Online Appendix

A1. Derivation of selection gradient and stability with genetic assortment

At each generation, an individual is identical by descent to some number in its group including itself (among the total group members N), which is a random variable K. We define Pr(k) as the probability that a rare mutant in the population has exactly $k-1$ other genetically identical mutants in its group. Thus, $\langle K \rangle = \sum_{k=1}^{N} k\mathrm{Pr}(k)$ is the expectation of the number of mutants in a random mutant's group, and $\langle K \rangle / N$ is the likelihood that a recipient of the mutant's act is also a mutant, again assuming mutants are rare in the population. So with no assortment $\mathrm{Pr}(1) = 1$, and with assortment $\mathrm{Pr}(1) < 1$. Our expression for relatedness of a rare mutant to a random individual in its group is just the expected genotype of a recipient's genotype to that of the actor (unity) (Grafen 1985) which is:

$$r_w = \frac{\langle K \rangle}{N} = \sum_{k=1}^{N} \frac{k}{N} \mathrm{Pr}(k). \tag{A1}$$

This is whole-group relatedness (r_w) which is of interest since cooperators get a direct benefit from their own cooperation. (Others-only relatedness, or r_o, is relatedness to recipients other than the focal cooperator itself with $r_w = r_o \frac{N-1}{N} + \frac{1}{N}$ (Pepper 2000). This value is the same as $\langle \rho \rangle$ in the main text). The expected payoff of an individual with trait y in a monomorphic population with trait value x is given by

$$P(y,x) = \sum_{k=1}^{N} \frac{\mathrm{Pr}(k) B(ky + (N-k)x)}{N} - C(y). \tag{A2}$$

From this, the selection gradient becomes:

$$D(x) = \sum_{k=1}^{N} \frac{k\mathrm{Pr}(k)B'(Nx)}{N} - C'(x) \tag{A3}$$

$$= r_w B'(Nx) - C'(x) = \frac{B'(Nx)(1 + r_o(N-1))}{N} - C'(x), \tag{A4}$$

The solutions of $D(x^*) = 0$ are the singular strategies. The condition for convergence stability of a singular strategy is:

$$\frac{\partial D}{\partial x}|_{x=x^*} = NB''(Nx^*)r_w - C''(x^*) \tag{A5}$$

$$= B''(Nx^*)(1 + r_o(N-1)) - C''(x^*) < 0, \tag{A6}$$

Using ρ from the main text, this condition is simply:

$$\frac{\partial D}{\partial x}|_{x=x^*} = NB''(Nx^*)r_w - C''(x^*) \tag{A7}$$

$$= B''(Nx^*)(1 + \langle\rho\rangle(N-1)) - C''(x^*) < 0, \tag{A8}$$

and for evolutionary stability:

$$\frac{\partial^2 P}{\partial y^2}|_{y=x^*} = \frac{1}{N} \sum_{k=1}^{N} \mathrm{Pr}(k)k^2 B''(Nx^*) - C''(x^*) \tag{A9}$$

$$= \frac{B''(Nx^*)}{N} \sum_{k=1}^{N} \mathrm{Pr}(k)k^2 - C''(x^*) < 0. \tag{A10}$$

Now

$$\mathrm{Var}(K) = \sum_{k=1}^{N} \mathrm{Pr}(k)(k - \langle K\rangle)^2 \tag{A11}$$

$$\sum_{k=1}^{N} \mathrm{Pr}(k)k^2 = 2\sum_{k=1}^{N} \mathrm{Pr}(k)k\langle K\rangle - \sum_{k=1}^{N} \mathrm{Pr}(k)\langle K\rangle^2 + \mathrm{Var}(K) \tag{A12}$$

$$\sum_{k=1}^{N} \Pr(k)k^2 = \langle K \rangle^2 + \mathrm{Var}(K) \tag{A13}$$

By using ρ from the main text, the condition for evolutionary stability becomes:

$$\frac{B''(Nx^*)}{N}[(1+(N-1)\langle\rho\rangle)^2 + \mathrm{Var}(\rho)(N-1)^2] - C''(x^*) < 0. \tag{A14}$$

A2. Assuming linear costs and nondecreasing benefits, if a positive level of cooperation evolves without assortment at some group size N0, then a positive level of cooperation evolves in arbitrarily large group sizes when $r_o > 1/N_0$:

For some x_o, N_0,the selection gradient is positive:

$$\frac{B'(N_0x_0)}{N_0} - c > 0. \tag{A15}$$

Now for some arbitrarily high N_1, let $x_1 = \frac{N_0x_0}{N_1}$. The selection gradient now is

$$\frac{B'(N_1x_1)}{N_1} + \frac{r_o(N_1-1)B'(N_1x_1)}{N_1} - c > 0. \tag{A16}$$

which is equivalent to

$$r_oB'(N_1x_1) + \frac{B'(N_1x_1)(1-r_o)}{N_1} - c \tag{A17}$$

Now, so long as $r_o > 1/N_o$, this gradient is positive since the first term itself is greater than c.

A3. Variance can affect evolutionary branching

Here we show an example (fig A3) of a situation where $\langle\rho\rangle$ is constant but $\mathrm{Var}(\rho)$ is not, and this distinction determines whether evolutionary branching occurs. When $N = 4$, variance may be zero if every group has two individuals of one genotype and two of another, or it may follow a binomial distribution with $\langle\rho\rangle = 1/3$. Both of these give $\langle\, r_w\rangle = 1/2$.

A4. Quorum sensing model

The payoff to an individual with trait values x and s in two distinct group sizes N_1 and N_2 is:

$$P_1 = \frac{B\left(Q(s_1,N_1)x_1 + \ldots + Q(s_{N_1},N_1)x_{N_1}\right)}{N_1} - C\left(Q(s_i,N_1)x_i\right) \quad \text{(A18)}$$

$$P_2 = \frac{B\left(Q(s_2,N_2)x_2 + \ldots + Q(s_{N_2},N_2)x_{N_2}\right)}{N_2} - C\left(Q(s_i,N_2)x_i\right) \quad \text{(A19)}$$

where $Q(s_i,N) = 1/(1+e^{s_i - N})$ as in the main text. The expected payoff then for a pair (x,s), given that the individual randomly joins one of the two group sizes is $P = (P_1 + P_2)/2$. The singular strategies occur when $\frac{\partial P}{\partial x} = \frac{\partial P}{\partial s} = 0$, and these zeros of these selection gradients define the isoclines of figure 5. See (Geritz et al. 1997; Brown and Taylor 2010) for details of analyzing systems of multiple trait evolution.