PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Båvik LM, Mehta RS, Weissman DB. 2023 Fifty shades of greenbeard: robust evolution of altruism based on similarity of complex phenotypes. *Proc. R. Soc. B* **290**: 20222579. https://doi.org/10.1098/rspb.2022.2579

Received: 24 December 2022
Accepted: 19 May 2023

Subject Category:
Evolution

Subject Areas:
evolution, theoretical biology, computational biology

Keywords:
altruism, cooperation, evolutionary dynamics, evolutionary game theory, mathematical biology

Author for correspondence:
Linnéa M. Båvik
e-mail: lbavik@emory.edu

†Denotes equal contribution.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6672279.

THE ROYAL SOCIETY PUBLISHING

Fifty shades of greenbeard: robust evolution of altruism based on similarity of complex phenotypes

Linnéa M. Båvik†, Rohan S. Mehta† and Daniel B. Weissman

Department of Physics, Emory University, Atlanta, GA, USA

RSM, 0000-0002-6244-9968

We study the evolution of altruistic behaviour under a model where individuals choose to cooperate by comparing a set of continuous phenotype tags. Individuals play a donation game and only donate to other individuals that are sufficiently similar to themselves in a multidimensional phenotype space. We find the generic maintenance of robust altruism when phenotypes are multidimensional. Selection for altruism is driven by the coevolution of individual strategy and phenotype; altruism levels shape the distribution of individuals in phenotype space. Low donation rates induce a phenotype distribution that renders the population vulnerable to the invasion of altruists, whereas high donation rates prime a population for cheater invasion, resulting in cyclic dynamics that maintain substantial levels of altruism. Altruism is therefore robust to invasion by cheaters in the long term in this model. Furthermore, the shape of the phenotype distribution in high phenotypic dimension allows altruists to better resist the invasion by cheaters, and as a result the amount of donation increases with increasing phenotype dimension. We also generalize previous results in the regime of weak selection to two competing strategies in continuous phenotype space, and show that success under weak selection is crucial to success under strong selection in our model. Our results support the viability of a simple similarity-based mechanism for altruism in a well-mixed population.

1. Introduction

An altruistic trait does not benefit the carrier of the trait but benefits other individuals with which the carrier interacts. Altruism can succeed under a wide variety of specific conditions, but the spectre of a cheating individual who fails to reciprocate is constantly looming, and long-term maintenance of altruism requires some sort of mechanism to prevent cheaters from winning. This tension between altruism and cheating is present in many behavioural strategies in the natural world, including in alarm calls [1], the sharing of secondary metabolites [2], eusociality [3], slime mould life cycles [4], bacterial mat formation [5] and the major transitions of evolution [6].

Through decades of research, a general rule has emerged [7–13]: altruism can be present in the long-term average state of the population if the interaction structure between individuals sufficiently disproportionately confers cooperative benefits upon cooperators. Potential mechanisms of this interaction structure include interacting with relatives [9,14], repeatedly interacting with the same individuals [15], playing the game multiple times in a row [16], and interacting with individuals that are nearby in space [17,18]. These mechanisms are by no means mutually exclusive; many of them can be in operation at the same time, and they can interact with each other in non-trivial ways [19–21].

One class of interactions that can lead to altruistic success is interaction with phenotypically similar individuals, a specific form of which is called the 'greenbeard' effect [22]. This effect requires a gene that codes both for a particular distinctive 'tag' trait—such as a green beard—as well as the tendency to cooperate with others who have that trait. If such a gene were to exist, then

the effects of the altruistic behaviour would be completely distributed to altruists, and altruism can succeed. Greenbeard models generally have two limitations. First, it seems unlikely that there would exist a gene that encodes for both the tag trait and the behavioural trait, though some examples have been found in nature [e.g. 23–26]. Secondly, and more importantly, greenbeards are highly susceptible to invasion by cheaters who have the tag trait but do not themselves cooperate [27]; once these individuals are introduced to the population by mutation, it is unclear how the cooperative behaviour would survive in the long term.

One solution to these problems with greenbeards is to decouple the tag and the behaviour. A simple mechanism proposed by Riolo *et al*. [28] allowed individuals to choose to cooperate based on how similar they were to each other with respect to a one-dimensional, continuous tag. This tag evolves separately from the altruistic strategy that determines the minimum similarity required to cooperate. Their model demonstrated the ability of such a system to maintain altruistic behaviour over the long term by exhibiting a periodic co-invasion dynamic of more cooperative individuals (with larger minimum similarity thresholds) being invaded by less cooperative individuals (with smaller minimum similarity thresholds) which are then eventually re-invaded by more cooperative individuals based around a different set of tags. These cycles of altruistic behaviour have been termed 'tides of tolerance' [29].

Controversy around the particulars of Riolo *et al.*'s [28] model [30] led to the development of a minimal two-tag, two-strategy model by Traulsen & Schuster [31]. This and subsequent models [32,33] have confirmed the existence of this 'tides of tolerance' behaviour while providing restricted conditions for when to expect it—including an upward mutational bias towards altruism or under a narrow range of mutation rate values.

Contrary to the traditional view of the greenbeard effect as associated with a single phenotypic trait or tag, we note that organisms in nature simultaneously observe many traits in conspecifics, and use those observations to inform their interactions. For example, many organisms use chemical signalling to identify interaction partners. In this case, each different compound used for signalling can be considered a phenotype dimension, with a direct mapping between an organism's chemical profile and a multidimensional phenotype. Kin recognition in *Hymenoptera* involves comparisons of complex cuticular hydrocarbon profiles [34]. The characterization of these hydrocarbon profiles is still incomplete, but these interactions can involve up to 10–100 different hydrocarbons [35]. Signalling between plants takes place in an even larger space, with thousands of possible volatile organic compounds in use [36–38]. Other organisms have less well-characterized discriminatory mechanisms, but there is evidence for complex chemical-based interactions in birds [39] and plant-fungal symbioses [40], for example. In addition, organisms use complex non-chemical phenotypes such as vocalization to determine interaction partners (e.g. primates [41] and birds [42]). These complex phenotypes are clearly 'multidimensional', but explicitly characterizing the dimensionality is non-trivial.

