nature ecology & evolution

SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41559-018-0564-9

In the format provided by the authors and unedited.

Division of labour and the evolution of extreme specialization

Guy A. Cooper* **and Stuart A. West**

Department of Zoology, University of Oxford, Oxford, UK. *e-mail: guy.cooper@zoo.ox.ac.uk

Supplementary information:
Division of labour and the evolution of extreme specialisation

Guy Alexander Cooper* and Stuart Andrew West
Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom

*Correspondence:
guy.cooper@zoo.ox.ac.uk

This supplementary information contains further analysis and results. In section 1, we derive the equilibrium conditions that were used to show when each of the possible strategies was the ESS in our analysis. This includes separate conditions for the two different generalist types (section 1.1), for generalist and pure reproductive division of labour (section 1.2), for sterile helper and generalist division of labour (section 1.3), for sterile helper and pure reproductive division of labour (section 1.4) as well as for uniform non-cooperation (section 1.5) and uniform cooperation (section 1.6). In sections 2 and 3, we show that these equilibria are indeed uninvadable and convergent stable and thus ESSs with respect to a haystack model. In our model we assumed that the cost of cooperation is linear with respect to helper investment. In section 4, we show that our results hold if the cost of cooperation is allowed to be non-linear. In section 5, we test our hypothesis that extreme forms of specialisation are favoured by an evolutionary feedback loop between the levels of cooperation of each phenotype via dynamic, individual-based simulations. We did not specify the demographic processes that generate relatedness in our model and instead assumed that R was an independent parameter (an 'open' model). In section 6, we 'close' a simplification of our model by including limited dispersal and overlapping generations as processes that lead to the build-up of genetic correlations between interacting individuals. In section 7, we discuss further predictions of our model. This includes further discussion of how the efficiency benefits of specialisation (section 7.1), the sociality of the trait (section 7.2) and the essentiality of the trait (section 7.3) favour division of labour. We also further discuss predictions for the optimal proportion of helpers in a division of labour (section 7.4) and how our results compare to a previous, population genetic model of sterility in haplodiploid insects (section 7.5). In supplementary tables 1 and 2, we summarise how our model compares and links to previous theoretical work on the evolution of division of labour.

1 Supplementary information: equilibrium analysis

We discretised parameter space and checked, for all combinations of parameter values, whether each of the possible strategies was an equilibrium of the system. We considered R={1 (clonal), 1/3 (non-clonal; not presented) 1/7 (non-clonal)}, e={1 (essential), 9/10 (non-essenital), 8/10 (non-essential; not presented)}, $\log(\alpha)=\{-2.5,\ldots,2.5\}$ and $\lambda=\{1/100,\ldots,1\}$. The discretisations of $\log(\alpha)$ and λ were each evenly spaced over 151 nodes. The results of this analysis are depicted in Figure 3.

Here, we derive the conditions that were used to show when each of the possible strategies of the model are equilibria of the fitness landscape.

An *equilibrium* strategy is a fixed-point of the system: natural selection within a monomorphic population, to a first order-approximation, does not act to 'push' a mutant strain in any permissible direction locally. An equilibrium strategy is *convergent stable* if it is an attractor of the dynamical system: a monomorphic population close to the equilibrium strategy always experiences directional selection that favours a rare mutant lineage with trait value closer to the equilibrium[1]. An equilibrium is *uninvadable* if is a local maximum of the fitness landscape[2]. If this is the case, we say that the equilibrium is an Evolutionarily Stable Strategy (ESS), where an ESS is the joint strategy employed by all cells in the population such that no mutant lineage with an alternate strategy can successfully invade the population[3,4]. We expect the ESS to be the long-term strategy that a system will evolve to.

In subsequent sections, we show how convergence stability and uninvadability were evaluated. In particular, we found that all equilibria considered by our numerical discretization were convergent stable and uninvadable. As such, which strategy is an ESS of a system is determined entirely by the equilibrium conditions of the model. Overall, we found that one and only one strategy was the ESS for all parameter combinations except those for which $\alpha = 1$ and $\lambda = 1$, where there was no ESS.

Let the fitness of an arbitrary individual from a rare mutant strain be given by:

$$\begin{aligned} W = p(1-q_1)\Big[(1-e) + e\big((1-\lambda)q_1^\alpha + \lambda(PQ_1^\alpha + (1-P)Q_2^\alpha)\big)\Big] \\ + (1-p)(1-q_2)\Big[(1-e) + e\big((1-\lambda)q_2^\alpha + \lambda(PQ_1^\alpha + (1-P)Q_2^\alpha)\big)\Big]. \end{aligned}$$

Let $W_p(p, q_1, q_2) = {}^{\partial W}/_{\partial p} + R^{\partial W}/_{\partial P}$, $W_{q_1}(p, q_1, q_2) = {}^{\partial W}/_{\partial q_1} + R^{\partial W}/_{\partial Q_1}$ and $W_{q_2}(p, q_1, q_2) = {}^{\partial W}/_{\partial q_2} + R^{\partial W}/_{\partial Q_2}$ be the respective directional selection terms for p, q_1 and q_2, where all partial derivatives are evaluated at $p = P, q_1 = Q_1, q_2 = Q_2$[5,6]. These are given here:

$$\begin{aligned} W_p = \quad & (1-q_1)\big[(1-e) + e\big((1-\lambda)q_1^\alpha + \lambda(pq_1^\alpha + (1-p)q_2^\alpha)\big)\big] \\ & -(1-q_2)\big[(1-e) + e\big((1-\lambda)q_2^\alpha + \lambda(pq_1^\alpha + (1-p)q_2^\alpha)\big)\big] \\ & +p(1-q_1)e\lambda(q_1^\alpha - q_2^\alpha)R \\ & +(1-p)(1-q_2)e\lambda(q_1^\alpha - q_2^\alpha)R \end{aligned}$$

$$\begin{aligned} W_{q_1} = \quad & -p\big[(1-e) + e\big((1-\lambda)q_1^\alpha + \lambda(pq_1^\alpha + (1-p)q_2^\alpha)\big)\big] \\ & +p(1-q_1)e(1-\lambda)\alpha q_1^{\alpha-1} \\ & +p(1-q_1)e\lambda p\alpha q_1^{\alpha-1}R \\ & +(1-p)(1-q_2)e\lambda p\alpha q_1^{\alpha-1}R \end{aligned}$$

$$\begin{aligned} W_{q_2} \quad & = -(1-p)\big[(1-e) + e\big((1-\lambda)q_2^\alpha + \lambda(pq_1^\alpha + (1-p)q_2^\alpha)\big)\big] \\ & +(1-p)(1-q_2)e(1-\lambda)\alpha q_2^{\alpha-1} \\ & +p(1-q_1)e\lambda(1-p)\alpha q_2^{\alpha-1}R \\ & +(1-p)(1-q_2)e\lambda(1-p)\alpha q_2^{\alpha-1}R \end{aligned}$$

For each trait, directional selection quantifies the direction in which a rare mutant with a slight change in the trait value may increase its fitness when in a population that is otherwise monomorphic for all three trait values (increase in trait value favoured if directional selection is positive and decrease in trait value favoured if directional selection is negative).

In the following, we solve for the equilibrium conditions of each strategy in turn.

1.1 Two different generalist types

Division of labour with two different generalist types occurs in the interior of the state-space, in the domain: $\mathbf{D1} = (0 < p < 1, 0 < q_1 < 1, 0 < q_2 < 0; q_1 \neq q_2)$. The last condition states that the cooperative investment of each phenotype must be distinct, otherwise the strategy corresponds to uniform cooperation. We employ the classic method developed by Taylor and Frank and Brown and Taylor in this case[5,6]. We solve first for the equilibrium value (p^*, q_1^*, q_2^*) such that directional selection in each trait is zero: $W_p(p^*, q_1^*, q_2^*) = W_{q_1}(p^*, q_1^*, q_2^*) = W_{q_2}(p^*, q_1^*, q_2^*)=0$. We do this by solving for the null-planes as a function of (q_1, q_2) for each trait and finding the joint value (q_1^*, q_2^*) for which the three planes intersect. This is done numerically for the full model. If we assume that the cooperative trait is essential ($e = 1$) and others-only ($\lambda = 1$), then we have the following nullplanes:

$$\begin{aligned} q_1\text{-nullplane:} \qquad & p(q_1, q_2) = \frac{q_2^\alpha - (1-q_2)\alpha R q_1^{\alpha-1}}{q_2^\alpha - q_1^\alpha + \alpha R q_1^{\alpha-1}(q_2 - q_1)} \\ q_2\text{-nullplane:} \qquad & p(q_1, q_2) = \frac{q_2^\alpha - (1-q_2)\alpha R q_2^{\alpha-1}}{q_2^\alpha - q_1^\alpha + \alpha R q_2^{\alpha-1}(q_2 - q_1)} \end{aligned}$$

By setting these equal and rearranging, we arrive at the equation:

$$\alpha = \frac{\log(1-q_1) - \log(1-q_2)}{\log(q_1) - \log(q_2)}$$

If $q_1 > q_2$ then the numerator of the above equation is negative, the denominator is positive, and thus $\alpha < 0$ which is not permitted in our model. The same result holds if we stipulate that $q_2 > q_1$. We conclude that there are no intersections of the two nullplanes in the simplified model.

In the full model, we solved numerically for an intersection point of the three nullplaces and found no such point within the state space for any of the considered parameter combinations. As such, we conclude that two different generalist types is never a form of division of labour that is an ESS in our analysis.

1.2 Generalist and pure reproductive division of labour

Generalist and pure reproductive division of labour occurs in the symmetrically equivalent domains **D2** $= (0 < p < 1, 0 < q_1 < 1, q_2 = 0)$ or **D3** $= (0 < p < 1, q_1 = 0, 0 < q_2 < 1)$. Without loss of generality let us consider only the domain **D2**. For this strategy, one of the traits, q_2, is now a boundary trait ($q_2 = 0$; in contrast to an interior trait: $0 < q_2 < 1$) and as such the methodology for evaluating whether the strategy is an equilibrium with respect to q_2 must be adapted. There are two sources of instability for this strategy: loss of partial division of labour and loss of pure reproduction.

1.2.1 Stability to loss of partial division of labour

If we consider only the interior traits, p and q_1, then the analysis proceeds with the approach developed by Brown and Taylor[6]. We first solve for the nullclines of p and q_1 as functions of q_1. We then solve for the equilibrium values (p^*, q_1^*) such that the null-clines intersect (no directional selection in either trait). We do this numerically for the full model. If no such strategy, (p^*, q_1^*), exists then we conclude that generalist and pure reproductive division of labour is not an ESS. For the simplified model in which the trait is essential ($e = 1$), we have the following nullclines:

$$p\text{-nullcline:} \qquad p(q_1) = \frac{\lambda R + (1-\lambda)(1-q_1)}{\lambda(R+1)}\frac{1}{q_1}$$

$$q_1\text{-nullcline:} \qquad p(q_1) = \frac{\alpha(\lambda R + (1-\lambda)) - q_1(1-\lambda)(1+\alpha)}{\lambda(\alpha R+1)}\frac{1}{q_1}$$

These are monotonically decreasing functions from positive infinity at $q_1 = 0$. At $q_1 = 1$, the p-nullcline is equal to $R/(R+1)$, which is strictly contained in $]0,1[$ and the q_1-nullcline is strictly less than 1. Thus, both nullclines traverse the domain $[0,1] \times [0,1]$. Directional selection in both traits is negative at the point, $(p = 1, q_1 = 1)$. If $\alpha < 1$ the nullclines do not intersect in the positive quadrant and the q_1 nullcline is strictly less than the p-nullcline for all value of $0 \leq q_1 \leq 1$ and thus the equilibrium strategy is the intersection point of the q_1-nullcline with the boundary $p = 1$ (uniform cooperation; under the constraint that $q_2 = 0$). If $\alpha > 1$, then the nullclines intersect in the positive quadrant (q_1 nullcline crosses p-nullcline from above) and the equilibrium strategy (ignoring directional selection in q_2) depends on where the intersection point of the nullclines occurs. If the nullclines intersect before entering the domain, $(0 \leq p \leq 1, 0 \leq q_1 \leq 1)$, then the equillibrium is once again the intersection point of the q_1-nullcline with the boundary $p = 1$ (uniform cooperation; under the constraint that $q_2 = 0$). If the nullclines intersect after departing the domain $q_1^* > 1$, then the equilibrium strategy is the intersection point of the p-nullcline with the upper boundary of $q_1 = 1$ (sterile helper and pure reproductive division of labour; under the constraint that $q_2 = 0$). By exclusion, we have that generalist and pure reproductive is the equilibrium strategy (under the constraint that $q_2 = 0$) if the nullclines intersect within the domain.

