

# Ecological scaffolding and the evolution of individuality

Andrew J. Black <sup>1\*</sup>, Pierrick Bourrat <sup>2,3</sup> and Paul B. Rainey <sup>4,5\*</sup>

**Evolutionary transitions in individuality are central to the emergence of biological complexity. Recent experiments provide glimpses of processes underpinning the transition from single cells to multicellular life and draw attention to the critical role of ecology. Here, we emphasize this ecological dimension and argue that its current absence from theoretical frameworks hampers development of general explanatory solutions. Using mechanistic mathematical models, we show how a minimal ecological structure comprising patchily distributed resources and between-patch dispersal can scaffold Darwinian-like properties on collectives of cells. This scaffolding causes cells to participate directly in the process of evolution by natural selection as if they were members of multicellular collectives, with collectives participating in a death-birth process arising from the interplay between the timing of dispersal events and the rate of resource use by cells. When this timescale is sufficiently long and new collectives are founded by single cells, collectives experience conditions that favour evolution of a reproductive division of labour. Together our simple model makes explicit key events in the major evolutionary transition to multicellularity. It also makes predictions concerning the life history of certain pathogens and serves as an ecological recipe for experimental realization of evolutionary transitions.**

Evolutionary transitions in individuality are central to the emergence of biological complexity<sup>1–3</sup>. Each transition involved the formation of collective-level entities from the interaction of particles<sup>4,5</sup>. For example, chromosomes evolved from the joining of once independently replicating genes. Multicellular life evolved from independently replicating cells. In certain insects, eusociality evolved from independently replicating multicellular types.

Central to each of these transitions was the emergence of properties at the newly formed level that allowed individuals—at the newly formed level—to participate directly in the process of evolution by natural selection<sup>5–9</sup>. This required newly formed collectives to be discrete and vary one to another, to reproduce and to generate offspring that resemble parental types<sup>10</sup>. These essential and intertwined Darwinian properties of variation, differential reproduction and heredity are such fundamental features of living systems that it is easy to overlook that individuality is a derived state and in need of evolutionary explanation<sup>3,7–9,11–13</sup>.

With focus on multicellular life, it is evident that reproduction, in even simple multicellular forms, is a complex process<sup>9,11,12,14</sup>. It is tempting to invoke selection as its cause. However, this is problematic because the earliest collectives lacked capacity for collective-level reproduction. To invoke selection—at the collective level—as the cause of collective-level reproduction is to invoke the trait requiring explanation as the cause of its own evolution. Clearly such an explanation is unsatisfactory.

One way to avoid this dilemma is to recognize opportunities for co-option of pre-existing cellular traits. For example, in the colonial volvocine algae, group formation evolved by co-option and expansion of cell-cycle regulation evident in unicellular *Chlamydomonas*<sup>15</sup>. In experimentally evolved snowflake yeast, collective-level reproduction emerged via co-option of apoptotic capacity already apparent in single-cell precursors<sup>16</sup>.

We do not wish to downplay the importance of co-option but there is conceivable value in asking whether Darwinian properties—at the collective level—might emerge in the absence of co-option. Such a take-nothing-for-granted line of enquiry presents a challenge as it requires conceiving possibilities for the emergence of properties essential for collectives to participate in the process of evolution by natural selection from starting states that lack any manifestation of collective-level Darwinian properties. In essence, it begs explanations for how Darwinian properties might emerge from non-Darwinian entities and, therefore, by non-Darwinian means. Solutions stand to inform not only how multicellular states arise from single cells but also how Darwinian properties might materialize during each of the major evolutionary transitions, including that from non-living matter.

A solution that we advance draws heavily on ecology, the significance of which we suggest has been overlooked, even though the importance of population structure has been emphasized in literature on the levels of selection<sup>17</sup>. It recognizes that Darwinian properties can be ‘scaffolded’ by the environment: that these properties can be exogenously imposed in such a way as to cause lower-level entities (for example, cells) to become unwitting participants in a selective process that occurs over a longer timescale than the timescale over which cell-level selection occurs and as part of a larger (collective-level) entity. For development of general views on scaffolding processes see ref. <sup>18</sup>.

Ecological scaffolding underpinned a recent (and ongoing) experimental exploration of the evolution of multicellularity<sup>19,20</sup>. Discrete lineages established from the bacterium *Pseudomonas fluorescens* were propagated under conditions that required, for long-term persistence, repeated completion of a two-phase life cycle involving soma- and germ-like states. In the experiment, variation was discretized using glass microcosms but the design is

<sup>1</sup>School of Mathematical Sciences, University of Adelaide, Adelaide, South Australia, Australia. <sup>2</sup>Department of Philosophy, Macquarie University, Sydney, New South Wales, Australia. <sup>3</sup>Department of Philosophy & Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia.

<sup>4</sup>Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany. <sup>5</sup>Laboratoire de Génétique de l'Évolution, Chemistry, Biology and Innovation (CBI) UMR8231, ESPCI Paris, CNRS, PSL Research University, Paris, France. \*e-mail: [andrew.black@adelaide.edu.au](mailto:andrew.black@adelaide.edu.au); [rainey@evolbio.mpg.de](mailto:rainey@evolbio.mpg.de)

loosely analogous with an environment such as a pond in which reeds extend from the water<sup>19,21</sup>. Each reed allows establishment of a single microbial mat (the soma-like phase), with the spacing of reeds ensuring variation at the level of mats. Mats that collapse, for example, through physical disturbance, allow the possibility that an extant mat might, via production of a dispersing (germline-like) phase, increase its representation among the population of mats. The possibility of a selective process thus unfolds at the level of mats. After ten life-cycle generations, the fitness of derived mats significantly improved, with the most successful lineage having evolved a simple genetic switch that ensured quasi-reliable developmental change between soma- and germline-like phases<sup>19</sup>. Not only does this study demonstrate that scaffolding works but it also shows that externally imposed Darwinian properties can begin the shift toward endogenization exemplifying Van Valen's view that 'evolution is the control of development by ecology'<sup>22</sup>.

Our goal here is to show how a nominal set of ecological conditions (and ensuing evolutionary responses) can cause evolutionary transitions in individuality. We take inspiration from the experimental *Pseudomonas* studies but simplify ecological context to produce a minimal mechanistic model. Although focus is on the transition from single cells to multicellular life, we argue that the concept of ecological scaffolding is relevant to other transitions. It also makes predictions concerning the life history of certain pathogens and serves as an ecological recipe for top-down engineering of populations and communities.

## Results

**Scaffolding Darwinian properties.** We begin with a simple example of a population structure that confers Darwinian-like properties on collectives of particles (cells). Consider an environment in which resources are distributed across patches. A single cell founds a patch. Available resources allow exponential growth of the founding type; however, because resources are finite, the population eventually declines. Long-term persistence of cells requires dispersal to a new patch. Dispersal occurs at a regular time interval via, for example, some external factor such as wind, water splash or tidal flow.

Cell fate within the environment of patches depends on performance over two timescales. The first timescale is defined by the doubling time of cells. The second is defined by the timing of dispersal. To make apparent the impact of the second timescale, consider a second variant cell. This type replicates faster than the former (cells consume resources more rapidly), which means that, in a patch founded by both types, faster replicating cells rapidly exclude slower replicating types. In the following, we therefore limit the number of colonizers to a single founding cell, limiting within-patch competition.

Consider single slow-replicating (depicted in Fig. 1 as green) and fast-replicating (blue) cells that colonize separate patches (Fig. 1). Cells of both types grow and divide but different replication rates mean that blue cells deplete resources more rapidly than green cells. If dispersal occurs early (Fig. 1a) when cells are in exponential growth, then the number of extant blue cells exceeds the number of green cells and future recursions of patches are dominated by blue cells (Fig. 1b). If, however, dispersal occurs at a later time point—for example, once resources are depleted and population size is in decline, as in Fig. 1c—then future patch recursions are dominated by green types despite the fact that, within a patch, green types lose in competition with blue types (Fig. 1d).

Emerging from the dynamics of single cells and interaction with the timing of dispersal is a coupled evolutionary dynamic that unfolds at the level of patches (Fig. 1c,d). Patches manifest Darwinian-like properties of variation (spatially distributed resources ensure that variation is discretized and that patches vary one to another), differential reproduction (successful patches give rise to patches via dispersal) and heredity (offspring patches are

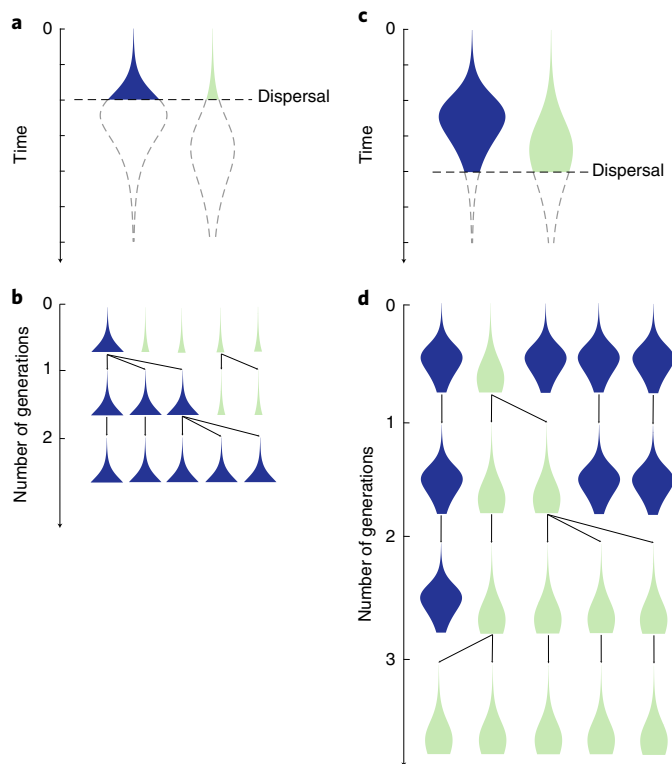
likely to resemble parental patches because new patches are founded by single cells) that are also features of the founding cells. These properties are externally imposed (scaffolded) on patches by virtue of the structure of the environment.

Note that we refer to the properties of patches as 'Darwinian-like'. It makes no sense to think of patches as multicellular organisms (they are not): if the ecological scaffold was removed, the Darwinian-like properties of the patches would instantly disappear. Yet, under the scenario outlined, cell fate is determined by selective conditions operating over the second (longer) timescale, just as if the cells themselves were members of multicellular collectives. Such a scaffolded framework, based on nothing other than patchily distributed resources and a means of dispersal between patches, establishes conditions sufficient for the evolution of traits that are adaptive at the level of patches. We elaborate the mechanistic bases using models developed in the following section but first comment on the connection between the heuristic model outlined above and existing theoretical frameworks.

