The Demographic Benefits of Belligerence and Bravery: Defeated Group Repopulation or Victorious Group Size Expansion?

Laurent Lehmann*

Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

Abstract

Intraspecific coalitional aggression between groups of individuals is a widespread trait in the animal world. It occurs in invertebrates and vertebrates, and is prevalent in humans. What are the conditions under which coalitional aggression evolves in natural populations? In this article, I develop a mathematical model delineating conditions where natural selection can favor the coevolution of belligerence and bravery between small-scale societies. Belligerence increases an actor's group probability of trying to conquer another group and bravery increase the actors's group probability of defeating an attacked group. The model takes into account two different types of demographic scenarios that may lead to the coevolution of belligerence and bravery. Under the first, the fitness benefits driving the coevolution of belligerence and bravery come through the repopulation of defeated groups by fission of victorious ones. Under the second demographic scenario, the fitness benefits come through a temporary increase in the local carrying capacity of victorious groups, after transfer of resources from defeated groups to victorious ones. The analysis of the model suggests that the selective pressures on belligerence and bravery are stronger when defeated groups can be repopulated by victorious ones. The analysis also suggests that, depending on the shape of the contest success function, costly bravery can evolve in groups of any size.

Introduction

Coalitional aggression between groups of individuals occurs not only in humans but is prevalent in other vertebrates and invertebrates as well. For instance, colonies of army ants battle against each other to increase the size of their territory [1] and termites satisfy the condition of eusociality by the existence of sterile soldiers, not because they produce sterile workers [2]. Groups of chimpanzees engage in coordinated raids to obtain additional mates and territory [3,4], and prehistoric groups of hunter-gatherers fought against each others for resources by throwing spears and boomerangs. It has even been suggested that a typical tribal society lost about 5 percent of its population in combat each year [5].

Coalitional aggression is a group level strategy, which allows individuals within groups to acquire reproduction enhancing resources from other groups. This strategy might have played a central role in human evolution [6–12]. While modern warfare between large scale societies is unlikely to be directly driven by motives of reproduction, it is important to try to understand the conditions under which coalitional aggression may have been selected for in natural populations [12,13]. This strategy is individually costly in time and energy. What then are the rewards that make this enterprise worthwhile so that it can be selectively favored over an evolutionary time scale?

While the analytics of the proximate causes of raiding is well developed [14–18], there are few analyses of the demographic benefits leading to the evolution of intraspecific coalitional aggression between small-scale societies, and existing ones consider a saturated habitat and assume that the benefits of warfare come through the repopulation of defeated groups. In this case, individuals from a victorious group of a pairwise contest migrate into the defeated group and partially or completely replace individuals from the defeated group, thereby adopting its territory and/or its females [19–22]. This scenario may be consistent with a situation where victorious groups take over specific resource points that were occupied by defeated ones. Fights in social insects may occur over colonies lying in suitable habitats [1] and primitive warfare in hominids occurred over crucial water points and shelters [12,23].

But repopulation of defeated groups is not the only plausible demographic benefit that may lead to the evolution of coalitional aggression. An alternative situation is that resources are captured from defeated groups by victorious ones, which, when used by the the individuals of victorious groups may result in an increase in their local carrying capacity. For instance, due to its high nutritional value, meat of all sorts was contested among hunter-gatherers [12,23]. Groups winning contests over animal populations may have obtained higher shares of the contested resources. The local use of such resources may then lead to different

demographic outcomes than repopulation of defeated groups. For instance, an increase in size of victorious groups may not only increase their contribution to the ancestry of the population because they are larger, but may also increase the probability that these groups win further battles. A self-enforcing demographic loop leading to higher levels of warfare between groups is thus a plausible outcome of this process.

The aim of this paper is to analyze the co-evolutionary dynamics of belligerence and bravery under this alternative demographic scenario, where victorious groups can expand in size as an outcome of warfare. These two traits are likely to be involved in coalitional aggression as belligerence increases an actor's group probability of trying to defeat another group and bravery increases an actors's group ability to defeat another group. In order to be able to contrast the selective pressure on belligerence and bravery under the two aforementioned demographic scenarios, I present a mathematical model, which takes both scenarios into account and allows me to contrast how different assumptions may lead to similar or different selective pressures on the evolution of these two traits when interaction occurs between small-scale, pre-state societies ("hunter-gatherer" warfare). The analysis of the model suggests that the selective pressure on belligerence and bravery are stronger when defeated groups can be repopulated by victorious ones.

Analysis

Life cycle assumptions

The population is assumed to consist of an infinite number of groups, which are connected by random dispersal (i.e., infinite island model of dispersal; [24]). The natural environment is assumed to allow only N individuals to reach adulthood in each

group. The life cycle of individuals living in this population is assumed to be similar in baseline structure to an earlier analyses of the evolution of coalitional aggression [22], but it is augmented with the possibility that the sizes of groups can vary as an outcome of warfare; namely, the groups can become larger or smaller than the baseline size N determined by the environmental condition. Because this feature adds more complexity to the model, individuals within groups are assumed, for simplicity, to be haploid and follow clonal reproduction. For ease of presentation of the demography underlying the model, I begin by introducing the basic demographic assumptions and model parameters by assuming a monomorphic population, where all individuals express the same phenotypes. I then introduce variation in behaviors and evolutionary dynamics.

In each generation, the following events occur in sequence. (1) Each adult individual in a group produces a large number of juveniles. Juveniles mature and become subadults. (2) Each subadult disperses with probability m to a new randomly chosen group. (3) War occurs between groups. With probability ϕ the subadults in each group engage into a fight with another group from the population and try to defeat it in order to gain its resources. The attacker group wins the ensuing battle with probability v. The individuals from each group might also engage into a fight locally, to defend their group because it is attacked. This also occurs with probability ϕ, in which case the attacker wins the fight with probability v. Each group is assumed to engage only in a single fight locally (to defend its resources), but a focal group can be both an attacker and a defender in the same generation (see Fig. 1). (4) Adults die and density-dependent competition (regulation) occurs in each group between subadult individuals to form the next generation. The number of individuals that reach adulthood in each group depends on the outcomes of warfare and
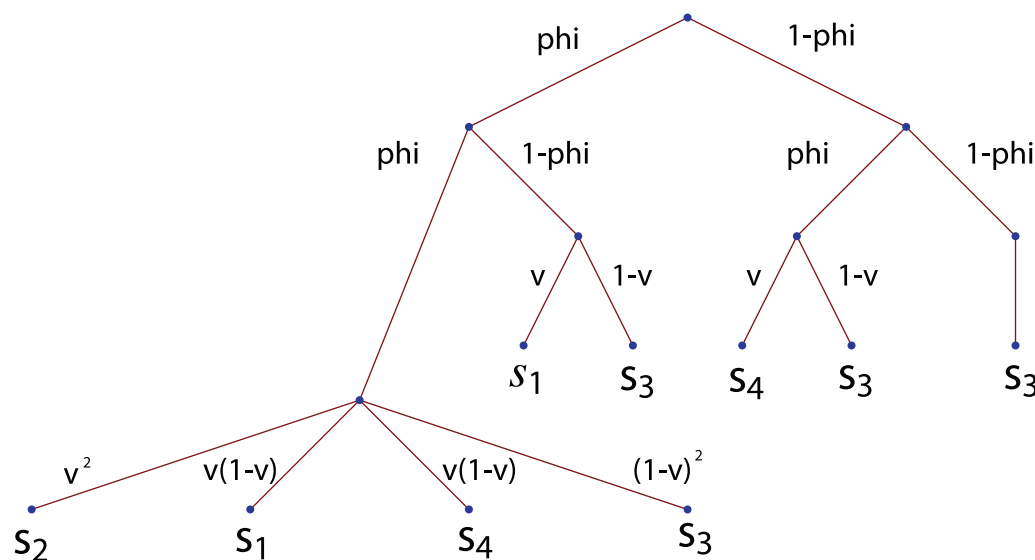

Figure 1. Demographic outcomes and events faced by the individuals in a focal group (the terms "outcomes" and "events" are used as defined in probability theory, e.g., p. 18, chapter 1 of [62]). Each different connected series of edges starting at the root node (top of the figure) provides a demographic outcome. Different outcomes may lead to similar demographic events, which are denoted by s_i and described in the text. The first series of edges, directly issued from the root node, represents the probabilities that the focal group fights or does not fight to conquer another group from the population. The second series of edges represent the probabilities that the focal group fights or does not fights locally. The third series of edges represent the winning probabilities of the various battles, which occur only if there is a fight. For instance, the left most series of connected edges represent the outcome where the focal group fights against another group upon attacking (probability ϕ), it fights locally because it is attacked (probability ϕ) and wins the two battles (probability v^2). In order to obtain the probabilities of occurrence of each demographic event describes in the text, one has to sum the probabilities of occurrence of outcomes where that event obtains, which gives $\Pr(s_1) = \phi(1-\phi)v + \phi^2 v(1-v) = (1-v\phi)v\phi$, $\Pr(s_2) = (v\phi)^2$, $\Pr(s_3) = (1-\phi v)^2$, and $\Pr(s_4) = \phi v(1-\phi v)$.
doi:10.1371/journal.pone.0021437.g001

two different but complementary types of demographic scenarios will be analyzed:

(i) Defeated group repopulation scenario (DGR). Here, in groups that have not been defeated, only subadults from that group compete against each other for access to breeding spots vacated by the death of adults. By contrast, in groups that have been defeated, subadults from both the victorious and defeated group compete against each other for vacated breeding spots. With probability h an individual randomly sampled from a defeated group after competition is assumed to be a member from that group before competition. Hence, defeated groups are partially repopulated by individuals from the victorious group and $1-h$ can be though of as the fraction of the individuals in a defeated group that are replaced by those of the victorious group. In both defeated and non-defeated groups exactly N individuals are assumed to reach adulthood.

This DGR scenario is essentially a haploid version a the model mentioned above [22], but different assumptions will be employed below for the behavior of individuals in the population. Further, the interactions between hostile groups will be made more mechanistic here, which complements the previous analysis and will also clarify its results.

(ii) Victorious group size expansion scenario (VGE). Here, instead of repopulating a fraction $1-h$ of defeated groups, individuals from the victorious group extract a share $1-h$ of the resources of defeated groups and use them in their natal group to produce offspring. Alternatively, this scenario could be interpreted as a situation where two groups contest for a given prize (e.g., some resources found in the environment) and where the winner of the contest obtains a fraction $1-h$ of the prize.

In order to be able to directly compare this VGE scenario to the DGR scenario described above, I assume that the extraction of resources from defeated groups leads to an increase in the number of individuals reaching adulthood in victorious groups by $(1-h)N$ individuals, while only hN individuals survive to adulthood in defeated groups. The increase (or reduction) in group size is assumed to last only a single generation because it is reasonable to postulate that the additional resources gained through warfare can only sustain additional individuals over a single generation. Gained resources are thus temporary and there is no inheritance of resources across generations.

One may also suppose that warfare occurs at different timings in the life cycle of the individuals. For instance, subadults within groups may wage war before their dispersal instead of after dispersal so that stages (2) and (3) in the life cycle are interchanged. This case will also be investigated below.

Demographic events

Because a focal group in a focal generation may either attack or not attack another group, fight or not fight a battle when it attacks another group, be attacked or not be attacked by another group, and win or loose any of the possible battles, several different outcomes may affect the fitness of an adult individual from a focal group (Fig. 1). It follows from the assumptions of scenario (i) and (ii) that the different outcomes can be gathered into four different demographic events (Fig. 1). The first, denoted s_1, is when the focal group keeps all its resources and its individuals obtain a share of the "resources" of another group (be it through repopulation or by taking a share of the group's material resources). This event occurs when the focal group attacks another one, defeats the attacked group and is not defeated locally. The second demographic event, denoted s_2, is when the focal group loses a share of it resources because it is defeated locally but obtains some resources from another group after attacking and winning the

battle. The third event, denoted s_3, is when the focal group does not lose any resources locally and does not gain any resources from another group. Finally, the fourth event, denoted s_4, occurs when the focal group loose resources locally (as it is defeated) and does not obtain any resources from another group. This is the worst event for a group.

Under the DGR scenario, a demographic event determines the expected number of groups in which the individuals of a focal group may reproduce. For the events s_1, s_2, s_3, and s_4, this number is, respectively, $1+(1-h), 1, 1$, and h and thus varies between zero and one [for event s_2, the mean of one is obtained as $h+(1-h)$]. For the VGE scenario, a demographic event determines the number of individuals that reach adulthood in the focal group. For the events s_1, s_2, s_3, and s_4, this number is, respectively, $N+(1-h)N, N, N$, and hN [for event s_2, N is obtained as $(s+[1-h])N$]. Because groups are of size N under the DGR scenario, the above assumptions imply that the additional number of descendants accruing to a group when it wins a battle is the same under the two scenarios and is given in both cases by $(1-h)N$. This increase in group level benefit only lasts a single generation, because, as assumed above, offspring do not inherit resources from the parental generation.

The demographic events experienced by each group may also change across generations. A focal group may experience a certain demographic event, say s, in a parental generation, and after one iteration of the life cycle that group may experience a different demographic event, say s', in the offspring generation. The changes in the demographic events between parental and offspring generation lead to a demographic dynamics, which is described by a stochastic process whose transition probabilities are affected by the behaviors of the individuals in the population.

Evolving behaviors and selection gradient

Following earlier models for the evolution of warfare in small scale societies, individuals within groups are assumed to express two different traits [20,22]. The first trait, denoted x, is called "belligerence". The average level of belligerence in a focal group is assumed to vary between zero and one and it affects the probability ϕ that a focal group fights another one. The second trait, denoted y, is called "bravery" and it is assumed to increase the probability v that a group fighting another one wins the battle. Bravery should be understood metaphorically as it may also represent the production of a sterile soldier caste in social insects or, for hominids, the investment into some technology of appropriation.