We can determine whether this is the outcome by considering directional selection at the aforementioned intersection points of each nullcline with the upper boundary of the opposing trait. The q_1-nullcline intersects the upper boundary, $p = 1$, at $q_1^* = \alpha(\lambda R+(1-\lambda))/(1+\alpha(\lambda R+(1-\lambda))$. If $w_p(1, q_1^*, 0) = (1-q^*)q^{*\alpha} - \lambda q^{*\alpha} + (1-q^*)\lambda R q^{*\alpha} < 0$, then the intersection point of the nullclines occurs after they have entered the domain. Meanwhile, the p-nullcline intersects the upper boundary, $q_1 = 1$, at $p^* = R/(R+1)$. If $w_{q_1}(p^*, 1, 0) = -p^*((1-\lambda+p^*\lambda)+(1-p^*)\alpha\lambda R p^* < 0$, then the intersection point of the nullclines occurs before exiting the domain. If the first condition is not satisfied then the equilibrium strategy is uniform cooperation under the constraint that $q_2 = 0$. If the second condition is not satisfied then the equilibrium strategy is sterile helper and pure reproductive under the constraint that $q_2 = 0$. In contrast, generalist and pure reproductive is stable to loss of partial division of labour mutations if both constraints are satisfied, giving the broader condition: $(1/\lambda < \alpha < (R+(1-\lambda))/\lambda R)$. If we further assume that $\lambda = 1$, then we find that these constraints collapse both to 1 and that thus that generalist and pure reproductive is never stable (if $\alpha = 1$, the nullclines intersect fully and the population may evolve neutrally along a spectrum of strategies).

We may now consider the conditions for which the strategy is an equilibrium along the q_2 axis.

1.2.2 Stability to loss of pure reproduction

Generalist and pure reproductive division of labour is composed of a boundary trait ($q_2 = 0$). As such, the condition to show that the boundary trait is at equilibrium is different. For an interior trait, we need to establish that directional selection of the trait is equal to zero at the putative equilibrium. For a boundary trait, we instead must show that the directional selection acts to 'push' the trait into the boundary. For our present purposes, this amounts to the condition that $W_{q_2} < 0$ at the putative equilibrium $(p^*, q_1^*, 0)$. Evaluating directional selection in q_2 we find a dependence on α:

$$W_{q_2}(p, q_1^*, 0) = \begin{cases} +\infty, & \text{if } \alpha < 1 \\ (1-p)(-(1-e) - e\lambda p q_1 + e((1-\lambda) + p(1-q_1)\lambda R + \lambda R(1-p))), & \text{if } \alpha = 1 \\ -(1-p)((1-e) + e\lambda p q_1^{\alpha}), & \text{if } \alpha > 1. \end{cases}$$

We thus have that a generalist and pure reproductive division of labour will always be stable to loss of pure reproduction mutations if $\alpha > 1$ and never stable if $\alpha < 1$. Stability if $\alpha = 1$ needs to be evaluated numerically.

1.2.3 Numerical results

We find in our numerical analysis that generalist and pure reproductive is only ever stable to loss of partial division of labour if there are efficiency benefits to specialisation ($\alpha > 1$). We note that stability to loss of pure reproduction is then automatically guaranteed by the above condition for loss of partial division of labour ($\alpha > 1$). We thus have that stability to loss of partial division of labour *guarantees* the stability of generalist and pure reproductive division of labour. As such, this intermediate form of division of labour will be stable so long as uniform cooperation, uniform non-cooperation or sterile helper and pure reproductive division are all unstable.

1.3 Sterile helper and generalist division of labour

Sterile helper and generalist division of labour occurs in the symmetrically equivalent domains **D4** $= (0 < p < 1, q_1 = 1, 0 < q_2 < 1)$ and **D5** $= (0 < p < 1, 0 < q_1 < 1, q_2 = 1)$. Without loss of generality we only consider the **D4** domain. Just as for the previous strategy, one of the traits here, q_1, is a boundary trait ($q_1 = 1$) and thus is treated differently. There are again two sources of instability for this strategy: loss of partial division of labour and loss of sterile helping.

1.3.1 Stability to loss of partial division of labour

If we hold $q_1 = 1$ fixed and focus on the interior traits, p and q_2, we may determine whether the strategy is stable to loss of partial division of labour. Analysis proceeds similarly to that for generalist and pure reproductive division of labour. We seek joint values (p^*, q_2^*) in the state space for which directional selection in both traits is zero. We do this by solving the value of q_2 such that the p and q_2 nullclines intersect. In the general model this is done numerically. If we assume that $e = 1$ and $\lambda = 1$, we find the following nullclines:

$$\begin{aligned} &p\text{-nullcline:} && p(q_2) = (q_2^{\alpha}(1+R) - R)/((1-q_2^{\alpha})(1+R)) \\ &q_2\text{-nullcline:} && p(q_2) = q_2^{\alpha-1}(q_2(1+\alpha R) - \alpha R)/((1 - q_2^{\alpha-1}(q_2(1+\alpha R) - \alpha R)) \end{aligned}$$

Solving for an intersection point leads to the following equation: $q_2^{\alpha-1}(\alpha + q_2(1-\alpha)) - 1 = 0$. If $\alpha = 1$ then the equation is trivially satisfied. In this scenario, the above nullclines intersect fully and thus a spectrum of equilibria exist along which the population may evolve neutrally, including strategies that are not sterile helper and generalist division of labour. We conclude that this strategy is then not an ESS in this case. We also have that $q_2 = 1$ solves the above equation. As this is also not a sterile helper and generalist strategy we may disregard it. If we assume that $\alpha \neq 1$, we then have that the left hand side of the above equation is either monotonically increasing or decreasing function of q_2 depending on the sign of α. In either case, we have that $q_2 = 1$ is the unique solution to the equation and thus that there exists no interior intersection of the nullclines in the simplified model. Thus, if the trait is essential ($e = 1$) and others-only ($\lambda = 1$), we predict that sterile helper and pure reproductive division of labour will never be an ESS. For the more general model, we determine whether there exists interior equilibria (p^*, q_2^*) numerically.

We now consider whether the equilibrium point is stable to mutations in the boundary trait, q_1.

1.3.2 Stability to loss of sterile helping

Similarly to generalist and pure reproductive division of labour, the strategy is an equilibrium in q_1 as long as $W_{q_1} > 0$ at the putative equilibrium $(p^*, 1, q_2^*)$. Evaluating, we find that $W_{q_1}(p^*, 1, q_2^*) = -p^*((1-e) + e((1-\lambda) + \lambda(p^* + (1-p^*)q_2^{*\alpha}))) + (1-p^*)p^*(1-q_2^*)e\lambda\alpha R$, which needs to be solved numerically in order to check that $W_{q_1} > 0$.

1.3.3 Numerical results

We find that sterile helper and generalist is only ever stable to loss of partial division of labour if there is an efficiency benefit to specialisation ($\alpha > 1$). However, we also find that the equilibrium values (p^*, q_2^*) for which this occurs are never stable to loss of sterile helping at the considered parameter values. Therefore, we find no parameter combinations such that sterile helper and generalist is a stable strategy.

1.4 Sterile helper and pure reproductive division of labour

Sterile helper and pure reproductive division of labour strategies occur in the following domains: **D6** $= (0 < p < 1, q_1 = 1, q_2 = 0)$ and **D7** $= (0 < p < 1, q_1 = 0, q_2 = 1)$, where the two domains are symmetrically equivalent. Without loss of generality, we consider only the domain **D6** where phenotype 1 is the sterile helper ($q_1 = 1$) and phenotype 2 is the pure reproductive ($q_2 = 0$). In this scenario, there are now two boundary traits: the levels of cooperation of each phenotype ($q_1 = 1, q_2 = 0$). We therefore treat each trait separately in our analysis. This leads to three sources of instability: loss of cooperation, loss of pure reproduction and loss of sterile helping. These can all be evaluated analytically.

1.4.1 Stability to loss of cooperation

The trait p is the one interior trait of the putative equilibrium and so is treated using the classic, 1 dimensional approach. We seek to determine whether there is a value p^*, for which directional selection in p is zero ($W_p(p^*, 1, 0) = 0$). We note that $W_p(1, 1, 0) < 0$ always and thus if $W_p(0, 1, 0) > 0$, then we may predict that a $0 < p^* < 1$ exists and thus that the putative equilibrium is stable to loss of cooperation ($p^* = 0$) mutations. Evaluated we have that, $W_p(0, 1, 0) = -(1-e) + e\lambda R$. Solving the inequality gives us:

$$e > 1/(1 + \lambda R).$$

Therefore, a sterile helper and pure reproductive division of labour strategy is always stable to loss of cooperation if the trait is essential ($e = 1$). Otherwise, the strategy is more likely to be stable to loss of cooperation if λ, e and R are high. In this case, solving for $W_p(p^*, 1, 0) = 0$, gives: $p^* = (e\lambda R - (1-e))/(e\lambda(R+1)))$. We may then consider whether the strategy is a stable equilibrium in the other two traits.

1.4.2 Stability to loss of sterile helping

Similarly to loss of sterile helping in sterile helper and generalist, a sterile helper and pure reproductive division of labour will be an equilibrium with respect to q_1 if $W_{q_1}(p^*, 1, 0) > 0$. Evaluating gives the condition:

$$W_{q_1}(p^*, 1, 0) = -p^*((1-e) + e((1-\lambda) + \lambda p^*)) + (1-p^*)e\lambda\alpha R p^* > 0.$$

Substituting in the equilibrium value p^* from above then gives the following condition for stability to loss of sterile helping:

$$\alpha > (R + e(1-\lambda))/(R((1-e) + e\lambda)).$$

The right hand side of this threshold condition in α is strictly greater than 1 so long as $\lambda \neq 1$. If $\lambda = 1$ then the condition holds that the returns need to be accelerating $\alpha > 1$. Otherwise, we have that stability to loss of sterile helping is more likely as α, R and λ increase and as e decreases.

1.4.3 Stability to loss of pure reproduction

Similarly to loss of pure reproduction in generalist and pure reproductive, a sterile helper and pure reproductive division of labour will be an equilibrium with respect to q_2 if $W_{q_2}(p^*, 1, 0) < 0$. Evaluating gives:

$$W_{q_2}(p^*, 1, 0) = \begin{cases} +\infty, & \text{if } \alpha < 1 \\ (1-p^*)(-(1-e) - e\lambda p^* + e((1-\lambda) + \lambda R(1-p^*))), & \text{if } \alpha = 1 \\ -(1-p^*)((1-e) + e\lambda p^*), & \text{if } \alpha > 1. \end{cases}$$

We see that sterile helper and pure reproductive division is stable to loss of pure reproduction if $\alpha > 1$. If $\alpha < 1$ then the strategy is unstable to loss of pure reproduction. By substituting p^* above, we find that the strategy is not stable to loss of pure reproduction when $\alpha = 1$. We thus find that, if the condition for stability to loss of sterile helping is verified, then the population will be trivially stable to loss of pure reproduction.

1.4.4 Numerical results

We have found analytically that loss of sterile helping is the only form of instability if the social trait is essential ($e = 1$). For non-essential traits ($e \neq 1$), stability to loss of cooperation needs to be satisfied first and then stability to loss of sterile helping. Stability to loss of pure reproduction is never a deciding factor as it is only ever a source of instability if loss of sterile helping is already a factor. For non-essential traits ($e \neq 1$), we find numerically that stability to loss of cooperation is the determining source of instability at the border of parameter space between the sterile helper and pure reproductive and uniform non-cooperation strategies. In contrast, stability to loss of sterile helping is always the determining source of instability at the border of parameter space between sterile helper and pure reproductive and generalist and pure reproductive (there are no other borders in this case).