On smaller scales, interaction choice is often determined by molecular affinity to cell surface receptors. Phenotype dimensionality can occur in two different ways in these cases. First, the amino acid sequence of the proteins involved occupies a hypercube of dimension equal to the length of the sequence (usually on the order of 10–100). A multidimensional phenotype would involve either the binding affinity of a protein to various potential polymorphic receptors or the binding affinity of a receptor to various polymorphic proteins. Many bacteria use these kinds of systems [43]. Examples of these types of interactions among polymorphic proteins in bacteria include two-component quorum sensing systems in Gram-positive bacteria [44–46], the *TibA* adhesin system in *Escherichia coli* [47], the contact-dependent inhibition (CDI) systems in many bacteria [48,49], and pilus-induced aggregation in *Vibrio cholerae* [50]. Second, these molecular interactions can actually involve multiple different protein-receptor pathways (e.g. two in quorum sensing in *Vibrio cholerae* and three in *Pseudomonas aeruginosa* [51]), which adds additional phenotype dimensionality. Other non-quorum-sensing examples of multigene interaction include two to five genes involved in kin recognition in *Proteus mirabilis* [52] and in outer membrane exchange in Myxobacteria, where two different systems exist [53]. These examples suggest that interaction across multiple phenotype dimensions is a widespread phenomenon across the tree of life and should be considered in models involving phenotype-based behavioural interaction such as this.

Here, we analyse a model in the spirit of [28,31–33] but taking into account the fact that individuals can recognize multidimensional continuous tags—or phenotypes. Using this model, we demonstrate long-term maintenance of altruistic behaviour that is robust to invasion by cheaters over a wide range of parameter values. Our model is similar to that of Spector & Klein [54], which found an increase in donation with increasing phenotype dimension; crucially, however, our model differs from theirs in that we assume a well-mixed population, isolating the effects of phenotypic space from those of physical space. The robustness of altruism in our model is fundamentally driven by the shape of the population phenotype distribution in multidimensional phenotype space. In particular, in high-dimensional phenotype space, it is extremely rare for individuals to be similar across every phenotype dimension unless they share recent common ancestry. This phenotype distribution makes it unlikely for non-altruistic individuals to be born similar enough to more than a small number of altruistic individuals off of which to cheat and thereby dominate. In addition, we extend the weak-selection results of [55,56] to two competing strategies with non-zero, continuous thresholds, and we demonstrate that altruistic success in this weak selection setting provides a sufficient condition for altruistic success in our strong selection setting. Our results suggest that the high dimensionality of phenotype space is sufficient to select for similarity-based altruism in the absence of additional mechanisms such as mutation bias, spatial structure or reciprocation.

2. Model

We first describe our model heuristically. Individuals have two evolving traits: a phenotype that is observable to other members of the population, and an unobservable strategy, illustrated in figure 1. Strategy is wholly determined by an internal 'threshold' value. A focal individual observes another individual's phenotype and determines the difference between their own and the other individual's coordinates in

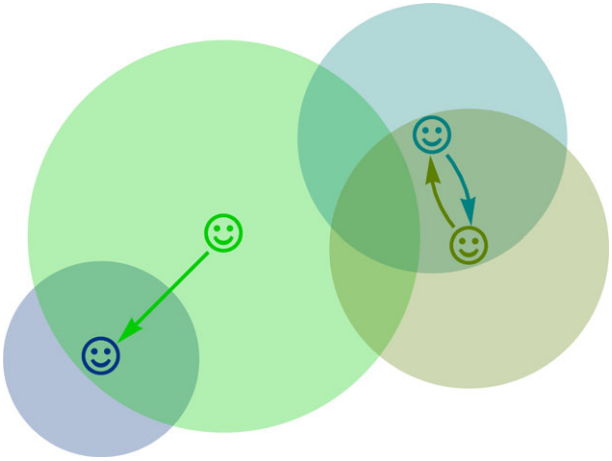

Figure 1. Schematic of the model for similarity-based altruism. Individuals (smiley faces) are shown as points in an abstract two-dimensional space that represents the components of their phenotype visible to other members of the population. Each individual also has an invisible strategy trait, a threshold for altruism, represented by a circle around them. They donate to all individuals whose phenotypes lie within their circle, signified by arrows pointing from donors to recipients.

phenotype space. They then compare this difference to their internal threshold, depicted in figure 1 as the radius of the circle that extends around each individual. This focal individual now has a choice: produce a small fitness benefit for themselves, or help the other individual to produce a bigger benefit for the other individual instead (i.e. perform the altruistic behaviour). If the other individual's phenotype lies within the focal individual's altruism threshold, the focal individual performs the altruistic behaviour. If not, then the focal individual instead chooses to increase its own fitness. Interactions need not be two-way. An individual with a larger threshold may donate to an individual with a smaller threshold, and thus receive no reciprocation for their efforts; this event is our model's version of 'cheating'. Thus, the threshold is the inherited quantity that determines the population's altruistic tendencies. The results of multiple interactions per individual per generation then determine the cumulative payoffs of each individual, which is directly proportional to the number of offspring they will have in the next generation. Offspring inherit their parent's threshold identically most of the time, with a small chance of mutation, reflecting the assumption that this strategy is a function of a small number of loci. Offspring inherit their parent's phenotype with Gaussian noise added to each coordinate, reflecting the assumption that the observable phenotype is a function of a large number of loci and can be treated as a quantitative trait.

We now describe our model in detail. The population is well-mixed and consists of N individuals, where each individual i has a phenotype $\Phi_i \in \mathbb{R}^D$—a point in a D-dimensional continuous phenotype space that can be observed by other individuals in the population—and an unobservable strategy—called its threshold $T_i \in \mathbb{R}_0^+$—which determines whether or not it will donate to another individual. Each generation consists of two steps, interaction and reproduction.

