimental observations [3]: $\Delta\bar{p}$ was greater than zero while Δp_g was less than zero for all g (Supplementary Figure S1). The model also qualitatively reproduces the dependence of growth of producers and nonproducers on p_g (Supplementary Figure S2).

Supplementary references

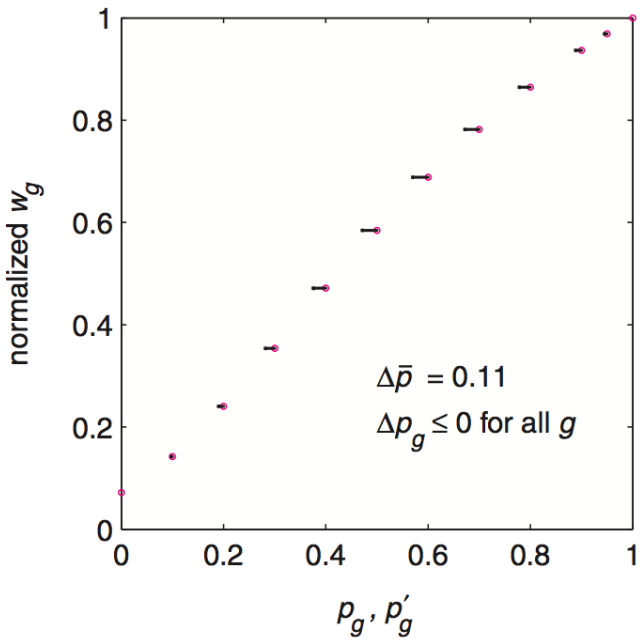
[1] Price G.R., *Selection and covariance*, Nature 227, 520-521 (1970).

[2] Hamilton W.D., *Innate social aptitudes of man: an approach from evolutionary genetics*, in ASA Studies 4: Biosocial Anthropology, edited by R. Fox (Malaby Press, London, 1975), pp. 133-153.

[3] Chuang J.S., Rivoire O., Leibler S., *Simpson's paradox in a synthetic microbial system*, Science 323, 272-275 (2009)

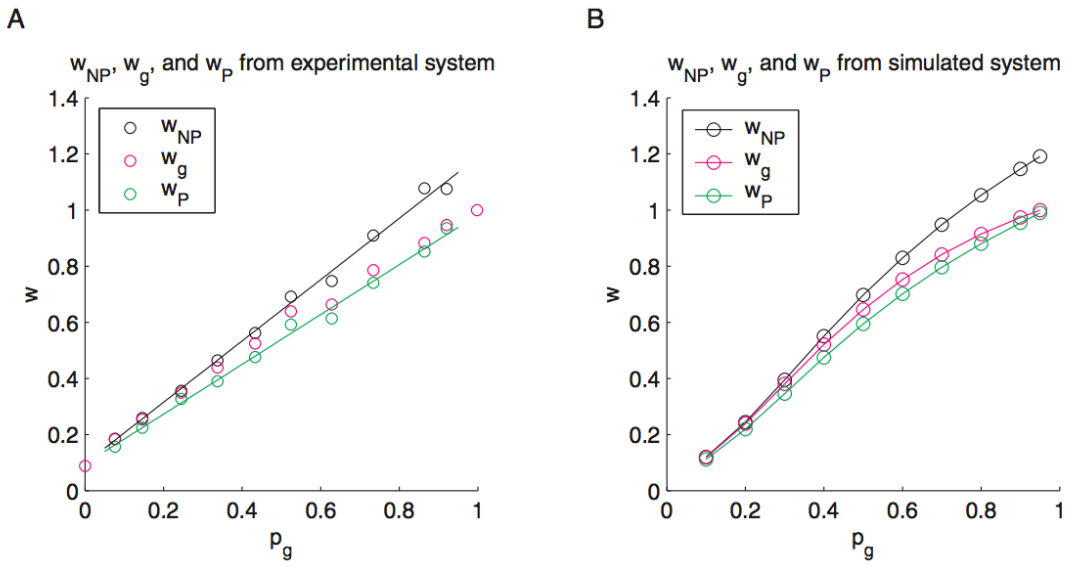
Supplementary figures

Figure S1: Simpsons paradox in a simulated population of producers and nonproducers



The model described in Supplementary Text S2 was numerically integrated in Matlab using parameters $K = 4 \times 10^9$ cells/ml, $K_M = 3$ μM, $s_{\max} = 0.0075$ min^{-1}, $s_{\min} = s_{\max}/10$, $\kappa = 1.05$, $\alpha = 3 \times 10^{-16}$ mmol cell^{-1} min^{-1}. Ten mixed subpopulations, each initially containing 4×10^7 cells/ml (representing the 100-fold dilution made at the beginning of an experiment with living cells), were formed with $p_g = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$. This distribution is the same as that used in our experiments with living cells. Since the amount of autoinducer in a saturated culture of pure producers is approximately 30 μM, the initial autoinducer concentration in each subpopulation was $30 \times (p_g/100)$ μM. The system was simulated for 780 min (13 hours), corresponding to the length (12 to 13 hours) of a typical experiment with living cells. The initial producer proportion p_g of each subpopulation is plotted as a magenta circle and a black line segment connects p_g to the final producer proportion p_g. In agreement with experiments with living cells, Simpsons paradox is observed in the simulation, since $\Delta p_g \leq 0$ for all g, but $\Delta\bar{p} > 0$ globally.

Figure S2: Growth of producers and nonproducers as a function of initial producer proportion



w_{NP} (black) is the same as w_- of Box 1. w_P (green) is the same as w_+ of Box 1. w_g (magenta) is the growth of the subpopulation composed of p_g producers and $(1 - p_g)$ nonproducers.

(A) Representative sample for experimental data, with lines representing fits from linear regression.

(B) Simulated system of Supplementary Figure S1.