1.5 Uniform non-cooperation

The strategy of uniform non-cooperation occurs within the domains: **D8** $= (0 \leq p \leq 1 q_1 = 0, q_2 = 0)$, **D9** $= (p = 0, 0 \leq q_1 \leq 1, q_2 = 0)$ and **D10** $= (p = 1, q_1 = 0, 0 \leq q_2 \leq 1)$, where **D9** and **D10** are symmetrically equivalent. We note that p is a neutral trait in **D8**, q_2 is a neutral trait in **D9** and q_1 is a neutral trait in **D10**. All other traits are boundary traits in the three domains. Since the domains are connected, a population may evolve neutrally along all three domains and thus stability to any mutations in the boundary traits needs to be satisfied for all possible strategies in the super-domain. Without loss of generality we consider only the domains **D8** and **D9**. There are then three sources of instability to uniform non-cooperation: instability to uniform cooperation, instability to rare specialism and instability to weak specialism.

1.6 Stability to invasion by weak specialism

Uniform non-cooperation is stable to weak specialism if the boundary trait q_1 is stable in the domain **D8**. This translates to the condition: $W_{q_1}(p, 0, 0) < 0$ for all values of $p \neq 0$ (at $p = 0$, the population may evolve neutrally onto the **D9** domain). We have that directional selection in q_1 depends on the value of α:

$$W_{q_1}(p,0,0) = \begin{cases} +\infty, & \text{if } \alpha < 1 \\ p(-(1-e) + e(1-\lambda) + e\lambda R), & \text{if } \alpha = 1 \\ -p(1-e), & \text{if } \alpha > 1. \end{cases}$$

Thus, uniform non-cooperation in this case is trivially stable if $\alpha > 1$ and $e \neq 1$ and is unstable if $\alpha < 1$. If $\alpha = 1$, we have that uniform non-cooperation is stable if $e < 1/(2+\lambda(R-1))$.

1.6.1 Stability to invasion by uniform cooperation

Uniform non-cooperation is stable to uniform cooperation if the boundary trait q_2, is at equilibrium, uninvadable and convergent stable in the domain **D9**. This translated to the condition $W_{q_2}(0, q_1, 0) < 0$ for all values of q_1. Evaluated, we find a dependence on α:

$$W_{q_2}(0,q_1,0) = \begin{cases} +\infty, & \text{if } \alpha < 1 \\ -(1-e) + e(1-\lambda) + e\lambda R, & \text{if } \alpha = 1 \\ -(1-e), & \text{if } \alpha > 1. \end{cases}$$

Stability along this axis is thus satisfied whenever $\alpha > 1$ and $e \neq 1$. If $\alpha = 1$, then stability holds so long as $e < 1/(2 + \lambda(R-1))$. We thus have, formally, that uniform non-cooperation is never stable if the trait is essential ($e = 1$).

1.6.2 Stability to invasion by to rare specialism

Uniform non-cooperation is stable to rare specialism if the boundary trait p is at equilibrium, uninvadable and convergent stable in the domain **D9**. This translates to the condition: $W_p(0, q_1, 0) < 0$ for all values of $q_1 \neq 0$ (at $q_1 = 0$ the population may evolve neutrally onto the **D10** axis). Evaluating directional selection in p in this domain gives the following inequality for stability to rare specialism:

$$W_p(0, q_1, 0) = -(1-e) + (1-q_1)((1-e) + e(1-\lambda)q_1^{\alpha}) + e\lambda R q_1^{\alpha} < 0.$$

Directional selection in p equals $-(1-e) + e\lambda R$ if $q_1 = 1$ and thus stability holds at this point if $e < 1/(1+\lambda R)$. Trivially, uniform non-cooperation is always unstable at this point in the simplified model ($e = 1, \lambda = 1$). In the full model, we must then determine if uniform non-cooperation is stable at all internal values of q_1. To do this we solve for all q_1 such that $\partial W_p/\partial q_1(0, q_1, 0) = 0$ and check that $W_p(0, q_1, 0) > 0$ at all these points. This is done numerically.

1.6.3 Numerical results

We find that uniform non-cooperation is unstable to all sources of instability if the cooperative trait is essential ($e = 1$). For non-essential traits ($0 < e < 1$), we find that stability to invasion by rare specialism is always the determining source of instability at the parameter space border with all other strategies.

1.7 Uniform cooperation

Uniform cooperation occurs in the following connected domains: **D11** $= (p = 1, 0 < q_1 < 1, 0 \leq q_2 \leq 1)$, **D12** $= (p = 0, 0 \leq q_1 \leq 1, 0 < q_2 < 1)$ and **D13** $= (0 \leq p \leq 1, 0 < q_1 < 1, q_2 = q_1)$, where **D11** and **D12** are symmetrically equivalent domains. We have that q_2 is a neutral trait in **D11**, q_1 is a neutral trait in **D12** and p is a neutral trait in **D13**. As such, the population may evolve neutrally along all three domains at a fixed cooperative investment, q^*, and stability thus needs to be shown for all points along the super-domain. There are then three sources of instability that need be considered, loss of cooperation (a stable cooperative investment, q^*, does not exists), instability to rare specialism and instability to weak specialism.

1.7.1 Stability to loss of cooperation

Uniform cooperation is stable to loss of cooperation if there exists a non-zero cooperative investment, q^*, for a population of monomorphic generalist helpers that is both an equilibrium, uninvadable and convergent stable. This may be calculated as the value q^* for which $W_{q_1}(1, q_1 = q^*, 0 \leq q_2 \leq 1) = -((1-e)+eq^{*\alpha})+(1-q^*)e(1-\lambda)\alpha q^{*\alpha-1}+(1-q^*)e\lambda R\alpha q^{*\alpha-1} = 0$. If this q^* is non zero (and less than one) then the strategy is stable to loss of cooperation. This needs to be calculated numerically in the full model. If we make the simplifying assumptions that $e = 1$ and $\lambda = 1$ then we find the following analytical result: $q^* = \alpha R/(1 + \alpha R)$, which is strictly less than 1, greater than 0 as long as $a < 1$.

1.7.2 Stability to invasion by rare specialism

Uniform cooperation is stable to rare specialism if the boundary trait, p, is an equilibrium in the domain **D11** ($p = 1$), for all possible values of $q_2 \neq q_1$ and where $q_1 = q^*$ is the equilibrium cooperative investment calculated numerically in the full model. At the point $q_2 = q_1$, the population is allowed to evolve neutrally onto the **D13** domain. This translates to the condition $W_p(1, q^*, q_2) > 0$ for all $q_2 \neq q^*$. Evaluating we find: $W_p(1, q^*, q_2) = (1 - q^*)((1-e) + e{q^*}^{\alpha}) - (1-q_2)((1-e) + e((1-\lambda)q_2^{\alpha} + \lambda {q^*}^{\alpha})) + (1-q^*)e\lambda R({q^*}^{\alpha} - q_2^{\alpha})$ We check numerically that $W_p(1, q^*, q_2) > 0$ at the boundary values $q_2 = 0$ and $q_2 = 1$ as well as at any internal points, $0 < q_2 < 1$ that may minimise $W_p(1, q^*, q_2)$ (solve for q_2 values such that $\partial W_p/\partial q_2(1, q^*, q_2) = 0$). For example, the following is the stability condition at the boundary $q_2 = 0$:

$$W_p(1, q^*, 0) = (1 - q^*)((1-e) + e{q^*}^{\alpha}) - ((1-e) + e\lambda {q^*}^{\alpha}) + (1-q^*)e\lambda R {q^*}^{\alpha} > 0$$

In the simplified model ($e = 1, \lambda = 1$), we find that the unique value of q_2 that solves $\partial W_p/\partial q_2(1, q^*, q_2) = 0$ is $q_2 = q^*$ and that $W_p(1, q^*, q^*) = 0$. This is simply the neutral instability that is allowed by the construction the model. In order to determine whether the population is unstable at other values of q_2 (including the boundaries) we need simply determine whether $q_2 = q^*$ represents a minimum or maximum of directional selection in p. If it is a minimum, then all other values of q_2 satisfy $W_p(1, q, q2) > 0$ and thus the strategy is stable and, if it is a maximum, then all other values of q_2 lead to negative directional selection in p and thus the population is unstable to weak specialism. We determine this by evaluating the sign of $\partial^2 W_p/\partial q_2^2(1, q^*, q^*)$ and find that uniform cooperation is stable in the case that $\alpha < 1$ and unstable if $\alpha \geq 1$ (for the simplified model).

1.7.3 Stability to invasion by weak specialism

Stability to weak specialism occurs if the interior traits $q_1 = q^*$ and $q_2 = q^*$ are jointly an equilibrium in the domain **D12** for all possible values of $0 < p < 1$. If we can show that $W_{q_1}(p, q^*, q^*) = 0$ for all values of p then both q_1 and q_2 are an equilibrium (by symmetry of the levels of cooperation of each phenotype). If we assume that the trait is essential ($e = 1$) and others-only ($\lambda = 1$), then we find that $W_{q_1}(p, q^*, q^*) = 0$ for all p.

1.7.4 Numerical results

At the parameter space border with other possible strategies, we find that the determining source of instability for whether uniform cooperation is favoured is always instability to invasion by rare specialism. In particular, uniform cooperation is stable so long as it is uninvadable by a rare mutant strain with a division of labour between a helper reproductive ($q_1 = q^*$) and a pure reproductive ($q_2 = 0$). We found previously that uniform cooperation is the only non-division of labour strategy that may be stable if the cooperative trait is essential ($e = 1$). Thus, if uniform cooperation is unstable in this regime ($e = 1$), then a division of labour strategy is the equilibrium strategy (by exclusion.) If the trait is non-essential ($e < 1$) and uniform cooperation is not stable, then either a division of labour strategy or uniform non-cooperation are the equilibrium.

2 Supplementary information: uninvadability of the equilibria

In order for an equilibrium strategy to be an ESS, we must show that a population otherwise at equilibrium is uninvadable by a rare mutant lineage with an alternate strategy. It is sometimes claimed that the uninvadability of an equilibrium point cannot be mathematically demonstrated in kin selection models without specifying the genetic details of the system or without recourse to a dynamical model or computer simulation[7,8]. Here we used an ESS condition for haystack models derived elsewhere (Cooper and West, in preparation) that does match the ESS condition derived using the simplest genetic model (asexual, haploid) for patch-structured populations as found in Taylor and Day[7].

We begin by evaluating the analytical uninvadability conditions for each trait in our model and show that all of our equilibria satisfy these conditions for each trait that may vary. However, we also verify computationally that our division of labour equilibria are uninvadable. First, we check numerically for each of the division of labour equilibria that mutant lineages with respect to each trait would indeed have lower fitness than that of an arbitrary individual in the equilibrium population. Second, we employ individual based simulations to show that a sampling of our division of labour equilibria are indeed stable points of the dynamical system. We note that this entire analysis hinges on the assumptions that (a) relatedness is an independent parameter of the model and (b) competition for niches is global.

2.1 Analytical conditions

The condition for uninvadability is that is is a local maximum of the fitness landscape[2,3]. Let x be the genic value of a rare mutant. In a population with equilibrium genic value x^*, we then have that the uninvadability condition amounts to evaluating the second order of fitness with respect to the mutant genic value[2,7–9]:

$$\frac{d^2W}{dx^2}\Big|_{x=x^*} < 0.$$

If this second derivative is negative, then we have that the fitness landscape is concave around the equilibrium and therefore that rare mutants have lower fitness than individuals at equilibrium.