During the interaction step, every individual chooses whether or not to donate to ten other individuals chosen uniformly at random from the population. An individual i will choose to donate if the potential recipient j has a phenotype Φ_j that is sufficiently similar, i.e. the distance between the two individuals in phenotype space is less than T_i: $\|\Phi_i - \Phi_j\| \leq T_i$. The sum of the payoffs from each of these interactions yields P_i, the total payoff for individual i. If the distance between the individual and a potential recipient is less than the individual's similarity threshold, the individual increases the recipient's payoff by the donation value b. If the potential recipient is outside of the individual's similarity threshold, the individual declines to donate, instead keeping their resource for themselves, thereby increasing their own payoff value by an increment of c, where $c < b$. The fitness of each individual is proportional to the total payoff acquired by these donation events as outlined in the subsequent reproduction step. Note that $P_i \geq 0$ in all cases, as the lowest payoff any individual can get from any given interaction is 0 (which occurs when they donate but are not donated to: they do not get the donation b and they do not get to keep the resource c for themselves). For large s, selection is strong, and for $s \to 0$, we obtain the weak selection limit (see electronic supplementary material). In the main text, we set $s = 1$ for simplicity so that the difference in weight is equal to the difference in payoff.

During the reproduction step, we assume uniparental reproduction, with offspring drawn from a multinomial distribution in which the probability W_i of having parent i is determined by the parent's payoff P_i

$$W_i = \frac{1 + sP_i}{\sum_{j=1}^{N}(1 + sP_j)} \tag{2.1}$$

Offspring inherit their parent's phenotype and strategy, plus slight deviations from additive mutations. This mutation scheme could describe genetic inheritance in an asexually reproducing population, or cultural inheritance from one parent or other individual in the parental generation. Phenotype mutations occur with every reproduction. For each dimension of the phenotype, mutations are independently drawn from a standard normal distribution $\mathcal{N}(0, 1)$ so that the phenotype in the jth dimension of individual i is $\Phi_{j,i} + \mathcal{N}(0, 1)$. In other words, we measure phenotypes on the natural scale set by mutation. Strategy mutations occur with probability $u \ll 1$ per reproduction and are drawn from a truncated normal distribution $T\mathcal{N}(0, \sigma_S, -T_i, \infty)$ so that the threshold of the offspring of individual i, given a strategy mutation, is $T_i + T\mathcal{N}(0, \sigma_S, -T_i, \infty)$. We use a truncated distribution to avoid unbiological negative thresholds. The truncation creates a small upward mutational bias on the order of σ_S when the threshold T_i approaches 0, so we interpret thresholds on the order of σ_S to be non-cooperative.

3. Results

(a) High phenotypic complexity leads to the evolution of frequent altruism

We find that for low phenotypic dimension $D \lesssim 10$, altruism is vulnerable to cheater invasion and donations are on average rare. But for high phenotypic dimension $D \gtrsim 10$, altruism consistently evolves and donations are maintained at a substantial rate of about 10–25%, so that in each generation, an individual donates to about one to two of the ten others they interact with (figure 2). Further, we find for each particular phenotype dimension D, the average threshold evolves to a narrow evolutionarily stable range of non-zero values (figure 3). For lower-dimensional phenotypes, there are regular oscillations within this range, but for higher-dimensional phenotypes these

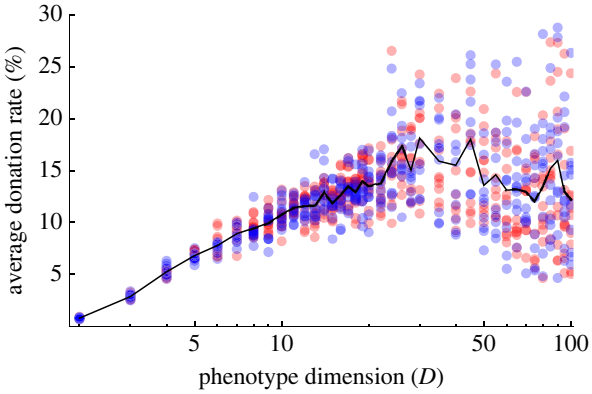

Figure 2. High-dimensional phenotypes lead to the evolution of altruism. Donation rate is the fraction of possible donations that occur. Each point represents the average donation rate of the last 250 000 generations of a single simulation, sampled every 100 generations. For each simulated dimension, two different initial conditions were run for 10 iterations each, for a total of 20 simulations per dimension. Red points show simulations initialized with smaller thresholds: 75% of the predicted 'cohesion threshold' shown in figure 4. Blue points show simulations initialized at a larger threshold: 125% of the cohesion threshold. The black line indicates the average across all simulations with the same phenotype dimension. Phenotype mutation size, strategy mutation rate, population size and other parameters were fixed according to the specifications in electronic supplementary material, table S1.

disappear (figure 3). This pattern is robust to changes in population size, payoff values, mutation rate, initial conditions and phenotype dimension, although the specific value of the evolutionarily stable threshold is a function of phenotype dimension and population size (figure 3; electronic supplementary material).

(b) Altruism shapes the distribution of phenotypes

By simulating populations with no strategy mutation, where each individual has the same fixed threshold, we find that there is a characteristic 'cohesion threshold' at which phenotypic diversity collapses (figure 4). For populations fixed at this threshold, there is strong selection against individuals who are too phenotypically different from the rest of the population and therefore receive fewer donations. This causes the population to cohere in a small region of phenotype space (figure 4*a*).

Figure 4*b* illustrates that for thresholds smaller than the cohesion threshold, donation is very rare, and so there is little variation in individual fitness. Thus, the phenotypic variation is nearly neutral, with a variance predicted by coalescent theory (see the dashed horizontal line in figure 4*a*; electronic supplementary material. When thresholds are far above the cohesion threshold, donation is universal and phenotypic variation is again neutral. But at and somewhat above the cohesion threshold, there is an intermediate amount of donation, creating a large variance in fitness sufficient to counteract mutation pressure and greatly reduce phenotypic variation. While the cohesion threshold appears for all dimensions, donation is much more prevalent at the cohesion threshold for high-dimensional phenotypes than for low-dimensional ones (figure 4*b*).

For high-dimensional phenotypes, the cohesion threshold increases with increasing population size (electronic supplementary material). This effect appears to be less pronounced as the population size increases, suggesting that there is an infinite-population-size asymptotic value to the cohesion threshold.