The basic structure of our model, with patchy environments and a dispersal process, bears similarity to various models of group selection, including models of Wright's shifting balance theory<sup>23–25</sup>, Maynard Smith's haystack model<sup>26</sup> and others<sup>27</sup>. However, our model differs from these in motivation, in what it attempts to explain and in emphasis on mechanism. Unlike standard trait-group models<sup>27</sup>, our focus is on ecological factors affecting the formation of groups and their recurrence, and not the effect of population structure on the evolution of behaviours that are costly to individuals (such as altruism<sup>27,28</sup>). Thus, we are concerned with the formation of boundaries (because boundaries discretize variation), modes of group reproduction and genealogical connections between groups. Our goal is an explanatory framework for major evolutionary transitions. With some exceptions<sup>29,30</sup>, most models of group structure are phenomenological; they are constructed in a way that captures the relationship between variables and the problem under consideration but are compatible with a range of causal structures<sup>31</sup>. In contrast, our model is mechanistic and so evolutionary change can be understood in terms of the underlying nonlinear dynamics of particles and the feedback that arises from interaction with the timing of dispersal.

Further connections are also possible and we draw particular attention to the levels of selection literature<sup>3,32</sup>. Early stages of evolutionary transitions are encapsulated by multi-level selection 1 (MLS1) models<sup>27</sup>. In these models group fitness is the average (or sum) of the fitness of the cells that comprise collectives. A transition is said to complete once collectives become individuals and units of selection in their own right. At this point fitness of collectives is defined independently of cell fitness and relevant models fall within the multi-level selection 2 (MLS2) framework<sup>4</sup>. The shift between levels—involving transference of fitness from cells to collectives—has been difficult to capture<sup>3</sup>. An important insight came from Michod and colleagues<sup>4,33,34</sup> who articulated and modelled the concept of fitness decoupling: the notion that the shift between MLS1 and MLS2 involves collective-level fitness 'decoupling' from lower-level (cell) fitness. The heuristic model depicted in Fig. 1, and elaborated below, captures both MLS1 and MLS2 frameworks (with the former nested within the latter). It shows how the transition between levels can, in principle, arise from a simple interplay between particle-level properties, timescales and patch dynamics.

**Evolution in nested Darwinian populations.** To explore the evolutionary dynamics of the above ecological model we allow mutation to affect the growth rate of individual cells. With such a model it becomes possible to determine the effect of the timing of dispersal—the second timescale—on the dynamics of within- and between-patch competition. Mathematical details are provided in the Methods and Supplementary Information, and an interactive



**Fig. 1 | Scaffolding Darwinian properties.** Patchily distributed resources provide opportunity for two cell types (blue and green) to replicate (blue cells grow faster than green). Single cells of each type colonize discrete patches at time  $t = 0$  and consume resources. Difference in growth rate means the relationship between cell density and time differs for blue versus green populations. **a–d**, Should a dispersal event occur during exponential growth (dashed line), then more blue cells will be dispersed relative to green (**a**) and the blue population will be more successful over the long term (**b**). Conversely, should dispersal occur at a later stage and after resources are depleted (**c**), then the population of green cells outcompetes green over the long term (**d**). Note that throughout this paper we use a dimensionless unit of time.

animation is available at <https://observablehq.com/@andxblack/nested-darwinian-populations>.

The full evolutionary model consists of  $M$  patches that are each founded by a single cell of a single phenotype (growth rate  $\beta$ ). Cells within patches replicate and consume resources with mutation giving rise to types that vary in growth rate. Once resources are depleted, the population size within patches declines. After a fixed time interval,  $T$ , which defines the second timescale, dispersal takes place. Dispersal is effected by randomly selecting  $M$  patches (with replacement) in proportion to the number of cells within each patch and then randomly selecting single cells to establish the next generation of patches. In effect, the procedure is equivalent to pooling all viable cells from all patches at the time of dispersal and picking  $M$  cells at random. The dispersal regime rewards patches containing the greatest number of cells.

The bottleneck wrought at the moment of dispersal means that types founding new patches are freed from competition with faster growing types. Figure 2 shows the number of cells within a patch for a single realization with initial (arbitrarily chosen) growth rate  $\beta = 1.8$ . The bottleneck imposes a strong homogeneity on patch composition because the original population has to grow significantly before mutants start to arise. The peak (maximum) number of cells within the patch is reached at time 16.5; thus, for cells with this initial growth rate, setting a dispersal time of  $T = 10$  is fast

(within the exponential growth phase) and  $T = 30$  is slow (cells have significantly declined since their peak numbers).

Figure 3 shows the time resolved dynamics of 50 independent realizations of the full evolutionary model with 64 patches, where patches experience 200 recursions under slow ( $T = 30$ ) and fast ( $T = 10$ ) dispersal regimes. In these simulations, the maximum cell growth is set to rate of  $\beta = 2$ , which in a real system would arise from chemical or physical constraints to the rate of cell replication. The state of populations at the time of dispersal are shown in Fig. 4a,b. Single realizations of the model under slow and fast dispersal regimes are also shown in Supplementary Videos 1 and 2. Other parameters:  $\mu = 0.05$ ,  $p = 0.02$ .

Under both fast and slow dispersal regimes, patch fitness (the number of cells within patches at time of dispersal) increases rapidly before reaching a plateau (Fig. 3b,d). This is consistent with predictions arising from the logic of Darwinism: imposition of Darwinian-like properties on patches ensures patches participate in a selection process akin to evolution by natural selection, one that could be the starting point for patches to be units in their own right, provided they eventually acquire features classically associated with evolutionary individuals<sup>5,35</sup>. The plateau arises because, under the slow dispersal regime, growth rate evolves to maximize the number of particles available at the time of dispersal. Under the fast dispersal regime, the plateau is a consequence of reaching the maximum limit imposed on growth rate.

The cause of enhanced evolutionary success of patches resides in properties of individual cells. Under both fast and slow dispersal regimes, selection favours patches that harbour the greatest number of cells at the time of dispersal. Under both regimes, fast-growing cells outcompete slow-growing types within patches; however, under the slow dispersal regime, selection rewards patches containing slow-growing mutants and selects against patches dominated by fast-growing cells. The opposite is true of patches evolving under the fast dispersal regime.

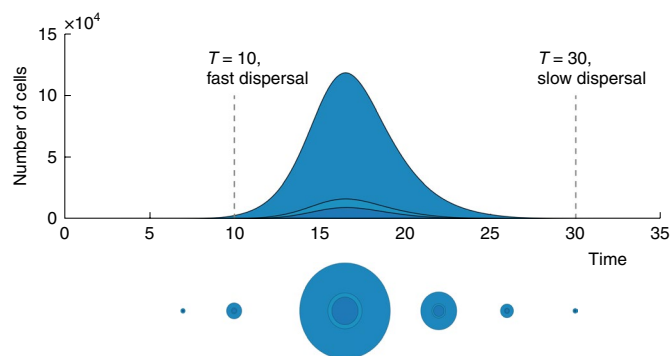
Under the slow dispersal regime, this results in the seemingly counterintuitive finding that patch fitness increases at the expense of cell fitness (Fig. 3a,b). Yet within our model, this is readily explained: fitness of a cell is measured over the short timescale while patch fitness is measured over the long timescale<sup>36,37</sup>. This captures precisely—and explains mechanistically—the notion of ‘fitness decoupling’ thought to occur during the earliest stages of the evolution of multicellular life but which has often been difficult to intuit<sup>3,38</sup>.

Under the fast dispersal regime, fast-growing cells are favoured both within patches and over the second timescale. From the perspective of the evolution of multicellular life, the selection regime imposes the same directionality at both timescales leading to the view that fitness at both timescales (levels) is ‘coupled’.

Extended Data Fig. 1a,b shows the evolutionary fate (genealogy) of ten lineages under the slow and fast dispersal regimes, respectively. Mapped on the phylogenies are changes in cell growth rate and patch size at time of dispersal. That a genealogical representation is possible derives from both the mechanistic nature of the model and the fact that patches are founded from single-cell types. Representations are shown in Supplementary Videos 3 and 4.

It is important to emphasize that the parameters and timescales chosen in the above simulations are arbitrary. For any initial growth rate  $> 1$ , it is always possible to choose fast and slow dispersal times relative to the time of the peak patch population that will result in selection over the timescale of dispersal feeding back to affect the growth rate of cells. In contrast, if the dispersal time is equal to the initial peak time, then no evolutionary change in cell growth rate will be observed.

**Synchronicity of dispersal.** The analysis thus far assumes a predictable environment: resources within patches are identical for all



**Fig. 2 | Within-patch dynamics.** The number of cells is plotted as a function of time starting from a single cell with growth rate  $\beta=1.8$  and resource  $N=10^6$ . Mutants arising during the course of growth are indicated by different shades of blue (darker and lighter colours correspond to faster and slower growing cells, respectively). Circles below the main plot show the number of cells within a patch at time points 7, 10, 16.5, 22, 26 and 30. The area of each region is proportional to the number of cells of each type within the patch (with the same colour scheme). The same representation is used in Fig. 4 and Supplementary Videos 1 and 2. The peak number of cells within the patch is reached at time  $\sim 16.5$ , thus for the initial growth rate of  $\beta=1.8$ , setting a dispersal time of  $T=10$  is considered fast, and  $T=30$  slow (shown by the dashed lines).

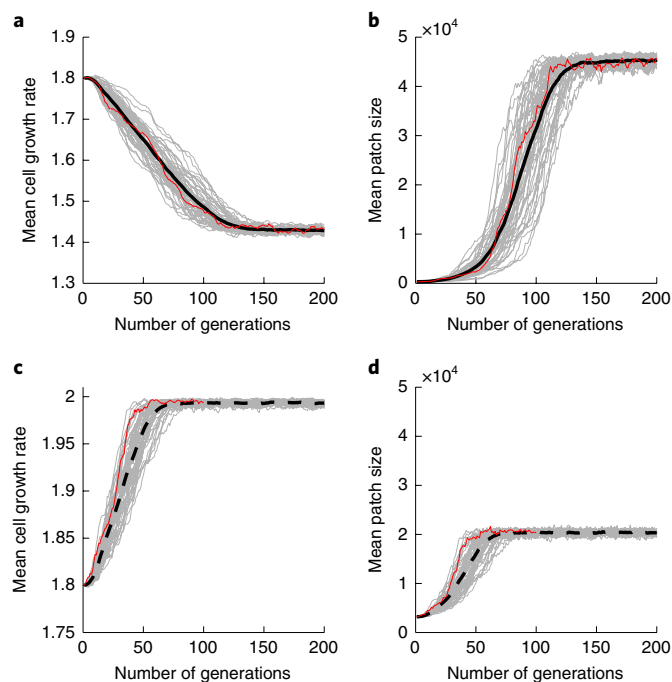
patches, the time at which patches are seeded by new colonists is fixed and so is the time interval until dispersal. In this section we explore relaxation of this strict ecology and do so by two adjustments to the model. In the first, variance in the time at which patches are seeded with new propagules is introduced. In the second, which is described in the Supplementary Information, the level of resource available in each patch is varied.