Although both belligerence and bravery increase group success, both traits are assumed to reduce the probability that an individual expressing them survives up to the stage of density-dependent competition (stage 4 of the life cycle). Since clonal reproduction is assumed, the average phenotype of an offspring is the same as that of its parent. One can then express the relative number of juveniles, which reach the stage where they compete for breeding spots, and that are produced by a focal adult individual with level of belligerence x and bravery y as $[1-C_x(x)][1-C_y(y)]$, where $C_x(x)$ is the cost of expressing belligerence and $C_y(y)$ the cost of expressing bravery (both cost functions are as assumed to be convex). By contrast to previous work, this formulation entails that belligerence and bravery are costly even if the group of the actor is not engaged into a contest. There are two reasons for this assumption. First, the model is simpler to analyze. Second, the model may also represent battles between colonies of social insects, in which case the production of a soldier caste is made independently of whether there is a battle between groups or not. More generally, one may also assume that investment into

belligerence and bravery may be costly to the actor throughout its life because these traits are probably costly in time and energy beyond the battlefield (e.g., belligerence may involve spending time convincing group members of raiding other groups and bravery developing the technologies of appropriation).

The evolutionary dynamics of belligerence (x) and bravery (y) are complicated to analyze in full form because the model is demographically explicit (groups may vary in size), evolution occurs in a subdivided population (fluctuations of gene frequencies between groups occurs), and individuals within groups are subject to a varying environment (warfare is a stochastic process, which further depends on the behavior of individuals in the population). I thus employ the assumptions of weak selection and additive gene action. With these assumptions, one can evaluate the evolutionary dynamics of a focal trait through a phenotypic gradient approach by holding the other trait constant. Each trait is then assumed to follow a gradual, step-by-step transformation caused by the successive invasion of mutant alleles having different phenotypic effects than resident alleles fixed in the population (e.g., adaptive dynamics approach or ESS method; [25–30]).

The change in the frequency p in the population of a given mutant allele, which codes for a small phenotypic deviation δ relative to the phenotype z expressed by individuals carrying a resident allele (where z stems either for belligerence, x, or bravery, y), can be written for the infinite island model as $\Delta p = p(1-p)\delta S(z) + O(\delta^2)$, where $S(z)$ is the force of directional selection on the mutant allele and $O(\delta^2)$ is a remainder that includes higher order terms [29]. Importantly, the selection gradient $S(z)$ on the trait is independent of the frequency of the mutant in the population, which allows one to evaluate a candidate evolutionary stable strategy $z^\star$ (ES level of belligerence or bravery) by solving $S(z) = 0$ for z. Evaluating the criteria of continuous stability [26,28,29] of the phenotype z requires one to evaluate the change of gene frequency Δp to the second order in δ [29,31], which is theoretically possible but a difficult task under the present demographic assumptions. However, the aim of this paper is to compare the forces of directional selection on belligerence and bravery under the VGE and DGR scenarios, which can be carried out by studying the gradient $S(z)$ since higher order terms will be less important.

Although such an ESS approach does thus not give a full picture of evolution, it allows one to obtain explicit analytical approximations, which would otherwise be out of reach, and that have repeatedly shown to provide good predictions for long-term phenotypic evolution and/or the direction of selection under a wide spectrum of biological applications involving local genetic drift under complex demographies with weak assumptions about the distribution of mutant's deviating phenotypic effects [28,32–37].

For the DGR scenario, group size is the same in each group in the population in each generation. There are also no correlations between the demographic events experienced by a group from one generation to the next, and the occurrence of warfare is independently distributed across generations. The selection gradient on the trait z can then be evaluated by a standard application of the direct fitness method [29,38,39], as

$$S(z) = \frac{\partial w}{\partial z_\bullet} + \frac{\partial w}{\partial z_0^R} R^R, \qquad (1)$$

where $w \equiv w(z_\bullet, z_0^R, z_1)$ is the fitness function giving the expected number of adult offspring of a focal adult individual (an individual taken at stage 1 of the life-cycle) bearing a mutant allele and R^R is

the probability of identity-by-descent between the focal individual and a randomly sampled individual (including himself) from the focal patch, which is a measure of relatedness between group members. The derivatives of w are the effects of actors on the fitness of a focal individual and are evaluated at the phenotypic value of the resident allele ($z_\bullet = z_0^R = z_1 = z$), where the actors are the focal individual itself with phenotype denoted $z_\bullet$, individuals from the focal group with average phenotype denoted z_0^R (the average is over all individuals from the focal group, thus including the focal individual itself), and individuals from different groups with average phenotype z_1.

Owing to the infinite island model assumptions, individuals from different groups (with phenotype z_1) have zero relatedness to the focal individual. One way of understanding this result is by noting that when there is an infinite number of groups, the ancestral lineages of two homologous genes sampled from two individuals within the same group have either stayed in the same group and coalesced, or have migrated to two different groups and can then be considered as being independent [40,41]. The relatedness coefficient R^R measures the increase in the probability of identity between genes sampled from individuals within the same group due to coalescence events within groups (e.g., "identity-by-descent"). By contrast, the ancestral lineages of two genes sampled from individuals from two different groups do not coalesce in a recent past and can be considered as being independent [40,41]. Hence, these individuals are unrelated as their homologous genes are not more identical than two genes sampled at random from the population.

For the VGE scenario, group size may fluctuate across generations and correlations between demographic states may develop over time. The selection gradient $S(z)$ will then also depend on the reproductive values of individuals facing different demographic events, which take into account the differences in expected future contribution to the population of these individuals according to the different types of demographic events that occur in the groups they settle in [34,42]. As a result, the form of the selection gradient $S(z)$ for the VGE scenario is more complicated than eq. 1 and is presented in Appendix S2 (eqs. B-1–B-2), but conceptually it is an extension of eq. 1, which includes demographic classes and where marginal fitness components are weighted by reproductive values.

Fitness components for the DGR scenario

I will present the derivations of the fitness function w and the selection gradient S for the DGR scenario directly in the main text and refer to Appendix S2 for the derivation of these quantities for the VGE scenario, as it is more complicated in the latter case but conceptually similar.

In order to evaluate w for the DGR scenario, or more generally the various components of the fitness function for the VGE scenario, we need to relate behaviors to fighting (ϕ) and winning (v) probabilities. These will determine the demographic events, which, in turn, determine fitness. The fighting and winning probabilities can be thought of as macro demographic variables, which depend on the mechanistic details relating trait value expression by individuals within groups to group level behavior.

Fighting probability. Owing to the infinite island model of dispersal assumptions introduced above, I assume that the probability that the focal group with subadults expressing average level of belligerence x_0^A enters into a fight with another group is given by some function $\phi(x_0^A, x_1^A)$, which also depends on the average level x_1^A of belligerence among subadults in other groups in the population. Because any class of individuals from other groups (immigrants, juveniles, or adults) have zero

relatedness to the focal individual, their average phenotypes can be considered to be equivalent and we can set $x_1^A = x_1$. Because warfare occurs after the dispersal of juveniles (or subadults), the average phenotype among subadults in the focal group is related to that of adults before dispersal by

$$x_0^A = (1-m)x_0^R + mx_1, \qquad (2)$$

since after dispersal a fraction m of offspring in the focal group are descendant from that group (with parents carrying average phenotype x_0^R), while a fraction m of offspring have immigrated from other groups (with parents carrying average phenotype x_1).

Alternatively to the situation described by eq. 2, one may also assume that warfare occurs right before the dispersal of juveniles, in which case the average phenotype among subadults is the same as that among adults so that $x_0^A = x_0^R$. These two different timings of warfare can be considered simultaneously by writing the average phenotype among subadults as

$$x_0^A = (1-m\epsilon)x_0^R + m\epsilon x_1, \qquad (3)$$

where the parameter ϵ is equal to zero when warfare occurs before the dispersal of offspring and equal to one if it occurs after the dispersal of offspring.

The exact functional relationship relating x_0^A and x_1 to the fighting probability does not affect the qualitative result reported below, but in order to obtain quantitative results one needs such a functional relationship. For this reason, I introduce the island model of warfare. Here, a group with average level of belligerence x attacks independently another random group from the population with probability $a(x)$, which is assumed to be an increasing function of its argument. Namely, the focal group attacks another group with probability $a(x_0^A)$, while another group from the population attacks a group sampled at random with probability $a(x_1)$. The assumption that each group attacks independently of each other another random group from the population implies that several groups may simultaneously attack the same group. Since resources are limited, I assume that during a single generation a focal group engages only into a single pairwise contest over its local resources with another attacker group. When several groups simultaneously attack the focal group (synchronously or asynchronously), one randomly sampled group among the attackers is assumed to engage into a fight with the focal group.

From these assumptions, the probability that the focal group engages into a fight locally is given by $1-\exp(-a(x_1))$, which increases with the level of belligerence x_1, while the probability that the focal group enters into a fight with another group from the population through attacking is

$$\phi(x_0^A, x_1) = a(x_0^A)\frac{1-\exp(-a(x_1))}{a(x_1)} \qquad (4)$$

(eqs. A-1–A-5 of Appendix S1). This equation says that the fighting probability of the focal group increases with the average level of belligerence x_0^A of its members, but decreases with the average level of belligerence in other groups in the population because, the higher the level of belligerence in the population, the lower the probability $[1-\exp(-a(x_1))]/a(x_1)$ that the focal groups enters into a fight with a given group conditional on it attacking. Note that in a monomorphic population, for instance when $x_1 = x_0^A = x$, the fighting probability $\phi(x,x) = 1-\exp(-a(x))$ increases with the overall level of belligerence.

Winning probability. Eq. 4 gives the probability that a focal group engages in a fight. But we also need an expression for the probability that it wins the ensuing battle. I assume that the probability that the focal group, where subadults express average level $y_0^A = (1-m\epsilon)y_0^R + m\epsilon y_1$ of bravery, and that attacks another group, where subadults express average level of bravery $y_1^A = y_1$, wins the ensuing battle is given by

$$v(y_0^A, y_1) = \frac{\omega g(y_0^A N)}{\omega g(y_0^A N) + (1-\omega)g(y_1 N)}. \qquad (5)$$

In this winning probability, N represents the number of individuals producing subadults, $g(\cdot)$ describes the power of contestants, which is a positive increasing function of its argument, which can be thought of as the total effort put by a group into the battle (level of bravery times relative number of combatants), and the parameter ω allows one to tune the advantage of being offensive. If $\omega > 1/2$, there is advantage to attacking if everything else among the two groups is held constant (e.g., $v(y_0^A, y_1) = \omega$ in a monomorphic population, where $y_0^A = y_1$). An offensive advantage may be justified by the fact that it has been suggested that most primitive warfare consisted of ambushes and ambuscades, which result from unilateral actions conducted under conditions where individuals from attacked group were caught helpless [12,43]. On the other hand, a defensive advantage ($\omega < 1/2$) has been justified by the fact that attackers often don't know the intruded area as well as defenders, which puts attackers at a disadvantage [44].

Equation 5 is a so-called contest success function [14,45,46], which depends on the ratio of the effort the two opposing parties put into winning a battle. Such a functional relationship follows from a series of assumptions about contests [46], the strongest of which is probably that the contest among a smaller number of combatants is qualitatively similar to those among a large number of combatants.

Probabilities of occurrence of demographic events. In order to evaluate w we need the probabilities of occurrences of the four demographic events. Let us use the shorthand notations $\phi_0 \equiv \phi(x_0^A, x_1)$ and $v_0 \equiv v(y_0^A, y_1)$, where the second expression is the probability that the subadults in the focal group win the battle upon entering into a fight with another group from the population. Likewise, call $\phi_1 \equiv \phi(x_1, x_1)$ the probability that the subadults from another group in the population try to conquer the focal group and enter into a fight with it, and let $v_1 \equiv v(y_1, y_0^A)$ be the probability that the attacking group wins the ensuing battle. With these notations and using the different outcomes in Fig. 1, the probabilities that the events s_1, s_2, s_3, and s_4 obtain in the focal group are, respectively, given by

$$\Pr(s_1) = (1-\phi_1 v_1)\phi_0 v_0$$

$$\Pr(s_2) = \phi_1 v_1 \phi_0 v_0$$

$$\Pr(s_3) = (1-\phi_1 v_1)(1-\phi_0 v_0)$$

$$\Pr(s_4) = \phi_1 v_1(1-\phi_0 v_0). \qquad (6)$$

Fitness function. It is convenient to write the fitness function as

$$w = w_\mathrm{p} + w_\mathrm{d}, \tag{7}$$

where w_p is the expected number of the focal individual's settled offspring that remained philopatric and w_d is the expected number of the focal individual's settled offspring in other groups after dispersal. Let us further call $w_\mathrm{p}(s)$ the expected number of philopatric offspring of a focal individual breeding in a group when demographic event s occurs. Averaging over all demographic events given in eq. 6, the philopatric component of fitness in eq. 7 becomes

$$\begin{aligned} w_\mathrm{p} = {}& w_\mathrm{p}(s_1)[1-\phi_1 v_1]\phi_0 v_0 + w_\mathrm{p}(s_2)\phi_1 v_1 \phi_0 v_0 + \\ & w_\mathrm{p}(s_3)[1-\phi_1 v_1][1-\phi_0 v_0] + w_\mathrm{p}(s_4)\phi_1 v_1 [1-\phi_0 v_0]. \end{aligned} \tag{8}$$

The fitness functions corresponding to each demographic event are evaluated by following the life cycle assumptions presented in the main text and standard calculations for the infinite island model [29,41,42,47], which are as follows. A focal individual produces a relative number $b_\bullet = [1 - C_\mathrm{x}(x_\bullet)][1 - C_\mathrm{y}(y_\bullet)]$ of offspring (relative to that of an individual in a resident monomorphic population) that reach the regulation stage (stage 4 of the life cycle of the main text). When the focal group has not been conquered, the focal individual's offspring, which remained philopatric (fraction $1-m$ of $b_\bullet$), compete in the focal group with $(1-m)b_\mathrm{r}$ subadults produced in the focal group, where $b_\mathrm{r} = [1 - C_\mathrm{x}(x_0^\mathrm{R})][1 - C_\mathrm{y}(y_0^\mathrm{R})]$ is the average relative number of subadults produced in the focal group that reach the regulation stage. The focal individual's philopatric offspring also compete against mb_d immigrants, where $b_\mathrm{d} = [1 - C_\mathrm{x}(x_1)][1 - C_\mathrm{y}(y_1)]$ is the average relative number of subadults produced in other groups that reach the regulation stage.