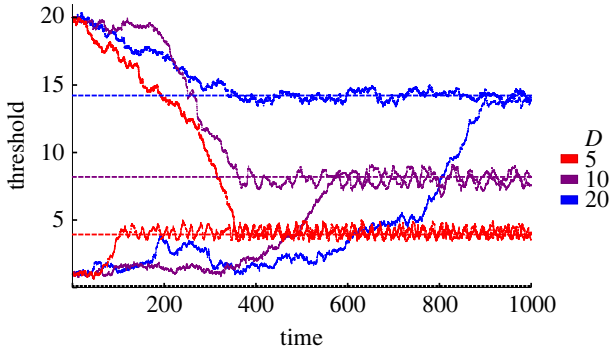

Figure 3. Altruism thresholds evolve to high values that increase with phenotype dimension, indicating the maintenance of altruism in the long term. Each solid trajectory is the average threshold of a single simulation with the same mutation parameters and population size as figure 2. Each line shows a single simulation, with the two lines for each colour starting from two different initial conditions. Dashed lines show the cohesion thresholds as identified in figure 4. The dashed black line just above the horizontal axis shows σ_S, the characteristic threshold scale in non-altruistic populations. Parameters used in these simulations are in electronic supplementary material, table S1. Time is in units of $N = 1000$ generations of the simulation.

(c) Coevolution of strategies and phenotypes maintains altruism

Remarkably, it appears that the two independently identified characteristic thresholds—the evolutionarily stable threshold and the cohesion threshold—are the same (figures 3 and 5). In other words, populations evolve to just the threshold at which their phenotype diversity is narrowest. In addition, for high-dimensional phenotypes ($D > 20$), there is an additional convergence, as the width of the population diversity in each phenotypic dimension also approaches the same value as the two thresholds (figure 5). As discussed below, this is because the population essentially always stays in the cohesive regime. This approach of the phenotype distribution to the evolved threshold is what gives rise to the substantial frequency of donations seen in figure 2.

In figure 5, we show the standard deviation for an individual phenotypic character, the cohesion threshold, and the evolutionarily stable threshold for each simulation in figure 2 (up to $D = 70$). Note that we use the average value of the standard deviation of all D single phenotype characters in order to eliminate the effect of dimension on the scale of this quantity. We also divide the threshold values by a factor of $\sqrt{D}$ in order to eliminate the effect of dimension on the scale of the thresholds.

To understand why the cohesion threshold is evolutionarily stable, first consider a population dominated by individuals with thresholds smaller than the cohesion threshold. Then donation is rare (figure 4*b*) and only slightly perturbs the distribution of phenotypes, which is close to neutral (figure 4*a*). This is the key assumption underlying the standard weak selection models in evolutionary game theory (see electronic supplementary material). Under weak selection, the larger threshold will succeed if the benefit/cost ratio of donation is greater than some critical value. For our simulations, we have chosen $b/c = 10$, as in Riolo *et al.* [28], which is high enough to produce this selective pressure under most circumstances, causing the population to evolve up towards the cohesion threshold. We conduct a mathematical analysis of this regime in the electronic supplementary material, generalizing the results of [55,56] to continuous strategies.

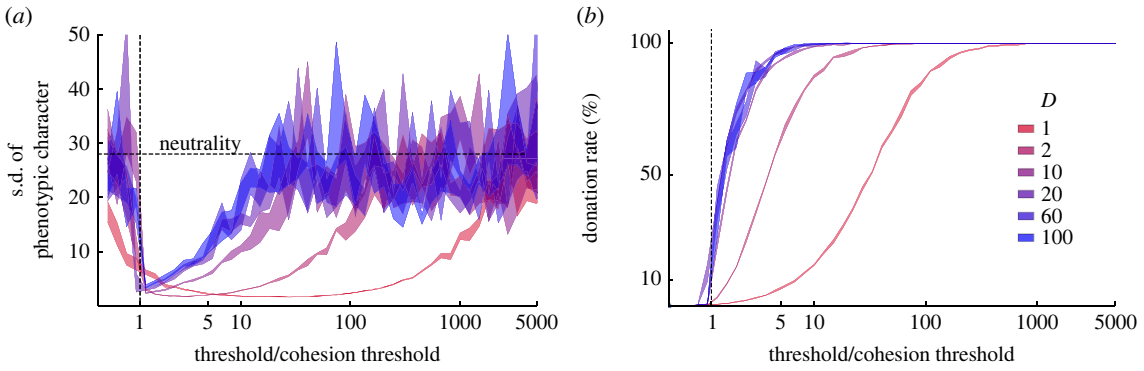

Figure 4. There is a critical 'cohesion threshold' at which altruism greatly reduces equilibrium phenotypic diversity. Curves show the results of simulations in which all individuals have a single fixed threshold, shown on the horizontal axis (scaled by the cohesion threshold). Strategy mutation rate is set to 0, and we start every individual with the same threshold, so that $T_i = T$ for all individuals i. (*a*) The standard deviation of each phenotypic dimension collapses to a small value at the cohesion threshold, and increases gradually as threshold increases beyond the cohesion threshold. Each ribbon curve is centred at the mean of the measured value: the standard deviation of a single phenotype character, with upper and lower bounds indicating one standard deviation of the measured value in each direction. Curves are coloured by the number of phenotype characters (phenotype dimension). For thresholds below and far above the cohesion threshold, the phenotype distribution approaches the distribution expected under neutrality (dashed horizontal line, electronic supplementary material). (*b*) The cohesion threshold is the threshold at which donation starts to become frequent. Each ribbon curve is centred at the mean of the measured value: the per-generation donation rate, with upper and lower bounds indicating one standard deviation of the measured value in each direction. Curves are coloured by the number of phenotype characters (phenotype dimension). Parameters used in these simulations are in electronic supplementary material, table S1.