Variance in the growth time until dispersal is introduced by allowing each patch to grow for a period of time  $\tau \sim \text{norm}(T, \sigma_\tau)$ . This leads to variance in the size of the population of cells within patches at the time of dispersal. We consider only mean dispersal times that result in tension between the short- and long-term interest of cells and patches.

Figure 5 shows evolution of the average growth rate of cells and patch size for a fixed mean dispersal time ( $T=30$ ) but with increasing variance in the growth time before dispersal. For small values of  $\sigma_\tau$ , the dynamics show little change but the effect is pronounced for larger values of  $\sigma_\tau$ . The average growth rate at equilibrium increases slightly but the main effect is an increase in the average time for the equilibrium state to be realized.

Differences in average patch size, both at the beginning and end of the process, are explained by distributions in patch size that are induced by the variance in growth period across the patches in a single generation. This is visualized in Fig. 5a,b. With no variance ( $\sigma_\tau=0$ ), the size of patches is homogeneous, with small size differences caused only by mutations within the patch. Variability in growth time leads directly to variability in patch size at dispersal time. As variance in  $\tau$  increases, so does variance in patch size.

The system starts out of equilibrium, with dispersal time being much longer than the average time for populations within patches to peak. As variance in growth time increases, the likelihood that patches with large numbers of individuals at the time of dispersal increases, skewing the size distribution of patches in the next generation of patches (Fig. 5a). Once at equilibrium (Fig. 5b), the average peak population time within a patch coincides approximately with the average growth time before dispersal. This means that patches in which cells grow for shorter or longer times harbour populations with reduced sizes compared to the average; hence, mean patch population size at equilibrium is reduced when



**Fig. 3 | Effect of dispersal timescale on properties of cells and patches.**

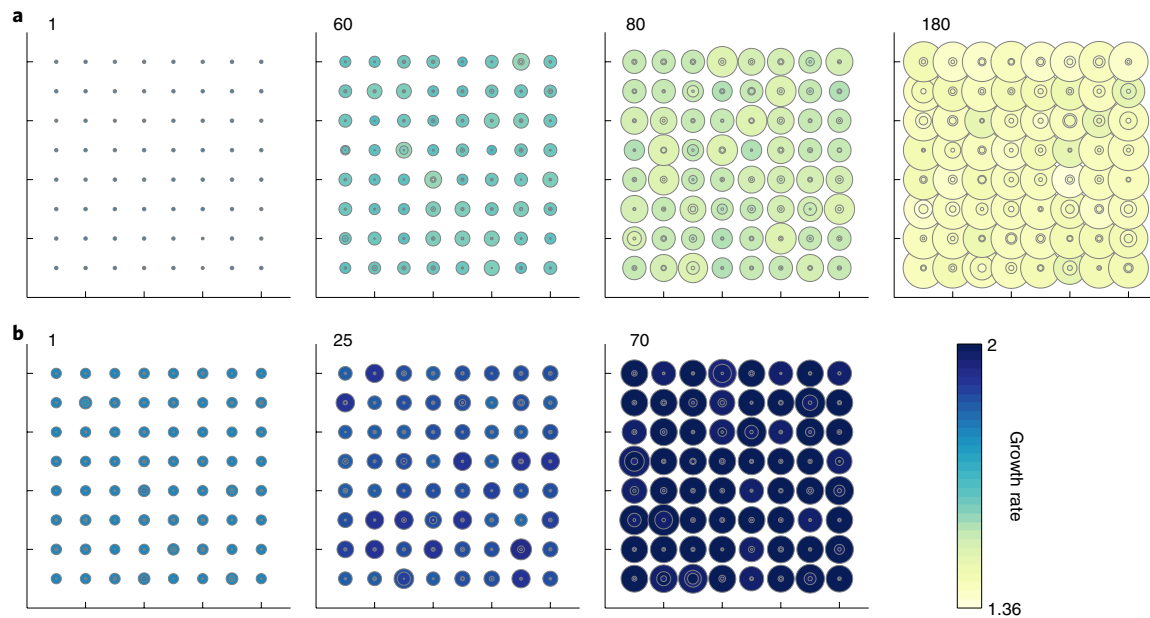
Each grey line is from an independent realization of the stochastic model with 64 patches. The solid black lines are averages over 50 realizations.

**a, b.** The evolution of the average growth rate (**a**) and patch size (**b**) over all patches under the slow dispersal ( $T=30$ ) regime. **c, d.** The evolution of the average growth rate (**c**) and patch size (**d**) under the fast dispersal ( $T=10$ ) regime. Both regimes start with a homogenous population of cells with  $\beta=1.8$  and, in both cases, the average patch size increases but for slow dispersal this is achieved by cells decreasing their average growth rate. Figure 4 shows the 64 patches at the moment of dispersal for two single realizations (shown in red above). Other parameters as in Fig. 1.

compared with the treatment in which patches are seeded at the same time. Also affected by increasing variance is average cell growth rate. This decreases even when the time of patch seeding shows maximal variance but both the rate of reduction is lowered and the equilibrium growth rate is higher. As long as average patch size correlates with the growth rate of the initial colonist cell, evolution in growth rate is observed. Quantitatively similar findings stem from stochastic alterations in the availability of resources within patches (see Supplementary Information).

**Evolution of patch traits.** Introduction of a second timescale, defined by dispersal events necessary for establishment of new patches, affects the evolution of cell growth rate, and leads to changes that affect the evolutionary dynamics of patches. From the patch perspective, derived patches are more fit than ancestral patches but this is not a consequence of traits adaptive at the patch level. Under both slow and fast dispersal regimes, selection favours cells whose growth rate maximizes the number of cells available for dispersal. Changes in cell growth rate fully explain the evolutionary dynamics of patches. This cell-level perspective further emphasizes the previous comment that patches are not to be confused with even the most basic manifestations of multicellular life forms.

Nonetheless, an ecological scaffold that couples short- and long-term timescale dynamics establishes kin groups<sup>39</sup>, conducive to the evolution of traits adaptive at the level of patches. By this we mean traits that would be difficult to explain from the viewpoint of cells. This prediction becomes intuitive upon switching perspectives, from cell level to patch level. Although patches are endowed



**Fig. 4 | Evolution of patch size under slow and fast regimes. a,b,** The dynamic of patch-size evolution and corresponding effects on cell growth rate under slow (a) and fast (b) dispersal regimes. Simulations are shown in Supplementary Videos 1 and 2. Colour corresponds to cell growth rate and patch size is proportional to the number of cells within patches at the moment immediately before dispersal. Generation number is indicated above each panel and corresponds to the realizations highlighted in red in Fig. 3. An interactive animation is available at <https://observablehq.com/@andxblack/nested-darwinian-populations>.

with Darwinian-like properties, there is scope for patches to evolve genuine Darwinian properties—in a ratchet-like manner<sup>40</sup>—so that patches participate in the process of evolution by natural selection and become bearers of adaptations at the patch level.

What might such patch-level adaptations entail and what might constitute their mechanistic (cell level) basis? A fundamental requirement given the need for patches to pass through single cell bottlenecks at each recursion is evolution of a stochastic epigenetic switch (a simple developmental programme), such as observed previously<sup>41–43</sup> including those arising from experimental explorations of the evolution of multicellular life<sup>19</sup>.

To investigate this possibility, the basic model is extended to include two types of cell. The first type, which we denote G, is essentially the same as in our first model, with the exception that at each reproduction event there is some probability,  $q$ , that instead of giving rise to another G cell, a different cell type, denoted S, is produced instead. S cells also consume resources but, unlike G cells, S cells cannot replicate or be dispersed. The production of S cells is thus costly. Full mathematical details of the model are given in Methods and Supplementary Information. The phenotype of G cells is quantified by their growth rate,  $\beta$ , and the probability of production of S cells,  $q$ . All other parameters are fixed.

In this formulation, S cells are a rough approximation for soma. Like soma, S cells are an evolutionary dead-end. In this switch to considering S cells as proxy for soma, it follows that G cells approximate germ cells: like germ cells, G cells found the next collective generation.

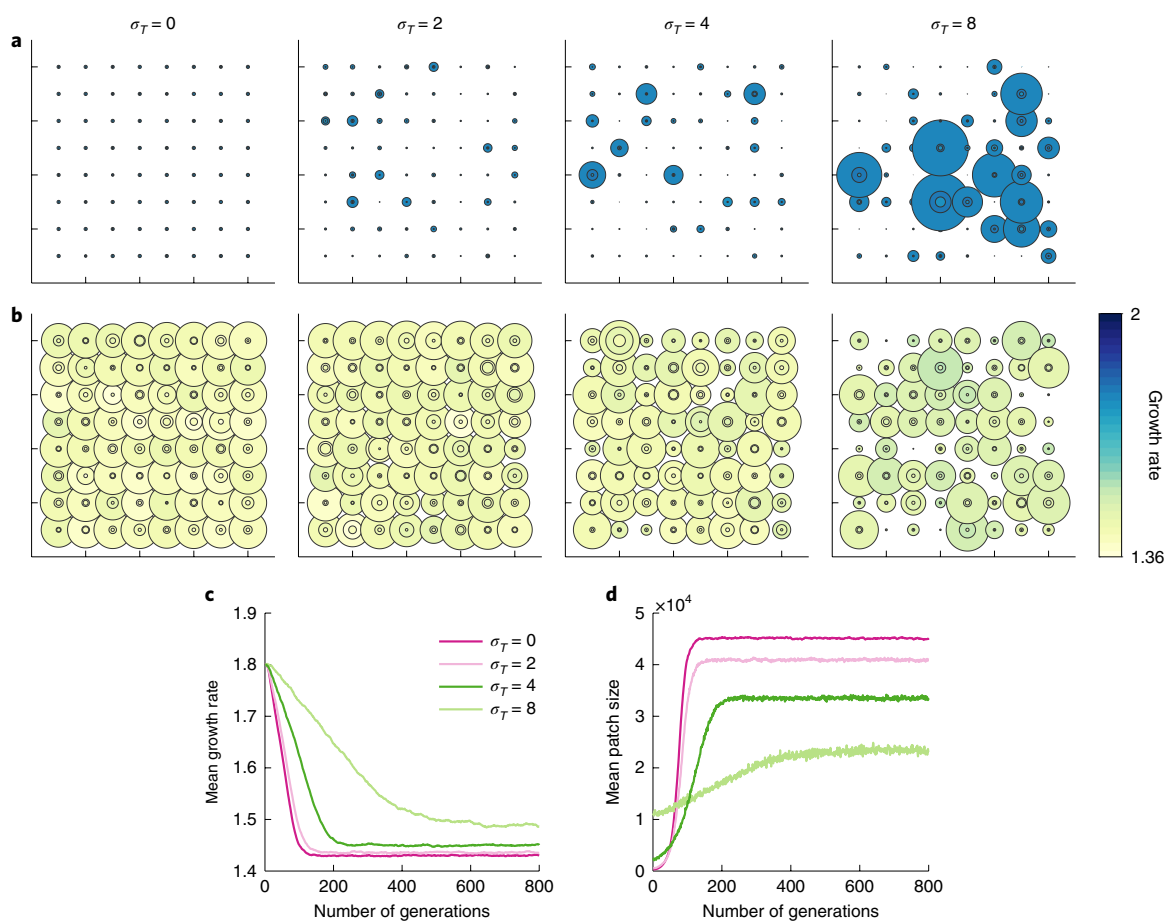
To connect with our previous results, simulations of the model were first performed with dispersal depending solely on the number of G cells within the patch at the time of dispersal; thus, patches that optimize the number of G cells maximize the number of descendent patches. As to be expected, in repeated simulations of the model in which mutation affects both growth rate and the probability of production of S cells, the rate of S cell production under both slow and fast dispersal regimes declines to zero (Fig. 6a–c). The equilibrium

fitness of both cells and patches tend to the same values as in the previous model.