We derive uninvadability by a simple extention of the Taylor-Frank approach[5,7]. Assume that the population is infinite and structured into finite patches of uniform size. Let the genic value x determine the phenotype of the mutant, y, and influence the average phenotype of the focal mutant's patch, Y, (via the phenotype of all patch-mates that are IBD to the focal mutant.) Let y^* then be the equilibrium phenotype that corresponds to x^*. In this case, the fitness of a rare mutant in a patch may be expressed as $W(x) = W(y, Y)$ and we may evaluate the second-order condition above by applying the chain rule. This gives the following uninvadability condition for a population at the equilibrium y^*:

$$\Big(\frac{\partial^2 W}{\partial y^2} + 2\frac{\partial^2 W}{\partial y \partial Y}R + \frac{\partial^2 W}{\partial Y^2}R^2\Big)\Big|_{y=Y=y^*} < 0,$$

where $R = \frac{dY}{dx}/\frac{dy}{dx}$ and higher order derivatives of genic value on phenotype are assumed to be negligible (Cooper and West, in preparation)[2,3,5,7–9]. This assumption is permissible because we are only considering global competition and do not consider the effect of trait mutations on relatedness in the population (R is fixed and independent).

Below, we evaluate this ESS condition for each of the traits of our model:

$$\text{Uninvadability in } p: \qquad 2R(q_2 - q_1)e\lambda(q_1^{\alpha} - q_2^{\alpha}) < 0$$

$$\begin{aligned}\text{Uninvadability in } q_1: \qquad & (1-\lambda)((1-q_1)(\alpha-1) - 2q_1) - 2Rp\lambda q_1 \\ & + R^2\lambda(\alpha-1)(p(1-q_1) + (1-p)(1-q_2)) < 0\end{aligned}$$

$$\begin{aligned}\text{Uninvadability in } q_2: \qquad & (1-\lambda)((1-q_2)(\alpha-1) - 2q_2) - 2R(1-p)\lambda q_2 \\ & + R^2\lambda(\alpha-1)(p(1-q_1) + (1-p)(1-q_2)) < 0\end{aligned}$$

We affirm that these conditions should only be applied to interior traits of the equilibrium. If a trait of an equilibrium lies at the boundary, then the equilibrium is negligibly uninvadable with respect to that trait due simply to the equilibrium condition that directional selection in that trait is directed into the boundary.

If a trait of an equilibrium is neutral, such that the population may evolve neutrally along that trait axis without altering the underlying strategy, the uninvadability with respect to that trait need not be shown.

Across all of the equilibria that we identify in our model, we found using the above conditions that each equilibria is also uninvadable and thus is an ESS of the system.

2.2 Numerical verification

In our analysis, we discretised parameter space into $205,209$ nodes. Amongst these, we identified $54,611$ division of labour equilibria ($35,276$ sterile helper and pure reproductive equilibria and $19,335$ generalist and pure reproductive equilibria.) We found that each of these division of labour equilibria satisfied our analytical conditions for uninvadability. However, we can also numerically verify that these equilibria are maxima of the fitness landscape (and not minima.)

We do this by considering each equilibrium in turn, (p^*, q_1^*, q_2^*) and its associated parameter values (α, λ, R, e), and evaluating the relative fitness of a series of mutant lineages that each deviate a given amount from only one of the traits at a time. In our analysis, we considered 20 mutant lineages for each trait that varied ($\pm\epsilon$, $\pm 2\epsilon$, $\ldots$, $\pm 10\epsilon$) from the equilibrium value of the trait where $\epsilon = 0.01$. However, we did not consider mutant lineages if the perturbed trait value falls outside of the boundaries of our model.

For a given equilibrium, if the relative fitness of all considered mutant lineages was less than the fitness of a neutral mutant (with no change in trait values) then we may conclude that the equilibrium strategy is a maximum of the local fitness landscape and is therefore uninvadable. In Supplementary figure 1, we illustrate this analysis for two such equilibria, a generalist and pure reproductive equilibrium and and sterile helper and pure reproductive equilibrium. We see that in both cases, all mutant lineages have lower fitness than the equilibrium strategy, which we thus deem uninvadable in both cases.

We repeated this analysis for all $54,611$ division of labour equilibria and found in each case that the equlibrium (neutral) strategy always had a higher fitness than any of the mutant lineages and thus affirm that they are all ESSs of the fitness landscape.

2.3 Dynamic individual-based simulations

We now show that a sampling of the division of labour equilibria are stable using dynamic individual-based simulations. For each equilibrium considered, we check whether a monomorphic population at that equilibrium is uninvadable by a rare mutant deviating from the equilibrium in a single trait value by an amount $\epsilon = 0.05$. We do this for each possible trait value that may vary, repeating the invasion simulation a total of $L = 10,000$ times for each possible mutant lineage. We then record the fraction of simulations for which the particular mutation either went extinct or evolved to fixation (or neither outcome) by the end of $T = 10,000$ generations.

We employ a haystack model for our simulations wherein we presume that relatedness is fixed. This simplifies the computation and means that we do not need to explicitly model the size that groups grow to after founding. As such, we need only model the founding individuals of each group.

Consider a population with $M = 10,000$ groups of $F = 1/R$ founding individuals. Let all founding individuals have the equilibrium strategy (p^*, q_1^*, q_2^*), save one founding individual in an arbitrary group with mutant strategy (p', q_1', q_2').

For T timesteps, repeat the following:

1. Calculate the average p, q_1 and q_2 in each group.
2. Calculate the expected fitness, W, of each individual in the population.
3. Draw $F \times M$ random individuals (with replacement) from the population with probabilities equal to their fitness, W.
4. Assort these individuals randomly into M groups of F individuals. These are then the founding individuals of the next generation.

At each step of the above iteration, we keep track of the number of mutant individuals in the population. If at some point, the number of mutant individuals falls to 0, then it will stay at 0 for the rest of the simulation and we say that the mutant lineage has gone extinct. On the other hand, if the number of mutants rises to $M \times F$ individuals, then it will stay at this number and we say that the mutation has evolved to fixation.

In Supplementary figure 2, we use simulation to show the uninvadability of a sterile helper and pure reproductive division of labour equilibrium to a mutant lineage with a perturbed helper probability. We see that over $L = 10,000$ simulated invasions, all mutant lineages have gone extinct by the end of $10,000$ generations.

We performed the above analysis for each of our 60 division of labour equilibria (42 sterile helper and pure reproductive equilibria and 18 generalist and pure reproductive equilibria), considering fixed mutations along each of the permissible trait axes. The results of this analysis are shown in Supplementary figure 3. We found that all $10,000 \times 4 \times 42$ simulations for the sterile helper and pure reproductive equilibria had mutant lineages that all went extinct. We also found that no single mutant invaded to fixation in the generalist and pure reproductive equilibria ($10,000 \times 5 \times 18$ simulations). However, 6 of the generalist and pure reproductive equilibria did have $1 - 2$ mutant lineages (out of $10,000$) that were not extinct by the end of the simulation. Strictly speaking, this does not meant that these equlibria are invadable by mutations but rather that they are weakly uninvadable (shallow maxima of the fitness landscape.) Indeed, we found that on average even neutral mutations would have about 2 mutant lineages invade to fixation out of $10,000$ simulations simply due to stochastic drift. In Supplementary figure 3, we see that all of these weakly uninvadable division of labour equilibria are near the parameter space boundary with non-division of labour strategies which is to be expected.

3 Supplementary information: convergence stability

An equilibrium strategy is *convergent stable* if it is an attractor of the dynamical system: a monomorphic population close to the equilibrium strategy always experiences directional selection that favours a rare mutant lineage with trait value closer to the equilibrium[1]. For each parameter combination considered in our numerical discretization, we also checked that the associated equilibrium is convergent stable. We use the methodology developed by Brown and Taylor (2010). That is, the first derivative of directional selection in each interior trait (non-boundary traits) that is non neutral must be negative and the jacobian of directional selection of all possible combinations of interior traits must be negative definite. The convergence stability of neutral traits need not be considered as the population is free to evolve along these axes without altering the underlying strategy. Boundary traits of an equilibrium, meanwhile are negligibly convergent stable due to the fact that directional selection in that trait is directed into the boundary (by the equilibrium condition.)

The strategy comprising two different generalist types is the only strategy for which all traits are interior traits and non-neutral. However, no equilibrium with respect to this strategy was identified in our analysis and so the conditions for convergence stability in this case are never applied. All other strategies contain at least one boundary or one neutral trait. For example, the strategy of generalist and pure reproductive consists of one boundary trait ($q_2 = 0$ without loss of generality) and two non-neutral, interior traits (p and q_1). As such, the first derivative of directional selection in p and q_1 (with respect to that trait) must be shown to be negative. In addition, the jacobian of directional selection in both traits must be shown to be negative definite, which amounts to showing that the determinant of the jacobian is negative[6]. The strategy of uniform non-cooperation consists of two boundary traits and one neutral trait and thus is negligibly convergent stable if the equilibrium condition is satisfied. The strategy of uniform cooperation is treated differently depending on which of the domains is considered. In the domain, **D11**, there is one boundary trait ($p = 1$), one neutral trait (q_2) and one interior trait that must be analysed singly (q_1). However, in the domain **D13**, p is a neutral trait whereas q_1 and q_2 are interior traits. As such the the convergence stability of both interior traits must be shown for all values of p. This is done numerically for a fine discretization of p for each uniform cooperation equilibrium.

Across all discretized parameter combinations of analysis, all of the associated equilibria were found to be convergent stable.

4 Supplementary information: non-linear costs to cooperation

Our model presumes that the costs of cooperative investment are linear; we consider only non-linear effects upon the beneficial returns from cooperation (α). Here we present a generalized fitness function that includes a non-linear cost to cooperation. We show that a simple change of variables can yield the same qualitative fitness function as was considered previously. Let the generalized fitness function be:

$$\begin{aligned} W &= p(1-q_1^{\beta})\Big[(1-e)+e\big((1-\lambda)q_1^{\alpha}+\lambda(PQ_1^{\alpha}+(1-P)Q_2^{\alpha})\big)\Big] \\ &\quad + (1-p)(1-q_2^{\beta})\Big[(1-e)+e\big((1-\lambda)q_2^{\alpha}+\lambda(PQ_1^{\alpha}+(1-P)Q_2^{\alpha})\big)\Big], \end{aligned}$$

where the parameter $\beta > 0$ is the shape of the costs of cooperation. We now propose the change of variables: $k_1 = q_1^{\beta}$ and $k_2 = q_2^{\beta}$. Thus we are now considering the cost off cooperation as an explicit variable in our model rather than the underlying cooperative investment. The variables Q_1 and Q_2 are the others-only averages of cooperative investment for each phenotype in the focal social group. If we consider that selection is weak, then the variance in cooperative investment across all individuals of the same phenotype in the group will be nearly negligible. As such, Q_1 and Q_2 may be calculated using the geometric mean rather than the arithmetic mean. Let m_1 and m_2 be the others-only number of individuals in the social group of phenotypes 1 and 2 respectively. Assume that individuals of each phenotype may be indexed by i. For example, q_{1i} (k_{1i}) is the cooperative investment (cost of cooperation) of the ith individual of phenotype 1. This gives the following for phenotype 1:

$$Q_1 = \Big(\prod_{i=1}^{m_1} q_{1i}\Big)^{1/m_1} = \Big(\prod_{i=1}^{m_1} k_{1i}\Big)^{1/m_1 \times 1/\beta} = K_1^{1/\beta},$$

where K_1 is the others-only average cooperative cost payed by individuals of phenotype 1 in the social group. The same logic can be used to show that $Q_2 = K_1^{1/\beta}$. If we then substitute these back into the fitness equation we arrive at:

$$\begin{aligned} W &= p(1-k_1)\Big[(1-e)+e\big((1-\lambda)q_1^{\alpha/\beta}+\lambda(PK_1^{\alpha/\beta}+(1-P)K_2^{\alpha/\beta})\big)\Big] \\ &\quad + (1-p)(1-k_2)\Big[(1-e)+e\big((1-\lambda)q_2^{\alpha/\beta}+\lambda(PK_1^{\alpha/\beta}+(1-P)K_2^{\alpha/\beta})\big)\Big], \end{aligned}$$

which has the same functional form as the original fitness function and thus yields the same qualitative results. In this generalized fitness function, α, is replaced by α/β which is the ratio of the shape of the returns from cooperation to that of the costs. As such, the condition $\alpha > 1$ no longer signifies that the returns must be accelerating but translates to: the returns from cooperation must accelerate more (or diminish less) than the costs of cooperation.