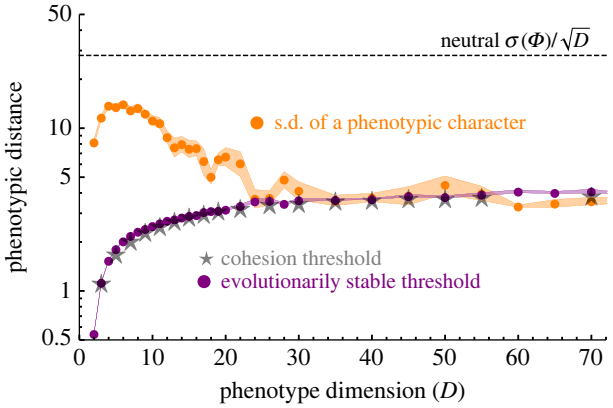

Figure 5. The cohesion threshold (grey) is the evolutionarily stable threshold (purple). For high-dimensional phenotypes, the standard deviation of phenotypic diversity in each phenotypic character (orange) converges to this value as well. Grey stars indicate the value of the cohesion threshold obtained by the simulations with no strategy mutation as seen in figure 4. Purple points indicate the long-term evolutionarily stable threshold obtained by allowing strategy mutation as in figure 3. The purple ribbon bounds plus or minus 1 s.d. of this value obtained over ten iterations each for two initial conditions (the same as in figure 2; 0.75 × the cohesion threshold and 1.25 × the cohesion threshold). Orange points indicate the average standard deviation of a single phenotypic character (as in figure 4) for the same simulations that were used to create the purple dots (i.e. with evolving strategies). The orange ribbon bounds plus or minus 1 s.d. of 10 iterations each for the same two initial conditions. The cohesion threshold and evolutionarily stable threshold values are divided by $\sqrt{D}$ to obtain their corresponding value along a single phenotype dimension, in order to match the scale of the standard deviation of a single phenotypic character. Parameters used in these simulations are in electronic supplementary material, table S1.

Now consider a population where most individuals have thresholds substantially above the cohesion threshold, so that donation is very high—most individuals donate to most individuals (figure 4). In this situation, individuals with smaller threshold values can invade, since most other individuals will still donate to them, driving the threshold down towards the cohesion threshold. This process can only stabilize once there is frequent discrimination, i.e. when thresholds are low enough that many interactions do not result in donation.

(d) High-dimensional phenotypes protect altruists from invasion by cheaters

Figures 2 and 5 show a pattern in which low-dimensional phenotypes produce modest donation rates, moderate phenotypic diversity, and cycling about the cohesion threshold, while high-dimensional phenotypes produce high donation rates, low phenotypic diversity and little cycling. To understand the origins of this pattern, we simulated populations in which individuals mutated between just two discrete thresholds, one below the cohesion threshold (strict) and one above it (generous).

These simulations showed that these different features are the result of a single process: low-dimensional phenotypes allow strict individuals to invade generous populations, because they are tightly clustered in phenotype space (figure 6, time 1). This then causes the population to decohere and spread out in phenotype space (figure 6, time 2), producing the elevated average phenotypic diversity seen in 5. The phenotypic diversity allows generous individuals to re-invade from a sparse region of phenotype space where they avoid strict cheaters (figure 6, time 3). But these generous individuals re-collapse the phenotype distribution, resetting the cycle (figure 6, time 4).

High-dimensional phenotypes, by contrast, do not allow strict individuals to easily invade. This is a direct consequence of their geometry: high dimensions leads to a high surface-to-volume ratio for the phenotype distribution, with essentially all individuals living on the sparse 'edge' of the distribution. While strict individuals can invade locally in phenotype space, in one part of the 'edge', they cannot reliably spread to parasitize the rest of the population, and so quickly become isolated and go extinct. The population thus stays cohesive with frequent donations.

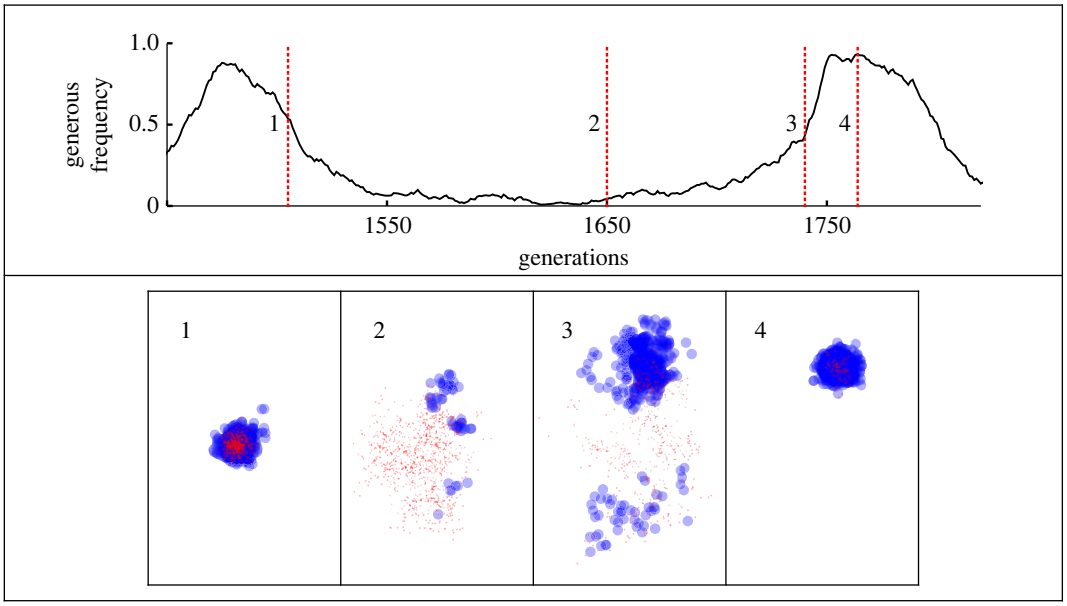

Figure 6. Low-dimensional phenotypes allow non-altruists to invade. Plots show results of simulations of a two-dimensional phenotype over the course of a cycle similar to those seen for $D = 5$ and $D = 10$ in figure 3. For clarity, individuals are restricted to only two possible thresholds, one strict and one generous—corresponding to thresholds below and above the cohesion threshold, respectively. The curve in the top panel represents the frequency of individuals with the larger threshold, or 'generous' individuals. The bottom four panels each correspond to a specific time point in the top panel, indicated by the dashed vertical lines. These four bottom panels plot each individual in two-dimensional phenotype space. Red circles are 'strict' individuals, and blue circles are generous individuals, with the radii of the circles equal to their thresholds. Time point 1: strict individuals have started to invade the cohesive cluster of generous individuals. Time point 2: the strict individuals have successfully invaded, causing the population to disperse in phenotype space. Time point 3: generous individuals from the sparse edge of the phenotype distribution re-invade. Time point 4: the population re-coheres in phenotype space and is again vulnerable to invasion by strict individuals. Parameters used in this simulation are: strict threshold 0.25, generous threshold 1.5, phenotype dimension $D = 2$ and strategy mutation rate 0.0001 (electronic supplementary material, table S1). The critical benefit-cost ratio from the weak selection analysis (electronic supplementary material) is 1.39, which is well below the value 10 used for this figure, for which we would expect the generous individuals to recover after invasion by the strict individuals.