To determine whether patchily distributed resources and dispersal between patches establishes conditions favouring evolutionary emergence of a division of labour, the model was re-run but with S cells now endowed with ability to aid dispersal of G types. Mathematically this was achieved by defining the probability that a cell within a patch chosen for dispersal be a function of both the number of G and S cells in the patch (see Methods and Supplementary Information). The strength of the additional contribution by the S cells to the probability of dispersal is controlled by the parameter  $\rho$  which is initially set at 0.02 per cell (see Methods).

As shown in Fig. 6d–f (and especially Fig. 6e), S cells are favoured under the slow dispersal regime (the probability of S cell production rapidly evolves away from zero and plateaus at 0.15). Under this scenario, the equilibrium cell growth rate is higher than when dispersal depends solely on the number of G cells (contrast this with the solid lines in Fig. 6a,d). This is because increased production of S cells slows the rate of production of G cells allowing the population to peak at a comparatively higher growth rate. Mean patch fitness depends on the contribution that S cells make toward dispersal of G types.

The time to fully reach equilibrium for the case where  $T=30$  and where both cell types contribute towards dispersal is much longer for this model than the previous version. This is because the fitness landscape is flat in the region of the equilibrium point. This means that a broad combination of values of  $\beta$  and  $q$  give very similar patch fitnesses. This also manifests in large fluctuations in the value of  $q$  seen for individual realizations in Fig. 6e. Simulations that run for 2,000 generations show eventual convergence (Extended Data Fig. 2). The strength of the dispersal assistance,  $\rho$ , affects the exact equilibrium but over a certain threshold becomes relatively insensitive to the precise value (Extended Data Fig. 3). This is because in the model there is an inherent tradeoff between the production of S and G cells.



**Fig. 5 | Asynchronicity in dispersal time.** Patches are founded with increasing variance in the time available for growth within each patch,  $\sigma_T$ . **a,b**, The size and composition of patches for realization of a single generation from the start of the simulation (**a**) and once equilibrium has been reached (**b**) for increasing  $\sigma_T$ . **c,d**, The evolution of average growth rate (**c**) and patch size (**d**) over 800 generations, respectively, from 50 realizations of the model. All simulations start from cells with growth rate of  $\beta = 1.8$ .

Under the fast regime, S cells are not favoured and the growth rate of G simply increases to its maximum limit. However, if the maximum allowable growth rate increases beyond the limit of  $\beta = 2$ , production of S cells under the fast dispersal regime can be favoured. The key point is that under the slow dispersal regime, production of S is always favoured. When dispersal time is fast, production of S cells is favoured only if cell growth rate can increase to the point at which peak population size is reduced (through early and rapid depletion of resources). In real systems it is likely that cells would be already close to their maximum growth rate and further increases would depend on rare beneficial mutations. In contrast, decreases in growth rate are readily achieved via deleterious mutations.

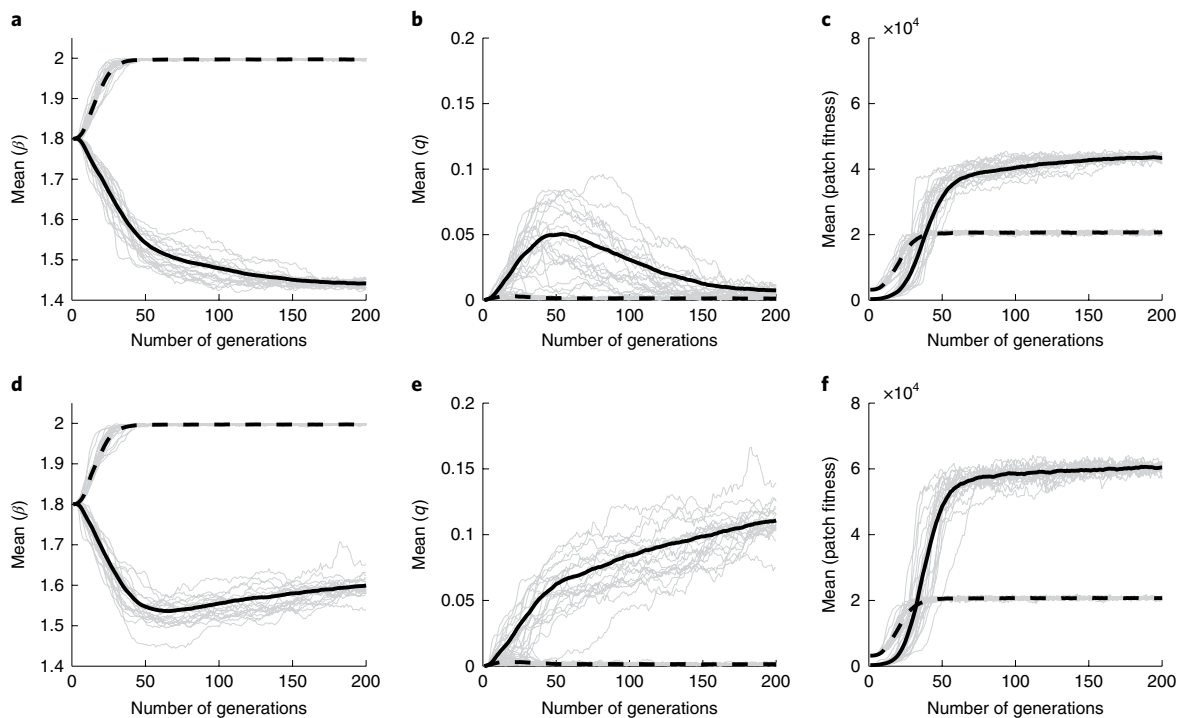
## Discussion

The major evolutionary transitions in individuality pose some of the most intriguing and complex problems in biology. Numerous perspectives have been offered, ranging from theoretical multi-level selection frameworks<sup>3,44–49</sup>, to views that give prominence to explanations for the evolution of cooperation<sup>50–52</sup>; from perspectives that emphasize the importance of specific mechanisms<sup>4,53–55</sup> through those, like ours, that emphasize the pivotal importance of the origins of collective-level Darwinian properties<sup>5,7,9,11,13,56,57</sup>.

Encompassed within these diverse views are central concepts that can appear ambiguous. This is particularly true of scenarios in which evolutionary transitions in individuality are described in terms of ‘shifts in levels of selection’ or, more specifically, shifts

between multi-level selection frameworks MLS1 (where individual cells are Darwinian) and MLS2 (where groups are Darwinian). A thorough analysis led Okasha<sup>3</sup> to propose the existence of a ‘grey area’ between early and later stages where both MLS1 and MLS2 perspectives can be taken. Closely allied is the notion of ‘fitness decoupling’<sup>38</sup>—a sense that, as selection shifts from a lower to a higher level, the fitness of the higher level decouples from that of the lower—and the related idea of ‘de-Darwinization’ of lower-level components<sup>5</sup>. While to the initiated all these terms convey meaning, they remain largely metaphorical and descriptive (but see Michod and Nedelcu<sup>34</sup> for a model). Discussion of issues surrounding evolutionary transitions in individuality needs to become more mechanistic.

Our model places emphasis on simplicity and causality and gives prominence to ecological factors. The ability of natural selection to act on collectives of cells depends on emergence of some manifestation of heritable variance in fitness at the collective level. In our approach, the possibility that this arises from co-option of pre-existing cell-level traits was recognized but put aside. While resulting in a high bar, it gives emphasis to the fact that reproduction, heredity and variation are derived traits and their existence should not be presumed<sup>7,11,14,58</sup>. It has also made transparent a genuine dilemma, namely, the need to explain how Darwinian properties emerge from non-Darwinian entities and by non-Darwinian means. If Darwinian properties do not pre-exist or cannot arise by co-option of pre-existing lower-level traits, then their earliest manifestation necessarily



**Fig. 6 | Simulations of the S-G model.** **a-c**, Simulations of the model where the probability of dispersal depends solely on the number of G cells in patches at the time of dispersal,  $T$ . **a**, Mean growth rate (averaged over generations). **b**, Mean probability of a reproduction event creating an S cell. **c**, Mean patch fitness (the weight assigned to each patch is described in the Methods, equation (4)). Each grey line is a single stochastic realization and the solid and dashed lines are averages over 100 independent realizations for slow ( $T=30$ ) and fast dispersal ( $T=10$ ) regimes, respectively. **d-f**, Similar simulations to **a-c** but where the probability of dispersal from a patch is a function of both the number of G and S cells ( $\rho=0.02$ ). In all scenarios, the maximum cell growth rate is limited to 2. Extended Data Fig. 3 shows the dynamics over 2,000 generations to confirm that equilibrium is reached for the case where dispersal is slow and both cell types contribute towards dispersal.

lies in some exogenous factor(s). The solution we have advocated involves recognizing the continuity between organisms and their environments; the idea that Darwinian-like properties can be scaffolded by the environment in much the same way that reproduction in viruses is scaffolded by the host cell<sup>5</sup> or that development can be scaffolded by overlap of parts between parents and offspring<sup>59</sup>.