5 Supplementary information: the evolution of extreme specialisation

We hypothesise that there may be an evolutionary feedback loop in which helper specialisation drives reproductives to help less and reproductive specialisation drives helpers to help more. In particular, if we consider the effect of a change in the cooperative investment of one phenotype upon directional selection in the cooperative investment of the other phenotype we find that:

$$\frac{\partial W_{q_1}}{\partial q_2} = -p\lambda(1-p)\alpha q_2^{\alpha-1} - (1-p)e\lambda p\alpha q_1^{\alpha-1}R,$$

which is always negative. Similarly, we can show that $\frac{\partial W_{q_2}}{\partial q_1} < 0$. As such, a monomorphic decrease in the cooperative investment of the less cooperative phenotype leads to selective pressure for an increase in the cooperative investment of the more cooperative phenotype and vice versa. However, this is not dynamically sufficient analysis and so is not conclusive.

In order to test this hypothesis, we performed a series of dynamic, individual based simulations. Our simulations were similar in form to those performed in the uninvadability analysis except that random mutations in the trait value of some of the traits (evolving traits) may accrue. The non-evolving traits are fixed at their starting values and dot not accrue mutations. Consider a population with M groups of $F = 1/R$ founding individuals. Let all individuals in the initial population have the starting strategy (p, q_1, q_2). Broadly, a single simulation is composed of the following iteration.

For T timesteps, repeat the following:

1. Calculate the average p, q_1 and q_2 in each group.
2. Calculate the expected fitness, W, of each individual in the population.
3. Draw $F \times M$ random individuals (with replacement) from the population with probabilities equal to their fitness, W.
4. With probability μ each individual may experience a mutation in one of the evolving traits (p, q_1 or q_2) where the mutation is randomly drawn from a normal distribution with variance σ.
5. Assort these individuals randomly into M groups of F individuals. These are then the founding individuals of the next generation.

With this setup we performed the following analyses. In Figure 4a, we fixed reproductive cooperation (q_2) at high and low values and allowed the level of cooperation of the helper (q_1) to evolve from a starting value of the ancestral expected value of cooperation $(R\lambda + (1 - \lambda))/(1 + \alpha(R\lambda + (1 - \lambda))$. In Figure 4b, we fixed helper cooperation (q_1) at high and low values and allowed the level of cooperation of the reproductive (q_2) to evolve from a starting value of the ancestral expected value of cooperation. In Figure 4d, we fixed reproductive cooperation to $q_2 = 0.25$ and allowed only helper cooperation (q_1) to evolve for $10,000$ generations after which the levels of both helper and reproductive cooperation (q_1 and q_2) were allowed to evolve for another $10,000$ generations. In Figures 4a, 4b & 4d, we repeated each simulation a total of 25 times, averaging results across trials and present 95 percent confidence intervals based on normal distribution. In Figure 4c, we fixed the level of cooperation of one phenotype over a series of values, allowed the level of cooperation of the other phenotype to evolve and recored the population average level of cooperation of the evolving phenotype at the end of 3000 generations. We repeated this simulation for each fixed level of cooperation at total of 10 times, averaged results across trials and presented 95 percent confidence intervals. In Figures 4a-d, we set the number of groups to $M = 5000$, we set the mutation rate to $\mu = 0.01$ per evolving trait per generation, we set the variance of mutation size to $\sigma = 0.01$ and we fixed the probability of being one phenotype over the other to $p = 0.5$. The same analysis we performed with an evolving phenotype probability (p) and the same qualitative results were found (both phenotypes always evolve to full specialisation but the speed at which they do so depends on the level of cooperation of the other phenotype; the phenotype probability evolves to a value that reflects the fixed level of cooperation). In all figures, we set $\alpha = 2$, $e = 1$ and $\lambda = 1$. For Figures 4a and 4b we set $R = 1/3$ and for Figures 4c and 4d we set $R = 1/2$. The parameter values of each analysis were chosen such that sterile helper-pure reproductive division of labour was the favoured strategy and for ease of interpretation.

6 Supplementary information: a closed model of division of labour

In our analysis, we have employed an open model approach wherein the demographic processes that generate relatedness have not been specified. Instead, we leave R as an independent parameter of the model and show how the evolution of division of labour depends on the value of R. The benefit of this approach is that it may lead to general predictions that hold regardless of the specific way that relatedness may arise within populations. However, an open model may fail to make accurate predictions for specific systems in which there are significant interactions between R and other factors modelled. This shortcoming is overcome by a closed model approach in which demographic detail is specified and and relatedness is solved for as a function of these processes. This allows for a more detailed model that can capture interactions between relatedness and other feature of the model. However, the downside of a closed model approach is that any predictions made will then only be applicable to biological systems that match the demographic assumptions of the closed model. As such, the trade-off between using an open or closed model approach is that of demographic precision against broad applicability.

The purpose of our work was to explain the evolution and diversity of division of labour across the tree of life, regardless of the specific mechanisms by which relatedness arises. As such, an open model was the more appropriate approach for our purposes. To illustrate the conclusions that may be drawn from a closed model, we now analyse a simplification of our division of labour model using a closed approach.

6.1 Extending the open model

For the purposes of analytical tractability, we first assume that al helpers are sterile (q_1=1) and that all reproductives are pure reproductives (q_2=0). As such we seek candidate ESS values of p, the probability of becoming a sterile helper. For a particular combination of model parameters, if we find that $p^* > 0$ then we conclude that division of labour between a sterile helper and pure reproductive is the ESS strategy and otherwise we conclude that uniform non-cooperation is the ESS. In the open model analysis, we found that sterile helper and pure reproductive division of labour was stable to invasion by uniform non-cooperation if a threshold condition in trait essentiality was met ($e > 1/(1 + \lambda R)$; Supplementary figure 4 depicts this threshold condition for $\lambda = 1$.)

Now suppose that we wanted to model the demographic processes that generate relatedness between interaction individuals. After each generation, if offspring individuals have a probability s of staying (and competing for niches) on their natal patch (social group) and a probability $1 - s$ of dispersing to a different patch in the population (and competing for niches there), then this limited dispersal will lead to a buildup of of genetic correlations over time amongst the individuals that stay on their natal patches. An additional way to generate relatedness between individuals is to have overlapping generations. If individuals have a probability k of surviving (and retaining their niches) from one generation to the next or of reproducing and dying (with probability $1-k$), then this means that IBD individuals from different generations may interact which represents an increase in social group relatedness. In combining these two processes (limited dispersal and non-overlapping generations), we have from Taylor and Irwin that the long-term whole-group relatedness is given by:

$$R_{wg} = \frac{1+k}{N + kn + 2ks - 2kns + s^2 - ks^2 - ns^2 + kns^2},$$

where N is the number of niches on a patch[10]. This may be used to calculate the others-only relatedness by the transformation: $R = (R_{wg} - 1)/(N-1)$. Broadly this formulation means that relatedness increases the more that individuals stay on their natal patches (high s), the more that generations overlap (high k) and for smaller social groups (low N). These effects are depicted in Supplementary figure 5.

6.2 Accounting for the demographic processes

We may now begin to reformulate our fitness equation for the closed model. We first have that the payoff to an individual with trait value p on a patch with (others-only) average trait value P is given by

$$G(p,P) = (1-p)\Big((1-e) + e\lambda P\Big)$$

We then have that the fitness of an individual will depend upon whether it survives from one generation to the next and on the competitive environment that offspring face (depending on whether they stay or disperse from the natal patch.) If we assume that p^* is the population wide average trait value and that ρ is the (group-wide) average trait value in the focal patch ($\rho = (p + (N-1)P)/N$) then the fitness of a focal individual is:

$$W(p,P) = k + (1-k)s\frac{G(p,P)}{sG(\rho,\rho) + (1-s)G(p^*,p^*)} + (1-k)(1-s)\frac{G(p,P)}{G(p^*,p^*)}$$

where the first term is the probability that the focal individual survives. Otherwise, the second term is the fitness due to offspring that compete on the natal patch and the third term is the fitness due to offspring that disperse and compete for niches on other patches in the population. We solve for a candidate ESS using the Taylor-Frank methodology[5]:

$$\begin{aligned}\Big(\frac{\partial W}{\partial p} + R\frac{\partial W}{\partial P}\Big)\Big|_{p=P=p^*} = \quad & \frac{1-k}{G(p^*,p^*)}\Big((1 - s^2/N)G_1(p^*,p^*) - (s^2/N)G_2(p^*,p^*))\Big) \\ + \quad & R\frac{1-k}{G(p^*,p^*)}\Big((1 - ((N-1)/N)s^2)G_2(p^*,p^*) \\ - \quad & ((N-1)/N)s^2 G_1(p^*,p^*))\Big) \\ = 0 \quad & ,\end{aligned}$$

where $G_1(p,P) = -((1-e) + e\lambda P)$ and $G_2(p,P) = (1-p)e\lambda$ are the first derivatives of $G(p,P)$ with respect to the first and second variables (p and P), respectively. We note that the above equation is still an open model. While further demographic detail has been included in the fitness equation, R is still present as an independent parameter in the model. In order to fully 'close' the model, we must substitute in R as a function of s, k and N as calculated earlier before solving for the equilibrium value of p^*.

6.3 Closed model results

Setting $\lambda = 1$, we numerically discretised the parameters e, s, k and N and solved for each combination of parameter values whether the equilibrium valued of $p^* > 0$. The results of this analysis are plotted in Supplementary figure 6, where we see that a similar threshold condition in e is recovered as was found for the open model (Supplementary figure 4.) We find that division of labour between a sterile helper and pure reproductive is favoured over uniform non-cooperation if the demographic parameters s and k are high and if N is low. Incidentally, limited dispersal (high s), overlapping generations (high k) and small group sizes (low N) are exactly the demographic conditions that lead to high social group relatedness and thus we have qualitatively recovered the same interaction between trait essentiality and social group relatedness as found in the open model (supplementary figure 4.) Additionally, we find that when there are non-overlapping generations ($k = 0$), division of labour is never favoured regardless of the values of s or N (not depicted). This recovers the well-known result that limited dispersal alone cannot favour cooperation in this way as the indirect fitness benefits due to an increase in relatedness is exactly cancelled by the competitive costs of related offspring competing for niches on the natal patch[11]. The inclusion of overlapping-generations has previously been shown to remove this effect[10,12].

7 Supplementary information: further predictions of the model

7.1 The efficiency benefits of specialisation.

We found in our model that an efficiency benefit to specialisation ($\alpha > 1$) is necessary, but not sufficient for division of labour (Figure 3)[13,14]. This occurs because other factors may hinder the favourability of division of labour. For example, if there is limited sharing of the benefits of cooperation (lower λ), then it may not pay to divide labour, despite high benefits of specialisation ($\alpha > 1$).

We find that, in most scenarios, division of labour and sterile helping are favoured by an increase in the efficiency benefits of specialisation (higher α; Figure 3). However, if the interplay with other factors (lower e, λ and R) means that the benefits of cooperation are relatively minor, then an increase in the efficiency benefits of specialisation (higher α) may actually disfavour division of labour. This prediction, in the opposite direction to that usually found, arises because uniform non-cooperation may be favoured instead (Figure 3). Indeed, a greater efficiency benefit to specialisation (higher $\alpha > 1$) means that it can pay to defect as significant benefits to cooperation can only be obtained at a high private cost (higher q, Figure 2b).

Empirically, there have been no formal tests of the influence of the efficiency benefit to specialisation on whether division of labour is favoured and the form it takes. Our predictions suggest that not only is such a test required, but that it would be useful to look at how the efficiency benefits (α) interplay with social group relatedness (R), the essentiality of the trait (e) and the trait sociality (λ).

7.2 How are the benefits of cooperation shared?

In all cases, we find that an increase in trait sociality (higher λ) favours division of labour and helper sterility (Figure 3). At higher trait sociality (higher λ), less of the benefits of cooperation are directed to those who cooperate, making it more efficient to divide between helpers and reproductives and less costly for helpers to become sterile.