When the focal group has been conquered, $(1-m)b_\bullet h$ of the focal individual's offspring compete against $h[(1-m)b_\mathrm{r} + mb_\mathrm{d}]$ subadults reaching the regulation stage in the focal group after dispersal and warfare. The focal individual's offspring also compete against $(1-h)b_\mathrm{d}$ subadults from the victorious group. When the focal group conquers another group, $(1-m)b_\bullet(1-h)$ of the focal individual's offspring compete against $(1-h)[(1-m)b_\mathrm{r} + mb_\mathrm{d}]$ subadults from the focal group and hb_d subadults from the conquered group.

If warfare occurs before the dispersal of subadults; that is, stages (2) and (3) of the life cycle are interchanged but every other assumption of the model is the same, the above fecundity functions still apply. These fecundity functions can be applied to both warfare occurring before or after the dispersal of subadults, because in both cases regulation occurs after dispersal and warfare and it is the number of offspring reaching the regulation stage that determines fitness. Hence, we can use the same fitness functions to analyze the two different timings of warfare by making use of eq. 3 and the analogous equation for bravery ($y_0^\mathrm{A} = (1-m\epsilon)y_0^\mathrm{R} + m\epsilon y_1$).

Gathering all the above terms and taking into account the fact that, when the focal group conquers another group, the focal individual has descendants in the focal group and in the conquered group gives

$$w_\mathrm{p}(s_1) = \frac{(1-m)b_\bullet}{(1-m)b_\mathrm{r} + mb_\mathrm{d}} + \frac{(1-m)b_\bullet(1-h)}{(1-h)[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + hb_\mathrm{d}}$$

$$w_\mathrm{p}(s_2) = \frac{(1-m)hb_\bullet}{h[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + (1-h)b_\mathrm{d}} + \frac{(1-m)b_\bullet(1-h)}{(1-h)[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + hb_\mathrm{d}}$$

$$w_\mathrm{p}(s_3) = \frac{(1-m)b_\bullet}{(1-m)b_\mathrm{r} + mb_\mathrm{d}}$$

$$w_\mathrm{p}(s_4) = \frac{(1-m)hb_\bullet}{h[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + (1-h)b_\mathrm{d}}. \tag{9}$$

In order to evaluate w_d, we need the probabilities that a random group from the population is in demographic states $s_1, s_2, s_3,$ and s_4. These probabilities are, respectively, $[1-\phi_1 v_{11}]\phi_1 v_{11}$, $(\phi_1 v_{11})^2$, $[1-\phi_1 v_{11}]^2$, and $\phi_1 v_{11}[1-\phi_1 v_{11}]$, where $v_{11} \equiv v(y_1, y_1) = \omega$ because the demographic events that obtain in other groups than the focal group depend only on the average level of belligerence and bravery in the population (x_1 and y_1). Averaging over all these cases, the immigrant component of fitness is

$$\begin{aligned} w_\mathrm{d} = {}& w_\mathrm{d}(s_1)[1-\phi_1 v_0]a_1\omega + w_\mathrm{d}(s_2)(\phi_1\omega)^2 + \\ & w_\mathrm{d}(s_3)[1-\phi_1\omega]^2 + w_\mathrm{d}(s_4)\phi_1\omega[1-\phi_1\omega], \end{aligned} \tag{10}$$

where

$$w_\mathrm{d}(s_1) = \frac{mb_\bullet}{b_\mathrm{d}} + \frac{mb_\bullet(1-h)}{b_\mathrm{d}}$$

$$w_\mathrm{d}(s_2) = \frac{mb_\bullet h}{b_\mathrm{d}} + \frac{mb_\bullet(1-h)}{b_\mathrm{d}}$$

$$w_\mathrm{d}(s_3) = \frac{mb_\bullet}{b_\mathrm{d}}$$

$$w_\mathrm{d}(s_4) = \frac{mb_\bullet h}{b_\mathrm{d}}. \tag{11}$$

Substituting eqs. 9–11 into eq. 7 and simplifying yields the fitness of a focal adult as

$$\begin{aligned} w = (1-m)b_\bullet & \left[\frac{(1-\phi_1 v_1)}{(1-m)b_\mathrm{r} + mb_\mathrm{d}} + \frac{\phi_1 v_1 h}{h[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + (1-h)b_\mathrm{d}} \right. \\ & \left. + \frac{\phi_0 v_0(1-h)}{(1-h)[(1-m)b_\mathrm{r} + mb_\mathrm{d}] + hb_\mathrm{d}} \right] + \frac{mb_\bullet}{b_\mathrm{d}}, \end{aligned} \tag{12}$$

where the philopatric component of fitness depends on three terms. The first term accounts for the fitness accruing to the focal individual when the focal group is not attacked plus when the group is attacked and wins the battle, the second term accounts for the fitness to the focal individuals when the focal group is attacked but loses the battle, and, finally, the third term accounts for the

fitness to the focal individuals when the focal group attacks another group and wins the battle.

In order to compare the demographic benefits that accrue to a focal group under the DGR scenario to that obtained under the VGE scenario, it is useful to note that the reproductive value $v(s)$ of a group in demographic state s, which is the asymptotic contribution of a group in state s to the growth rate of the population (see eqs. B-5–B-6 of Appendix S2 and [34]) is obtained in the DGR scenario as $v(s) = w_p(s) + w_d(s)$, where the fitness functions are evaluated at the phenotypic value of the resident ($x_\bullet = x_0^R = x_1$ and $y_\bullet = y_0^R = y_1$), which gives

$$v(s_1) = 1 + (1-h), v(s_2) = 1, v(s_3) = 1, \text{and} v(s_4) = h, \quad (13)$$

and where the average of these group reproductive values over the demographic events is equal to one.

Probability of identity-by-descent. Having an explicit expression for the fitness function w to substitute into eq. 1, it remains to evaluate the average probability of identity-by-descent R^R between two individuals sampled with replacement from the same group, which is carried out by using standard methods [29,47–49] applied to the island model with warfare [50]. We have

$$R^R = \frac{1}{N} + \left(\frac{N-1}{N}\right)R, \quad (14)$$

where R is the probability of identity between two individuals sampled without replacement from the same group and that satisfies the recursion

$$R = \left[1 - \phi\omega + \phi\omega\{h^2 + (1-h)^2\}\right](1-m)^2 R^R, \quad (15)$$

where ϕ and ω are, respectively, the fighting and winning probabilities in a population monomorphic for belligerence and bravery. Eq. 15 can be understood as follows. With probability $1 - \phi\omega$ the focal group where the two individuals are sampled has not been defeated, in which case the probability of identity between the two individuals is $(1-m)^2 R^R$: with probability $(1-m)^2$ the two offspring descend from the same group, in which case they descend from the same individual with probability $1/N$ and then carry identical genes or they descend from two different individuals with probability $1 - 1/N$ and then carry identical genes with probability R. With probability $\phi\omega$ the focal group has been defeated by another group in which case the two individuals descend from the same group before regulation with probability $h^2 + (1-h)^2$ (with probability h^2 from the defeated group and with probability $(1-h)^2$ from the victorious group), and their probability of identity is then equal to $(1-m)^2 R^R$ [50].

After simplification, eq. 15 reduces to

$$R = (1-m)^2[1 - 2\phi\omega h(1-h)]R^R \quad (16)$$

and on substitution of eq. 14 and solving the for R gives

$$R = \frac{(1-m)^2[1 - 2\phi\omega h(1-h)]}{N - (N-1)(1-m)^2[1 - 2\phi\omega h(1-h)]}, \quad (17)$$

which further yields

$$R^R = \frac{1}{N - (N-1)(1-m)^2[1 - 2\phi\omega h(1-h)]}. \quad (18)$$

Results

Defeated group repopulation

Belligerence. Substituting the direct fitness function (eq. 12) and the probability of identity-by-descent (eq. 18) into eq. 1, using belligerence as the focal trait ($z = x$), eq. 3 to describe the phenotype of subadults, holding bravery constant ($y_\bullet = y_0^R = y_1 = y$), and applying the chain rule at the neutrality point ($x_\bullet = x_0^R = x_1 = x$), the selection gradient on belligerence can be expressed as

$$S = \frac{-C_x'}{1 - C_x}\left(1 - (1-m)^2\left[1 - \phi\omega + \phi\omega h^2 + \phi\omega(1-h)^2\right]R^R\right) + \frac{\partial\phi_0}{\partial x_0^A}\frac{\partial x_0^A}{\partial x_0^R}\omega(1-h)(1-m)R^R, \quad (19)$$

where $C_x' \equiv dC_x(x_\bullet)/dx_\bullet$. Using $\partial x_0^A/\partial x_0^R = (1 - \epsilon m)$, one has

$$S = \frac{-C_x'}{1 - C_x}\left(1 - (1-m)^2[1 - 2\phi\omega h(1-h)]R^R\right) + \phi'\omega(1-h)(1-\epsilon m)(1-m)R^R, \quad (20)$$

where $\phi' \equiv \partial\phi_0/\partial x_0^A$ is the derivative of $\phi(\cdot,\cdot)$ with respect to its first argument, which, when using eq. 4, becomes $\phi' = \partial a(x_0^A) / \partial x_0^A \times [1 - \exp(-a(x))]/a(x)$.

Because $(1-m)^2[1 - 2\phi\omega h(1-h)]R^R$ is equal to R (eq. 15), the selection gradient reduces to

$$S = \frac{-C_x'}{1 - C_x}(1 - R) + \phi'\omega(1-h)(1-\epsilon m)(1-m)R^R. \quad (21)$$

Setting $S = 0$, the candidate ES level of belligerence then satisfies the equation

$$\frac{C_x'}{1 - C_x} = \phi'\omega(1-h)\left[\frac{(1-\epsilon m)R_p}{1 - R}\right], \quad (22)$$

where $R_p = (1-m)R^R$ is the average probability of identity between a gene sampled in a philopatric individual and a homologous gene sampled from a neighbour [$R_p = (1-m)(1/N + (N-1)R/N)$], which can also be interpreted as the probability of identity between a gene sampled in a focal individual and one in an individual sampled at random from the same group in the parental generation.

The left member of eq. 22 is the marginal decrease $C_x'/(1-C_x)$ in fitness stemming from a single individual increasing its level of belligerence as a result of expressing a mutant allele relative to that of expressing a resident allele. The term $\phi'\omega$ in the right member of eq. 22 is the marginal increase in the probability of conquering another group (change in contest probability times winning probability), which results from an individual expressing belligerence, and it tunes the fitness benefit of defeating an attacked group.

The increase in fitness stemming from defeating a group is given by $1-h$ in eq. 22, which can be thought of as the average increase in the reproductive value of a victorious group, and is weighted by a measure of relatedness $(1-\epsilon m)R_\mathrm{p}/(1-R)$ between group members, which depends on the model's parameters and functions (m, N, ϕ, ω, and h). The term $1-R$ in the denominator reflects the reduction in the genetic variance in the population because interactions occur in a spatially structured population and $(1-\epsilon m)R_\mathrm{p}$ reflects the increase in the variance between groups. Eq. 22 shows that the more individuals are related within groups, the higher the level of belligerence.

When the right-hand-side of eq. 22 is higher than the left-hand-side, a higher level of belligerence is selected for, because the relatedness times the fitness benefits exceeds the fitness costs. When the right-hand-side of eq. 22 is lower than the left-hand-side, a lower level of belligerence is selected for. At an equilibrium point both sides of eq. 22 balance each other out. Substituting the explicit value of R (eq. 17) into eq. 22, this equilibrium is given by

$$\frac{C_\mathrm{x}'}{1-C_\mathrm{x}} = \frac{\phi'\omega(1-h)(1-\epsilon m)(1-m)}{N\left[1-(1-m)^2\{1-2\phi\omega h(1-h)\}\right]}. \quad (23)$$

When migration is low, the right member is large provided provided N is not too big, in which case the selection pressure on costly belligerence is high. When migration increases, the right member decreases, which decrease the selection pressure on costly belligerence.

A special case of eq. 23 will be useful for comparing different results. This is the weak warfare benefit limit; namely, when h becomes close to one but nevertheless remains smaller than one ($h<1$), in which case eq. 23 reduces to

$$\frac{C_\mathrm{x}'}{1-C_\mathrm{x}} = \frac{\phi'\omega(1-h)(1-\epsilon m)(1-m)}{N\left[1-(1-m)^2\right]}, \quad (24)$$

where terms of order $(1-h)^2$ and of higher order in the right member of eq. 23 have been neglected.