The second timescale is of critical importance. While explicit in the *Pseudomonas* experiment, the ‘ingredient’ has been missing from theoretical frameworks. When included, the second timescale underpins a patch-level, death–birth process<sup>30,60</sup>. This establishes a feedback between patch-level process and the evolutionary dynamics of cells: patches fail or succeed based on properties of the cells. That slow-growing cells are favoured when dispersal time is long is a direct consequence of this feedback. Although within-patch selection favours fast-growing cells, patches dominated by fast-growing cells contain few viable cells for dispersal. Long-term success of cells comes from alignment of cell and patch fitness. The model thus explicates further the concept of fitness decoupling<sup>34</sup> and captures in a single structure the transition between MLS1 and MLS2 frameworks.

An at first unexpected, albeit important, subtlety surrounding the second timescale arises from its frequency of occurrence relative to the initial growth rate of cells. Beginning from a position where the growth rate of cells leads to suboptimal patch occupancy at the time of dispersal, as in Fig. 3, a dispersal time that coincides with the exponential growth phase of cells drives an increase in cell growth rate (Fig. 3c,d), while also marginally increasing patch fitness. More significant though is that the fast dispersal regime is not conducive to the evolution of a reproductive division of labour. Under the fast dispersal regime, growth rate of cells is the sole factor governing patch

success: any reduction in total yield of G cells due to production of S cells is not offset by contributions that S cells make to dispersal.

For the evolutionary emergence of a reproductive division of labour, the second timescale needs to occur when cells are not in exponential growth phase (Fig. 3a,b). When evolutionary success depends on being the fastest, limited opportunities exist for exploration of phenotypic novelty<sup>61–63</sup>. This stems from the simple fact that manifestations of phenotypic novelty typically tradeoff against growth rate.

There are additional aspects of the longer timescale that are illuminating. It is usual when discussing the transition from cells to multicellular life to consider the cell as the ‘lower level’ and the group as the ‘higher level’. Accordingly, the transition from cells to multicellular life is often referred to as a ‘levels of selection problem’<sup>3,5,44,47</sup>. The same is true for all the major evolutionary transitions in individuality. It is apparent from our patch model that the evolution of multicellular life might be better articulated as a problem to be solved by understanding conditions leading to the emergence of a second timescale over which a birth–death process operates on discretely packaged variation. This shift from levels to timescales does much to clarify the kinds of conditions necessary to effect transitions in individuality<sup>22,36,37</sup>.

Parallels exist in models of virulence evolution in pathogens<sup>64–69</sup>. These are evident in the use of mechanistic models but also in the model structure and broader findings. The tradeoff between growth rate and dispersal has been previously studied but differently framed. For example, in models of the evolution of virulence, high external mortality of the host is known to favour the evolution of virulence<sup>69</sup>. In our model this is equivalent to the case where dispersal acts on a fast timescale compared to peak population time.

On the other hand, models on the tradeoff between growth rate and growth yield have shown longevity of the host to be critical for the evolution of growth rate<sup>70</sup>, which echoes our findings when the dispersal time is long. Given the similarity of our models and assumptions, it is likely that other findings from this field may be important for understanding the consequences of ecological scaffolding and the development of further models. In this direction, our model can be mapped directly to a multi-strain susceptible–infected–recovered epidemic model<sup>71,72</sup>, where cell types correspond to infected individuals with different pathogen strains and susceptible types are analogous to within-patch resources. This opens the possibility for modelling and even formulation of predictions regarding viral division of labour<sup>73</sup>.

Once attention focusses on dispersal events, along with recognition that such events may cause collective reproduction, then attention turns to emergence of simple primordial life cycles. Those involving more than a single phase, for example those involving soma- and germline-like phases, establish, by virtue of the life cycle, a second timescale over which selection acts<sup>19</sup>. A significant aspect of such life cycles is that birth–death events depend on the efficacy of developmental processes that become the focus of selection. Indeed, the evolution of life cycles is intimately connected to the transition to multicellular life<sup>9,13,74</sup>. Even in our simple conceptual model, the moment that S cells have a selective advantage, then a rudimentary life cycle manifests, with selection able to set the developmental programme via effects on  $q$ , the rate at which S cells are produced. Arguably this marks an early step in the process of endogenization: the process by which externally imposed Darwinian-like properties become integral features of the new entity<sup>3,14</sup>. Another possibility would be for S cells to provision new environments with resources, as does the cotyledon in a plant seed, freeing evolving collectives from dependence on patchily distributed resources. In the *Pseudomonas* experiment, as generations of selection have passed, dependence of the evolving lineages on the scaffold has lessened. This is especially noticeable via evolution of a simple developmental programme controlling the switch between soma-like and germ-like phases<sup>19</sup>.

So far we have been silent on cooperation and the causes of cooperation that typically feature so prominently in discussions on the evolution of multicellular life<sup>51,75</sup>. There is no doubt that cooperation and integration are basic features of multicellular organisms but the scaffolding perspective does not presuppose cooperation among cells as a necessary first step. Nonetheless behaviours (interactions) that are clearly cooperative stand to evolve given appropriate ecological scaffolds. For example, the slow-growing cells favoured when dispersal time is long (Figs. 3a, 4a and 5a) could be labelled cooperating types because they show restraint in the face of plentiful resources, whereas the fast-growing mutants arising within patches might be termed selfish types, but there is no need to use such labels. Applying these labels brings focus to the individual cell and detracts from the ecological context where causal processes lie.

One behaviour often labelled as an extreme form of cooperation—suicidal altruism—is encapsulated by the S cells. Such behaviours can be seen from the perspective of single cells with the temptation to invoke inclusive fitness<sup>76</sup> but to do so risks overlooking the importance of ecology<sup>77</sup> and population structure<sup>17</sup>. It is the meta population structure determined by ecological circumstances that ensures patches are founded by single cells and this limits within-patch conflict, while also being necessary for the emergence and maintenance of a reproductive division of labour. In our model, within-patch close kinship is a consequence of environmental structure, with both group structure and kinship being particularly conducive to evolutionary transitions in individuality<sup>39,78</sup>.

The concept of ecological scaffolding has a number of applications and implications. In concluding, we mention briefly three areas. The first concerns the emergence of the first self-replicating

chemistries at the moment life emerges from non-living material. Recognition that Darwinian-like properties might emerge from the interplay between chemistry and environment opens the door to conceptual and experimental scenarios whereby chemistries that lack capacity for autonomous replication might begin to transition, through a process of templated production of bioactive compounds, toward a replicative process. Alkaline thermal vents appear to offer as much: these highly compartmentalized structures sit at the interface between hydrogen-rich fluids arising from the flow of heated water across serpentine and acidic ocean waters with the possibility that CO<sub>2</sub> is reduced to biologically active species<sup>79,80</sup>. Ensuing ‘growth’ of products within the porous compartments sets in place the possibility of a replicative process. Incorporation of such ecological structures in future experimental designs may provide Darwinian ingredients that are typically absent from explorations of the chemical origins of life.

The second area of relevance is infectious disease biology. To a pathogen, the eukaryotic host offers a discrete patch of resource. Pathogens that rely on transmission for long-term persistence experience selection over two timescales. Our model leads to the prediction that pathogens that pass through restrictive bottlenecks, such as HIV, are likely to have evolved more complex life histories than currently appreciated, involving, for example, a division of labour. This seems to be true of *Salmonella typhimurium* that switches stochastically between slow-growing virulent and fast-growing avirulent cell types, that invade the lumen or colonize the gut, respectively. Cells that colonize the lumen, express virulence factors and trigger the inflammatory response, which benefits the faster growing avirulent cells in the gut. However, unlike cells in the gut, lumen-colonizing cells are killed by the intestinal innate immune defences and are, like soma, an evolutionary dead-end<sup>81</sup>. A division of labour is also hinted at in the case of HIV and other chronic RNA viruses in humans that appear to escape the deleterious effects of short-sighted within-host evolution over prolonged time intervals. Growing evidence suggests that this may be attributable to establishment of germline-like lineages<sup>73</sup>.

The third example concerns application of ecological scaffolding for in vitro engineering of evolutionary transitions and particularly for top-down engineering of microbial communities in which communities eventually become a single symbiotic entity<sup>21,75,82,83</sup>. In the laboratory environment, and aided particularly by advances in micro/millifluidic technologies, it is a relatively trivial matter to confine populations and/or communities to thousands of discretized droplets that can then be subject to a death–birth process<sup>21,84</sup>. In a forthcoming paper, we detail this process and its outcome for the evolution of interactions that build integrated communities<sup>85</sup>.

## Methods

**Single phenotype model.** We construct the simplest possible model to demonstrate the dynamics of two nested Darwinian populations. We assume a fixed population of patches each provisioned with a fixed amount of resource that is consumed by the cells to divide. The dynamics within each patch are independent with cells undergoing a birth–death process for a fixed length of time,  $T$ , after which a dispersal event occurs leading to the colonization of a new generation of patches (with replenished resources).

We first describe the basic birth–death dynamics within a patch (with no mutations). Cells have a mean life time that, without loss of generality, is set to 1. We assume homogeneous mixing within the patch and hence mass-action like dynamics governing the rate of reproduction of cells, which each divide at a rate  $\beta$  multiplied by the proportion of resource in the patch. Under these dynamics, the number of cells and amount of resource within the patch follows modified Lotka–Volterra equations<sup>86,87</sup>, where there is no replenishment of the resource, that is

$$\begin{aligned}\frac{dx}{dt} &= \frac{\beta x}{N} - x \\ \frac{dy}{dt} &= -\frac{\beta xy}{N}\end{aligned}\quad (1)$$

where  $x(t)$  and  $y(t)$  are the number of cells and amount of resource at time  $t$  respectively and  $N$  is the initial amount of resource in the patch. The initial conditions for these equations are  $x(0) = 1$ , representing the initial founding cell, and  $y(0) = N$ . The population initially grows exponentially with the resource



being depleted at the same rate. At some point, the resource becomes significantly depleted leading to a reduction in growth rate. The population peaks once the rate of growth matches the rate that cells die and the population of cells declines (the trajectories shown in Fig. 1 are examples of these dynamics). The total number of possible cell divisions and hence the peak population size is limited by the initial amount of resource in the patch; high growth rates lead to large populations that peak early but then also decrease quickly.

In the simplest version of the model, we fix  $N=10^6$  for all patches. To investigate how environmental predictability affects the dynamics we can introduce variability into the within-patch dynamics in two ways. We can let  $N \sim \text{norm}(10^6, \sigma_N)$ ; that is, the initial amount of resource is sampled from a normal distribution independently for each patch in a generation. For the formulation of the model described above, where the growth rate is scaled by  $N^{-1}$ , this amounts to assuming that the patch volume scales with the initial resources, so the concentration remains fixed. Alternatively we can introduce variability into the initial condition,  $y(0) \sim \text{norm}(N, \sigma_y)$ , while keeping  $N$  fixed at  $10^6$ , which then introduces variability in the initial concentration and hence the initial growth rate.