Although this predicted role of sociality has not been directly tested, it is consistent with the pattern across microbes. In microbes, sociality is likely to vary in a predictable way across different types of traits. In species that form fruiting bodies to aid dispersal, such as slime moulds, the 'lifting up' of other cells will benefit other cells, and not the 'stalk' cells that do the lifting, leading to $\lambda \approx 1$ (others-only trait)[15]. In contrast, when bacterial cells produce and release public goods, the benefits are likely to be either shared equally amongst the local group (whole-group trait), or may go preferentially to the cell that produced them[16–19]. This would lead to $\lambda \leq (N-1)/N$, where N is the number of cells in the social group. Consequently, we predict that division of labour is less likely with public goods. Consistent with this prediction, bacteria produce many public goods without division of labour, and the example of intermediate division between a generalist and pure reproductive in *B. subtilis* is over a public good[16,18].

7.3 How essential is cooperation?

Many previous models of division of labour have assumed that cooperation is essential for survival and reproduction ($e = 1$; Supplementary table 1). However, this will often not be the case. While individuals in groups that lack cooperators may do less well, they do not necessarily have a fitness of zero[15,17,20–25]. The extent to which a trait is essential will vary depending on the environment of a system. Cooperative traits relating to resource acquisition or defense may be more or less essential depending on resource availability or the threat of predation in the local environment. In our model, we find that the essentiality of the trait (e) has multiple influences, such that a higher essentiality can either favour or disfavour division of labour (Figure 3).

On the one hand, in systems where cooperation is largely favoured (higher R and λ) a decrease in trait essentiality (lower e) favours division of labour and helper sterility (Figure 3a-3c). Division of labour is favoured in this scenario because it is then less costly to have pure reproductives that do not invest in cooperation. Helper sterility is favoured by a decrease in trait essentiality (lower e) because this leads to a lower proportion of helpers (lower p^*) which triggers higher levels of helper cooperation in compensation (higher q_1^*).

On the other hand, as the trait becomes less essential (lower e), it also makes it easier for the strategy of uniform non-cooperation to outcompete cooperation (Figure 3c and 3d). Consequently, with less essential traits (lower e), division of labour may be more favourable than uniform cooperation but uniform non-cooperation may become more stable than either strategy, particularly for low relatedness (Figure 3d). Overall, we find numerically that the more dominant effect of a lower trait essentiality (lower e) is to disfavour division of labour and helper sterility (Figure 3d).

However, these opposing effects makes it difficult to make broad predictions for what should be observed empirically. This problem can be overcome by focusing on biological systems in which uniform non-cooperation is never observed. In this case, we can make the clear prediction that both division of labour and a sterile helper are more likely with less essential traits (lower

e). The effect of the essentiality of the cooperative trait on division of labour has not been tested empirically in either animals or microbes.

7.4 What is the optimal proportion of helpers?

Across biological systems that employ division of labour, there is notable variation in the ratio of helpers to reproductives (p^*) and our model can predict the factors that contribute to this variation (Figure 5). In some cases, the predicted trend is consistent. A higher social group relatedness (higher R) and higher trait essentiality (higher e) both lead to a division of labour with a higher proportion of helpers (higher p^*). This occurs because higher relatedness and trait essentiality both increase the indirect benefits from helping relatives in the social group.

Consistent with our predicted influence of relatedness, species in which groups of cells are formed clonally ($R = 1$) have approximately three times the proportion of sterile helpers as species that form non-clonal groups ($R < 1$)[26]. However, this difference is based on a small number of phylogenetically independent comparisons, and hence lacks statistical power. Further data is required from phylogenetically diverse groups.

In contrast, the direction of the predicted relationship with other factors can depend upon the form of division of labour that is favoured. Considering the influence of trait sociality (λ), we find that a higher sociality (higher λ) leads to more helpers (higher p^*) when the helpers are generalists ($0 < q_1^* < 1$), but fewer helpers (lower p^*) when the helpers are fully specialised and sterile ($q_1^* = 1$; Figure 5b). Analogously, considering the influence of the efficiency benefit from greater cooperation (α), we find that a higher efficiency benefit (higher α) leads to fewer helpers (lower p^*) when the helpers are generalists ($0 < q_1^* < 1$) but has no influence on the fraction of helpers when the helpers are sterile ($q_1^* = 1$; Figure 5a). These different predictions arise because, when there are generalists, the amount that they help (q_1^*) also changes (Figures 5c and 5d). So for example, with a high efficiency benefit (higher α), we predict few generalists (lower p^*) but who help a lot (high q_1; Figure 5c).

These predictions about the relative investment into helpers and reproductives have not been tested. In microbes, the proportion of the different phenotypes varies widely, both between and within species, allowing numerous opportunities for such tests[27–31]. In the social insects, there is a rich theoretical and empirical literature examining the ratio of workers belonging to different behavioural castes[32–35]. The optimal ratio of different castes of helpers is a different problem from our focus on the division between helpers and reproductives. Nonetheless, a general pattern from the social insect work is that the proportion of a workers in a caste decreases as the caste becomes more specialised and thus fewer workers are needed to perform the associated tasks[32,35]. This is similar to our result that an increase in the benefits of specialisation or the sociality of the trait (higher α or λ) lead to both an increase in specialisation (higher q_1^*) and a decrease in representation in the group (lower p^*). In contrast, we found that when an increasing relatedness (R) drives increasing helper investment (higher q_1^*), then the proportion of helpers actually increases as well (higher p^*). This difference arises because increased efficiency benefits and trait sociality (higher α and λ) favours more specialised division of labour whereas increased relatedness (higher R) favours more cooperation overall.

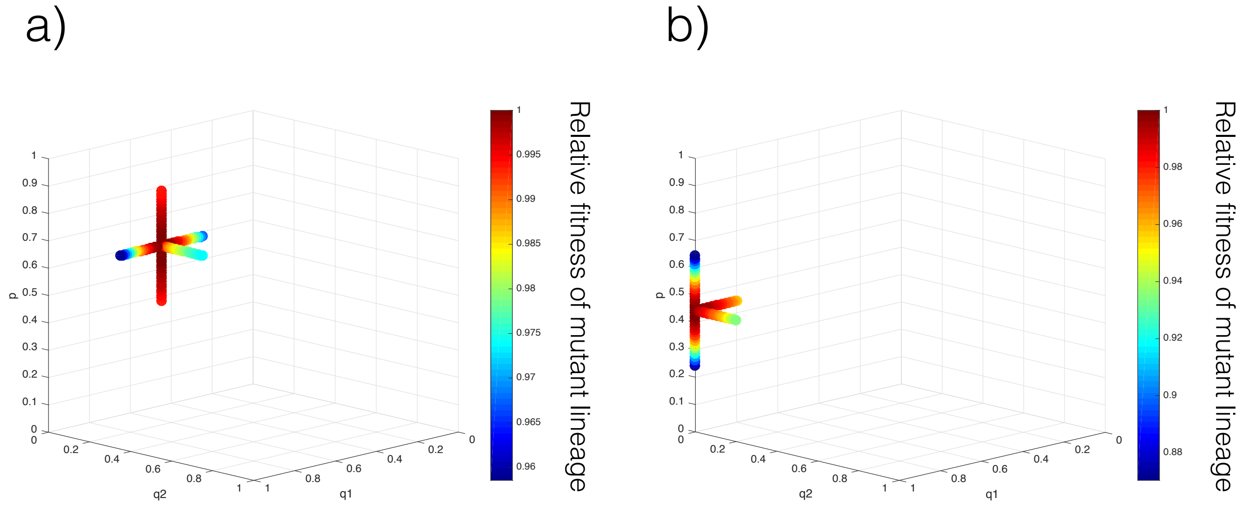
In the following we list the parameter values used to generate Figure 5. The results for these specific parameter values are representative of the broader pattern observed (not presented). The effect of the shape of the return (α; Figures 5a and 5c): $R = 1/3$, $e = 0.9$, $\lambda = 0.505$ (helper reproductives) and $\lambda = 1$ (sterile helpers). The effect of the sociality of the trait (λ;

Figures 5b and 5d): $R = 1/3$, $e = 0.9$, $\alpha = 2.2255$ (helper reproductives) and $\alpha = 7.3891$ (sterile helpers).

7.5 Conflicting results with a previous model

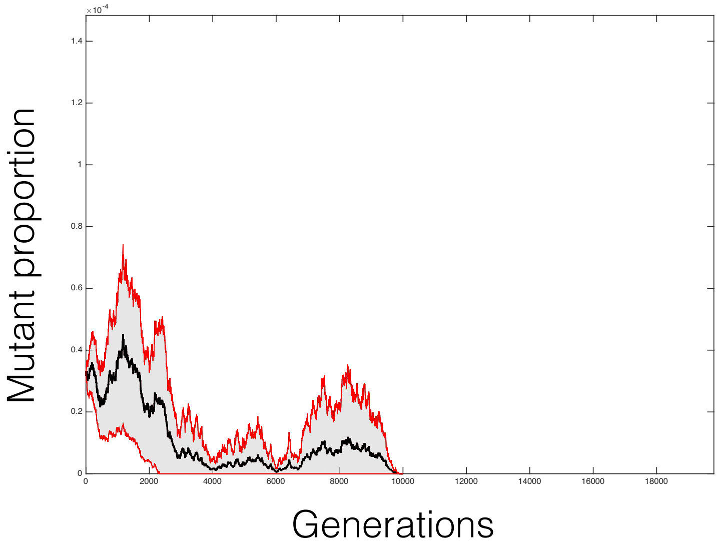
Olejars et al. constructed a population genetic model for the invasion of an allele for helper sterility ($q_1 = 1$) in a haplodiploid species[36]. In contrast to both our findings and empirical studies, their analysis found that, in some ecological conditions, queen promiscuity (lower R) actually promotes the genetic invasion of worker sterility (sterile helper and pure reproductive; $q_1 = 1, q_2 = 0$). Davies and Gardner have generalized this model to show that these findings arise for two reasons[37]. First, the analysis only considered sterility alleles that are always expressed by workers ($p \approx 1$). As such, the rarity of the gene and haplodiploid genetics means that half the workers in a monogamous colony will be sterile whereas only a quarter of workers will be sterile in a twice mated colony. If the parameters of the model are tuned such that the efficiency of a colony with 50 percent sterile workers is relatively low compared to that of a colony with 25 percent sterile workers, then worker sterility ($q_1 = 1$) may be more likely to spread in a population of twice mated colonies (lower R) then under monogamous mating (higher R). In contrast, Davies and Gardner found that queen monogamy (higher R) always favours the invasion of a worker sterility gene ($q_1 = 1$) if it is not unconditionally expressed ($0 < p < 1$.)

Secondly, Olejars et al. only performed an invasion analysis and did not further investigate the evolutionary outcome after successful invasion (an equilibrium analysis.) Davies and Gardner found that an equilibrium analysis of the model reveals that queen monogamy (higher R) always promotes worker sterility ($q_1 = 1$), as corroborated by our results[36,37].