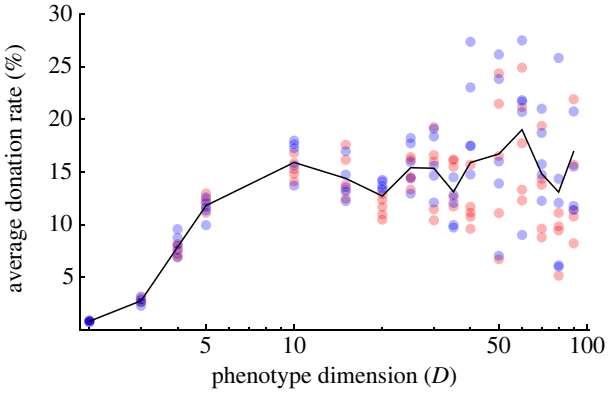

Figure 7. Even in the presence of obligate defectors, high-dimensional phenotypes lead to the evolution of altruism. Each point represents the average donation rate of the last 250 000 generations of a single simulation, sampled every 100 generations. For each simulated dimension, two different initial conditions were run for five iterations each, for a total of 10 simulations. Red points indicate a smaller threshold initialized in the simulation, blue points indicate a larger threshold initialized (75% and 125% of the cohesion threshold value, respectively). The black line connects the average across all simulations with the same phenotype dimension. Parameters used in these simulations are in electronic supplementary material, table S1, with the additional mutation rate into and out of the obligate defector state of 0.001 per generation.

(e) Obligate defectors do not affect long-term donation rates

In our base model, individuals with threshold $T = 0$ only cooperate with individuals of identical phenotype. Because phenotype mutation occurs every generation, the probability of phenotype identity is zero, so these individuals are obligate defectors. However, reaching this state requires a very rare strategy mutation (or series of strategy mutations) to push T all the way to 0. To further investigate the effect of obligate defectors in our model, we therefore conducted additional simulations in which individuals have a second strategy trait, a 'switch'. Individuals with the 'on' state of the switch donate according to their threshold, as in our base model, but individuals with the 'off' state of the switch are obligate defectors and never donate regardless of the value of their threshold trait. We assumed that the switch mutated between the on and off states at rate 0.001 per individual per generation.

Overall donation rates in the presence of obligate defectors (figure 7) match those without obligate defectors (figure 2). Once an individual becomes a obligate defector, selection no longer acts on their threshold parameter, so the strength of selection on strategy decreases and it takes longer for the system to get to a long-term steady state. However, the structure of the population in phenotype space cannot get more dispersed than the neutral/weak selection distribution, so the presence of obligate defectors at most hastens the population's evolution to a phenotype distribution where larger thresholds are favoured. Thus, the presence of obligate defectors does not affect the long-term viability of altruistic strategies in this model, and a higher-dimensional phenotype space has the same effect as it does without obligate defectors.

4. Discussion

Our model shows that high-dimensional phenotypes allow for the evolution of similarity-based altruism by mitigating invasion by cheaters: high-dimensional space is large enough to ensure that phenotypic similarity is an honest

signal of close relatedness. Previous work on the evolution of altruism by phenotypic similarity established relatively stringent conditions on mutation rates, the structure of phenotype space, or the structure of the population in physical space required for altruism to succeed. Our model demonstrates that this success can exist quite generally, independent of the particulars of the mutation rates. In addition, our model shows that with high-dimensional phenotypes, altruism does not need any particular population structure or individual memory or even reciprocal pairwise interactions between individuals to succeed.

It has often been difficult to determine the utility of weak selection models outside of the weak selection limit. Our model demonstrates that there are occasions where such analysis can provide insight outside of that limit. In particular, below the cohesion threshold, the population behaves as if it is under weak selection even though it is actually under strong selection, and we can use the weak selection limit to analyse and successfully describe the evolutionary outcomes.

The phenotypic space describing our simulated individuals is distinct from the physical space in which they would live in two primary ways. First, physical space is limited to no more than three dimensions, whereas phenotype space often has many more dimensions. The high number of dimensions is crucial to the evolution of altruism in our model, so we predict that phenotype-based altruism can be maintained even when space-based altruism cannot. Second, in the functionally neutral phenotype space of the type we consider, there is no local density regulation—individuals do not compete more for resources with phenotypically similar individuals than they do with phenotypically distant individuals. This distinction allows for the differences in local density that drive our results to be unaffected by considerations of local competition that influence models of altruism based on proximity in physical space (e.g. see Taylor's [57] famous cancellation result).

In actual biological populations, there are many possible phenotypic traits that organisms can observe or detect in making behavioural decisions towards other organisms. In many cases, the traits involved in processes like kin recognition involve multiple loci and are highly chemically diverse. We argue that 'realistically sized' phenotype spaces of many tens to hundreds of dimensions are a more appropriate setting for the understanding of the evolution of behaviour based on phenotypic similarity. In particular, the fact that high-dimensional phenotypes imply that everybody is atypical in some way is a general phenomenon and may be relevant to many more aspects of behavioural evolution is beyond the evolution of altruism. Thus, altruism based on phenotypic similarity is much less constrained by the properties of the space in which the strategy exists than altruism based on physical proximity. Simulations involving high-dimensional phenotype space as opposed to low-dimensional networks or lattices may lead to the discovery of new phenomena.

A potential avenue for further study would be to also allow the phenotype dimension to evolve to see if altruism can drive or limit the evolution of phenotypic complexity and the complexity of phenotypic sensing. In addition, in a multi-level selection setting, our results in figure 2 suggest that there may be an optimal phenotype dimensionality for the evolution of altruism—perhaps for average donation levels with low variation in outcome. It would also be interesting to consider strategies that do not use a fixed threshold but instead learn their threshold based on the current phenotype distribution. Finally, it is important to consider the potential effects of different forms of phenotypic transmission. Most obviously, one could consider a sexually reproducing population, possibly with assortative mating or inbreeding depression. It is also important to consider the possibility that the phenotype and strategy may be cultural traits with potential horizontal transmission.

Data accessibility. Data used for generating figures as well as code to generate figures are available at Dryad repository [58]. The electronic supplementary material is available at Figshare [59].

Authors' contributions. L.M.B.: conceptualization, formal analysis, investigation, methodology, software, visualization, writing—original draft, writing—review and editing; R.S.M.: conceptualization, formal analysis, investigation, methodology, supervision, visualization, writing—original draft, writing—review and editing; D.B.W.: conceptualization, funding acquisition, investigation, methodology, supervision, visualization, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by funding from a Simons Foundation Investigator Award in the Mathematical Modeling of Living Systems no. 508600, an Alfred P. Sloan Foundation Fellowship no. FG-2021-16667 and NSF award no. 1806833, the iPoLS Student Research Network.

Acknowledgements. We would like to thank Jeremy Van Cleve for helpful discussions about this manuscript.

References

1. Sherman PW. 1977 Nepotism and the evolution of alarm calls. *Science* **197**, 1246–1253. (doi:10.1126/science.197.4310.1246)
2. West SA, Diggle SP, Buckling A, Gardner A, Griffin AS. 2007 The social lives of microbes. *Annu. Rev. Ecol. Evol. Syst.* **38**, 53–77. (doi:10.1146/annurev.ecolsys.38.091206.095740)
3. Queller DC, Strassmann JE. 1998 Kin selection and social insects. *Bioscience* **48**, 165–175. (doi:10.2307/1313262)
4. Strassmann JE, Queller DC. 2011 Evolution of cooperation and control of cheating in a social microbe. *Proc. Natl Acad. Sci. USA* **108**, 10 855–10 862. (doi:10.1073/pnas.1102451108)
5. Rainey PB, Rainey K. 2003 Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**, 72–74. (doi:10.1038/nature01906)
6. Smith JM, Szathmáry E. 1997 *The major transitions in evolution*. Oxford, UK: Oxford University Press.
7. Eshel I, Cavalli-Sforza LL. 1982 Assortment of encounters and evolution of cooperativeness. *Proc. Natl Acad. Sci. USA* **79**, 1331–1335. (doi:10.1073/pnas.79.4.1331)
8. Fletcher JA, Doebeli M. 2008 A simple and general explanation for the evolution of altruism. *Proc. R. Soc. B* **276**, 13–19. (doi:10.1098/rspb.2008.0829)
9. Hamilton WD. 1964 The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–16. (doi:10.1016/0022-5193(64)90038-4)
10. Kay T, Keller L, Lehmann L. 2020 The evolution of altruism and the serial rediscovery of the role of relatedness. *Proc. Natl Acad. Sci. USA* **117**, 28 894–28 898. (doi:10.1073/pnas.2013596117)
11. Lehmann L, Keller L. 2006 The evolution of cooperation and altruism: a general framework and a

classification of models. *J. Evol. Biol.* **19**, 1365–1376. (doi:10.1111/j.1420-9101.2006.01119.x)

12. Queller DC. 1985 Kinship, reciprocity and synergism in the evolution of social behaviour. *Nature* **318**, 366–367. (doi:10.1038/318366a0)
13. Tarnita CE, Ohtsuki H, Antal T, Fu F, Nowak MA. 2009 Strategy selection in structured populations. *J. Theor. Biol.* **259**, 570–581. (doi:10.1016/j.jtbi.2009.03.035)
14. Cavalli-Sforza LL, Feldman MW. 1978 Darwinian selection and 'altruism'. *Theor. Popul. Biol.* **14**, 268–280. (doi:10.1016/0040-5809(78)90028-X)
15. Trivers RL. 1971 The evolution of reciprocal altruism. *Quar. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)
16. Axelrod R 1984 *The evolution of cooperation*. New York, NY: Basic Books.
17. Nowak MA, May RM. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)
18. Rousset F. 2004 *Genetic structure and selection in subdivided populations*, vol. 40. Princeton, NJ: Princeton University Press.
19. Akçay E, Van Cleve J. 2012 Behavioral responses in structured populations pave the way to group optimality. *Am. Nat.* **179**, 257–269. (doi:10.1086/663691)
20. García J, van Veelen M, Traulsen A. 2014 Evil green beards: tag recognition can also be used to withhold cooperation in structured populations. *J. Theor. Biol.* **360**, 181–186. (doi:10.1016/j.jtbi.2014.07.002)
21. Van Cleve J, Akçay E. 2014 Pathways to social evolution: reciprocity, relatedness, and synergy. *Evolution* **68**, 2245–2258. (doi:10.1111/evo.12438)
22. Dawkins R. 1976 *The selfish gene*. Oxford, UK: Oxford University Press.
23. Keller L, Ross KG. 1998 Selfish genes: a green beard in the red fire ant. *Nature* **394**, 573–575. (doi:10.1038/29064)
24. Queller DC, Ponte E, Bozzaro S, Strassmann JE. 2003 Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science* **299**, 105–106. (doi:10.1126/science.1077742)
25. Sinervo B *et al.* 2006 Self-recognition, color signals, and cycles of greenbeard mutualism and altruism. *Proc. Natl Acad. Sci. USA* **103**, 7372–7377. (doi:10.1073/pnas.0510260103)
26. Smukalla S *et al.* 2008 FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**, 726–737. (doi:10.1016/j.cell.2008.09.037)
27. Gardner A, West SA. 2010 Greenbeards. *Evolution* **64**, 25–38. (doi:10.1111/j.1558-5646.2009.00842.x)
28. Riolo RL, Cohen MD, Axelrod R. 2001 Evolution of cooperation without reciprocity. *Nature* **414**, 441–443. (doi:10.1038/35106555)
29. Sigmund K, Nowak MA. 2001 Evolution: tides of tolerance. *Nature* **414**, 403–405. (doi:10.1038/35106672)
30. Roberts G, Sherratt TN. 2002 Does similarity breed cooperation? *Nature* **418**, 499–500. (doi:10.1038/418499b)
31. Traulsen A, Schuster HG. 2003 Minimal model for tag-based cooperation. *Phys. Rev. E* **68**, 046129. (doi:10.1103/PhysRevE.68.046129)
32. Traulsen A. 2008 Mechanisms for similarity based cooperation. *Eur. Phys. J. B* **63**, 363–371. (doi:10.1140/epjb/e2008-00031-3)
33. Traulsen A, Nowak MA. 2007 Chromodynamics of cooperation in finite populations. *PLoS ONE* **2**, e270. (doi:10.1371/journal.pone.0000270)
34. Sturgis SJ, Gordon DM. 2012 Nestmate recognition in ants (Hymenoptera: Formicidae): a review. *Myrmecol. News* **16**, 101–110.
35. Sprenger PP, Menzel F. 2020 Cuticular hydrocarbons in ants (Hymenoptera: Formicidae) and other insects: how and why they differ among individuals, colonies, and species. *Myrmecol. News* **30**, 1–26.
36. Anten NP, Chen BJ. 2021 Detect thy family: mechanisms, ecology and agricultural aspects of kin recognition in plants. *Plant Cell Environ.* **44**, 1059–1071. (doi:10.1111/pce.14011)
37. Ninkovic V, Rensing M, Dahlin I, Markovic D. 2019 Who is my neighbor? Volatile cues in plant interactions. *Plant Signal. Behav.* **14**, 1634993. (doi:10.1080/15592324.2019.1634993)
38. Ninkovic V, Markovic D, Rensing M. 2021 Plant volatiles as cues and signals in plant communication. *Plant Cell Environ.* **44**, 1030–1043. (doi:10.1111/pce.13910)
39. Krause ET, Krüger O, Kohlmeier P, Caspers BA. 2012 Olfactory kin recognition in a songbird. *Biol. Lett.* **8**, 327–329. (doi:10.1098/rsbl.2011.1093)
40. Zeilinger S *et al.* 2016 Friends or foes? emerging insights from fungal interactions with plants. *FEMS Microbiol. Rev.* **40**, 182–207. (doi:10.1093/femsre/fuv045)
41. Levréro F, Carrete-Vega G, Herbert A, Lawabi I, Courtiol A, Willaume E, Kappeler PM, Charpentier MJE. 2015 Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nat. Commun.* **6**, 7609. (doi:10.1038/ncomms8609)
42. Riebel K, Langmore NE. 2023 Birdsong: not all contest but also cooperation? *Curr. Biol.* **33**, R67–R69. (doi:10.1016/j.cub.2022.12.015)
43. Wall D. 2016 Kin recognition in bacteria. *Annu. Rev. Microbiol.* **70**, 143–160. (doi:10.1146/annurev-micro-102215-095325)
44. Cook LC, Federle MJ. 2014 Peptide pheromone signaling in streptococcus and enterococcus. *FEMS Microbiol. Rev.* **38**, 473–492. (doi:10.1111/1574-6976.12046)
45. Novick RP, Geisinger E. 2008 Quorum sensing in staphylococci. *Annu. Rev. Genet.* **42**, 541–564. (doi:10.1146/annurev.genet.42.110807.091640)
46. Stefanic P, Mandic-Mulec I. 2009 Social interactions and distribution of *Bacillus subtilis* pherotypes at microscale. *J. Bacteriol.* **191**, 1756–1764. (doi:10.1128/JB.01290-08)
47. Sherlock O, Vejborg RM, Klemm P. 2005 The *TibA* adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect. Immun.* **73**, 1954–1963. (doi:10.1128/IAI.73.4.1954-1963.2005)
48. Aoki SK *et al.* 2010 A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**, 439–442. (doi:10.1038/nature09490)
49. Ruhe ZC, Low DA, Hayes CS. 2013 Bacterial contact-dependent growth inhibition. *Trends Microbiol.* **21**, 230–237. (doi:10.1016/j.tim.2013.02.003)
50. Adams DW, Stutzmann S, Stoudmann C, Blokesch M. 2019 DNA-uptake pili of *Vibrio cholerae* are required for chitin colonization and capable of kin recognition via sequence-specific self-interaction. *Nat. Microbiol.* **4**, 1545–1557. (doi:10.1038/s41564-019-0479-5)
51. Rutherford ST, Bassler BL. 2012 Bacterial quorum sensing: its role in virulence and possibilities for its control. *CSH Perspect. Med.* **2**, a012427. (doi:10.1101/cshperspect.a012427)
52. Gibbs KA, Urbanowski ML, Greenberg EP. 2008 Genetic determinants of self identity and social recognition in bacteria. *Science* **321**, 256–259. (doi:10.1126/science.1160033)
53. Wall D. 2014 Molecular recognition in myxobacterial outer membrane exchange: functional, social and evolutionary implications. *Mol. Microbiol.* **91**, 209–220. (doi:10.1111/mmi.12450)
54. Spector L, Klein J. 2007 Multidimensional tags, cooperative populations, and genetic programming. In *Genetic programming theory and practice IV*, pp. 97–112. Berlin, Germany: Springer.
55. Antal T, Ohtsuki H, Wakeley J, Taylor PD, Nowak MA. 2009 Evolution of cooperation by phenotypic similarity. *Proc. Natl Acad. Sci. USA* **106**, 8597–8600. (doi:10.1073/pnas.0902528106)
56. Kroumi D, Lessard S. 2015 Evolution of cooperation in a multidimensional phenotype space. *Theor. Popul. Biol.* **102**, 60–75. (doi:10.1016/j.tpb.2015.03.007)
57. Taylor PD. 1992 Altruism in viscous populations' an inclusive fitness model. *Evol. Ecol.* **6**, 352–356. (doi:10.1007/BF02270971)
58. Mehta RS, Båvik LM, Weissman DB. 2023 50 shades of greenbeard: Robust evolution of altruism based on similarity of complex phenotypes. Dryad Digital Repository. (doi:10.5061/dryad.qbzkh18p2)
59. Båvik LM, Mehta RS, Weissman DB. 2023 Fifty shades of greenbeard: robust evolution of altruism based on similarity of complex phenotypes. Figshare. (doi:10.6084/m9.figshare.c.6672279)