The possibility of mutations is added to the basic model, which is assumed to only affect the growth rate of cells. For simplicity, possible growth rates are discretized (with step size  $\mu$ ) so that the model tracks the populations of each type with a particular growth rate. As before, the mean life-span of all types is fixed at 1. The single-cell bottleneck and limit to the amount of growth within a patch imposes a strong homogeneity on the composition of a patch. This is because mutants of the founding cell type cannot reach appreciable numbers within a single growth phase and mutants produced by mutants are correspondingly much rarer and can be ignored to a first approximation. Thus in our model we limit the number of mutant types tracked to the first two, one step higher and lower than the founders growth rate (hence each patch has at most three strains within it). If the founding cell has growth rate  $\beta_0$ , the mutant strains will have growth rates  $\beta_0 \pm \mu$ .

To model the dynamics of this expanded system we adopt a piecewise-deterministic approach, where the times of the introduction of new types (with different growth rates) via mutations are modelled stochastically but the growth dynamics between these times are modelled deterministically<sup>88</sup>. At each division event we assume a probability,  $p$ , of creating a child cell with a different growth rate, and hence with probability  $1-p$  a cell of the same type is produced. The times at which mutants are introduced can then be stochastically simulated as outlined in the Supplementary Information.

Between the introduction of new types, the numbers of each type already in the patch and the amount of resource evolves via a set of coupled ordinary differential equations

$$\begin{aligned} \frac{dx_i}{dt} &= (N^{-1}\beta_i y - 1)x_i, \quad i = 0, \dots, m(t) \\ \frac{dy}{dt} &= -yN^{-1} \sum_{i=0}^m \beta_i x_i \end{aligned} \quad (2)$$

where  $\beta_i$  is the growth rate of the  $i$ th cell type and  $m(t)$  is the number mutant types currently in the patch (so if  $m(t)=0$ , only the initial colonizing type is present). This is equivalent to Lotka–Volterra dynamics with competition but where resources are not replenished. An example trajectory from this model is shown in Fig. 2. The use of a piecewise-deterministic model, where times of new mutants arising are stochastic but the growth dynamics are otherwise deterministic is a pragmatic compromise between computational efficiency and realism but ignores other stochastic effects, such as the time for the cells to grow to an appreciable number before exponential growth is fully underway. Full stochastic models<sup>91</sup> can account for this and display a distribution of patch sizes with a much larger variance similar to simply including variance in concentration of resource as done in this paper. Forthcoming work shows that the evolutionary dynamics are similar (manuscript in preparation).

Simulation of the full model over multiple generations proceeds as follows. Each of the  $M$  patches is seeded with a single cell, with growth rates determined from the previous dispersal step. In the first model, all patches experience the same fixed growth time until dispersal,  $T$ . We can relax this by allowing growth periods for individual patches within a generation to be sampled from a normal distribution,  $\tau \sim \text{norm}(T, \sigma_\tau)$ , where  $\sigma_\tau$  is the variance and  $T$  is now the mean. As dispersal is taken to occur simultaneously across the patches, this variance in growth time is taken to arise in the time for the cells to be first deposited in the patches after dispersal.

For the first model, the dispersal dynamics only depend on the numbers of cells within each patch. A new generation of patches is founded by randomly selecting a patch in proportion to the number of cells within it and then randomly selecting a cell, within the chosen patch, again in proportion to its number within the patch. This procedure is equivalent to simply picking particles randomly from the whole population of patches. The larger the number of a given type within the population, the more likely it is to be dispersed. This two-step procedure is simulated for a given number of generations and quantities, such as the average growth rate within a generation, can be calculated. Because the model is mechanistic, it is possible to track the genealogies of both the cells and patches, as shown in Extended Data Fig. 1.

**Sterile cell model.** We take a similar approach to simulating this version of the model, using a piecewise-deterministic approximation where the times of the

introduction of new types due to mutations are stochastic and the birth–death dynamics are deterministic. As the S cells only interact with the G cells via competition for resources and influence dispersal when they have reached an appreciable number, the growth of the S types is assumed deterministic rather than stochastic. Hence between mutations, the dynamics evolve as

$$\begin{aligned} \frac{dx_i}{dt} &= N^{-1}\beta_i(1 - q_i)yx_i - x_i, \quad i = 0, \dots, m(t) \\ \frac{dy}{dt} &= -yN^{-1} \sum_{i=0}^m \beta_i x_i - dyz \\ \frac{dz}{dt} &= yN^{-1} \sum_{i=0}^m q_i \beta_i x_i - z \end{aligned} \quad (3)$$

where  $z(t)$  is the number of S cells in the patch. The initial conditions will be  $x(0) = (1, 0, 0)$ ,  $y(0) = N$  and  $z(0) = 0$ . This model introduces two new parameters:  $q_0$ , which is the per event probability of producing an S cell and  $d$ , which is the rate at which S cells consume the resource. The parameter  $d$  remains fixed but  $q_i$  is subject to mutation. As the phenotype space is now two-dimensional, the scheme for generating mutants is different from the first model but the number of new mutants remains limited to the first two produced by the founding type. We assume that only G cells can be dispersed; hence, the bottleneck enforced by the dispersal mechanism means a patch is always seeded from a single G type cell. More details of the simulation procedure and mutational process are given in the Supplementary Information.

For dispersal we assume the system to be composed of  $k = 1, \dots, K$  patches. Then let  $x_i^{(k)}$  be the number of type G, and  $z^{(k)}$  be the number of type S in patch  $k$ . Each patch is founded by a single G cell with phenotype  $(\beta, q)^{(k)}$ . These cells reproduce, mutate and create S cells until the dispersal time,  $T$ , at which point a sample is taken from the resulting populations to create the next generation. This occurs in two steps:

- (1) Randomly select a patch  $k$ , in proportion to its weight,  $w_k$ , which is a function of the proportion of its constituents,  $x_i^{(k)}$   $i=0, 1, 2$  and  $z^{(k)}$ .
- (2) From the patch selected in step (1), randomly select a G cell from the total patch population.

To assign the weight to patches in step (1), we use the function

$$w_k = \left(1 + \rho z^{(k)}\right) \sum_{i=0}^m x_i^{(k)} \quad (4)$$

where  $\rho$  is the additional benefit to the dispersal process per S cell in the patch. This can be interpreted as the S cells aiding dispersal from the patch, for example by attracting the dispersal agent. If  $\rho=0$ , then the dispersal process is as in the first model; that is, the probability of choosing a patch is proportional only to the number of G cells, so patches with more G cells at the time of dispersal are more likely to be sampled.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All data are available within the manuscript.

## Code availability

Simulation codes for the models presented in this paper are available at the GitHub repository <https://github.com/andxblack/eco-scaff-paper-code> under an MIT licence.

Received: 7 June 2019; Accepted: 17 December 2019;

Published online: 10 February 2020

## References

1. Buss, L. W. *The Evolution of Individuality* (Princeton Univ. Press, 1987).
2. Maynard Smith, J. & Szathmáry, E. *The Major Transitions in Evolution* (Freeman, 1995).
3. Okasha, S. *Evolution and the Levels of Selection* (Oxford Univ. Press, 2006).
4. Michod, R. E. *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality* (Princeton Univ. Press, 1999).
5. Godfrey-Smith, P. *Darwinian Populations and Natural Selection* (Oxford Univ. Press, 2009).
6. Hull, D. L. Individuality and selection. *Annu. Rev. Ecol. Syst.* **11**, 311–332 (1980).
7. Rainey, P. B. & Kerr, B. Cheats as first propagules: a new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *Bioessays* **32**, 872–880 (2010).
8. Libby, E. & Rainey, P. B. A conceptual framework for the evolutionary origins of multicellularity. *Phys. Biol.* **10**, 035001 (2013).
9. De Monte, S. & Rainey, P. B. Nascent multicellular life and the emergence of individuality. *J. Biosci.* **39**, 237–248 (2014).