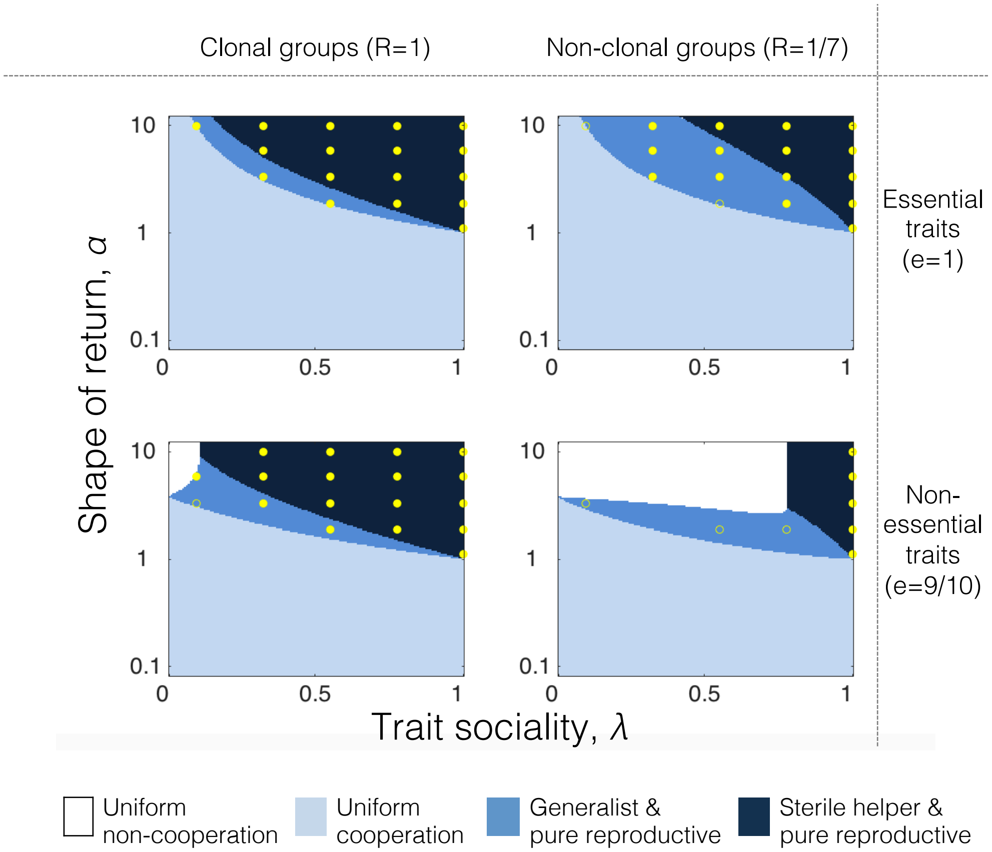
Supplementary figure 1: numerical check of uninvadability

We showed that all of the equilibria in our model are uninvadable and thus ESSs of the system using a novel analytical methodology. In order to corroborate this, we can show numerically that all of the division of labour equilibria are maxima of the local fitness landscape. Above are two division of labour equilibria: a) a generalist and pure reproductive equilibrium (parameters: $\alpha = 1.3956$, $\lambda = 0.703$, $R = 1/3$, $e = 0.9$) and b) a sterile helper and pure reproductive equilibrium (parameters: $\alpha = 1.1052$, $\lambda = 1$, $R = 1$, $e = 0.9$). In both cases, we plot the relative fitness of rare mutant lineages with respect to each of the traits at fixed distances from the equilibrium strategy. We see that in all cases, the equilibrium strategy (at the meeting point of the considered mutant lineages) is the maximum of the local fitness landscape and is thus an ESS. This results is found for all division of labour equilibria considered in the analysis.



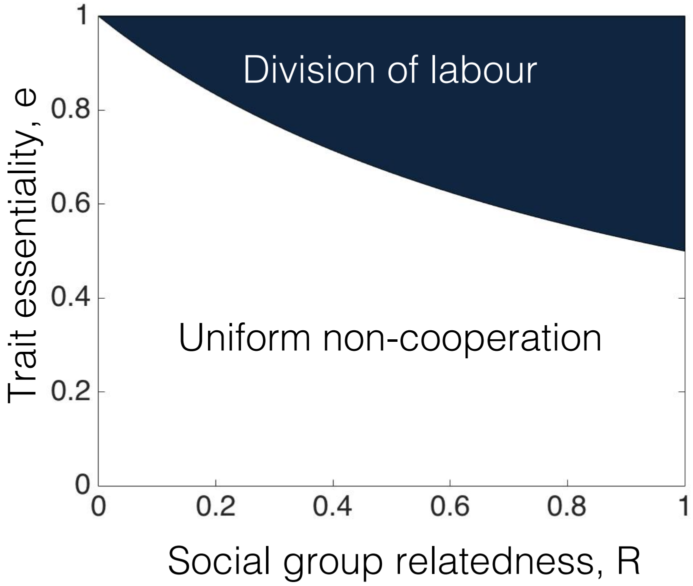
Supplementary figure 2: mutant invasion simulations

We also used individual based-simulations to show that division of labour equilibria are stable under a dynamic process. We show here the results of $10,000$ simulated invasions of a sterile helper and pure reproductive equilibrium (parameters: $\alpha = 9.9742$, $\lambda = 1$, $R = 1/3$, $e = 1$; traits: $p = 0.25$; $q_1 = 1$; $q_2 = 0$) by a mutant lineage with perturbed phenotype probability ($p' = 0.2$). We plot the mean (black line) and 95 percent confidence intervals (red lines) of mutant proportion in the population over time. We see that at the end of the simulation, the upper boundary of the confidence interval is 0 and therefore all mutant lineages have gone extinct. The number of social groups was $M = 10,000$. This invasion analysis was repeated for 60 division of labour equilibria and for perturbations along all possible trait axes.



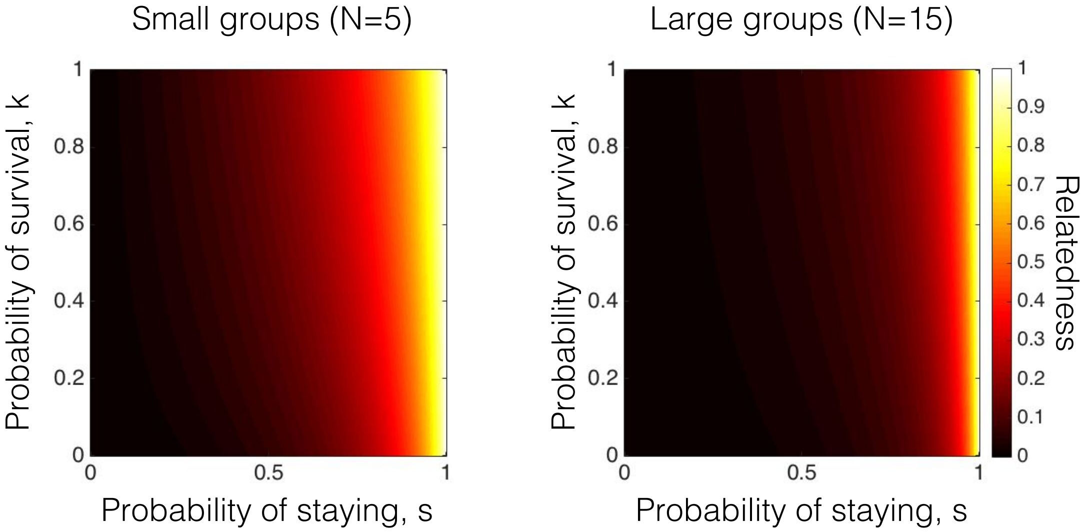
Supplementary figure 3: the stability of division of labour

We used individual-based simulations to show that 60 division of labour equilibria are stable under a dynamic process. For each equilibrium, we considered mutant lineages that deviated by an amount $\pm\epsilon = 0.05$ in each possible trait value. Each such mutant invasion was simulated $L = 10,000$ times for $T = 10,000$ generations in a population consisting of $M = 10,000$ social groups. The yellow circles show the parameter space values of each of the division of labour equilibria that were considered. A filled circle signifies that all mutant lineages went extinct for that equilibrium. An unfilled circle signifies that $1-2$ mutant lineages did not go extinct (out of $10,000$) for at least one of the trait perturbations. We note that no mutant lineage in any of the simulations invaded to fixation and that the unfilled circles are still ESSs albeit weakly uninvadable.



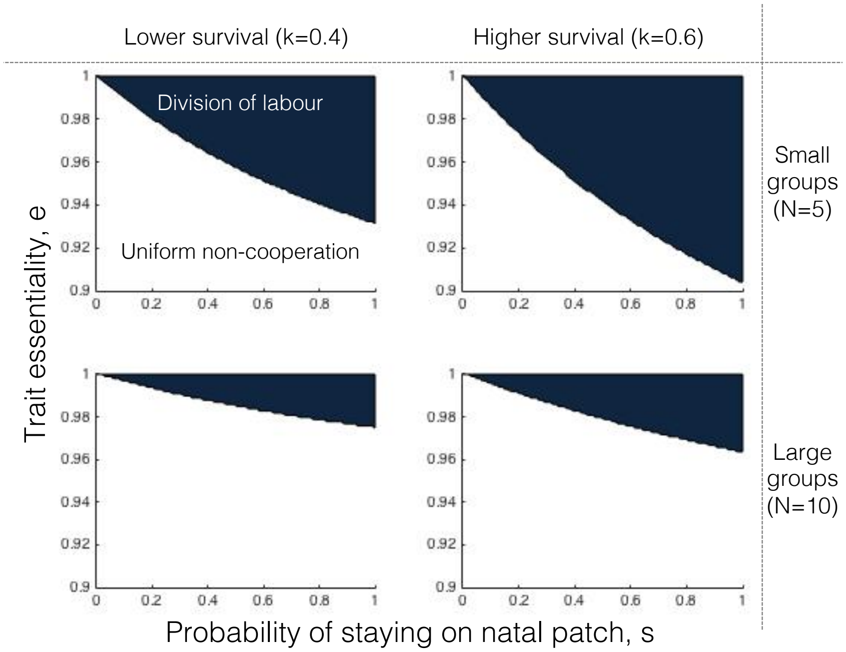
Supplementary figure 4: open model result

In our open model analysis, in which relatedness, R is left as a tuneable, independent parameter, we found that division of labour between a sterile helper ($q_1 = 1$) and a pure reproductive ($q_2 = 0$) is favoured if a threshold condition in trait essentiality is satisfied ($e > 1/(1 + \lambda R)$). Here we depict this threshold condition for $\lambda = 1$ (others-only trait). This is intended for comparison with a closed model analysis.



Supplementary figure 5: the demographic parameters that determine relatedness

In our closed model analysis, relatedness is generated by the interaction of two processes. First, limited dispersal means that a fraction of offspring (s) stay and compete for niches on the natal patch (focal social group.) Secondly, a fraction of individuals (k) survive from one generation to the next leading to overlapping generations. Group size (N) is modelled explicitly due to its interaction with the demographic processes. Here, we depict how each of these terms in our model impact others-only relatedness. We find that limited dispersal (high s), overlapping generations (high k) and small group sizes (low N) lead to high social group relatedness.



Supplementary figure 6: closed model result

For our closed model, we recover the threshold condition in trait essentiality (e) found in our open model analysis (Supplementary figure 4.) We have that division of labour between a sterile helper and pure reproductive is favoured under limited dispersal (high s), overlapping generations (high k) and small group sizes (low N), which are all factors that lead to high relatedness (Supplementary figure 5.)

Previous models	**System**	**Within-group conflict?**	**All forms of division allowed?**	**Conditions for division?**
Hamilton (1964); Trivers and Hare (1976) and others.	Social insects	✓	✗	✓*
Oster and Wilson (1978); Wilson (1968, 1976) and others.	Social insects	✗	✗	✗**
Michod et al. (2006)	Cell groups	✗	✓	✓
Ackermann et al. (2008)	Cell groups	✓	✗	✓
Willensdorfer (2009)	Cell groups	✗	✗	✓
Rossetti et al. (2010)	Cell groups	✓	✗	✓
Ispolatov et al. (2011)	Cell groups	✗†	✗	✓
Solari et al. (2013)	Cell groups	✗	✗	✓
Tannenbaum (2007)	General	✗	✗	✓
Gavrilets (2010)	General	✗††	✓	✓
Rueffler et al. (2012)	General	✗	✓	✓

Supplementary table 1: **Summary of previous theoretical models of division of labour.** We list the biological system that each model was based upon, inspired by, or applied to. We specify whether each model allows for the potential of within-group conflict. We also include whether the model is free to predict all forms of division of labour (Figure 1) or whether there are any built-in constraints in this respect. Finally, we state whether the model makes explicit predictions for the conditions that favour division of labour. *The conditions required for the evolution of sterile helpers in the social insects has been examined with a number of techniques, including Hamilton's rule, inclusive fitness theory and population genetics. Here we cite the earliest theoretical works. However, this has been a very productive field of research with a large body of literature as the past 5 years alone can attest[36,48–55]. **These models of division of labour in social insects focus on the conditions that favour multiple worker castes and the ratios thereof. †This is an aggregation model and so within-group conflict is allowed in principle. However, the model is solved by assuming perfect asymmetry between partners that could only arise through pleiotropy on the same genotype. ††Social groups are formed by a unicellular bottleneck but genetic variation within groups may arise due to mutations. Thus, while there is scope for within-group conflict, it will tend to be very small.