Bravery. Substituting eq. 12 and eq. 18 into eq. 1, using bravery as the focal trait ($z=y$), holding belligerence constant ($x_\bullet=x_0^\mathrm{R}=x_1=x$), and following similar calculations as above, the selection gradient on bravery can be written as

$$\begin{aligned} S = &-\frac{C_\mathrm{y}'}{1-C_\mathrm{y}}\left(1-(1-m)^2[1-2\phi\omega h(1-h)]R^\mathrm{R}\right)\\ &+(1-m)\phi\left[-(1-h)\frac{\partial v_1}{\partial y_0^\mathrm{A}}\frac{\partial y_0^\mathrm{A}}{\partial y_0^\mathrm{R}}+(1-h)\frac{\partial v_0}{\partial y_0^\mathrm{A}}\frac{\partial y_0^\mathrm{A}}{\partial y_0^\mathrm{R}}\right]R^\mathrm{R}, \end{aligned} \quad (25)$$

where $C_\mathrm{y}' \equiv \mathrm{d}C_\mathrm{y}(y_\bullet)/\mathrm{d}y_\bullet$, ϕ is the value of the fighting probability in a population monomorphic for the resident trait value. Using $\partial y_0^\mathrm{A}/\partial y_0^\mathrm{R}=(1-\epsilon m)$, $\partial v_0/\partial y_0^\mathrm{A}=-\partial v_1/\partial y_0^\mathrm{A}=N\omega(1-\omega)g'/g$, where $g'\equiv \mathrm{d}g(u)/\mathrm{d}u$, and defining

$$v' \equiv \frac{N\omega(1-\omega)g'}{g}, \quad (26)$$

which is the derivative of eq. 5 with respect to its first argument, the selection gradient can be further simplified to

$$S = -\frac{C_\mathrm{y}'}{1-C_\mathrm{y}}(1-R)+2\phi v'(1-h)(1-\epsilon m)(1-m)R^\mathrm{R}. \quad (27)$$

At an evolutionary rest point $S=0$, the candidate ES level of bravery then satisfies

$$\frac{C_\mathrm{y}'}{1-C_\mathrm{y}} = \phi v'2(1-h)\left[\frac{(1-\epsilon m)R_\mathrm{p}}{1-R}\right]$$

$$= \frac{\phi v'2(1-h)(1-\epsilon m)(1-m)}{N\left[1-(1-m)^2\{1-2\phi\omega h(1-h)\}\right]}, \quad (28)$$

which is similar in form to eq. 22. The left member of eq. 28 is the marginal decrease $C_\mathrm{y}'/(1-C_\mathrm{y})$ in fitness and $\phi v'$ is the increase in the probability of conquering another group, which both stem from an individual expressing a mutant bravery allele. The benefit in the right member of eq. 28 takes the same interpretation as that in eq. 23 but is twice as large because bravery also increases the probability that the focal group wins a battle when attacked by another one and is thus not taken over.

Bravery in groups of very large size. The marginal change in the wining probability given by eq. 26 has the interesting feature of increasing with group size, N, which suggests that, depending on the shape of the function g describing the power of contestants, the ES condition of bravery may become independent of groups size (N cancels out in the right member of eq. 28). Individually costly bravery may then evolve under very large group size. In order to see when this might be the case, it is useful to note that there are at least two meaningful ways to relate effort committed into fighting to the power function g [14,45,46]. The first is to assume that the winning probability predicts contest outcome from the ratio of each side's effort e committed into fighting, in which case one can use $g(e)=e^\gamma$, where γ scales the decisiveness of fighting effort disparities [14,45]. The second is to assume that the winning probability predicts the outcome of a battle from the difference of each side's effort committed into fighting, a situation that is described by the function $g(e)=\exp(\gamma e)$, which leads to the contest success function being logistic. This difference form of the contest success function characterizes a situation where the outcome of a battle depends critically on just a little difference between the opponents force [45].

Using the ratio form of the contest success function gives $g'=\gamma(yN)^{\gamma-1}$ so that $v'=\gamma\omega(1-\omega)/y$, in which case the selection pressure on bravery decreases with group size and vanishes when groups become very large. For the difference form of the contest success function, one has $g'/g=\gamma$ so that $v'=\gamma N\omega(1-\omega)$, in which case the selection pressure on bravery becomes independent of N as this parameter cancels out in eq. 28. It follows that bravery may now evolve in groups of any size.

Relation to previous analytical models. An analytical model of selective groups extinction was developed in [19] (see also [51]), where altruists within groups increase the probability that a group wins a battle during pairwise contests. In this previous model, a parameter κ is used, which is defined as the likelihood that groups engage in a contest and is equivalent to the fighting probability, ϕ, whereby we can set $\kappa=\phi$. [19] also introduces a parameter λ_A, which is the change in a group's survival probability due to its members expressing altruism should a contest occur, so that one can write $\lambda_\mathrm{A}=v'$. The cost c in [19] is also equivalent to $C_\mathrm{y}'/(1-C_\mathrm{y})$ here and since complete repopulation of defeated

groups was assumed in [19], one can also set $h = 0$. With all these stipulations, we can write eq. 28 as $c = \kappa\lambda_A(1-\epsilon m)R_p/(1-R)$, which agrees with eq. 6 of [19] in the absence of reproductive leveling ($\tau = 0$) and if one makes the additional assumption that warfare occurs after dispersal, in which case $\epsilon = 1$ and $(1-\epsilon m)R_p = R$ in eq. 28 if $h = 0$. The result of [19] thus agrees qualitatively with those derived above.

Eq. 28 is also related to eq. 3.3 of [22], where the cost C_a is given here by $C_y'/(1-C_y)$, while the change ωB_w in winning probability is given here by v'. With this, we can write eq. 28 as $C_a/(\phi\omega B_w) = 2(1-h)(1-\epsilon m)R_p/(1-R)$. The right hand side of eq. 3.3 of [22] can be simplified to $4(1-h)R_p/(1-R)$ [50], so that this eq. 3.3 reads $C_a/(\omega B_w) = 4(1-h)R_p/(1-R)$. The difference in the factor 2 between the two results stems from the fact that in our earlier analysis we assumed a two-sex rather than a one-sex model and assumed that only males expressed belligerence, while both sexes benefited from it so that the cost of expressing belligerence is halved relative to the present model.

In addition, the term ϕ appearing in the denominator of the left-hand side of eq. 28 is missing in $C_a/(\omega B_w) = 4(1-h)R_p/(1-R)$ because we assumed previously [22] that bravery was costly only conditional on there being a contest between groups, which cancels the ϕ in the denominator of eq. 28. Finally, the absence of the term $1-\epsilon m$ in eq. 3.3 of [22] shows that our earlier model involves warfare occurring before the dispersal of juveniles ($\epsilon = 0$), although we described in our life cycle section that warfare occurred after the dispersal of juveniles. Hence, in order to match the life cycle description, the benefit ωB_w must be taken as being of order $1-m$ and the same reasoning applies to the invasion condition of belligerence [22, eq. 3.1]. Alternatively, our earlier model can be interpreted as a model of warfare before the dispersal of offspring.

These considerations show that DGR demographic scenario analyzed here is consistent with previous work endorsing closely related assumptions. The main difference with previous work so far is that the qualitative results obtained earlier apply to warfare occurring before and after the dispersal of subadults (the warriors in the model) and the winning and fighting probabilities have been made more mechanistic (e.g., eq. 4 and eq. 5). Further, making these functions more mechanistic has revealed that, depending on the form of the contest success function, costly bravery may actually evolve in large-scale societies and that the fighting probabilities are likely to depend on the level of belligerence in other groups from the population, which was not clear from previous work.

Victorious group size expansion

Under the VGE scenario, the selection gradients on belligerence and bravery involves more terms because the reproductive value of offspring settling in groups facing different demographic events must be taken into account (compare eq. 1 to eqs. B-1–B-2 of Appendix S2). In addition, the stochastic process determining the probabilities of occurrences of the demographic events (s_1, s_2, s_3, and s_4) of a focal group depends on its size and the sizes of other groups because the contest success functions depend on the sizes of pairs of opposing groups in that case (eqs. B-14–B-15 of Appendix S2). With all these complications, I was unable to obtain explicit analytical expressions for the selection gradients on belligerence and bravery for all parameter values.

Nevertheless, it was observed in numerical work that the stationary probabilities of occurrences of the demographic states often take values very close to that of independently distributed probabilities of occurrences across generations, which are the stationary values of the demographic process when h is equal to

one. Hence, an approximation for this model is to consider that h takes values close to one (a defeated groups keeps a large fraction of its resources), in which case one can obtain analytical expressions for the selection gradients on belligerence and bravery for the VGE model, which are now detailed.

Belligerence. Assuming that the fitness benefits of defeating other groups are small, we can evaluate the selection strengths on belligerence to the first order in h (i.e., neglecting terms of order $(1-h)^2$ and higher order; eqs. B-23–B-32 of Appendix S2). The candidate ES level of belligerence is then found to satisfy

$$\frac{C_x'}{1-C_x} = \phi'\omega\left[(1-h)\left(1-(1-m)^2\right)\right]\left[\frac{(1-\epsilon m)R_p}{1-R}\right], \quad (29)$$

where the left member has the same form as that in eq. 22. The term in square brackets in the right member of eq. 29 measures the average increase in a group's reproductive value from defeating another group, where the average is over the cases where the focal group is and is not defeated locally by another group. The second term in the right member of eq. 29 is a measure of relatedness, which has the same interpretation as that in the DGR scenario (eq. 22) and which is given by

$$\frac{(1-\epsilon m)R_p}{1-R} = \frac{(1-\epsilon m)(1-m)}{\left(1-(1-m)^2\right)N} \quad (30)$$

(eq. B-38 of Appendix S2), whereby eq. 29 reduces to

$$\frac{C_x'}{1-C_x} = \frac{\phi'\omega(1-h)(1-\epsilon m)(1-m)}{N}. \quad (31)$$

Comparing eq. 31 to eq. 24 highlights that the demographics benefits of belligerence are generally lower under the VGE scenario than under the DGR scenario and that this difference decreases when the level of migration, m increases. The selection gradients on belligerence for the two demographic situations become equivalent when migration becomes very strong (set m close to one in eq. 24 and eq. 31). The difference between the two selection gradients for small m values stems from the fact that under the DGR model the benefits of obtaining additional resources from other groups are partially cancelled out as this also increases local competition. Indeed, an increase in the carrying capacity of victorious groups implies that there are more individuals competing to gain access to adulthood in such groups in the next generation. This is why the benefit of conquering other groups in eq. 29, $(1-h)\left(1-(1-m)^2\right)$, decreases as the level of philopatry increases.

The probability of identity R_p is higher under the VGE scenario than under the DGR scenario for most of the range of parameter values (Fig. 2). This stems from the fact that under the VGE model, defeated groups are repopulated by individuals from two groups, which increases the diversity within groups unless $h = 0$ in which case groups are completely repopulated so that the relatedness within groups is not affected by the repopulation event. But by comparing eq. 31 to eq. 24, we see that the selective pressure on belligerence under the DGR scenarios is generally larger than that under the VGE scenario.

Bravery. Assuming that the benefits of defeating other groups are small (e.g., h is close to one), the candidate ES level of bravery is found to satisfy
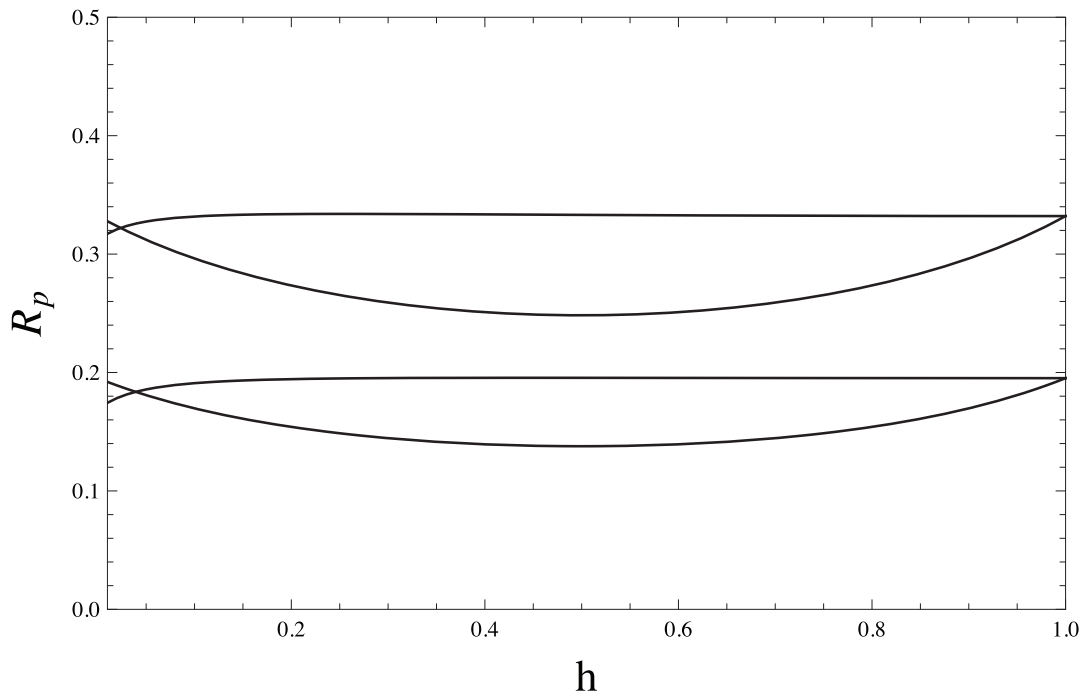
Figure 2. Graph comparing the probability of identity R_p in the DGR scenario to that obtained in the VGE scenario. The quasi flat line is the the value of R_p for the VGE scenario obtained without making any assumption on the value h can take, while the downward bent line is the relatedness for the DGR scenario. The first set of two lines in the graph is for $N = 10$, while the second set of two lines is for $N = 20$, while the other parameter values are $m = 0.1$, $\phi = 0.5$, $\omega = 0.5$, and $g(e) = e^{\gamma}$, where the value γ takes does not affect R_p.
doi:10.1371/journal.pone.0021437.g002

$$\frac{C_y'}{1 - C_y} = \frac{\phi v' 2 (1 - h)(1 - \epsilon m)(1 - m)}{N} \qquad (32)$$

(eq. B-43 of Appendix S2). Eq. 32 is similar in form to the selection gradient on bravery under the DGR model (eq. 22). But here again, comparing these two equations shows that the demographic benefits of bravery are generally lower under the VGE scenario than under the DGR scenario, which is again due to the fact that the benefits of obtaining resources from other groups are partially cancelled out as they also increase local competition.

Stronger warfare benefits. The analytical results presented in the last two section (eq. 31 and eq. 32) suggest that the selective pressure on belligerence and bravery is generally stronger under the DGR than under the VGE scenario, especially when migration is limited. However, these analytical results rely on the assumption that the benefits of warfare are small and one may wonder if these results still hold if the benefits of warfare are strong; that is, if h is low.

Fig. 3 compares the selective pressure on belligerence under the VGE scenario to that obtained under the DGR scenario without making any assumption on the values h can take. Fig. 3 illustrates that the selective pressure on belligerence is always larger under the DGR than under the VGE scenario for all values of h. This suggests that the qualitative results obtained by assuming large h values hold more generally, and that the weak warfare benefit approximation (large h values) is actually often good for values as high as $h = 1/2$, and the same results apply to the evolution of bravery. Similar patterns were observed for the selection pressure on bravery and using both the difference and the ratio forms of the contest success function.

Discussion

Evolutionary models of social interactions occurring within and between groups of individuals have often emphasized the benefits and conflicts associated with resource exchange within groups [52–56]. But since resources come in total finite supply in a population, struggle and conflict between groups is also likely to be an evolutionary outcome [4,6,7,10,16,19]. Here, the evolutionary dynamics of belligerence and bravery in the island model of warfare has been analyzed under two distinct but complementary types of demographic benefits that can drive group conflict: defeated group repopulation and victorious group size expansion.

The analysis of the models shows that the condition under which both belligerence and bravery spread in the population under the two demographic scenarios depends on four quantities (see eq. 22, eq. 28, eq. 29, and eq. 32). The first quantity is the relative marginal cost to a focal individual of expressing the trait under scrutiny [$C_x'/(1 - C_x)$ for bravery and $C_y'/(1 - C_y)$ for belligerence]. The second is the increase in the probability of defeating another group from the population ($\phi' \omega$ for belligerence and $\phi v'$ for bravery). The third quantity is the increase in the reproductive value of a group, which defeats another one, relative to that of not defeating it ($1 - h$ for the DGR scenario and $(1 - h)[1 - (1 - m)^2]$ for the VGE scenario). This quantifies the demographic benefits of belligerence and bravery, which will affect the selection gradients on both traits proportionally to the relatedness between group members. The higher the relatedness within groups, the higher the reproductive value of a group belongs to the gene lineage carried by a focal individual expressing belligerence or bravery. The fourth quantity is thus a measure of relatedness between group members [$(1 - \epsilon m) R_p / (1 - R)$]. Kinship, therefore, not only plays its classical effect in determining the
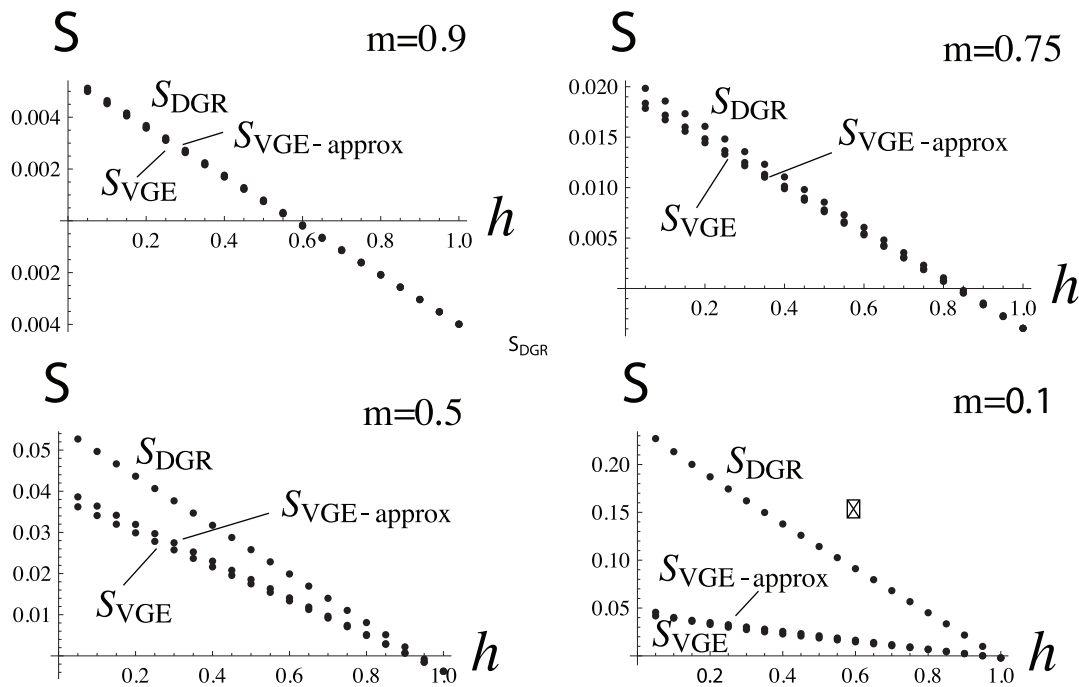
Figure 3. Selection gradient S on belligerence for the DGR and VGE scenarios as a function of h. The functional relationships used are $C_x(x) = x^4$, ϕ given by eq. 4 with $a(x) = x$, and $g(e) = e^\gamma$, and the resident trait value was set to $x = 1$. The parameter values are $N = 5$, $\omega = 0.5$, and the value γ takes does not affect the selection pressure on belligerence. Further, $m = 0.9$ in the top right panel, $m = 0.75$ in the top left panel, $m = 0.5$ in the lower right panel, and $m = 0.1$ in the lower left panel. The top line in each panel is eq. 21; that is, the selection gradient for the DGR scenario and denoted by S_{DGR}. The second line in each panel is $S_f + S_{Pr}$ with eq. B-18 of Appendix S2 using $z = x$ and eq. B-20 of Appendix S2; the selection gradient for the VGE without making any assumption on the value h can take, which is denoted by S_{VGE}. The last line in each panel is eq. B-36 of Appendix S2; the selection gradient for the VGE assuming that h is small and is denoted by $S_{VGE-approx}$. Three observations follow from this figure, which were also observed under a wider range of numerical exploration. First, for strong migration rates all three selective pressures agree. Second, for low migration rates the selective pressure is stronger under the DGR than under the VGE scenario for all values of h. Third, the small h approximation of the VGE scenario underestimates the strength of selection on belligerence when migration is low, but the approximation works quiet will for $h > 0.5$. Similar results were observed if the contest success function is of the difference form.
doi:10.1371/journal.pone.0021437.g003

individual's sacrifice in the promotion of its group [6,53,55,56], but it also markedly affects the tendency of individuals within groups to try to take over other ones.

The increase in the probability of defeating other groups from the population ($\phi' \omega$ and $\phi v'$) depends on the mechanistic details underlying the interactions between groups (e.g., contest success functions, how attack rates convert into fighting probabilities). Different ecological and environmental situations may result in different functional relationships mapping belligerence and bravery into fighting and winning probabilities. The strength of selection on bravery turns out to critically depend on such mechanistic details because effort committed into fighting may translate into very different wining probabilities [14,45,46]. For the case where contest outcome depends on the ratio of each side's fighting effort, the selection pressure on bravery decreases with group size and vanishes in large groups. By contrast, when contest outcome depends on the difference of each side's fighting effort, which may be a common situation [14], the selection pressure on bravery turns out to be independent of group size. This is interesting because it allows for the evolution of costly bravery in large-scale societies, without the need of invoking any special forces for that case.

How different are the selection on belligerence and bravery under the two demographic scenarios? It is worthwhile to emphasize at this point that in order to be initially able to compare the DGR to the VGE scenario, I assumed that the demographic benefits of warfare accruing to a focal group are

comparable in magnitude. Although the demographic processes underlying these two models are different, the reproductive value $v(s)$ of a group in demographic state s, which provides the far distant future contribution of a group in state s to the growth rate of the population (see eq. B-5 of Appendix S2 and [34,57]), are exactly the same for the two scenarios when the migration rate of juveniles is complete and the distribution of demographic states is independent across generations [$v(s_1) = 1 + (1 - h), v(s_2) = 1, v(s_3) = 1, v(s_4) = h$, eq. 13 and eq. B-12 of Appendix S2]. In this case, the long-term fitness benefits accruing to an individual in a group that has won a war is exactly the same under the two scenarios, but the selective pressure on belligerence and bravery depends on the four quantities discussed above. It is thus a priori not clear which scenario is more conducive to belligerence and bravery, and by how much the selection pressure on the traits will differ between them.

Under the life-cycle assumptions used in this paper, the main quantitative difference in the selective pressure on belligerence and bravery under the two scenarios stems from the fitness benefits of defeating other groups and relatedness taking different values. Under the DGR scenario, the fitness benefits of defeating another group; namely, the increase in group reproductive value, is equal to $1 - h$, but this value is lower under the VGE and equal to $(1 - h)\left[1 - (1 - m)^2\right]$ when the benefit of warfare are small. Hence, the benefit of defeating a group decreases as migration decreases under the VGE model. The main reason for this difference is that when the benefits of warfare increase local

carrying capacity and there is small migration, a local increase in group size does not contribute much to the far distant future growth rate of the population because the benefits of warfare are not exported but increase competition locally, which reduces group reproductive value (eq. B-29 of Appendix S2). That local competition between individuals within groups can partially or even completely inhibit the benefits obtained by expressing social behaviors is a classical result [29,42,56,58,59]. By contrast, under the DGR scenario, groups contribute more to the far distant future growth rate of the population by repopulating other groups because this is a form of exportation of the demographics benefits of belligerence and bravery, which are not destroyed locally through an increase in competition. Another way to interpret this is to say that competition is more of the "hard" type under the DGR scenario than under the VGE scenario [60,61], and thus more conducive to selection on costly social behaviors.

The relatedness coefficient tends to be smaller under the DGR scenario than under the VGE scenario, which now tips the balance in favor of stronger selection on belligerence and bravery under the VGE scenario. This follows from the fact that genetic mixing occurs during repopulation, which decreases the probability of identity within groups, unless groups are completely repopulated (Fig. 2). By contrast, no such mixing occurs under the VGE scenario, where groups remain the same unit after warfare, although an increase in groups size due to gaining more resources also tends to decrease the relatedness within groups because coalescence within groups (probability that two individuals descend from the parent) decreases (Fig. 2). A consequence is that the difference in the relatedness between the two scenarios is not so strong (Fig. 2).

Combining the effect of relatedness and the fitness benefits of defeating other groups on the selective pressure suggests that selection is stronger on both traits under the DGR than under the VGE scenario. From an evolutionary perspective, this suggests that the possibility to repopulate defeated groups by fission of a victorious one or by its members fertilizing the females of defeated groups leads to higher fitness benefits than the gain of additional resources by conquest, which increase local carrying capacity.

The two demographic scenarios of warfare analyzed here may also occur under different ecological conditions. The DGR scenario may be more consistent with a situation where groups live on or around resources that are concentrated and have static positions in space and time (such as patchily distributed resources, water points, or shelters) and fight over them for monopolization. This may be the case in social insects [1] or in hominids living in arid or semiarid environments, where population density is low and water holes were often the main cause of competition [12,23]. The DGR scenario may also be consistent with situations where bands of males aimed at gaining additional mates by raiding and mating with female from other groups. By contrast, the VGE scenario is more likely to represent a situation where groups fight for animal resources, which were important for meat, clothing, and raw materials used in tool-making [12,23]. Here, the gained resources can be shared among combatants and put in use in the focal group.

As a final point, the DGR and VGE scenarios are mainly relevant for understanding the fitness benefits of warfare in pre-state societies, where interactions occur between individuals that are more likely to share a recent common ancestor than are individuals sampled at random from the population. But the role of and motivations behind belligerence and bravery in the transition from small-scale to large-scale societies remains to be better understood. In the absence of benefits to self or relatives, there should be no naturally selected motivations in the world causing an individual to lay down its life.

Supporting Information

Appendix S1 Derivation of the fighting probabilities.
(PDF)

Appendix S2 Derivation of the selective pressures on belligerence and bravery under the VGE demographic scenario.
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: LL. Performed the experiments: LL. Analyzed the data: LL. Contributed reagents/materials/analysis tools: LL. Wrote the paper: LL.

References

1. Wilson EO, Hölldobler B (1990) The Ants. Harvard: The Belkap Press of Harvard University Press.
2. Horiuchi S, Yamamura N, T A (2002) Soldier production strategy in lower termites: from young instars or old instars? Journal of Theoretical Biology 218: 195–205.
3. Goodall J (1986) The Chimpanzees of Gombe: Patterns of Behavior. CambridgeMA: Harvard University Press.
4. Crofoot M, Wrangham RW (2010) Intergroup aggression in primates and humans: the case for a unified theory. In: Kappeler PM, Silk JB, eds. Mind the Gap: Tracing the Origins of Human Universals. Berlin-Heidelberg: Springer-Verlag. pp 171–195.
5. Keeley L (1997) War Before Civilization: The Myth of the Peaceful Savage. New York: Oxford University Press.
6. Hamilton WD (1975) Innate social aptitudes in man, an approach from evolutionary genetics. In: Fox R, ed. Biosocial Anthropology. London: Malaby Press. pp 133–157.
7. Durham WH (1976) Resource competition and human aggression, part i: a review of primitive war. The Quarterly Review of Biology 51: 385–415.
8. Chagnon NA (1988) Life histories, blood revenge, and warfare in a tribal population. Science 239: 985–992.
9. Tooby J, Cosmides L (1988) The evolution of war and its cognitive foundations. Institute for Evolutionary Studies Technical Report 88-1.
10. Alexander RD (1990) Epigenetic rules and darwinian algorithms: the adaptive study of learning and development. Ethology and Sociobiology 11: 241–303.
11. Turchin P (2003) Historical Dynamics: Why States Rise and Fall. NJ: Princeton University Press.
12. Gat A (2006) War in human civilization. Oxford: Oxford University Press.
13. Gat A (2009) So why do people fight? evolutionary theory and the causes of war. European Journal of International Relations 15: 571–599.
14. Hirshleifer J (1991) The technology of conict as an economic activity. The American Economic Review 81: 130–134.
15. Hausken K (1995) The dynamics of within-group and between-group interaction. Journal of Matematical Economics 24: 655–657.
16. Hirshleifer J (2001) The Dark Side of the Force: Economic Foundations of Conict Theory. Cambridge: Cambridge University Press.
17. Hausken K (2004) Mutual raiding of production and the emergence of exchange. Economic Inquiry 42: 572–586.
18. Garfinkel MR, Skaperdas S (2007) Economics of conict: an overview. In: T S, Hartley K, eds. Handbook of Defense Economics, Volume 2. Amsterdam: North-Holland. pp 649–709.
19. Bowles S (2006) Group competition, reproductive leveling, and the evolution of human altruism. Science 314: 1569–1572.
20. Choi JK, Bowles S (2007) The coevolution of parochial altruism and war. Science 318: 636–640.
21. Smirnov O, Arrow H, Kennett D, Orbell J (2007) Ancestral war and the evolutionary origins of "heroism". The Journal of Politics 69: 927–940.
22. Lehmann L, Feldman MW (2008) War and the evolution of belligerence and bravery. Proceedings of the Royal Society of London Series B-Biological Sciences 275: 2877–2885.
23. Gat A (2000) The human motivational complex: evolutionary theory and the causes of huntergatherer fighting. part i. primary somatic and reproductive causes. Anthropological Quarterly 73: 20–34.

24. Wright S (1931) Evolution in mendelian populations. Genetics 16: 97–159.
25. Maynard-Smith J (1982) Evolution and the Theory of Games. Cambridge: Cambridge University Press.
26. Eshel I (1996) On the changing concept of evolutionary population stability as a reection of a changing point of view in the quantitative theory of evolution. Journal of Mathematical Biology 34: 485–510.
27. Geritz SAH, Metz JAJ, Kisdi E, Meszéna G (1997) Dynamics of adaptation and evolutionary branching. Physical Review Letter 78: 2024–2027.
28. Geritz SAH, Kisdi E, Meszéna G, Metz JAJ (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology 12: 35–57.
29. Rousset F (2004) Genetic Structure and Selection in Subdivided Populations. PrincetonNJ: Princeton University Press.
30. Vincent TL, Brown JS (2005) Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics. Cambridge: Cambridge University Press.
31. Ajar E (2003) Analysis of disruptive selection in subdivided populations. Bmc Evolutionary Biology 3: 22.
32. Pen I (2000) Reproductive effort in viscous populations. Evolution 54: 293–297.
33. Leturque H, Rousset F (2002) Dispersal, kin competition, and the ideal free distribution in a spatially heterogeneous population. Theoretical Population Biology 62: 169–180.
34. Rousset F, Ronce O (2004) Inclusive fitness for traits affecting metapopulation demography. Theoretical Population Biology 65: 127–141.
35. Guillaume F, Perrin N (2006) Joint evolution of dispersal and inbreeding load. Genetics 173: 497–509.
36. Lehmann L, Feldman MW, Rousset F (2009) On the evolution of harming and recognition in finite panmictic and infinite structured populations. Evolution 63: 2896–2913.
37. Lion S, Gandon S (2009) Habitat saturation and the spatial evolutionary ecology of altruism. Journal of Evolutionary Biology 22: 1487–1502.
38. Taylor PD, Frank SA (1996) How to make a kin selection model. Journal of Theoretical Biology 180: 27–37.
39. Rousset F, Billiard S (2000) A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. Journal of Evolutionary Biology 13: 814–825.
40. Rousset F (2002) Inbreeding and relatedness coefficients: what do they measure? Heredity 88: 371–380.
41. Roze D, Rousset F (2003) Selection and drift in subdivided populations: A straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics 165: 2153–2166.
42. Lehmann L, Perrin N, Rousset F (2006) Population demography and the evolution of helping behaviors. Evolution 60: 1137–1151.
43. Gat A (1999) The pattern of fighting in simple, small-scale, prestate societies. Journal of Anthropological Research 55: 563–583.
44. Kelly RC (2005) The evolution of lethal intergroup violence. Proceedings of the National Academy of Sciences of the United States of America 102: 15294–15298.
45. Hirshleifer J (1987) Conict and rent-seeking success functions: ratio vs. difference models of relative success. Public Choice 63: 101–112.
46. Skaperdas S (1996) Contest success functions. Economic Theory 7: 283–290.
47. Gandon S (1999) Kin competition, the cost of inbreeding and the evolution of dispersal. Journal of Theoretical Biology 200: 345–364.
48. Karlin S (1968) Equilibrium behavior of population genetic models with non-random mating: part ii: pedigrees, homozygosity and stochastic models. Journal of Applied Probability 5: 487–566.
49. Gillespie JH (2004) Population Genetics: a Concise Guide. Baltimore & London: Johns Hopkins.
50. Lehmann L, Feldman MW (2009) Corrigendum for the paper "war and the evolution of belligerence and bravery". Proceedings of the Royal Society B-Biological Sciences 276: 4083–4085.
51. Bowles S (2009) Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? Science 324: 1293–1298.
52. Levin BR, Kilmer WL (1974) Interdemic selection and the evolution of altruism: a computer simulation study. Evolution 28: 527–545.
53. Maynard Smith J, Szathmary E (1995) The Major Transitions in Evolution. Oxford: Oxford University Press.
54. Michod RE (1999) Darwinian Dynamics. Princeton: Princeton University Press.
55. Keller L (1999) Levels of Selection in Evolution. PrincetonNJ: Princeton University Press.
56. West SA, Griffin AS, Gardner A (2007) Evolutionary explanations for cooperation. Current Biology 17: 661–672.
57. Taylor P (1990) Allele-frequency change in a class-structured population. American Naturalist 135: 95–106.
58. Taylor PD (1992) Altruism in viscous populations - an inclusive fitness model. Evolutionary Ecology 6: 352–356.
59. Johnstone RA, Cant MA (2008) Sex differences in dispersal and the evolution of helping and harming. American Naturalist 172: 318–330.
60. Wade MJ (1985) Soft selection, hard selection, kin selection, and group selection. American Naturalist 125: 61–73.
61. Whitlock MC (2002) Selection, load and inbreeding depression in a large metapopulation. Genetics 60: 1191–1202.
62. Grinstead CM, Snell JL (1997) Introduction to Probability. ProvidenceRI: American Mathematical Society, second edition.

Appendix S1: from attack rates to pairwise fighting probabilities

In this appendix, I present the derivation of the probability that a focal group engages in a fight with another one under the assumption that each group attacks another random group from the population independently of each other. Further, during a single generation each group is assumed to engage at most into a single pairwise contest when it is in the role of a defender (see section "Fighting probability" of the main text); there is thus only a one-shot local contest when a given group is attacked.

I first consider the case where the total population consists of a finite number n_{d} of groups and then let the number of groups grow very large ($n_{\mathrm{d}} \to \infty$). With the finite number of groups assumption, the probability that a focal group is not attacked when the attack rate is a for each group in the population is $[1 - a/(n_{\mathrm{d}} - 1)]^{n_{\mathrm{d}}-1}$, where $a/(n_{\mathrm{d}} - 1)$ is the probability that a random group from the population attacks the focal group. In the limit of an infinite number of groups, the probability that the focal group is not attacked is then given by $\exp(-a)$, which is obtained by using the standard relation $\lim_{n \to \infty}(1 - x/n)^n \to \exp(-x)$. The probability that a focal group is attacked by another group and that a fight occurs when it is in the role of a defender is then $1 - \exp(-a)$.

With probability $a/(n_{\mathrm{d}}-1)$, the focal group attacks another random group from the population. When this event occurs and k other groups from the population attack the same group as the focal group, the probability that the focal group gets into a fight with the attacked group is given by $1/(1 + k)$, where $1 + k$ is the total number of groups (including the focal) attacking the same group. Hence, conditional on the focal group attacking another random group from the population, the probability that it will engage into a contest with the attacked group is given by the average of $1/(1 + k)$ over a Binomial distribution, which gives the number of simultaneous attacks on the group attacked by the focal group. This yields the average

$$q = \sum_{k=0}^{n_{\mathrm{d}}-2} \frac{1}{1+k} \frac{(n_{\mathrm{d}}-2)!}{k!(n_{\mathrm{d}}-2-k)!} \left(\frac{a}{n_{\mathrm{d}}-1}\right)^k \left(1 - \frac{a}{n_{\mathrm{d}}-1}\right)^{n_{\mathrm{d}}-2-k}, \tag{A-1}$$

where $n_{\mathrm{d}} - 2$ is the number of groups in the population without the focal group and the group it has attacked.

Using the fact that $(n_{\mathrm{d}} - 2)! = (n_{\mathrm{d}} - 1)!/(n_{\mathrm{d}} - 1)$, eq. A-1 can be written as

$$q = \frac{1}{a} \sum_{k=0}^{n_{\mathrm{d}}-2} \frac{(n_{\mathrm{d}}-1)!}{(1+k)![n_{\mathrm{d}}-1-(1+k)]!} \left(\frac{a}{n_{\mathrm{d}}-1}\right)^{1+k} \left(1 - \frac{a}{n_{\mathrm{d}}-1}\right)^{n_{\mathrm{d}}-1-(1+k)}. \tag{A-2}$$

I now make the change of variable $h = k + 1$ in the summation sign, which gives

$$q = \frac{1}{a} \sum_{h=1}^{n_{\mathrm{d}}-1} \frac{(n_{\mathrm{d}}-1)!}{h!(n_{\mathrm{d}}-1-h)!} \left(\frac{a}{n_{\mathrm{d}}-1}\right)^h \left(1 - \frac{a}{n_{\mathrm{d}}-1}\right)^{n_{\mathrm{d}}-1-h}, \tag{A-3}$$

and shows that the sum is over a Binomial distribution with probability of success $a/(n_{\mathrm{d}}-1)$ and number of draws $n_{\mathrm{d}} - 1$. Since the sum runs from $h = 1$ to $h = n_{\mathrm{d}} - 1$, the sum gives the probability that there is at least one success in $n_{\mathrm{d}} - 1$ draws, whereby

$$q = \frac{1}{a} \left[1 - \left(1 - \frac{a}{n_{\mathrm{d}}-1}\right)^{n_{\mathrm{d}}-1}\right]. \tag{A-4}$$

By letting the number of groups grow large ($n_{\mathrm{d}} \to \infty$) and using again $\lim_{n \to \infty}(1 - x/n)^n \to \exp(-x)$ the probability that a focal group, conditional on it attacking another group, enters into a contest with the attacked group is given by $[1 - \exp(-a)]/a$. Since the focal group can attack $n_{\mathrm{d}} - 1$ groups, the probability that the focal groups enters into a fight when it is in the role of an attacker in a monorophic population is $(n_{\mathrm{d}} - 1) \times a/(n_{\mathrm{d}} - 1) \times [1 - \exp(-a)]/a = 1 - \exp(-a)$.

If we now call $\phi(i,j)$ the probability that a focal group with average level of belligerence i, and thus attack rate $a(i)$, enters into a fight with another group when j is the average level of belligerence in the remaining groups in the population, which thus have attack rate $a(j)$, then the above results give

$$\phi(i,j) = a(i)\frac{1 - \exp(-a(j))}{a(j)}. \tag{A-5}$$

Appendix S2: VGE model

Selection gradient

In this appendix, I present the derivation of the selective pressures on belligerence and bravery under the VGE demographic scenario. For this scenario, the demographic events s_1, s_2, s_3, and s_4, determine the number of individuals that reach adulthood in a group after stage 4 of the life cycle. For the events s_1, s_2, s_3, and s_4, the group size that obtains is, respectively, $N + (1 - h)N, N, N$, and hN (see section "Demographic events" of the main text). A focal group is thus likely to fluctuate between different sizes across generations.

These demographic fluctuations make the calculations of the selection gradient on an evolving phenotype z more complicated. In the presence of such demographic fluctuations and if the number of juveniles produced by each individual is assumed to be large (or is Poisson distributed), the selection gradient can be expressed as

$$S = S_{\mathrm{f}} + S_{\mathrm{Pr}} \tag{B-1}$$

[1], where

$$S_{\mathrm{f}} = \sum_s \sum_{s'} \nu(s') \left[\frac{\partial f_{\mathrm{p}}(s)}{\partial z_\bullet} + \frac{\partial f_{\mathrm{p}}(s)}{\partial z_0^{\mathrm{R}}} R^{\mathrm{R}}(s) \right.$$
$$\left. + \sum_{s^*} \Pr(s^*) \left(\frac{\partial f_{\mathrm{d}}(s, s^*)}{\partial z_\bullet} + \frac{\partial f_{\mathrm{d}}(s, s^*)}{\partial z_0^{\mathrm{R}}} R^{\mathrm{R}}(s^*) \right) \right] \Pr(s' \mid s) \Pr(s) \tag{B-2}$$

and

$$S_{\mathrm{Pr}} = \sum_s \sum_{s'} \nu(s') \frac{\partial \Pr(s' \mid s)}{\partial z_0^{\mathrm{R}}} f_{\mathrm{p}}(s) R^{\mathrm{R}}(s) \Pr(s), \tag{B-3}$$

where all sums are over all elements s of the set $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$ of demographic events (see Fig. 1), and where the components of selection follow from rearranging eqs. 26–27 of [1] as given by eqs. A-48–A-49 of [2]. All partial derivatives in eqs. B-2–B-3 are evaluated at the phenotypic value of the resident allele $(z_\bullet = z_0^{\mathrm{R}} = z_1 = z)$ and all other quantities in S are also evaluated at this neutrality point.

Because the four demographic events determine the state of a focal group (its size), I will from now on, and for ease of presentation, refer to the four events $s1$, $s2$, $s3$, and $s4$, as the demographic states of a group. The size of a focal group is the main variable that will affect the components of selection given by eq. B-2 and eq. B-3, and which depend on four type of quantities, the justification of which are given in [1].