10. Lewontin, R. C. The units of selection. *Annu. Rev. Ecol. Syst.* **1**, 1–18 (1970).
11. Griesemer, J. The units of evolutionary transition. *Selection* **1**, 67–80 (2000).
12. Bourrat, P. From survivors to replicators: evolution by natural selection revisited. *Biol. Phil.* **29**, 517–538 (2014).
13. Rainey, P. B. & De Monte, S. Resolving conflicts during the evolutionary transition to multicellular life. *Annu. Rev. Ecol. Syst.* **45**, 599–620 (2014).
14. Okasha, S. The strategy of endogenization in evolutionary biology. *Synthese* <https://doi.org/10.1007/s11229-018-1832-6> (2018).
15. Hanschen, E. R. et al. The *Gonium pectorale* genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. *Nat. Commun.* **7**, 11370 (2016).
16. Ratcliff, W. C., Denison, R. F., Borrello, M. & Travisano, M. Experimental evolution of multicellularity. *Proc. Natl Acad. Sci. USA* **109**, 1595–1600 (2012).
17. Wade, M. J. *Adaptation in Metapopulations: How Interactions Change Evolution* (Univ. Chicago Press, 2016).
18. Caporael, L. R., Griesemer, J. R. & Wimsatt, W. C. (eds) *Developing Scaffolds in Evolution, Culture, and Cognition* (Vienna Series in Theoretical Biology, MIT Press, 2014).
19. Hammerschmidt, K., Rose, C., Kerr, B. & Rainey, P. B. Life cycles, fitness decoupling and the evolution of multicellularity. *Nature* **515**, 75–79 (2014).
20. Rose, C. J., Hammerschmidt, K., Pichugin, Y. & Rainey, P. B. Meta-population structure and the evolutionary transition to multicellularity. Preprint at *bioRxiv* <https://doi.org/10.1101/407163> (2019).
21. Rainey, P. B., Remigi, P., Farr, A. D. & Lind, P. A. Darwin was right: where now for experimental evolution? *Curr. Opin. Genet. Dev.* **47**, 102–109 (2017).
22. Van Valen, L. Energy and evolution. *Evol. Theory* **1**, 179–229 (1976).
23. Wright, S. Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931).
24. Wright, S. Tempo and mode in evolution: a critical review. *Ecology* **26**, 415–419 (1945).
25. Wright, S. *Evolution and the Genetics of Populations: Experimental Results and Evolutionary Deductions* Vol. 3 (Univ. Chicago Press, 1977).
26. Maynard Smith, J. Group selection and kin selection. *Nature* **201**, 1145–1146 (1964).
27. Wilson, D. S. A theory of group selection. *Proc. Natl Acad. Sci. USA* **72**, 143–146 (1975).
28. Goodnight, C. J. & Stevens, L. Experimental studies of group selection: what do they tell us about group selection in nature? *Am. Nat.* **150**, S59–S79 (1997).
29. Traulsen, A. & Nowak, M. A. Evolution of cooperation by multilevel selection. *Proc. Natl Acad. Sci. USA* **103**, 10952–10955 (2006).
30. Simon, B., Fletcher, J. A. & Doebeli, M. Towards a general theory of group selection. *Evolution* **67**, 1561–1572 (2012).
31. Black, A. J. & McKane, A. J. Stochastic formulation of ecological models and their applications. *Trends Ecol. Evol.* **27**, 337–345 (2012).
32. Heisler, I. L. & Damuth, J. A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**, 582–602 (1987).
33. Michod, R. E. & Roze, D. Transitions in individuality. *Proc. R. Soc. B* **264**, 853–857 (1997).
34. Michod, R. E. & Nedelcu, A. M. On the reorganization of fitness during evolutionary transitions in individuality. *Integr. Comp. Biol.* **43**, 64–73 (2003).
35. Clarke, E. The problem of biological individuality. *Biol. Theory* **5**, 312–325 (2010).
36. Bourrat, P. Levels of selection are artefacts of different temporal fitness measures. *Ratio* **28**, 40–50 (2015).
37. Bourrat, P. Levels, time and fitness in evolutionary transitions in individuality. *Philos. Theor. Biol.* **7**, e601 (2015).
38. Michod, R. E. & Roze, D. in *Mathematical and Computational Biology: Computational Morphogenesis, Hierarchical Complexity, and Digital Evolution* (ed. Nehaniv, C.L.) 47–92 (American Mathematical Society, 1999).
39. Birch, J. Kin selection, group selection, and the varieties of population structure. *Br. J. Philos. Sci.* <https://doi.org/10.1093/bjps/axx028> (2018).
40. Libby, E. & Ratcliff, W. C. Ratcheting the evolution of multicellularity. *Science* **346**, 426–427 (2014).
41. Gallie, J. et al. Bistability in a metabolic network underpins the de novo evolution of colony switching in *Pseudomonas fluorescens*. *PLoS Biol.* **13**, e1002109 (2015).
42. Gallie, J. et al. Repeated phenotypic evolution by different genetic routes in *Pseudomonas fluorescens* SBW25. *Mol. Biol. Evol.* **36**, 1071–1085 (2019).
43. Remigi, P. et al. Ribosome provisioning activates a bistable switch coupled to fast exit from stationary phase. *Mol. Biol. Evol.* **36**, 1056–1070 (2019).
44. Calcott, B. & Sterelny, K. (eds) *The Major Transitions in Evolution Revisited* (Vienna Series in Theoretical Biology, MIT Press, 2011).
45. Godfrey-Smith, P. & Kerr, B. Gestalt-switching and the evolutionary transitions. *Br. J. Philos. Sci.* **64**, 205–222 (2013).
46. Clarke, E. Origins of evolutionary transitions. *J. Biosci.* **39**, 1–14 (2014).
47. Shelton, D. E. & Michod, R. E. Group selection and group adaptation during a major evolutionary transition: insights from the evolution of multicellularity in the volvocine algae. *Biol. Theory* **9**, 452–469 (2014).
48. Clarke, E. A levels-of-selection approach to evolutionary individuality. *Biol. Philos.* **31**, 893–911 (2016).
49. Bourrat, P. Evolutionary transitions in individuality: a formal analysis. *Synthese* <https://doi.org/10.1007/s11229-019-02307-5> (2019).
50. Queller, D. C. & Strassmann, J. E. Beyond society: the evolution of organismality. *Phil. Trans. R. Soc. Lond. B* **364**, 3143–3155 (2009).
51. Bourke, A. F. *Principles of Social Evolution* (Oxford Univ. Press, 2011).
52. West, S. A., Fisher, R. M., Gardner, A. & Kiers, E. T. Major evolutionary transitions in individuality. *Proc. Natl Acad. Sci. USA* **112**, 10112–10119 (2015).
53. Boraas, M. E., Seale, D. B. & Boxhorn, J. E. Phagotrophy by a flagellate selects for colonial prey: a possible origin of multicellularity. *Evol. Ecol.* **12**, 153–164 (1998).
54. van Gestel, J. & Tarnita, C. E. On the origin of biological construction, with a focus on multicellularity. *Proc. Natl Acad. Sci. USA* **114**, 11018–11026 (2017).
55. Herron, M. D. et al. De novo origins of multicellularity in response to predation. *Sci. Rep.* **9**, 2328 (2019).
56. Rainey, P. B. Unity from conflict. *Nature* **446**, 616 (2007).
57. Bourrat, P. Evolutionary transitions in heritability and individuality. *Theory Biosci.* **138**, 305–323 (2019).
58. Maliet, O., Shelton, D. E. & Michod, R. E. A model for the origin of group reproduction during the evolutionary transition to multicellularity. *Biol. Lett.* **11**, 20150157 (2015).
59. Griesemer, J. in *Towards a Theory of Development* (eds Minelli, A. & Pradeu, T.) 183–202 (Oxford Univ. Press, 2014).
60. Doebeli, M., Ispolatov, Y. & Simon, B. Towards a mechanistic foundation of evolutionary theory. *eLife* **6**, e23804 (2017).
61. Muller, G. B. & Wagner, G. P. Novelty in evolution—restructuring the concept. *Annu. Rev. Ecol. Syst.* **22**, 229–256 (1991).
62. Pfeiffer, T., Schuster, S. & Bonhoeffer, S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**, 504–507 (2001).
63. Pfeiffer, T. & Bonhoeffer, S. An evolutionary scenario for the transition to undifferentiated multicellularity. *Proc. Natl Acad. Sci. USA* **100**, 1095–1098 (2003).
64. Ewald, P. W. *Evolution of Infectious Disease* (Oxford Univ. Press, 1994).
65. Levin, B. R. & Bull, J. J. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* **2**, 76–81 (1994).
66. Frank, S. A. Models of parasite virulence. *Q. Rev. Biol.* **71**, 37–78 (1996).
67. Frank, S. A. Natural selection. III. Selection versus transmission and the levels of selection. *J. Evol. Biol.* **25**, 227–243 (2012).
68. Lythgoe, K. A., Pellis, L. & Fraser, C. Is HIV short-sighted? insights from a multistrain nested model. *Evolution* **67**, 2769–2782 (2013).
69. Cressler, C. E., McLeod, D. V., Rozins, C., van den Hoogen, J. & Day, T. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology* **143**, 915–930 (2016).
70. Frank, S. A. Demography and the tragedy of the commons. *J. Evol. Biol.* **23**, 32–39 (2010).
71. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans* (Oxford Univ. Press, 1991).
72. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton Univ. Press, 2007).
73. Lythgoe, K. A., Gardner, A., Pybus, O. G. & Grove, J. Short-sighted virus evolution and a germline hypothesis for chronic viral infections. *Trends Microbiol.* **25**, 336–348 (2017).
74. Staps, M., van Gestel, J. & Tarnita, C. E. Emergence of diverse life cycles and life histories at the origin of multicellularity. *Nat. Ecol. Evol.* **3**, 1197 (2019).
75. Wilson, D. S. & Sober, E. Reviving the superorganism. *J. Theor. Biol.* **136**, 337–356 (1989).
76. Hamilton, W. D. The genetical evolution of social behavior, I & II. *J. Theor. Biol.* **7**, 1–52 (1964).
77. Wilson, E. O. & Holldobler, B. Eusociality: origin and consequences. *Proc. Natl Acad. Sci. USA* **102**, 13367–13371 (2005).
78. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22**, 659–671 (2009).
79. Martin, W. & Russell, M. J. On the origin of biochemistry at an alkaline hydrothermal vent. *Phil. Trans. R. Soc. Lond. B* **362**, 1887–1925 (2007).
80. Lane, N. *The Vital Question: Why is Life the Way it is?* (Profile Books, 2015).
81. Diard, M. et al. Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature* **494**, 353–356 (2013).
82. Swenson, W., Wilson, D. S. & Elias, R. Artificial ecosystem selection. *Proc. Natl Acad. Sci. USA* **97**, 9110–9114 (2000).
83. Xie, L., Yuan, A. E. & Shou, W. Y. Simulations reveal challenges to artificial community selection and possible strategies for success. *PLoS Biol.* **17**, e3000295 (2019).
84. Cottinet, D. et al. Lineage tracking for probing heritable phenotypes at single-cell resolution. *PLoS ONE* **11**, e0152395 (2016).
85. Doulier, G., Lambert, A. J., De Monte, S. & Rainey, P. B. Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity. Preprint at *bioRxiv* <https://doi.org/10.1101/827592> (2019).

86. Renshaw, E. *Modelling Biological Populations in Space and Time* (Cambridge Univ. Press, 1993).
87. Brauer, F. & Castillo-Chávez, C. *Mathematical Models in Population Biology and Epidemiology* (Springer, 2001).
88. Davis, M. H. A. Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. *J. R. Stat. Soc. B* **46**, 353–388 (1984).

### Acknowledgements

We thank J. Wietz for critical review of the manuscript and valuable comment. We thank S. De Monte, G. Doucier, M. Diard and members of our respective teams for lively discussion. A.J.B. acknowledges an Australian Research Council (ARC) DECRA fellowship (grant no. DE160100690) and support from both the ARC Centre of Excellence for Mathematical and Statistical Frontiers (CoE ACEMS) and the Australian Government NHMRC Centre for Research Excellence in Policy Relevant Infectious diseases Simulation and Mathematical Modelling (CRE PRISM2). P.B. acknowledges a Macquarie University Research Fellowship and a grant from the John Templeton Foundation (grant no. ID 60811). P.B.R. acknowledges generous financial support from MPG core funding and previously from the Marsden Fund Council from New Zealand Government funding, administered by the Royal Society of New Zealand.

### Author contributions

P.B.R. and P.B. developed the main concepts. A.J.B. contributed the mathematical models and ran simulations to clarify initial ideas. All authors wrote and revised the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41559-019-1086-9>.