Previous models	**System**	$\boldsymbol{\alpha}$	$\boldsymbol{R}$	$\boldsymbol{\lambda}^{\dagger}$	$\boldsymbol{e}$	**Form(s) of division**	**Condition for division:**
Michod et al. (2006)	Volvocine algae	-	1	$\frac{N-1}{N}$	1	all	if returns accelerating ($\alpha > 1$)
Tannenbaum (2007)	General	> 1	1	$\frac{N-1}{N}$	1	only sterile helper and pure reproductive	No conditions relevant here
Ackermann et al. (2008)	*Salmonella typhimurium*	1	-	1	-	only sterile helper and pure reproductive	if high relatedness (R)
Willensdorfer (2009)	Volvocine algae, *D. discoideum*	-	1	$\frac{N-1}{N}$	1	assumes helper is sterile	for many shapes of return (α)
Rossetti et al. (2010)	Cyanobacteria	-	0 or 1	$\frac{N-1}{N}$	-	only sterile helper and pure reproductive	only if $R = 1$
Gavrilets (2010)	General	-	≈ 1	$\frac{N-1}{N}$	1	all	if α is high
Ispolatov et al. (2011)	Volvocine algae, cyanobacteria	-	1	$\frac{N-1}{N}$	1	not specified	if returns accelerating ($\alpha > 1$)
Rueffler et al. (2012)	General	-	1	$\frac{N-1}{N}$	-	not specified	if returns accelerating ($\alpha > 1$)
Solari et al. (2013)	Volvocine algae	-	1	$\frac{N-1}{N}$	1	assumes helper is sterile	if returns accelerating ($\alpha > 1$)

Supplementary table 2: **Specific links to some previous models of division of labour.** Only the microbial and general models of division of labour have attempted to derive conditions for which division of labour is of adaptive value. We list the biological systems that each model was based upon or inspired by. For ease of comparison we only consider the model assumptions and predictions of each that are relevant to the parameter space of our theoretical model. The parameters listed are: the shape of the non-linear return to cooperation, α; relatedness to social group neighbours, R; the sociality of the trait, λ and how essential the trait is, e. N is the fixed size of the social group and the symbol '$-$' signifies that the parameter (or closest approximation thereof) is left free to vary. We also include whether the model is free to predict all forms of division of labour or whether there are any built-in constraints in this respect. Each model may make further assumptions or include further parameters not captured by our model (i.e. group size, rate of mutation, cost of differentiation) and predictions for how division of labour may depend on these factors is not presented here. † For many of these models, group fitness maximisation is assumed, and so how the benefits of cooperation are shared within the group is not made explicit. For these models, we say that the cooperative trait is 'whole-group' ($\lambda = (N/(N-1))$ as all individuals receive the same benefits from cooperation.

References

[1] Eshel, I. & Motro, U. Kin selection and strong evolutionary stability of mutual help. *Theoretical population biology* **19**, 420–433 (1981).

[2] Taylor, P. D. Evolutionary stability in one-parameter models under weak selection. *Theoretical Population Biology* **36**, 125–143 (1989).

[3] Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge university press, 1982).

[4] Maynard Smith, J. & Price, G. R. The logic of animal conflict. *Nature* **246**, 15–18 (1973).

[5] Taylor, P. D. & Frank, S. A. How to make a kin selection model. *Journal of Theoretical Biology* **180**, 27–37 (1996).

[6] Brown, S. P. & Taylor, P. D. Joint evolution of multiple social traits: a kin selection analysis. *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 415–422 (2010).

[7] Day, T. & Taylor, P. D. Unifying genetic and game theoretic models of kin selection for continuous traits. *Journal of Theoretical Biology* **194**, 391–407 (1998).

[8] Mullon, C., Keller, L. & Lehmann, L. Evolutionary stability of jointly evolving traits in subdivided populations. *The American Naturalist* **188**, 175–195 (2016).

[9] Geritz, S. A., Mesze, G., Metz, J. A. *et al.* Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology* **12**, 35–57 (1998).

[10] Taylor, P. D. & Irwin, A. J. Overlapping generations can promote altruistic behavior. *Evolution* **54**, 1135–1141 (2000).

[11] Taylor, P. D. Inclusive fitness in a homogeneous environment. *Proceedings of the Royal Society of London B: Biological Sciences* **249**, 299–302 (1992).

[12] Levin, S. R. & West, S. A. The evolution of cooperation in simple molecular replicators. In *Proc. R. Soc. B*, vol. 284, 20171967 (The Royal Society, 2017).

[13] Gavrilets, S. Rapid transition towards the division of labor via evolution of developmental plasticity. *PLoS Comput Biol* **6**, e1000805 (2010).

[14] Willensdorfer, M. On the evolution of differentiated multicellularity. *Evolution* **63**, 306–323 (2009).

[15] Strassmann, J. E., Zhu, Y. & Queller, D. C. Altruism and social cheating in the social amoeba dictyostelium discoideum. *Nature* **408**, 965–967 (2000).

[16] Veening, J.-W. *et al.* Transient heterogeneity in extracellular protease production by bacillus subtilis. *Molecular systems biology* **4**, 184 (2008).

[17] Griffin, A. S., West, S. A. & Buckling, A. Cooperation and competition in pathogenic bacteria. *Nature* **430**, 1024–1027 (2004).

[18] West, S. A., Diggle, S. P., Buckling, A., Gardner, A. & Griffin, A. S. The social lives of microbes. *Annual Review of Ecology, Evolution, and Systematics* 53–77 (2007).

[19] Kümmerli, R., Jiricny, N., Clarke, L., West, S. & Griffin, A. Phenotypic plasticity of a cooperative behaviour in bacteria. *Journal of evolutionary biology* **22**, 589–598 (2009).

[20] Foster, W. Experimental evidence for effective and altruistic colony defence against natural predators by soldiers of the gall-forming aphid pemphigus spyrothecae (hemiptera: Pemphigidae). *Behavioral Ecology and Sociobiology* **27**, 421–430 (1990).

[21] Jiricny, N. *et al.* Fitness correlates with the extent of cheating in a bacterium. *Journal of evolutionary biology* **23**, 738–747 (2010).

[22] Rumbaugh, K. P. *et al.* Quorum sensing and the social evolution of bacterial virulence. *Current Biology* **19**, 341–345 (2009).

[23] Sandoz, K. M., Mitzimberg, S. M. & Schuster, M. Social cheating in pseudomonas aeruginosa quorum sensing. *Proceedings of the National Academy of Sciences* **104**, 15876–15881 (2007).

[24] Strassmann, J. E. & Queller, D. C. Evolution of cooperation and control of cheating in a social microbe. *Proceedings of the National Academy of Sciences* **108**, 10855–10862 (2011).

[25] Queller, D. C., Strassmann, J. E. *et al.* Fruiting bodies of the social amoeba dictyostelium discoideum increase spore transport by drosophila. *BMC evolutionary biology* **14**, 105 (2014).

[26] Fisher, R. M., Cornwallis, C. K. & West, S. A. Group formation, relatedness, and the evolution of multicellularity. *Current Biology* **23**, 1120–1125 (2013).

[27] West, S. A. & Cooper, G. A. Division of labour in microorganisms: an evolutionary perspective. *Nature Reviews Microbiology* **14**, 716 (2016).

[28] Bonner, J. T. *Cellular Slime Molds* (Princeton University Press, 2015).

[29] Votaw, H. R. & Ostrowski, E. Stalk size and altruism investment within and among populations of the social amoeba. *Journal of Evolutionary Biology* (2017).

[30] Koufopanou, V. The evolution of soma in the volvocales. *American Naturalist* 907–931 (1994).

[31] Koufopanou, V. & Bell, G. Soma and germ: an experimental approach using volvox. *Proceedings of the Royal Society of London B: Biological Sciences* **254**, 107–113 (1993).

[32] Oster, G. F. & Wilson, E. O. *Caste and ecology in the social insects* (Princeton University Press, 1978).

[33] Wilson, E. O. The ergonomics of caste in the social insects. *The American Naturalist* **102**, 41–66 (1968).

[34] Gordon, D. M. The organization of work in social insect colonies. *Nature* **380**, 121–124 (1996).

[35] Wilson, E. O. The relation between caste ratios and division of labor in the ant genus pheidole (hymenoptera: Formicidae). *Behavioral Ecology and Sociobiology* **16**, 89–98 (1984).

[36] Olejarz, J. W., Allen, B., Veller, C. & Nowak, M. A. The evolution of non-reproductive workers in insect colonies with haplodiploid genetics. *Elife* **4**, e08918 (2015).

[37] Davies, N. G. & Gardner, A. Monogamy promotes worker sterility in insect societies. *bioRxiv* 059154 (2016).

[38] Hamilton, W. D. The genetical evolution of social behaviour. i and ii. *Journal of theoretical biology* **7**, 1–52 (1964).

[39] Trivers, R. L. & Hare, H. Haploidploidy and the evolution of the social insect. *Science* **191**, 249–263 (1976).

[40] Wilson, E. O. Behavioral discretization and the number of castes in an ant species. *Behavioral Ecology and Sociobiology* **1**, 141–154 (1976).

[41] Michod, R. E., Viossat, Y., Solari, C. A., Hurand, M. & Nedelcu, A. M. Life-history evolution and the origin of multicellularity. *Journal of theoretical Biology* **239**, 257–272 (2006).

[42] Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**, 987–990 (2008).

[43] Rossetti, V., Schirrmeister, B. E., Bernasconi, M. V. & Bagheri, H. C. The evolutionary path to terminal differentiation and division of labor in cyanobacteria. *Journal of Theoretical Biology* **262**, 23–34 (2010).

[44] Ispolatov, I., Ackermann, M. & Doebeli, M. Division of labour and the evolution of multicellularity. *Proceedings of the Royal Society of London B: Biological Sciences* rspb20111999 (2011).

[45] Solari, C. A., Kessler, J. O. & Goldstein, R. E. A general allometric and life-history model for cellular differentiation in the transition to multicellularity. *The American Naturalist* **181**, 369–380 (2013).

[46] Tannenbaum, E. When does division of labor lead to increased system output? *Journal of theoretical biology* **247**, 413–425 (2007).

[47] Rueffler, C., Hermisson, J. & Wagner, G. P. Evolution of functional specialization and division of labor. *Proceedings of the National Academy of Sciences* **109**, E326–E335 (2012).

[48] Gardner, A., Alpedrinha, J. & West, S. A. Haplodiploidy and the evolution of eusociality: split sex ratios. *The American Naturalist* **179**, 240–256 (2011).

[49] Alpedrinha, J., West, S. A. & Gardner, A. Haplodiploidy and the evolution of eusociality: worker reproduction. *The American Naturalist* **182**, 421–438 (2013).

[50] Rautiala, P., Helanterä, H. & Puurtinen, M. Unmatedness promotes the evolution of helping more in diplodiploids than in haplodiploids. *The American Naturalist* **184**, 318–325 (2014).

[51] Rautiala, P., Helanterä, H. & Puurtinen, M. The evolutionary dynamics of adaptive virginity, sex-allocation, and altruistic helping in haplodiploid animals. *Evolution* (2017).

[52] Avila, P. & Fromhage, L. No synergy needed: ecological constraints favor the evolution of eusociality. *The American Naturalist* **186**, 31–40 (2015).

[53] Liao, X., Rong, S. & Queller, D. C. Relatedness, conflict, and the evolution of eusociality. *PLoS biology* **13**, e1002098 (2015).

[54] Nowak, M. A. & Allen, B. Inclusive fitness theorizing invokes phenomena that are not relevant for the evolution of eusociality. *PLoS biology* **13**, e1002134 (2015).

[55] Davies, N., Krebs, J. & West, S. An introduction to behavioural ecology 4th edition (2012).