(1) Two frequency functions. First, the probability $f_{\mathrm{p}}(s)$ that a gene sampled in a group, which is in a given state s' in the offspring generation, descends from an individual from that group that was in state s in the parental generation. This probability is independent of the states of the group in the offspring generation [1], which follows from the assumption that fecundity is Poisson distributed (or very large). The second frequency function is the probability $f_{\mathrm{d}}(s, s^*)$ that a gene sampled in a group, which is in a given state s' in the offspring generation and was itself in state s in the parental generation, descends through migration from a group in state s^* in the parental generation. For completeness, I mention that these frequency functions can be related to individual fitness functions, which measure the expected number of offspring reaching adulthood of a focal parent conditional on the same demographic events described for the frequency functions, by $w_{\mathrm{p}}(s, s') = f_{\mathrm{p}}(s)N_{s'}/N_s$, $w_{\mathrm{d}}(s, s', s^*) = f_{\mathrm{d}}(s, s^*)N_{s'}/N_{s^*}$ [1], where N_s is the number of individuals in a group in demographic state s.

(2) The stationary probability $\Pr(s)$ that a focal group of individuals is in demographic state s, which is the steady-state distribution of the Markov chain with forward transitions probabilities $\Pr(s' \mid s)$, where

s is the demographic state of the group in the parental generation and s' is its state in the offspring generation. This stationary distribution satisfies

$$\Pr(s') = \sum_s \Pr(s' \mid s)\Pr(s). \tag{B-4}$$

(3) The relative reproductive value $\nu(s)$ of all individuals within a group in demographic state s, which satisfies the recursion

$$\nu(s) = \sum_{s'} \left[f_{\mathrm{p}}(s)\Pr(s' \mid s) + \sum_{s^*} f_{\mathrm{d}}(s^*, s)\Pr(s' \mid s^*)\Pr(s^*) \right] \nu(s') \tag{B-5}$$

and $\sum_s \nu(s)\Pr(s) = 1$, where $\nu(s)\Pr(s)$ is the reproductive value of all groups in state s in the population. In the special case where the demographic states are independently distributed across generations $\Pr(s' \mid s^*) = \Pr(s')$, in which case eq. B-5 can be simplified by using $\nu(s)\Pr(s)$ so that the reproductive value of a group in state s is directly given by

$$\nu(s) = f_{\mathrm{p}}(s) + \sum_{s^*} f_{\mathrm{d}}(s^*, s)\Pr(s^*). \tag{B-6}$$

(4) The probability of identity by descent $R^{\mathrm{R}}(s)$ between two homologous genes sampled with replacement in a group in state s. This probability can be expressed as

$$R^{\mathrm{R}}(s) = \frac{1}{N_s} + \left(\frac{N_s - 1}{N_s} \right) R(s), \tag{B-7}$$

where N_s is the number of individuals in a group in demographic state s and $R(s)$ is the probability of identity between two homologous genes sampled without replacement in that group. This probability satisfies the recursion

$$R(s') = \sum_s \Pr(s \mid s') f_{\mathrm{p}}(s)^2 R^{\mathrm{R}}(s), \tag{B-8}$$

where $\Pr(s \mid s')$ is the probability that a group in state s' in the offspring generation derives from a group in state s in the parental generation [backward transition probability of the demographic states $\Pr(s \mid s') = \Pr(s' \mid s)\Pr(s)/\Pr(s')$], and $f_{\mathrm{p}}(s)^2$ is the backward migration probability that a pair of genes sampled in a group in a given state s' in the offspring generation both descend from the same group that was in state s in the parental generation.

Frequency functions

For the life-cycle assumptions described in the main text, the frequency functions are given by

$$f_{\mathrm{p}}(s) = \frac{(1-m)b_{\bullet}N_s}{(1-m)\,b_{\mathrm{r}}N_s + mb_{\mathrm{d}}N_{\mathrm{eq}}} \tag{B-9}$$

and

$$f_{\mathrm{d}}(s, s^*) = \frac{mb_{\bullet}N_{s^*}}{(1-m)\,b_{\mathrm{d}}N_s + mb_{\mathrm{d}}N_{\mathrm{eq}}}, \tag{B-10}$$

where $N_{\mathrm{eq}} = \sum_s N_s\Pr(s)$ is the average group size in the population [1, eqs. 33–34] with the number of adults in the different demographics states being given by $N_{s_1} = [1 + (1-h)]\,N$, $N_{s_3} = [h + (1-h)]\,N = N$, $N_{s_2} = N$, and $N_{s_4} = hN$.

On substitution of these numbers into eq. B-9 one has

$$f_{\mathrm{p}}(s_1) = \frac{(1-m)b_\bullet[1+(1-h)]}{(1-m)\,b_{\mathrm{r}}[1+(1-h)] + m\,b_{\mathrm{d}}\,(N_{\mathrm{eq}}/N)}$$

$$f_{\mathrm{p}}(s_3) = \frac{(1-m)b_\bullet}{(1-m)\,b_{\mathrm{r}} + m\,b_{\mathrm{d}}\,(N_{\mathrm{eq}}/N)}$$

$$f_{\mathrm{p}}(s_2) = \frac{(1-m)b_\bullet}{(1-m)\,b_{\mathrm{r}} + m\,b_{\mathrm{d}}\,(N_{\mathrm{eq}}/N)}$$

$$f_{\mathrm{p}}(s_4) = \frac{(1-m)b_\bullet h}{(1-m)\,b_{\mathrm{r}}h + m\,b_{\mathrm{d}}\,(N_{\mathrm{eq}}/N)}, \tag{B-11}$$

and similar calculations lead to explicit expressions for the frequency functions for the migrant offspring $[f_{\mathrm{d}}(\cdot,\cdot)]$.

Assuming a population of resident individuals and substituting the frequency functions into eq. B-5 results in complicated expressions but they can be handled with a symbolic algebra system like Mathematica [3]. Assuming that the demographic states are independently distributed across generations, eq. B-6 still results in complicated expressions but when dispersal is complete ($m = 1$), the group reproductive values reduce to

$$\nu(s_1) = 1 + (1-h), \nu(s_2) = 1, \nu(s_3) = 1, \text{ and } \nu(s_4) = h, \tag{B-12}$$

which is similar to those obtained under the DGR scenario (eq. 13 of the main text).

Transition probabilities

In order to evaluate S, we also need the transition probabilities of the demographic states for a focal group, which, as for the DGR scenario, will depend on the fighting and winning probabilities. The fighting probabilities between groups are assumed to be the same as those in the DGR scenario (eq. 1 of the main text or eq. A-5 of Appendix S1), with the only difference that the average phenotype among subadults in the focal group after dispersal may now depend on the demographic states because the frequency of individuals in the focal group after dispersal and descending from the focal group may depend on the size of that group in the parental generation.

When warfare occurs before the dispersal of subadults, we have $x_0^{\mathrm{A}} = x_0^{\mathrm{R}}$ and $y_0^{\mathrm{A}} = y_0^{\mathrm{R}}$, but when warfare occurs after the dispersal of subadults, we have to take into account that migration has changed the average phenotype in the focal group (see eq. 2 of the main text). The average level of belligerence in the focal group when the size of the focal group in the parental generation is N_s is $x_0^{\mathrm{A}}(s) = [(1-m)N_s]/[(1-m)\,N_s + mN_{\mathrm{eq}}]x_0^{\mathrm{R}} + [mN_{\mathrm{eq}}]/[(1-m)\,N_s + mN_{\mathrm{eq}}]x_1$, where $[(1-m)N_s]/[(1-m)\,N_s + mN_{\mathrm{eq}}]$ is the probability of sampling an individual of philopatric origin in the focal group after dispersal. Note that x_1 does not depend on demographic states because any class of individuals from other groups have zero relatedness to the focal individual so that their average phenotypes can be considered to be equivalent when viewed from the focal group and taken to be equal to x_1. In order to take simultaneously into account the case where warfare occurs before and after the dispersal of subadults (e.g. eq. 3 of the main text), I write the average phenotypes of subadults in the focal group as

$$x_0^{\mathrm{A}}(s) = \frac{(1-m\epsilon)N_s}{(1-m\epsilon)\,N_s + m\epsilon N_{\mathrm{eq}}}x_0^{\mathrm{R}} + \frac{m\epsilon N_{\mathrm{eq}}}{(1-m\epsilon)\,N_s + m\epsilon N_{\mathrm{eq}}}x_1$$

$$y_0^{\mathrm{A}}(s) = \frac{(1-m\epsilon)N_s}{(1-m\epsilon)\,N_s + m\epsilon N_{\mathrm{eq}}}y_0^{\mathrm{R}} + \frac{m\epsilon N_{\mathrm{eq}}}{(1-m\epsilon)\,N_s + m\epsilon N_{\mathrm{eq}}}y_1, \tag{B-13}$$

where the parameter ϵ is set to zero when warfare occurs before the dispersal of offspring and set to one if it occurs after the dispersal of offspring.

4

With probability $\phi_0(s) \equiv \phi(x_0^{\mathrm{A}}(s), x_1)$ the subadults in a focal group, which was in demographic state s in the parental generation, enter into a fight with another group by attacking. The focal group is assumed to win the ensuing battle with probability

$$v_0(s) = \sum_{s^*} \Pr(s^*) \frac{\omega g\big(y_0^{\mathrm{A}}(s)N_s\big)}{\omega g\big(y_0^{\mathrm{A}}(s)N_s\big) + (1-\omega)g(y_1 N_{s^*})}, \tag{B-14}$$

which is an average over all possible contests faced by the individuals in the focal group at steady-state. This equation is the direct extension of eq. 5 of the main text to the case where opposing groups vary in size, and is then averaged over the distribution $\Pr(s^*)$ of demographic states of the attacked group evaluated in the neutral process. Taking such an average can be justified by the fact that owing to the infinite island model assumptions, groups affect each other in a nonstochastic manner, only through limiting densities, which can be taken as those of the focal deme itself [1, 4].

With probability $\phi_1 \equiv \phi(x_1, x_1)$ a focal group in demographic state s is attacked by another group from the population, in which case the probability that the attacking group wins the ensuing battle is

$$v_1(s) = \sum_{s^*} \Pr(s^*) \frac{\omega g(y_1 N_{s^*})}{(1-\omega)g\big(y_0^{\mathrm{A}}(s)N_s\big) + \omega g(y_1 N_{s^*})}, \tag{B-15}$$

which is an average over all demographic states of the attacker group.

With these notations, the forward transition probabilities, from a focal group in demographic state s in the parental generation to any of the four other states in the offspring generation are given by

$$\begin{aligned}
\Pr(s_1 \mid s) &= [1 - \phi_1 v_1(s)]\,\phi_0(s)v_0(s) \\
\Pr(s_2 \mid s) &= \phi_1 v_1(s)\phi_0(s)v_0(s) \\
\Pr(s_3 \mid s) &= [1 - \phi_1 v_1(s)][1 - \phi_0(s)v_0(s)] \\
\Pr(s_4 \mid s) &= \phi_1 v_1(s)[1 - \phi_0(s)v_0(s)].
\end{aligned} \tag{B-16}$$

The stationary distribution (eq. B-4) induced by this Markov chain is difficult to analyze because the transition probabilities themselves depend on the stationary distribution (e.g., eq. B-14) so that the stationary distribution it is only implicitly determined and cannot be evaluated by using linear algebra. This is a non-homogeneous Markov chain that is homogeneous only asymptotically.

Selection: effect on settled offspring number S_{f}

The frequency functions (eq. B-9 and eq. B-10) describe a model of local group size fluctuations with Wright-Fisher reproduction [1,2]. For this case, it has been shown that the first component of the selection gradient on a phenotype z, S_{f}, can be simplified and expressed solely in terms of the perturbations of the relative fecundity of a focal individual (number of juveniles produced before a competition stage) as

$$S_{\mathrm{f}} = \sum_s \frac{1}{b}\left(\frac{\partial b_\bullet}{\partial z_\bullet} + \frac{\partial b_\bullet}{\partial z_0^{\mathrm{R}}}\frac{1}{N_s}\right)\nu(s)\left(1 - R(s)\right)\Pr(s) \tag{B-17}$$

[2, eq. 21 of the Appendix].

In the present model, we further have $\partial b_\bullet/\partial z_0^{\mathrm{R}} = 0$, whereby

$$S_{\mathrm{f}} = \frac{1}{b}\frac{\partial b_\bullet}{\partial z_\bullet}\sum_s \nu(s)\left(1 - R(s)\right)\Pr(s), \tag{B-18}$$

which can be used to evaluate both the selection gradient on belligerence and bravery by using $b_\bullet = [1 - C_{\mathrm{x}}(x_\bullet)]\,[1 - C_{\mathrm{y}}(y_\bullet)]$.

Selection: effect on offspring reproductive value S_{Pr}

Belligerence

Using eq. B-13 to describe the phenotype of subadult individuals and holding bravery constant ($y_0^{\mathrm{R}} = y_1 = y$), the perturbations of the transition probabilities (eq. B-16) due to expressing belligerence and evaluated at the neutrality point ($x_0^{\mathrm{R}} = x_1 = x$) are given by

$$
\begin{aligned}
\frac{\Pr(s_1 \mid s)}{\partial x_0^{\mathrm{R}}} &= \frac{\partial \phi_0(s)}{\partial x_0^{\mathrm{R}}} v_0(s) \left[(1 - \phi v_1(s)) \right] \\
\frac{\Pr(s_2 \mid s)}{\partial x_0^{\mathrm{R}}} &= \frac{\partial \phi_0(s)}{\partial x_0^{\mathrm{R}}} v_0(s) \phi v_1(s) \\
\frac{\Pr(s_3 \mid s)}{\partial x_0^{\mathrm{R}}} &= -\frac{\Pr(s_1 \mid s)}{\partial x_0^{\mathrm{R}}} \\
\frac{\Pr(s_4 \mid s)}{\partial x_0^{\mathrm{R}}} &= -\frac{\Pr(s_2 \mid s)}{\partial x_0^{\mathrm{R}}}.
\end{aligned}
\tag{B-19}
$$

Substituting these expressions into eq. B-3 and rearranging produces

$$
S_{\mathrm{Pr}} = \sum_s f_{\mathrm{p}}(s) R^{\mathrm{R}}(s) \Pr(s)
$$
$$
\times \frac{\partial \phi_0(s)}{\partial x_0^{\mathrm{R}}} v_0(s) \Big[\{1 - \phi v_1(s)\} \{\nu(s_1) - \nu(s_3)\} + \phi v_1(s) \{\nu(s_2) - \nu(s_4)\} \Big]. \tag{B-20}
$$

Bravery

Using eq. B-13 in eq. B-16 and holding belligerence constant, the perturbations of the transition probabilities evaluated at the neutrality point can be written as

$$
\begin{aligned}
\frac{\Pr(s_1 \mid s)}{\partial y_0^{\mathrm{R}}} &= \phi \left[\{1 - \phi v_1(s)\} \frac{\partial v_0(s)}{\partial y_0^{\mathrm{R}}} - \phi v_0(s) \frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}} \right] \\
\frac{\Pr(s_2 \mid s)}{\partial y_0^{\mathrm{R}}} &= \phi^2 \left[v_1(s) \frac{\partial v_0(s)}{\partial y_0^{\mathrm{R}}} + v_0(s) \frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}} \right] \\
\frac{\Pr(s_3 \mid s)}{\partial y_0^{\mathrm{R}}} &= -\frac{\Pr(s_1 \mid s)}{\partial y_0^{\mathrm{R}}} - \phi \frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}} \\
\frac{\Pr(s_4 \mid s)}{\partial y_0^{\mathrm{R}}} &= -\frac{\Pr(s_2 \mid s)}{\partial y_0^{\mathrm{R}}} + \phi \frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}}.
\end{aligned}
\tag{B-21}
$$

Substituting these expressions into eq. B-3 and rearranging produces

$$
S_{\mathrm{Pr}} = \sum_s f_{\mathrm{p}}(s) R^{\mathrm{R}}(s) \Pr(s) \left[\{\nu(s_1) - \nu(s_3)\} \frac{\Pr(s_1 \mid s)}{\partial y_0^{\mathrm{R}}} \right.
$$
$$
\left. + \{\nu(s_2) - \nu(s_4)\} \frac{\Pr(s_2 \mid s)}{\partial y_0^{\mathrm{R}}} + \{\nu(s_4) - \nu(s_3)\} \phi \frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}} \right]. \tag{B-22}
$$

Approximate functionals

Stationary distribution

By using eq. B-16, I was unable to obtain analytical expression for the stationary demographic distribution $\Pr(s)$ under neutrality, which we need in order to evaluate eq. B-2 and eq. B-3. Without an analytical

expression of the stationary distribution there is no hope to evaluate explicitly the selection gradients on belligerence and bravery.

However, numerical exploration of the stationary distribution suggests that it is weakly affected by the parameter h when $h > 0.5$ and that the probabilities of occurrences of the demographic states take values very close to that if h were equal to one in that case. An approximation for the Markov chain described by eq. B-16 for large h values can then be obtained by carrying out a first order Taylor expansion around $h = 1$ of the transition probabilities given by eq. B-16, which was done with Mathematica [3]. Substituting the resulting transition probabilities into eq. B-4 and solving for the stationary distribution gives

$$
\begin{aligned}
\Pr(s_1) &= (1 - \phi\omega)\,\phi\omega + O\left((1-h)^2\right) \\
\Pr(s_2) &= \phi^2\omega^2 + O\left((1-h)^2\right) \\
\Pr(s_3) &= (1 - \phi\omega)^2 + O\left((1-h)^2\right) \\
\Pr(s_4) &= \phi\omega(1 - \phi\omega) + O\left((1-h)^2\right),
\end{aligned}
\tag{B-23}
$$

where $O\left((1-h)^2\right)$ is a remainder of order $(1-h)^2$. In the rest of this Appendix, I evaluate the components of the selection gradients on belligerence and bravery to the first order in $1 - h$ so that eq. B-23 can be used in order to evaluate eq. B-2 and eq. B-3 to the first order in h around $h = 1$.

Winning and fighting probabilities

A first order Taylor expansion of the fighting probability around $h = 1$ gives

$$
\begin{aligned}
\frac{\partial \phi_0(s)}{\partial x_0^{\mathrm{R}}} &= \frac{\partial \phi_0(s)}{\partial x_0^{\mathrm{A}}(s)} \frac{\partial x_0^{\mathrm{A}}(s)}{\partial x_0^{\mathrm{R}}} \\
&= (1 - \epsilon m)\phi' + O\left(1 - h\right),
\end{aligned}
\tag{B-24}
$$

where $\phi' \equiv \partial \phi_0(s)/\partial x_0^{\mathrm{A}}(s)$.

Substituting eq. B-23 into eqs. B-14–B-15 gives

$$
\begin{aligned}
\frac{\partial v_0(s)}{\partial y_0^{\mathrm{R}}} &= \frac{\partial v_0(s)}{\partial y_0^{\mathrm{A}}(s)} \frac{\partial y_0^{\mathrm{A}}(s)}{\partial y_0^{\mathrm{R}}} \\
&= (1 - \epsilon m)v' + O\left(1 - h\right),
\end{aligned}
\tag{B-25}
$$

where $v' \equiv N\omega(1 - \omega)g'/g$, and

$$
\begin{aligned}
\frac{\partial v_1(s)}{\partial y_0^{\mathrm{R}}} &= \frac{\partial v_1(s)}{\partial y_0^{\mathrm{A}}(s)} \frac{\partial y_0^{\mathrm{A}}(s)}{\partial y_0^{\mathrm{R}}} \\
&= -(1 - \epsilon m)v' + O\left(1 - h\right).
\end{aligned}
\tag{B-26}
$$

In a monormorphic population we also have

$$
\begin{aligned}
v_0(s) &= \omega + O\left(1 - h\right) \\
v_1(s) &= \omega + O\left(1 - h\right).
\end{aligned}
\tag{B-27}
$$

With these results, eq. B-21 reduces to

$$
\begin{aligned}
\frac{\Pr(s_1 \mid s)}{\partial y_0^{\mathrm{R}}} &= \phi v'(1 - \epsilon m) + O(1 - h) \\
\frac{\Pr(s_2 \mid s)}{\partial y_0^{\mathrm{R}}} &= O(1 - h) \\
\frac{\Pr(s_3 \mid s)}{\partial y_0^{\mathrm{R}}} &= O(1 - h) \\
\frac{\Pr(s_4 \mid s)}{\partial y_0^{\mathrm{R}}} &= -\phi v'(1 - \epsilon m) + O(1 - h).
\end{aligned}
\tag{B-28}
$$

Reproductive values

Substituting eqs. B-9–B-10, eq. B-16, and eq. B-23 into eq. B-5 and solving for the reproductive values provides the following first order Taylor expansion around $h = 1$ at neutrality:

$$
\begin{aligned}
\nu(s_1) &= 1 + (1 - h)\left(1 - (1 - m)^2\right) + O\left((1 - h)^2\right) \\
\nu(s_2) &= 1 + O\left((1 - h)^2\right) \\
\nu(s_3) &= 1 + O\left((1 - h)^2\right) \\
\nu(s_4) &= 1 - (1 - h)\left(1 - (1 - m)^2\right) + O\left((1 - h)^2\right),
\end{aligned}
\tag{B-29}
$$

which gives

$$
\begin{aligned}
\nu(s_1) - \nu(s_3) &= (1 - h)\left(1 - (1 - m)^2\right) + O\left((1 - h)^2\right) \\
\nu(s_2) - \nu(s_4) &= (1 - h)\left(1 - (1 - m)^2\right) + O\left((1 - h)^2\right) \\
\nu(s_4) - \nu(s_3) &= -(1 - h)\left(1 - (1 - m)^2\right) + O\left((1 - h)^2\right).
\end{aligned}
\tag{B-30}
$$

These expressions are needed to evaluate the selection pressure S_{Pr} for both belligerence and bravery (eq. B-20 and eq. B-22). Eq. B-30 illustrates that the changes in reproductive value involve terms of order $(1 - h)$. Hence, in order to evaluate S_{Pr} to the first order around $h = 1$, it is sufficient to evaluate all other functionals ($f_{\mathrm{p}}(s)$, $R^{\mathrm{R}}(s)$, $\Pr(s)$, etc.) to the zero's order as any term of order $(1 - h)$ in these functionals would result in second order terms in the selection gradient.

Probabilities of identity by descent

Substituting eq. B-10, eq. B-16, and eq. B-23 into eq. B-8 and solving for relatedness provides the following first order Taylor expansions

$$
R(s) = \frac{(1 - m)^2}{N - (N - 1)(1 - m)^2} + O\left((1 - h)^2\right)
\tag{B-31}
$$

and

$$
\begin{aligned}
R_{\mathrm{p}} &\equiv \sum_s f_{\mathrm{p}}(s) R^{\mathrm{R}}(s) \Pr(s) \\
&= \frac{(1 - m)}{N - (N - 1)(1 - m)^2} + O\left((1 - h)^2\right).
\end{aligned}
\tag{B-32}
$$

Approximate selection gradient

We now have all the elements to evaluate the selection gradients on belligerence and bravery to the first order around $h = 1$.

Selection on belligerence

Holding bravery constant and using belligerence as the evolving trait, eq. B-18 reduces to

$$S_{\mathrm{f}} = -\frac{C_{\mathrm{x}}'}{1 - C_{\mathrm{x}}}(1 - R) + O\left((1 - h)^2\right), \tag{B-33}$$

because $\sum_s \nu(s)\Pr(s) = 1$ and $R(s)$ is the same for all s (see eq. B-31) and is denoted R.

Substituting eq. B-30, eq. B-24, and eq. B-27 into eq. B-20 gives

$$S_{\mathrm{Pr}} = \sum_s f_{\mathrm{p}}(s) R^{\mathrm{R}}(s)\Pr(s)$$

$$\times (1 - \epsilon m)\phi'\omega\left[(1 - \phi\omega)(1 - h)\left(1 - (1 - m)^2\right) + \phi\omega(1 - h)\left(1 - (1 - m)^2\right)\right], \tag{B-34}$$

which can be simplified to

$$S_{\mathrm{Pr}} = \phi'\omega(1 - h)\left(1 - (1 - m)^2\right)(1 - \epsilon m)R_{\mathrm{p}} + O\left((1 - h)^2\right) \tag{B-35}$$

by using eq. B-32. The coefficient R_{p} is the average probability of identity between a gene sampled in a philopatric individual (individual that has not dispersed) and a homologous gene sampled in a neighbour.

Adding up S_{f} and S_{Pr}, and taking into account only first order terms, one obtains the inclusive fitness effect from expressing belligerence as

$$S = -\frac{C_{\mathrm{x}}'}{1 - C_{\mathrm{x}}}(1 - R) + \phi'\omega(1 - h)\left(1 - (1 - m)^2\right)(1 - \epsilon m)R_{\mathrm{p}}. \tag{B-36}$$

Setting the selective pressure to zero and neglecting terms of order $(1 - h)^2$ and of higher order, the cost-to-benefit ratio under which belligerence spreads can be written as

$$\frac{C_{\mathrm{x}}'}{1 - C_{\mathrm{x}}} = \frac{\phi'\omega(1 - h)\left(1 - (1 - m)^2\right)(1 - \epsilon m)R_{\mathrm{p}}}{1 - R}, \tag{B-37}$$

where using eqs. B-31–B-32 produces

$$\frac{R_{\mathrm{p}}}{1 - R} = \frac{(1 - m)}{\left(1 - (1 - m)^2\right)N}. \tag{B-38}$$

Selection on bravery

Following similar calculations as in the last section, we have from eq. B-18 for bravery that

$$S_{\mathrm{f}} = -\frac{C_{\mathrm{y}}'}{1 - C_{\mathrm{y}}}(1 - R). \tag{B-39}$$

Using eq. B-30 and eqs. B-26–B-28 in eq. B-20, and holding belligerence constant produces

$$S_{\mathrm{Pr}} = \sum_s f_{\mathrm{p}}(s) R^{\mathrm{R}}(s)\Pr(s)\left[(1 - h)\left(1 - (1 - m)^2\right)\phi(1 - \epsilon m)v'\right.$$

$$\left. + (1 - h)\left(1 - (1 - m)^2\right)\phi(1 - \epsilon m)v'\right] + O\left((1 - h)^2\right), \tag{B-40}$$

whereby

$$S_{\mathrm{Pr}} = \phi v' 2(1-h)\left(1-(1-m)^2\right)(1-\epsilon m)R_{\mathrm{p}} + O\left((1-h)^2\right). \tag{B-41}$$

Adding up the selective pressures one obtains

$$S = -\frac{C'_{\mathrm{y}}}{1-C_{\mathrm{y}}}(1-R) + v'2(1-h)\left(1-(1-m)^2\right)(1-\epsilon m)R_{\mathrm{p}} + O\left((1-h)^2\right). \tag{B-42}$$

Using eq. B-38, eq. B-39, and eq. B-42 in this selection gradient, neglecting the remainder, and setting $S = 0$ finally gives

$$\frac{C'_{\mathrm{y}}}{1-C_{\mathrm{y}}} = \frac{\phi v' 2(1-h)(1-\epsilon m)(1-m)}{N}. \tag{B-43}$$

References

1. Rousset F, Ronce O (2004) Inclusive fitness for traits affecting metapopulation demography. Theoretical Population Biology 65: 127-141.

2. Lehmann L, Perrin N, Rousset F (2006) Population demography and the evolution of helping behaviors. Evolution 60: 1137-1151.

3. Wolfram S (2003) Mathematica, 5th ed. Cambridge: Cambridge University Press.

4. Chesson PL (1981) Models for spatially distributed populations: the effect of within-patch variability. Theoretical Population Biology 19: 288-325.