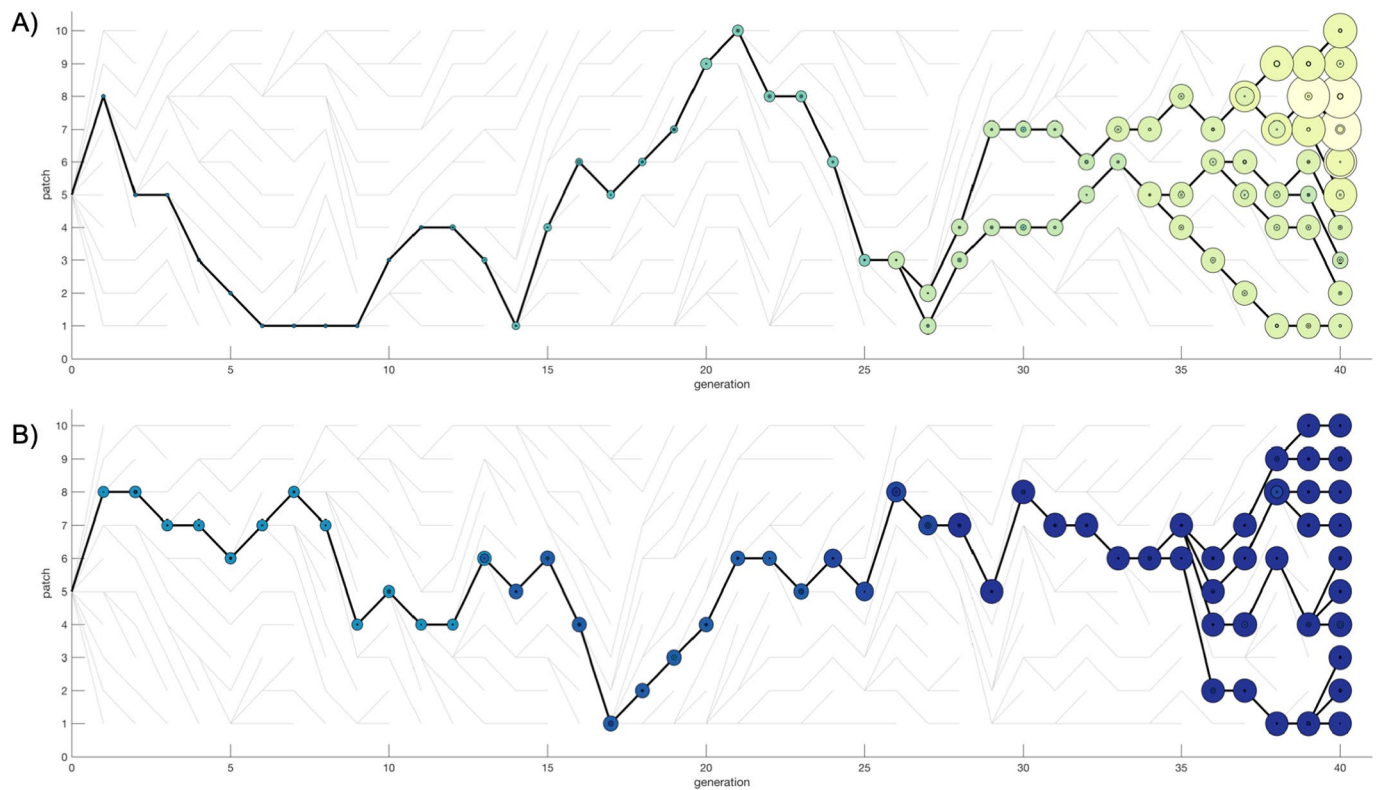
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41559-019-1086-9>.

**Correspondence and requests for materials** should be addressed to A.J.B. or P.B.R.

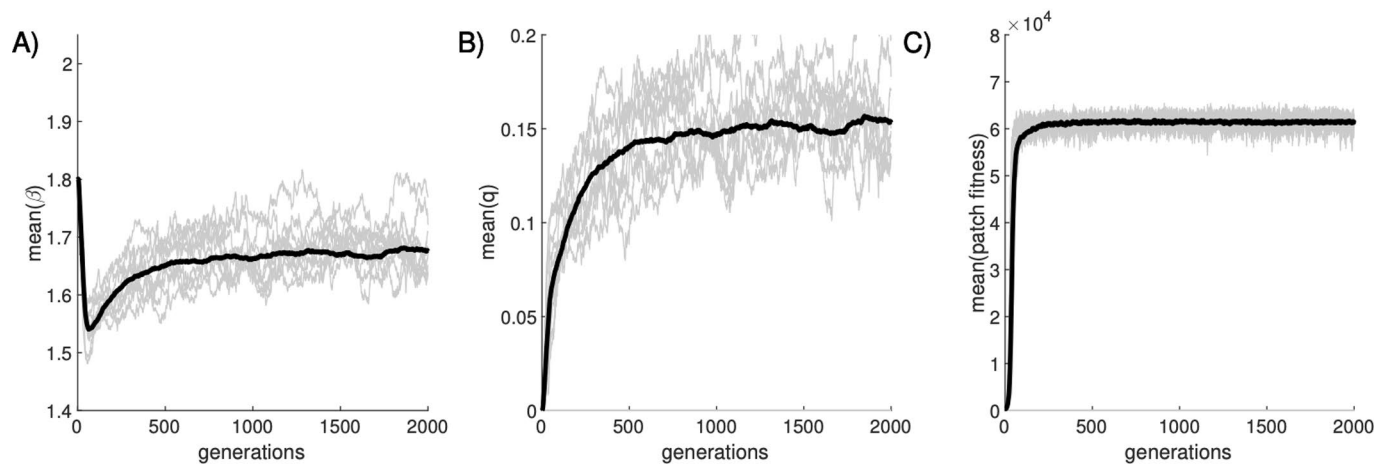
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

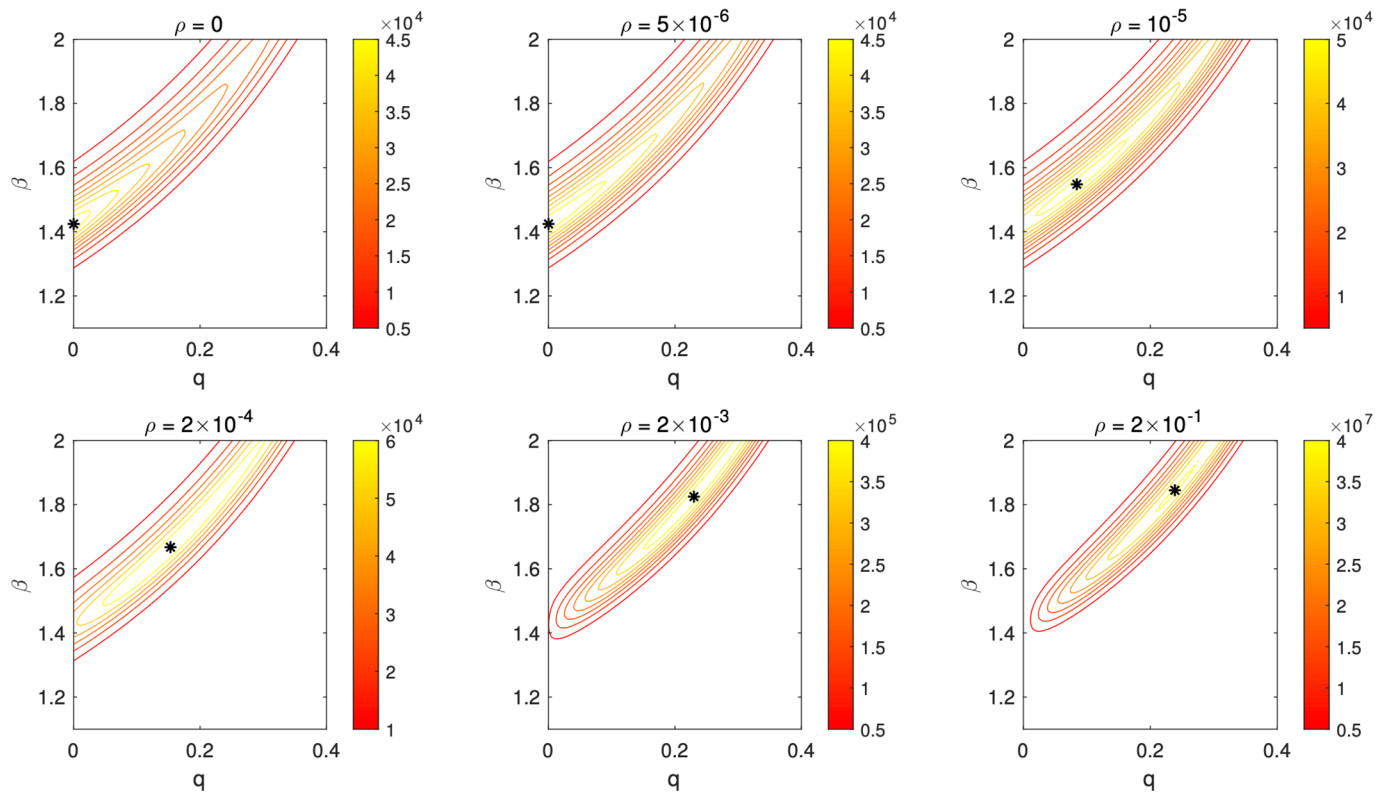
© The Author(s), under exclusive licence to Springer Nature Limited 2020



**Extended Data Fig. 1 | Genealogy of patches under slow (A) and fast (B) dispersal regimes.** The simulations have only 10 patches and modified mutational parameters compared with those in Fig. 2 of the main text. This is to allow a clearer visualization of the process, which otherwise requires many more generations for change to be apparent. Video versions of these are also included in the supplementary material. As in Figs. 4 and 5 of the main text, the cell numbers in each patch are proportional to the area of the circles and the growth rates are indicated by the colours, as shown by the colour bar in Fig. 4 of the main text. The mutational parameters are larger for these simulations ( $\mu=0.05$ ,  $p=0.05$ ) so evolution occurs on a quicker timescale as compared with the results shown in Fig. 3 of the main text.



**Extended Data Fig. 2 | Simulations of the model with sterile types over 2000 generations show convergence to equilibrium.** The comparative slowness of this convergence, as well as the large fluctuations in the mean value of  $q$  for single realizations, can be attributed to the flatness of the fitness landscape about the equilibrium as shown in the Supplementary Information.



**Extended Data Fig. 3 | Change in the position of the equilibrium as a function of the dispersal assistance provided by S cells assistance  $\rho$  for slow dispersal.** Fitness landscapes showing the effect of changes in the assistance,  $\rho$ , given to dispersing G cells by non-dispersing S cells on the equilibrium (\*) relationship between cell growth rate,  $\beta$ , and the rate of production of S cells,  $q$ , under the slow ( $T=30$ ) dispersal regime.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

MATLAB software was used to simulate the mathematical models and analyse the resulting data.

Data analysis

GitHub repository: <https://github.com/andxblack/eco-scaff-paper-code>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data are available within the manuscript.

### Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<input type="text" value="The study is a theoretical analysis of the evolution of individuality"/>
Research sample	<input type="text" value="NA"/>
Sampling strategy	<input type="text" value="NA"/>
Data collection	<input type="text" value="NA"/>
Timing and spatial scale	<input type="text" value="NA"/>
Data exclusions	<input type="text" value="NA"/>
Reproducibility	<input type="text" value="Simulations of the mathematical model can be rerun using code provided."/>
Randomization	<input type="text" value="NA"/>
Blinding	<input type="text" value="NA"/>

Did the study involve field work?  Yes  No

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging