RESEARCH ARTICLE

Fixation probabilities in graph-structured populations under weak selection

Benjamin Allen ⓘ *, **Christine Sample** ⓘ, **Patricia Steinhagen, Julia Shapiro**¤,
Matthew King, Timothy Hedspeth, Megan Goncalves

Department of Mathematics, Emmanuel College, Boston, Massachusetts, United States of America

¤ Current address: Department of Biostatistics, University of Florida, Gainesville, Florida, United States of America
* allenb@emmanuel.edu

Abstract

A population's spatial structure affects the rate of genetic change and the outcome of natural selection. These effects can be modeled mathematically using the Birth-death process on graphs. Individuals occupy the vertices of a weighted graph, and reproduce into neighboring vertices based on fitness. A key quantity is the probability that a mutant type will sweep to fixation, as a function of the mutant's fitness. Graphs that increase the fixation probability of beneficial mutations, and decrease that of deleterious mutations, are said to amplify selection. However, fixation probabilities are difficult to compute for an arbitrary graph. Here we derive an expression for the fixation probability, of a weakly-selected mutation, in terms of the time for two lineages to coalesce. This expression enables weak-selection fixation probabilities to be computed, for an arbitrary weighted graph, in polynomial time. Applying this method, we explore the range of possible effects of graph structure on natural selection, genetic drift, and the balance between the two. Using exhaustive analysis of small graphs and a genetic search algorithm, we identify families of graphs with striking effects on fixation probability, and we analyze these families mathematically. Our work reveals the nuanced effects of graph structure on natural selection and neutral drift. In particular, we show how these notions depend critically on the process by which mutations arise.

Author summary

When a new mutation appears in a population, it may ultimately spread to all individuals, or it may go extinct. Which outcome occurs depends on how the mutation affects the organism's fitness (i.e., natural selection), but also on random chance. The spatial arrangement of organisms in the population can alter the balance between selection and random chance: amplifying one, suppressing the other. However, these effects can be difficult to predict or compute, even in simple, idealized mathematical models. We develop a method to efficiently calculate the effects of spatial structure on natural selection, in the case that mutations have only a weak effect on fitness. We use this method to comb through millions of distinct spatial structures, identifying those with the most extreme effects on natural selection. The question we study has applications to cancer research, with an

individual's cells considered as a population, of which cancer cells are an invading mutant type.

Introduction

Evolution proceeds by the arrival and fixation of mutations. The fate of each new mutation depends on selection (how the mutation affects the organism's fitness) as well as drift (random chance). The combined effects of selection and drift determine how a population evolves, with selection driving adaptation to the environment, and drift maintaining genetic variety.

Spatial population structure can alter the balance between these two forces [1–13]. Some spatial structures *amplify* selection, so that fitness plays a larger role in which mutations become fixed. Other structures *suppress* selection, reducing the role of fitness and increasing the role of random drift. Spatial structure can also change the accumulation rate of neutral mutations, which do not affect fitness [14]. The effects of spatial structure on selection have consequences for microbial evolution [15], cancer [16–20], aging [19, 20], and infectious disease [21].

These effects can be probed mathematically by modeling population structure as a graph. Each vertex is occupied by a single haploid individual. Reproduction occurs along the graph's edges, according to a specified update rule (see below). The key quantity of interest is the fixation probability $\rho(r)$, defined as the probability that a new mutation of fitness r will take over a population of wild-type fitness 1.

The update rules most commonly considered in this context are Birth-death (Bd) and death-Birth (dB). In this naming convention, the ordering indicates which event (birth or death) occurs first, while the capitalization indicates which event(s) are affected by fitness (birth, in this case). Most studies [1, 3–5, 8–11, 22–26] focus on Birth-death updating, in which an individual is first selected to reproduce, proportionally to its fitness. The offspring then replaces a neighbor chosen at random (independently of fitness). A minority of works [2, 6, 12, 13, 27–29] have considered death-Birth updating, in which an individual is first selected for death, uniformly at random. A neighbor is then chosen, proportionally to fitness, to produce an offspring, which fills the vacancy.

Determining fixation probability on an arbitrary graph is computationally intensive. Current methods [6, 9, 30, 31] require solving a system of linear equations whose size grows exponentially with the graph size. This is prohibitive except for graphs that are small [6, 7, 9–11, 31], highly symmetric [1, 3–5, 22–25], or have other special properties [32].

However, most nonlethal biological mutations are either neutral ($r = 1$) or weakly selected ($r \approx 1$). Our previous work [12, 14, 33] has shown that, in these cases, fixation probabilities can be computed in polynomial time. For neutral mutations, the fixation probability determines the population's "molecular clock"—the rate at which neutral genetic substitutions accrue over time. Allen et al. [14] showed that spatial structure can either accelerate or slow this molecular clock rate.

For weakly selected mutations, fixation probabilities can be computed by combining perturbative methods [33–37] with coalescing random walks [38–41]. Allen et al. [12] applied this method to dB updating, and showed how it allows for the efficient identification of amplifiers and suppressors of weak selection.

Here we apply these methods to Birth-death updating on arbitrary weighted graphs. We show that fixation probabilities under weak selection can be expressed in terms of *coalescence times*—the expected times for two independent random walks to meet. These coalescence

times—and hence the fixation probability for weak selection—can be computed in polynomial time. While our methods apply to arbitrary placement of the initial mutant, we focus in particular on temperature initialization (mutations arise only in new offspring) and uniform initialization (mutations arise uniformly in all individuals).

Using these methods, we compute weak-selection fixation probabilities for all simple connected graphs up to size 10. For larger sizes, we employ a previously-developed genetic algorithm [10] to identify graphs with extreme effects on fixation probability. These investigations reveal a family of "Cartwheel" graphs, which strongly amplify selection under temperature initialization. They also show that a family of "Detour" graphs [10] can significantly decrease the ratio of beneficial to neutral mutations accruing over time. Our results highlight previously-unexamined subtleties in the notions of amplifier and suppressor, and in the way that spatial structure affects neutral and selective genetic change.

Methods

Graph structure

We represent spatial structure by a weighted graph G. The edge weight from vertex i to j is $w_{ij} \geq 0$. The graph may be directed (w_{ij} not necessarily equal to w_{ji}) and may contain self-loops (w_{ii} may be positive). We require that the graph is *strongly connected*, meaning that there is a path of directed edges with nonzero weight from any vertex to any other; for undirected graphs ($w_{ij} = w_{ji}$), this reduces to the usual notion of connected.

We define the *weighted (out-)degree* of vertex i as $w_i = \sum_{j \in G} w_{ij}$. The probability that a random walker at vertex i will step to vertex j is $p_{ij} = w_{ij}/w_i$. We also define the *temperature* of vertex i as $T_i = \sum_{j \in G} p_{ji}$. We note that the total temperature is equal to the total population size: $\sum_{i \in G} T_i = \sum_{i,j \in G} p_{ji} = N$.

Birth-death process

We consider a well-studied model of natural selection [1–11, 22–30]. There are two types of individuals: residents (or wild-types), which have fitness 1, and mutants, which have fitness r. The mutant is advantageous if $r > 1$, deleterious if $r < 1$, and neutral if $r = 1$.

In each state of the process, each vertex is occupied by a single individual, either mutant or resident. Selection proceeds according to the Birth-death (Bd) update rule. First, an individual i is selected at random, proportionally to its fitness. The chosen individual i produces an offspring; this offspring replaces another individual j, chosen with probability p_{ij}. Offspring inherit the type of the parent.

This process is a finite Markov chain with two absorbing states: one in which only residents are present, and one in which only mutants are present.

Initialization

The key biological question is, if a new mutation arises in a single individual, how likely is this type to take over the resident population? The answer depends on where the initial mutant arises. In the general case, we can consider an arbitrary probability distribution $\{\mu_i\}$ over the vertices of G, such that the initial mutation arises at vertex i with probability μ_i.

Most previous works [1, 3, 4, 6, 7, 9, 10, 22–25, 30] consider only *uniform initialization*, meaning that the mutant is equally likely to appear at each vertex, $\mu_i = 1/N$ for all i. Uniform initialization corresponds to a biological assumption that heritable mutations arise primarily in adult individuals, with constant probability per unit time.

If we instead suppose that mutations occur primarily in new offspring, then mutations will be more likely to arise in sites that are replaced more often. This leads to *temperature initialization* [5, 8, 11, 14, 42], meaning that the initial mutant's location i is chosen proportionally to the temperature T_i, $\mu_i = T_i/N$.

Here we consider both initialization schemes, with a particular focus on temperature initialization because it leads to a rich interplay of selection and drift. More generally, the methods we present apply to any probability distribution $\{\mu_i\}$ of initial mutant locations.

Fixation probability

We define the *fixation probability* $\rho_G(r)$, for a mutant of fitness r on a graph G, as the probability that the state of all mutants is reached, from an initial state chosen according to the specified initialization scheme.

For the complete graph K_N, representing a well-mixed population of size N, the fixation probability is [43]

$$\rho_{K_N}(r) = \frac{1 - r^{-1}}{1 - r^{-N}}.$$ (1)

More generally, Eq (1) holds for any graph that is *isothermal*, meaning that $T_i = 1$ for each vertex i. This is known as the Isothermal Theorem [1]. For Bd updating, the isothermal condition means that each vertex has the same probability of being replaced per time-step. For any isothermal graph G of size N, the fixation probability is given by [1]:

$$\rho_G(r) = \frac{1 - r^{-1}}{1 - r^{-N}}.$$ (2)

This result is valid for both uniform and temperature initialization, since the two schemes are equivalent on isothermal graphs.

Although calculating fixation probability is computationally intensive in general [6, 7, 30, 31], it simplifies when selection is weak; that is when mutations are nearly neutral ($r \approx 1$). In this case, we may form the Taylor expansion

$$\rho_G(1 + \delta) = \rho^\circ + \delta\rho' + \mathcal{O}(\delta^2).$$ (3)

The constant term, $\rho^\circ = \rho_G(1)$, is the fixation probability of neutral mutations, while the linear term, $\rho' = \frac{d}{dr}\rho_G(r)|_{r=1}$, represents the effect of weak selection (Fig 1A).

For isothermal graphs, Taylor expansion of Eq (2) gives

$$\rho_G(1 + \delta) = \frac{1}{N} + \delta\frac{N - 1}{2N} + \mathcal{O}(\delta^2).$$ (4)

So $\rho^\circ = 1/N$ and $\rho' = (N - 1)/(2N)$ for all isothermal graphs.

Results of Allen et al. [14] imply that, for uniform initialization, $\rho^\circ = 1/N$ for all graphs. Thus spatial structure does not affect the rate of neutral drift when mutations appear uniformly. In contrast, for temperature initialization, Result 3 of Allen et al. [14] implies that $\rho^\circ \leq 1/N$, with equality if and only if the graph is isothermal.

Amplifiers and suppressors

Our goal is to understand how spatial structure affects selection, neutral drift, and the relationship between the two. More specifically, we are interested whether a graph amplifies or suppresses selection, compared to the baseline of a well-mixed population. However, there are (at least) two distinct ways of making this comparison.
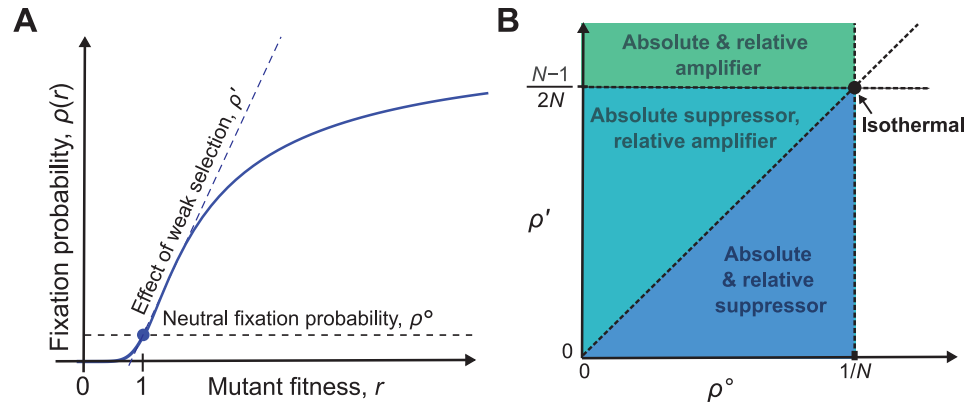
Fig 1. Classifying the effects of graph structure on weak selection. (**A**) The fixation probability $\rho_G(r)$, for a mutation of fitness r on a graph G, can be expanded under weak selection as $\rho_G(1 + \delta) = \rho^\circ + \delta\rho' + \mathcal{O}(\delta^2)$. The zeroth-order coefficient, ρ°, is the fixation probability of a neutral mutation, while the first-order coefficient, ρ', determines the effect of weak selection. (**B**) The effects of graph structure on weak selection can be classified using ρ° and ρ'. All isothermal graphs (black dot) have $\rho^\circ = 1/N$ and $\rho' = (N-1)/(2N)$. We say a graph is an *absolute amplifier* if $\rho' > (N-1)/(2N)$, and an *absolute suppressor* if $\rho' < (N-1)/(2N)$. We say a graph is a *relative amplifier* if $\rho'/\rho^\circ > (N-1)/2$ and a *relative suppressor* if $\rho'/\rho^\circ < (N-1)/2$. The three possible combinations for temperature initialization are shown in the colored regions.

First, we can examine how the probability of fixation increases with fitness, as quantified by the first-order term, ρ'. It is intuitive that the rate of increase should be greater for amplifiers, and less for suppressors, relative to the well-mixed value of $\rho' = (N-1)/(2N)$. This leads us to define a graph G as an *absolute amplifier of weak selection* if $\rho' > (N-1)/(2N)$, and an *absolute suppressor of weak selection* if $\rho' < (N-1)/(2N)$.

Second, to quantify the relationship between selection and neutral drift, we can compute the ratio of fixation probabilities for selected versus neutral mutations:

$$\frac{\rho_G(1+\delta)}{\rho_G(1)} = \frac{\rho^\circ + \delta\rho'}{\rho^\circ} + \mathcal{O}(\delta^2) = 1 + \delta\frac{\rho'}{\rho^\circ} + \mathcal{O}(\delta^2). \qquad (5)$$

We see that the ratio of fixation probabilities, for weakly selected versus neutral mutations, is determined by ρ'/ρ°. Consequently, ρ'/ρ° determines the ratio of beneficial to neutral mutations that accumulate over time. (This idea is reminiscent of the dN/dS or K_a/K_s ratios that are used to study genetic sequence evolution [44].) For a well-mixed population (or any isothermal graph), $\rho'/\rho^\circ = (N-1)/2$. We therefore define a graph G to be a *relative amplifier of weak selection* if $\rho'/\rho^\circ > (N-1)/2$, and a *relative suppressor of weak selection* if $\rho'/\rho^\circ < (N-1)/2$.

For a family of graphs of unbounded size ($N \to \infty$), we can compare $\lim_{N\to\infty} \rho'$ to $1/2$ to determine absolute amplifiers or suppressors, and $\lim_{N\to\infty} \rho'/(N\rho^\circ)$ to $1/2$ to determine relative amplifiers or suppressors, in the large-population limit.

For uniform initialization, there is no distinction between the relative and absolute definitions, since $\rho^\circ = 1/N$ for all graphs. But for temperature initialization, the two notions are distinct. Moreover, since $\rho^\circ \leq 1/N$ for all graphs [14], we have $\rho'/\rho^\circ \geq N\rho'$. This rules out the possibility that a graph can be simultaneously an absolute amplifier and a relative suppressor, leaving three possible classifications for non-isothermal graphs: (i) absolute and relative amplifier; (ii) absolute suppressor and relative amplifier; and (iii) absolute and relative suppressor. We illustrate these cases in Fig 1B.

Results

Here we present our analytical and numerical results. Derivations are given in S1 Text

Calculating fixation probability

To calculate the zeroth- and first-order coefficients, ρ° and ρ' respectively, in the weak-selection expansion, Eq (3), we apply methods developed in previous works [14, 33, 45]. We address the zeroth-order (neutral drift) and first-order (weak selection) terms separately.

Neutral drift. To obtain the neutral fixation probability, ρ°, we must first derive the fixation probability π_i of a neutral mutation arising at a particular vertex i. This probability π_i can also be understood as the reproductive value of vertex i [45, 46]. For Birth-death updating, the reproductive values π_i are the unique solution to the system of equations

$$T_i\pi_i \ = \sum_{j \in G}p_{ij}\pi_j \quad \text{for all } i \in G, \tag{6a}$$

$$\sum_{i \in G}\pi_i = 1. \tag{6b}$$

In the undirected case ($w_{ij} = w_{ji}$), Eq (6) has an explicit solution in which reproductive value is inversely proportional to weighted degree: $\pi_i = w_i^{-1}/\tilde{W}$, where $\tilde{W} = \sum_{i \in G}w_i^{-1}$. In the directed case, Eq (6) appears not to have an explicit solution, but can be solved straightforwardly as a system of $N + 1$ linear equations (one redundant) in N variables.

The overall neutral fixation probability ρ°, from an arbitrary probability distribution $\{\mu_i\}$ of initial mutant locations, is then given by

$$\rho^\circ = \sum_{i \in G}\mu_i\pi_i. \tag{7}$$

For uniform initialization, we have $\rho^\circ = \frac{1}{N}\sum_{i \in G}\pi_i = 1/N$ by Eq (6b), in agreement with previous work [14]. For temperature initialization ($\mu_i = T_i/N$) on an undirected graph ($w_{ij} = w_{ji}$), substituting $\mu_i = T_i/N$ and $\pi_i = w_i^{-1}/\tilde{W}$ gives

$$\rho^\circ = \frac{1}{N\tilde{W}} \sum_{i \in G}\frac{T_i}{w_i} = \frac{1}{N\tilde{W}} \sum_{i,j \in G}\frac{w_{ij}}{w_iw_j}. \tag{8}$$

Weak selection. To obtain the first-selection coefficient, ρ', we turn to a method developed by McAvoy and Allen [33]. Consider an arbitrary initialization, characterized by a probability distribution $\{\mu_i\}$ of initial mutant locations. Let τ_{ij} be the expected time, from initialization to fixation, that vertices i and j have different types. We prove in S1 Text that these τ_{ij} are uniquely determined by the recurrence relation

$$\tau_{ij} = \begin{cases} \dfrac{N(\mu_i + \mu_j) + \sum_{k \in G}(p_{ki}\tau_{kj} + p_{kj}\tau_{ki})}{T_i + T_j} & i \neq j, \\ 0 & i = j. \end{cases} \tag{9}$$

In particular, for temperature initialization ($\mu_i = T_i/N$), we have

$$\tau_{ij} = \begin{cases} 1 + \dfrac{\sum_{k \in G}(p_{ki}\tau_{kj} + p_{kj}\tau_{ki})}{T_i + T_j} & i \neq j, \\ 0 & i = j. \end{cases} \tag{10}$$

In this case, τ_{ij} can be interpreted as the coalescence time from i and j [12, 36, 39–41, 47]; that is, the expected time for independent random walks from i and j to meet each other. These random walks can be taken to represent the ancestral lineages of the occupants of i and j, and τ_{ij} can be understood as the expected time to their most recent common ancestor [40, 41]. Eq (10) can be understood as follows: If i and j are the same vertex, the coalescence time is zero. Otherwise, we consider all the ways that either i or j could be replaced by the offspring of a neighbor, under neutral drift. Given that either i or j is replaced, the conditional probability that i is replaced by offspring of another vertex k is $p_{ki}/(T_i + T_j)$, and the conditional probability that j is replaced by k's offspring is $p_{kj}/(T_i + T_j)$. When either of these events occurs, we increment time by one and consider the coalescence time of k and the vertex that was not replaced.

In the more general case of Eq (9), τ_{ij} can be thought of as a rescaled coalescence time, with the time spent at each vertex $k \in G$ scaled proportionally to μ_k. For uniform initialization ($\mu_i = 1/N$) the recurrence becomes

$$\tau_{ij} = \begin{cases} \dfrac{2 + \sum_{k \in G}(p_{ki}\tau_{kj} + p_{kj}\tau_{ki})}{T_i + T_j} & i \neq j, \\ \\ 0 & i = j. \end{cases} \tag{11}$$

For any initialization, we show in S1 Text that the weak-selection coefficient, ρ', can be expressed in terms of the τ_{ij} as

$$\rho' = \frac{1}{2N\tilde{W}} \sum_{i,j} \frac{w_{ij}}{w_i w_j} \tau_{ij}. \tag{12}$$

Together, Eqs (7), (9), and (12) allow for fixation probabilities to be computed under weak selection, for Bd updating on an arbitrary weighted graph, in polynomial time. In this way, one can efficiently determine the amplification and suppression properties of any given graph.

Exhaustive analysis of small graphs

To explore the variety of possible effects of graph structure on fixation probabilities, we performed an exhaustive analysis of all connected simple graphs (unweighted, undirected, and with no self-loops) up to size 10, obtained from an online database [48]. For each graph, we calculated ρ' and ρ° numerically, for both temperature and uniform initialization, by solving Eqs (10) or (11) and applying Eqs (8) and (12). We provide code to compute ρ' and ρ° in Ref. [49].

Temperature initialization. Our results for temperature initialization are summarized in Table 1 and Fig 2. We find that the vast majority of small unweighted graphs (99.6%, for $N = 10$) are relative amplifiers but absolute suppressors of weak selection. After that, for $8 \leq N \leq 10$, absolute and relative suppressors are the next most abundant, followed by absolute and relative amplifiers, and finally isothermal graphs. (Note that an unweighted, undirected graph is isothermal if and only if it is regular.) Overall, there is a positive relationship between ρ° and ρ', meaning that the graphs that slow the neutral molecular clock are also likely to suppress the (absolute) effects of weak selection.

We are particularly interested in the graphs that maximize or minimize ρ' and ρ'/ρ° (Fig 3); these are the graphs with the most pronounced effects of spatial structure on natural selection. We find that the graphs that maximize ρ' consist of several "islands" that are joined by single edges.

Interestingly, for all graphs up to size 10, the Star, S_n, minimizes ρ' and also maximizes the ratio ρ'/ρ°. This suggests that in a star-structured population (with Bd updating and constant

Table 1. Classification of small graphs with temperature initialization.

Size, N	3	4	5	6	7	8	9	10
Absolute and relative suppressors	0	0	0	0	5	51	1,035	43,249
Absolute suppressors, relative amplifiers	1	4	19	106	838	11,006	259,776	11,671,038
Absolute and relative amplifiers	0	0	0	1	6	43	247	2,117
Isothermal (regular)	1	2	2	5	4	17	22	167
Total	2	6	21	112	853	11,117	261,080	11,716,571

probability of mutation per birth), both neutral and advantageous mutations accrue much more slowly—but the ratio of advantageous to neutral is much larger—than in a well-mixed population of the same size.

The graphs that minimize the ratio ρ'/ρ° belong to a particular family, termed "Detours" by Möller et al. [10]. Detours are formed by starting with a complete graph and replacing one of the edges with a path of length ≥ 2.

Uniform initialization. Our results for uniform initialization are presented in Table 2 and Fig 4. We find that the vast majority of small graphs (94%, in the case $N = 10$) are amplifiers of weak selection. This is consistent with previous numerical analyses [6, 9–11] which found that most small unweighted graphs are amplifiers for Bd with uniform initialization. In particular, the results for $N = 6$ and $N = 7$ agree with those from an exhaustive analysis performed by Cuesta et al. [9], who computed fixation probabilities for arbitrary mutant fitness r.

The Star emerged as the strongest amplifier of weak selection, again consistent with previous results [10, 11, 50]. Stronger unweighted amplifiers of selection have been found for much larger populations [25], but not for population sizes $N \leq 100$. Meanwhile, the strongest suppressors of weak selection have a well-connected portion joined by one or two edges to a tail—a structure we explore further below.
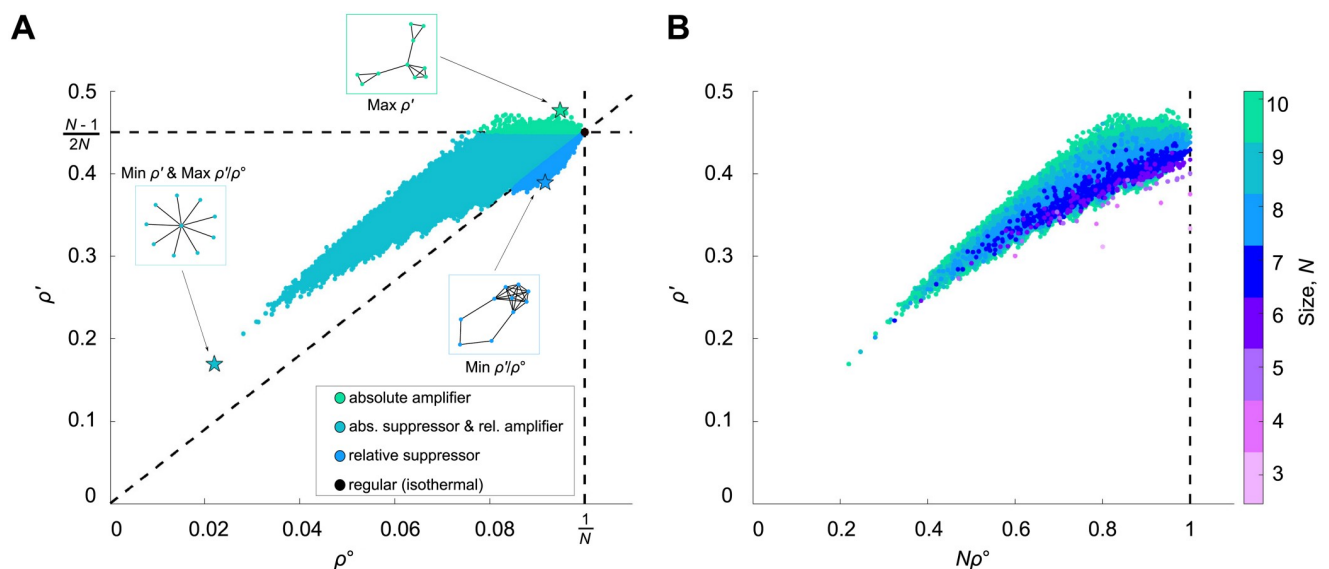

Fig 2. Exhaustive analysis of fixation probabilities under weak selection for small graphs. (A) The values of ρ° and ρ' are plotted for all 11,716,571 connected unweighted graphs of size 10. Colors correspond to the classification of graphs as shown in Fig 1B. (B) Scatter plot of ρ' versus $N\rho^\circ$ for all graphs up to size 10. Note that $N\rho^\circ \leq 1$ for all graphs, with equality if only if the graph is isothermal (or regular, in the context of unweighted graphs).
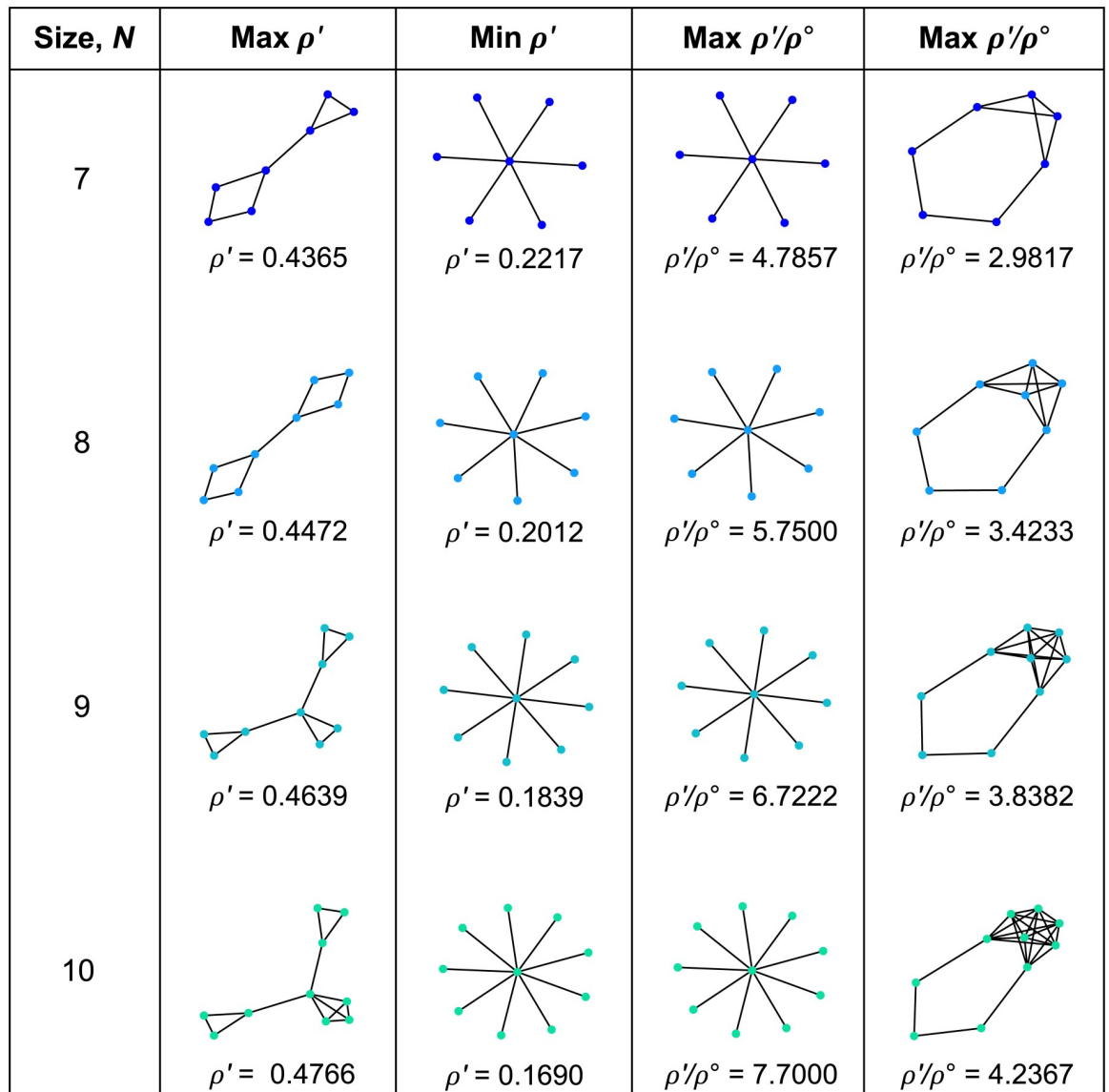
Size, N	Max ρ'	Min ρ'	Max $\rho'/\rho°$	Max $\rho'/\rho°$
7	$\rho' = 0.4365$	$\rho' = 0.2217$	$\rho'/\rho° = 4.7857$	$\rho'/\rho° = 2.9817$
8	$\rho' = 0.4472$	$\rho' = 0.2012$	$\rho'/\rho° = 5.7500$	$\rho'/\rho° = 3.4233$
9	$\rho' = 0.4639$	$\rho' = 0.1839$	$\rho'/\rho° = 6.7222$	$\rho'/\rho° = 3.8382$
10	$\rho' = 0.4766$	$\rho' = 0.1690$	$\rho'/\rho° = 7.7000$	$\rho'/\rho° = 4.2367$

Fig 3. Small graphs with extreme effects for temperature initialization. The graphs with the largest or smallest values of ρ' (which characterizes the likelihood of selected mutations to become fixed), and $\rho'/\rho°$ (which quantifies the balance of selection versus drift) are shown for sizes 7 to 10. The Star graph minimizes ρ' but maximizes $\rho'/\rho°$ for these sizes. The largest ρ' values arise for graphs with multiple components joined by single edges, while the smallest $\rho'/\rho°$ ratios occur for Detour graphs.

Table 2. Classification of small graphs with uniform initialization.

Size, N	3	4	5	6	7	8	9	10
Suppressors	0	0	1	7	55	671	14,890	659,784
Amplifiers	1	4	18	100	794	10,429	246,168	11,056,620
Isothermal (regular)	1	2	2	5	4	17	22	167
Total	2	6	21	112	853	11,117	261,080	11,716,571
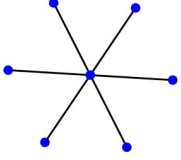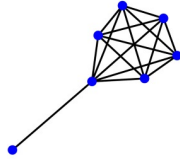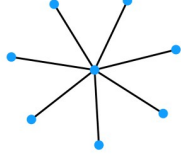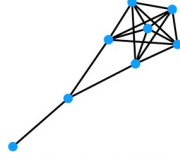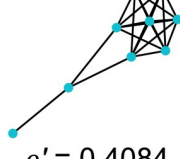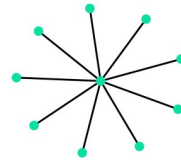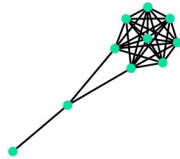
Size, N	Max ρ'	Min ρ'
7	$\rho' = 0.6354$	$\rho' = 0.4114$
8	$\rho' = 0.6737$	$\rho' = 0.4123$
9	$\rho' = 0.7050$	$\rho' = 0.4084$
10	$\rho' = 0.7310$	$\rho' = 0.4048$

Fig 4. Small graphs with extreme effects for uniform initialization. The graphs with the largest or smallest values of ρ' are shown for graphs of size 7 to 10. (Only ρ' is shown because, for uniform initialization, $\rho^\circ = 1/N$ for every graph.) Star graphs have the largest ρ', while the smallest ρ' values appear for graphs with a highly connected component linked to a tail.

Genetic algorithm

For graphs of size greater than 10, an exhaustive analysis is no longer feasible. To identify larger graphs that have extreme effects on fixation probability, we employed a genetic algorithm previously developed by Möller et al. [10]. The idea is to begin with a random ensemble of graphs, and select a subset with the largest or smallest values of a quantity of interest. These graphs are then "mated" with each other to produce an ensemble of "offspring" graphs, and the process is repeated. A formal description of the genetic algorithm and the parameters used is provided in S1 Text.

Based on our analysis of small graphs, we chose three targets to explore using the genetic algorithm: (i) maximal ρ' for temperature initialization, (ii) minimal $\rho'/\rho°$ for temperature initialization, and (iii) minimal ρ' for uniform initialization. The other possible targets (minimal ρ' and maximal $\rho'/\rho°$ for temperature initialization, maximal ρ' for uniform initialization) were optimized by the Star graph for all $N \leq 10$, and so are presumably less interesting to explore.

The results from the genetic algorithm extend and illuminate the patterns that were seen in the exhaustive analysis of smaller graphs. For temperature initialization, searching for large ρ' (strong absolute amplifiers; Fig 5) leads to the emergence of a densely connected "hub", joined by single edges to a number of "islands" consisting of two or more vertices each. Detour graphs, on the other hand, continue to appear when searching for small $\rho'/\rho°$ (strong relative suppressors; Fig 6). We analyze both of these structures in-depth in later sections.

For uniform initialization, searching for small ρ' (strong suppressors; Fig 7) led to graphs with a well-connected part and a tail. The tails are longer than those found in the exhaustive search for $N \leq 10$ (Fig 4). In some cases ($N = 11, 12, 13$, and also $N = 7$ in the exhaustive analysis) the tail connects at a single vertex to a clique, forming what is known as a Lollipop graph (also called a "standard kite" by Möller et al. [10]). Random walks on Lollipop graphs are known to have maximal hitting times [51], cover times [52], and commute times [53], so it is unsurprising that they have extreme effects on ρ', which Eq (12) expresses in terms of the closely-related notion of coalescence times. However, in other cases ($N = 14, 15$, and $N = 8, 9,$ 10 in the exhaustive search), the tail connects to two of the well-connected vertices, which are not connected to each other, forming what might be termed a "Balloon graph". Additionally, other exploratory analyses (not shown) revealed some Balloon-like graphs for which the tail ends in a star; we call this a "Balloon-star". We numerically computed the minimal ρ' for all three of these families (Lollipop, Balloon, and Balloon-Star); the results are compared to the genetic algorithm results in Fig 7.

Analysis of particular graph families

Here we analyze particular families of graphs that are found to have interesting properties. Derivations and proofs are given in S1 Text.

Star. The Star, S_n (Fig 8), consists of a single hub connected to each of $n \geq 2$ leaves by an edge of weight 1. The Star was one of the first-identified amplifiers of selection for uniform initialization [1, 22]. For graphs of size $N \leq 10$, our numerical investigation found Stars to be extremal in three different ways: the best absolute suppressors and relative amplifiers for temperature initialization, and the best amplifiers for uniform initialization.

For temperature initialization, solving Eqs (8), (10), and (12) yields

$$\rho° = \frac{2n}{(n+1)(n^2+1)} \tag{13a}$$

$$\rho' = \frac{n(2n^2 - n + 1)}{(n+1)^2(n^2+1)}. \tag{13b}$$

This agrees with Taylor expansion of previous results for the Star graph [22, 23, 54]. In S1 Text we prove that the Star is a relative amplifier and absolute suppressor for all population sizes. As the number of leaves n tends to infinity, the weak-selection coefficient ρ' behaves as $2/N$ while the ratio $\rho'/(N\rho°)$ converges to one. Thus, for large sizes N, the Star is an arbitrarily strong absolute suppressor, and a relative amplifier by a factor of two (relative to the well-mixed population, for which $\lim_{N\to\infty} \rho'/(N\rho°) = 1/2$).

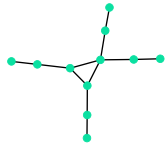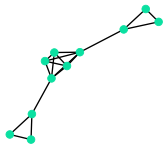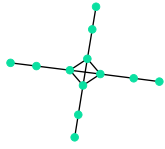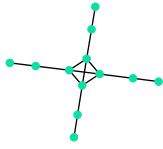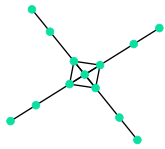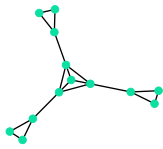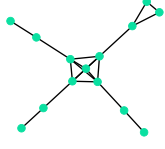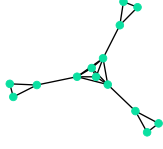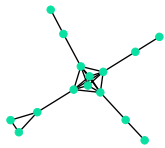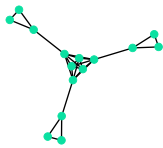| Size, N	Genetic Algorithm	Cartwheel
11	$\rho' = 0.4823$	$\rho' = 0.4835$
12 *	$\rho' = 0.4937$	$\rho' = 0.4937$
13	$\rho' = 0.4998$	$\rho' = 0.5030$
14	$\rho' = 0.5061$	$\rho' = 0.5171$
15	$\rho' = 0.5127$	$\rho' = 0.5273$

Fig 5. Discovering absolute amplifiers for temperature initialization. We used a genetic algorithm to discover graphs with large weak-selection effect ρ'. The resulting graphs (middle column) have a central "hub" joined by single links to outlying "islands". To formalize this structure, we introduce a family of "Cartwheel" graphs $CW_{n,m,h}$, consisting of a hub of size h and n islands of m vertices each (rightmost column). We find that the optimal Cartwheel graph has ρ' exceeding that found by the genetic algorithm, except for $N = 12$ for which the same graph was identified by both methods. All graphs found by both methods are absolute amplifiers of weak selection, meaning $\rho' > (N - 1)/(2N)$.

https://doi.org/10.1371/journal.pcbi.1008695.g005

For uniform initialization, $\rho^\circ = 1/N$ (as for any graph), and Eqs (11), and (12) yield

$$\rho' = \frac{n^2(n^2 - n + 2)}{(n + 1)^2(n^2 + 1)}. \tag{14}$$

This again agrees with Taylor expansion of previous results for the Star [22, 23, 54]. We
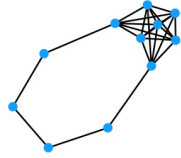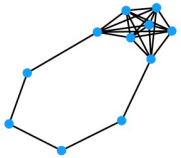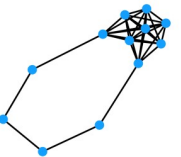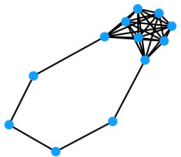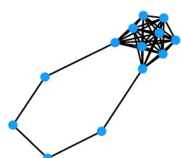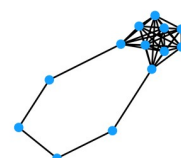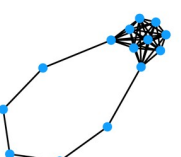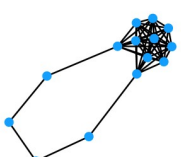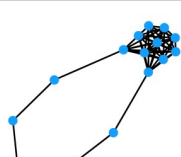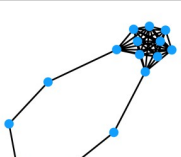
Size, N	Genetic Algorithm	Diamond Ring
11	$\rho'/\rho° = 4.6033$	$\rho'/\rho° = 4.6033$
12	$\rho'/\rho° = 4.9416$	$\rho'/\rho° = 4.9416$
13	$\rho'/\rho° = 5.2707$	$\rho'/\rho° = 5.2707$
14*	$\rho'/\rho° = 5.6041$	$\rho'/\rho° = 5.5935$
15	$\rho'/\rho° = 5.8887$	$\rho'/\rho° = 5.8887$

Fig 6. Discovering relative suppressors for temperature initialization. When seeking to minimize the ratio $\rho'/\rho°$, the genetic algorithm, in all cases, found Detour graphs [10], consisting of a complete graph with one edge replaced by a path. For comparison, we calculated $\rho'/\rho°$ for all Detour graphs of the given sizes. The results were identical to those of the genetic algorithm except in the case $N = 14$, for which the genetic algorithm found a Detour graph with a non-optimal number of ring vertices.

observe that $\lim_{n\to\infty} \rho' = 1$, meaning that the Star amplifies weak selection by a factor of two in this limit (relative to the well-mixed population, for which $\lim_{N\to\infty} \rho' = 1/2$). Although the Star is the strongest amplifier of weak selection for $N \leq 10$, it is eventually (for sufficiently large N) surpassed by graphs in the Cartwheel family, as we describe below.
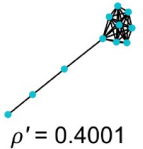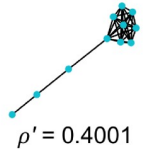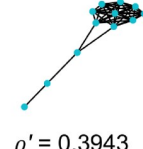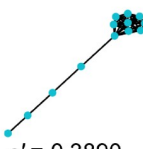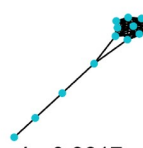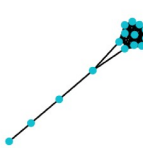
Fig 7. Discovering suppressors for uniform initialization. When seeking to minimize the ratio ρ'/ρ°, the genetic algorithm produced structures consisting of a well-connected part and a tail. We compared these to Lollipop graphs (known for their random walk properties [51–53]), and two new families, which we call Balloons and Balloon-Stars. The minimal ρ' from these families improved on the genetic algorithm results for $N = 12, 13, 14$ (albeit by less than 10^{-4} for $N = 14$), and matched the genetic algorithm results for $N = 11$ and $N = 15$.

https://doi.org/10.1371/journal.pcbi.1008695.g007

Complete bipartite graph. The complete bipartite graph, K_{n_A, n_B}, is formed by partitioning the vertex set into two subsets, of respective sizes n_A and n_B, and drawing an edge between each pair of vertices belonging to different subsets. No edges are drawn between vertices from the same subset. The Star, S_n, is the special case $K_{1,n}$, where one subset contains only the hub and the other subset contains all leaves. Tkadlec et al. [11] showed that complete bipartite graphs can amplify selection with only a relatively small increase in fixation time (relative to the well-mixed population).

In S1 Text we use our method to derive

$$\rho^\circ = \frac{2 n_A n_B}{(n_A + n_B)(n_A^2 + n_B^2)}, \qquad \rho' = \frac{n_A n_B (2(n_A^2 + n_B^2) - (n_A + n_B))}{(n_A + n_B)^2 (n_A^2 + n_B^2)}, \tag{15}$$
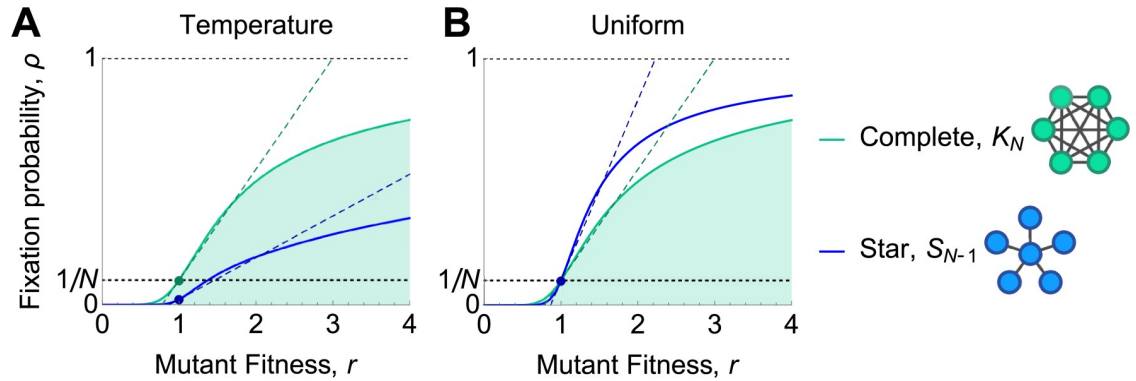
Fig 8. Star. Fixation probability, $\rho(r)$ is plotted against mutant fitness r, for the Star and the complete graph. Dotted lines show the linear approximation $\rho(1+\delta) \approx \rho^\circ + \delta\rho'$, accurate for weak selection ($\delta \ll 1$). **(A)** For temperature initialization, the Star is an absolute suppressor of weak selection, $\rho' < (N-1)/(2N)$, but a relative amplifier, $\rho'/\rho^\circ > (N-1)/2$. **(B)** For uniform initialization, the Star is an amplifier of weak selection, $\rho' > (N-1)/(2N)$. Note that $\rho(1) = \rho^\circ = 1/N$ for both graphs, as is true for any graph under uniform initialization.

https://doi.org/10.1371/journal.pcbi.1008695.g008

for temperature initialization, and

$$\rho^\circ = \frac{1}{N}, \qquad \rho' = \frac{\left(n_A^2 + n_B^2\right)^2 - \left(n_A^3 + n_B^3\right)}{\left(n_A + n_B\right)^2 \left(n_A^2 + n_B^2\right)}, \tag{16}$$

for uniform initialization. These agree with previously-obtained results for the complete bipartite graph [4, 23], as well as with Eqs (13) and (14) for the Star. We prove in S1 Text that, for $n_A \neq n_B$, K_{n_A,n_B} is a relative amplifier but absolute suppressor of weak selection for temperature initialization, and an amplifier of weak selection for uniform initialization. For $n_A = n_B$, K_{n_A,n_B} is isothermal, and so has the same fixation probabilities as a well-mixed population.

Cartwheel. Both the exhaustive search and the genetic algorithm identified strong absolute amplifiers (large ρ'), for temperature initialization, consisting of a highly intraconnected "hub" joined by single edges to a number of "islands". To formalize this pattern, we define a family of "Cartwheel" graphs $CW_{n,m,h}$ (Fig 9A), consisting of an h-vertex hub and n islands of m vertices each. The hub and each island are cliques [55], meaning that within each subpopulation, each vertex is connected to each other. Each island is connected, by a single edge, to a distinct hub vertex. We generalized to a weighted graph by setting the hub-to-island edge weight to be a free parameter $\epsilon > 0$; all other edge weights are 1. We consider only temperature initialization here, since this was the context in which Cartwheel graphs arose.

We have obtained closed-form expressions for ρ° and ρ' for the Cartwheel using Mathematica. The formula for ρ° is given in S1 Text, while the formula for ρ' is too lengthy to print. The behavior of ρ° and ρ' as ϵ varies is illustrated in Fig 9C. As shown in Fig 5, there are unweighted Cartwheel graphs whose ρ' values exceed those found by the genetic algorithm, suggesting that this family contains very strong absolute absolute amplifiers of weak selection. Table 3 shows the unweighted ($\epsilon = 1$) Cartwheel graphs with the largest values of ρ' for fixed sizes N. These ρ' values eventually surpass 1, which is an upper bound for ρ' values on the Star. We have not determined whether the ρ' values for the unweighted Cartwheel are bounded, or whether they diverge to infinity as N increases.
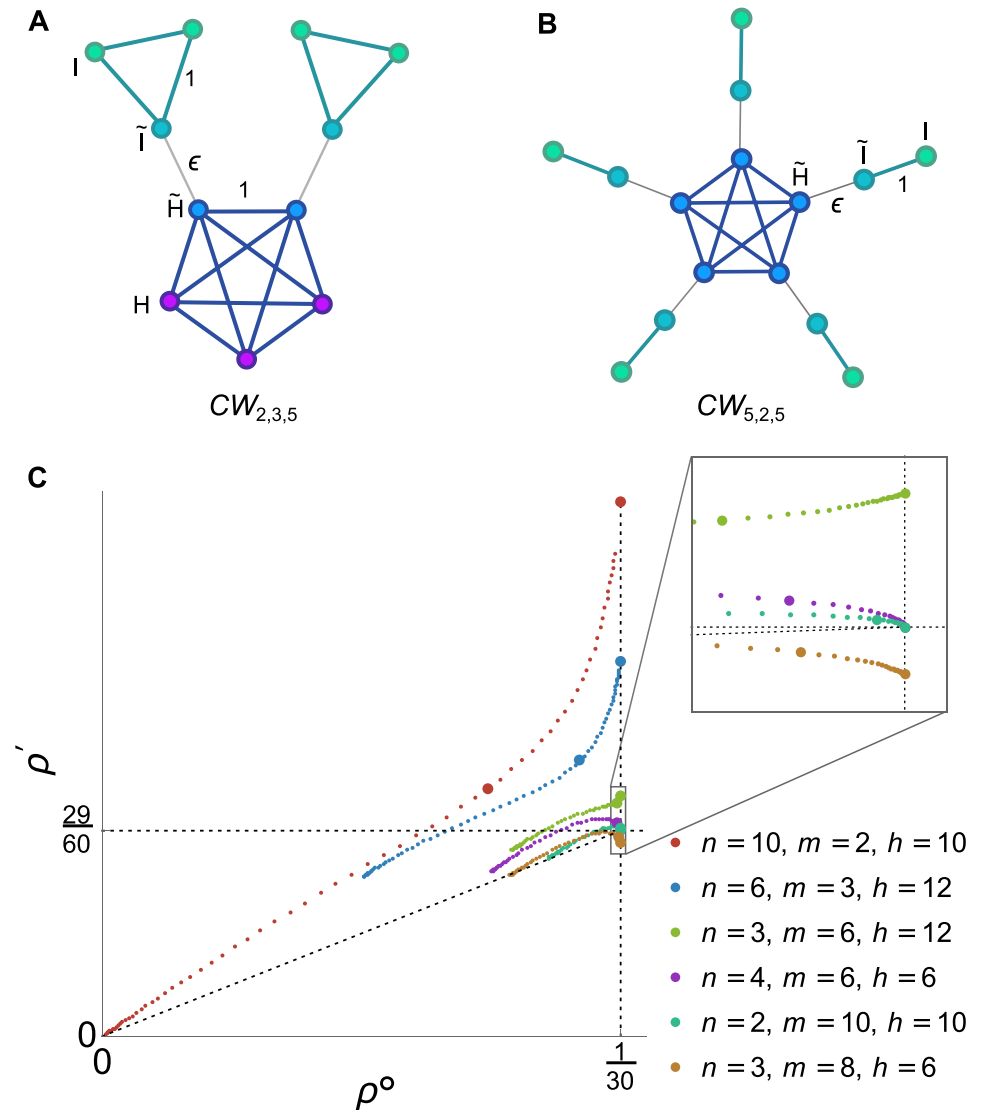
Fig 9. Cartwheel. (A) The Cartwheel graph $CW_{n,m,h}$ contains n islands of m vertices each and $h \geq n$ hub vertices. Each island is connected to a distinct hub vertex by an edge of weight ϵ; vertices within the hub and within each island are joined by edges of weight 1. (B) The special case $n = h$ and $m = 2$ has a "spider" structure; this graph has the largest ρ' in the $\epsilon \to 0$ limit. (C) Plot of ρ° vs ρ' for various Cartwheel graphs of size $N = 30$, with temperature initialization. Points are shown for each $\epsilon = 2^k$, where k varies from -5 to 9 in increments of 0.2. Larger points correspond to $\epsilon = 1$ and the $\epsilon \to 0$ limit, as derived in Eq (19). Note that $\lim_{\epsilon \to 0} \rho^\circ = 1/N$ for all Cartwheel graphs. $CW_{10,2,10}$ (the "spider" case) has by far the largest ρ' in the $\epsilon \to 0$ limit; it also has the largest ρ/ρ°, for all ϵ, among the graphs displayed. However, $CW_{6,3,12}$ has the largest ρ' for $\epsilon = 1$ among Cartwheels of size 30. $CW_{4,6,6}$ and $CW_{2,10,10}$ both have $h = m$ and therefore have the same fixation probability as a well-mixed population in the $\epsilon \to 0$ limit, according to Eq (18). $CW_{3,8,6}$ has $h < m$ and is therefore a suppressor of weak selection in the $\epsilon \to 0$ limit.

For arbitrary mutant fitness r, we derive (in S1 Text) the following closed-form expression for fixation probability in the $\epsilon \to 0$ limit:

$$
\lim_{\epsilon \to 0} \rho_{CW_{n,m,h}}(r) = \frac{(1 - r^{-1})(r^{m+h} - 1)}{(nm + h)(r^{m+h} - x^{n-1})}
$$

$$
\times \left(\frac{\left(\dfrac{nm}{m-1}\right)}{r^{-m}\left(\dfrac{1 - r^{-h}}{h-1}\right) + \left(\dfrac{1 - r^{-m}}{m-1}\right)} + \frac{\left(\dfrac{h}{h-1}\right)}{\left(\dfrac{1 - r^{-h}}{h-1}\right) + r^{-h}\left(\dfrac{1 - r^{-m}}{m-1}\right)} \right). \tag{17}
$$

Table 3. Unweighted Cartwheel graphs $CW_{n,m,h}$ that maximize ρ'.

Size, N	n	m	h	ρ'
9	3	2	3	0.4622
10	2	3	4	0.4744
11	2	3	5	0.4835
12	4	2	4	0.4937
13	3	3	4	0.5030
14	3	3	5	0.5171
15	3	3	6	0.5273
20	4	3	8	0.5732
40	7	3	19	0.7082
60	10	3	30	0.7999
80	9	4	44	0.8827
100	11	4	56	0.9569
120	13	4	68	1.0183
140	12	5	80	1.0775

https://doi.org/10.1371/journal.pcbi.1008695.t003

This result holds for both temperature and uniform initialization. In fact, the two initialization schemes become equivalent as $\epsilon \to 0$ since the temperature of each vertex converges to $1/N$. (However, the Isothermal Theorem does not apply in this limit, since the graph is not exactly isothermal for any nonzero value of ϵ.) Several cases of Eq (17) are illustrated in Fig 10A.

If the hub is the same size as the islands, $h = m$, then Eq (17) reduces to

$$\lim_{\epsilon \to 0} \rho_{CW_{n,m,m}}(r) = \frac{1 - r^{-1}}{1 - r^{-(mn+m)}}. \tag{18}$$

This is exactly the fixation probability for a well-mixed population of the same size, $N = mn + m$. Thus for $h = m$ and $\epsilon \to 0$, a mutant of arbitrary fitness r has the same fixation probability on the Cartwheel as in a well-mixed population.
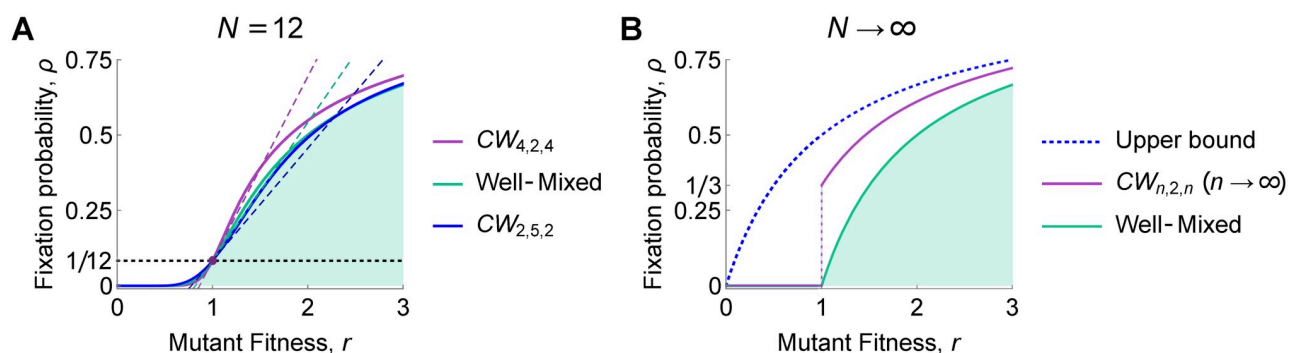
Fig 10. Nonweak selection on the Cartwheel graph. In the limit as the hub-to-island weight ϵ goes to zero, the fixation probability for arbitrary mutant fitness r is expressed in closed form by Eq (17). Temperature and uniform initialization are equivalent in this limit. Dashed lines show the weak-selection approximation, with slope ρ' given by Eq (19). **(A)** Fixation probability is plotted against r for two Cartwheels of size 12. $CW_{4,2,4}$ has $h > m$ and is therefore an amplifier of weak selection. $CW_{2,5,2}$ has $h < m$ and is therefore a suppressor of weak selection (but appears to amplify selection for $r > 2.4219$). Cartwheels with $h = m$ have the same fixation probability as the well-mixed population, for all r, in the $\epsilon \to 0$ limit. **(B)** In the limits $\epsilon \to 0$ and $n = h \to \infty$, with $m = 2$, the fixation probability jumps discontinuously from 0 to 1/3 as r crosses 1; the expression for $r > 1$ is given in Eq (20). For comparison we also show the upper bound $\rho(r) \leq 1 - (r+1)^{-1}$, derived by Pavlogiannis et al. [8], which applies to all weighted graphs with no self-loops under temperature initialization.

https://doi.org/10.1371/journal.pcbi.1008695.g010

Performing a Taylor expansion of Eq (17) around $r = 1$, we obtain $\rho^\circ = 1/N$ and

$$\rho' = \frac{N-1}{2N} + \frac{mnh(h-m)(m(n-2)(h-1) + h(m-1))}{2(h(m-1) + m(h-1))(mn+h)(mn(h-1) + h(m-1))}. \quad (19)$$

The second term has the sign of $h - m$. It follows that the Cartwheel, in the $\epsilon \to 0$ limit, is an amplifier of weak selection for $h > m$ and a suppressor of weak selection for $h < m$. These results hold in both the relative and absolute senses, for both temperature and uniform initialization.

The Cartwheel most strongly amplifies selection in the case $\epsilon \to 0$, $m = 2$, and $h = n \gg 1$. This results in a "spider" structure, in which a large number of fully interconnected hub vertices are each joined (by an edge of vanishingly small weight) to a two-vertex "leg". In this case, the fixation probability $\rho(r)$ becomes (in the $\epsilon \to 0$ limit) discontinuous as a function of r: $\rho(r) = 0$ for disadvantageous mutations ($r < 0$), but for advantageous mutations ($r > 0$) we have

$$\rho(r) = \frac{1 - r^{-2}/3}{1 + r^{-1}}. \quad (20)$$

In particular, the fixation probability jumps from zero to $1/3$ as r increases past 1. Any beneficial mutation has a fixation probability greater than one-third for such a structure, making this an especially powerful amplifier of natural selection.

We also show in S1 Text that the Cartwheel $CW_{n,2,n}$, with $n \geq 7$ and sufficiently small ϵ, is a stronger relative amplifier of weak selection than a Star of the same size.

Detour. In the exhaustive analysis and genetic algorithm, Detour graphs [10] (Fig 11A) emerged as the strongest relative suppressors of weak selection (minimal ρ'/ρ°) under temperature initialization. The Detour graph $D_{c,d}$ is constructed by starting with a complete graph of size c, and replacing one edge with a path containing d vertices. The total size is $N = c + d$.

We have numerically determined the minimal ρ'/ρ° ratio for all sizes up to 100 (Fig 11B and 11C). As N increases beyond 8, the minimal ρ'/ρ° decreases, making these graphs increasingly strong relative suppressors. We also observe that the minimizing number of detour vertices d increases sublinearly with N.

Minimal absolute amplifier. According to our small graph analysis in Table 1, there are no absolute amplifiers of weak selection of size ≤ 5, and only one of size 6. This graph of size 6 has a bowtie shape (Fig 12A) and is the minimal absolute amplifier of weak selection.
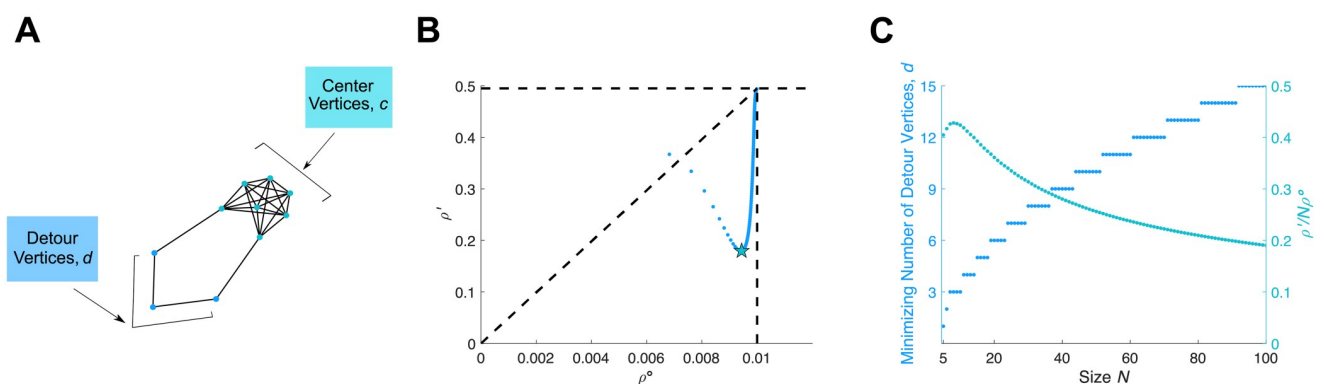
Fig 11. Detour. (A) The Detour graph $D_{c,d}$ begins with a complete graph of size c, and replaces the edge between two vertices by a path with d interior vertices. **(B)** Values of ρ° and ρ' are plotted as d varies, for $N = 100$. The Detour $D_{c,d}$ is a relative suppressor of weak selection except for $d = 1$. The minimal ratio ρ'/ρ° is achieved for $d = 15$ (marked with a star). **(C)** The minimal $\rho'/(N\rho^\circ)$ ratio, and the value of d achieving this ratio, is plotted for each N. Note that the minimizing d grows sub-linearly with N.

A

B
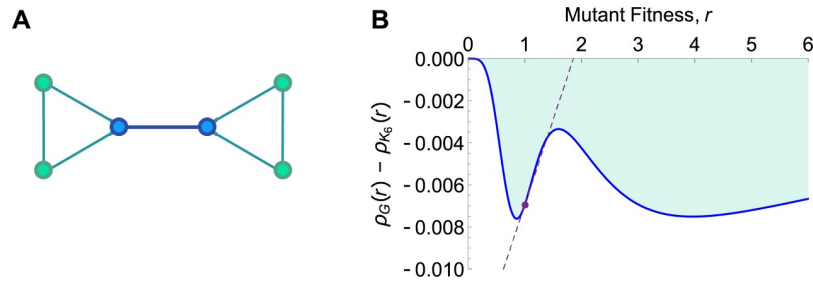
Fig 12. Minimal absolute amplifier. (A) This bowtie-shaped graph is the smallest absolute amplifier of weak selection under temperature initialization. **(B)** The difference in fixation probabilities between this graph and the complete graph K_6 is plotted against mutant fitness, r. The dashed line shows the linear approximation to this difference at $r = 1$, computed using our weak-selection methods. Although this difference is increasing at $r = 1$ (because the graph is an absolute amplifier of weak selection), the difference is negative for all values of r. Thus this graph does not amplify selection in the usual sense of $\rho_G(r) > \rho_{K_N}(r)$ for all $r > 1$.

Due to its small size and symmetry, the fixation probability for this graph can be computed in closed form for arbitrary mutant fitness r. Employing an algorithm developed by Cuesta et al. [7, 9], adapted for temperature initialization, we obtained the fixation probability $\rho_G(r)$ as a ratio of 17th-degree polynomials (shown in S1 Text).

We find (Fig 12B) that, despite being an absolute amplifier of weak selection, this graph has fixation probability less than that of the complete graph K_6 for all $r > 0$. This is possible because, although this graph increases the first-order term ρ', relative to K_6, it decreases the neutral term ρ°, leading to a smaller overall fixation probability.

Fan.   Finally, we exhibit a graph family that displays all possible classifications of behavior as a particular edge weight is varied. The Fan, $F_{n,m}$ [12] (or Windmill [56]), consists of a hub vertex attached to $n \geq 2$ blades (Fig 13A). Each blade is a clique of $m \geq 2$ vertices joined by edges of weight 1. The hub is connected to each blade vertex by an edge of weight ϵ. The total population size is $N = mn + 1$.

For the Fan $F_{n,m}$ with temperature initialization, we find a neutral fixation probability of

$$\rho^\circ = \frac{mn\epsilon(m - 1 + 2\epsilon)(mn + 2)}{(m - 1 + \epsilon)(mn + 1)(m - 1 + \epsilon(m^2n^2 + 1))}. \tag{21}$$

The first-order term is $\rho' = \text{num/denom}$, with

$$\begin{aligned}
\text{num} = mn\epsilon(2(m - 1 + \epsilon)^3 + 2mn(m - 1 + \epsilon)^2(2m - 1 - \epsilon) \\
+ 2m^2n^2(m - 1 + \epsilon)(2\epsilon^2 - \epsilon + m^2 - m) \\
+ 2m^4n^3\epsilon(m + 2\epsilon - 1 + nm^2 - n))),
\end{aligned} \tag{22a}$$

$$\begin{aligned}
\text{denom} = 2(m - 1 + \epsilon)(mn + 1)(m - 1 + \epsilon(m^2n^2 + 1))((m - 1 + \epsilon)^2 \\
+ mn(1 + \epsilon)(m - 1 + \epsilon) + m^3n^2\epsilon).
\end{aligned} \tag{22b}$$

We find that, as ϵ increases, $F_{n,m}$ displays all three classifications of behavior. First, for $0 < \epsilon < (m - 1)/(mn - 1)$, the Fan is both an absolute and relative suppressor. At $\epsilon = (m - 1)/(mn - 1)$, the graph is isothermal, and therefore has the same fixation probability as the well-mixed population. The graph is both an absolute and relative amplifier for $(m - 1)/(mn - 1) < \epsilon < \epsilon^*$, where ϵ^* is a particular cubic root for which $\rho' = (N - 1)/(2N)$. Then, for $\epsilon > \epsilon^*$, the graph is a relative amplifier but absolute suppressor. As $\epsilon \to \infty$, the values of ρ° and ρ' approach those of the star S_{nm}, as given by Eq (13).
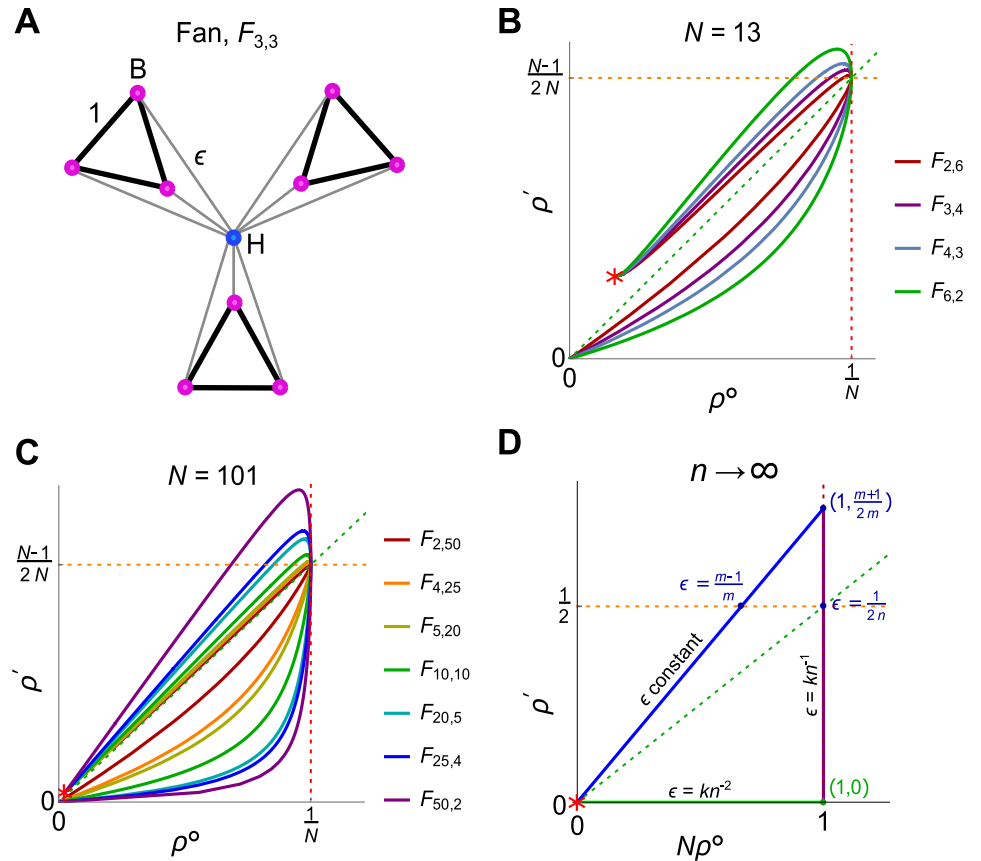
Fig 13. Fan. (A) The Fan $F_{n,m}$ (or Windmill [56]) consists of one hub and n blades containing m vertices each. We consider a weighted version, with edge weights as shown. Pictured here is the case $n = m = 3$. (B,C) The neutral fixation probability ρ° and weak selection coefficient ρ', plotted as ϵ varies from 0 to infinity, for all Fan graphs of sizes $N = 13$ and $N = 101$. As ϵ increases, the behavior changes from absolute and relative suppressor, to absolute and relative amplifier, to relative amplifier but absolute suppressor. For $\epsilon \to \infty$, the ρ° and ρ' values approach those of the Star graph S_{nm} (marked by a red star). (D) As $n \to \infty$, there are three regimes of behavior, depending whether ϵ is held constant (blue line) or scales as n^{-1} (purple line), or as n^{-2} (green line). The maximal ρ' for fixed m is $\rho' = (m + 1)/(2m)$, achieved for $n \to \infty$ with $\epsilon = (m - 1)/(m\sqrt{n})$.

In the limit of many islands, $n \to \infty$, the curve formed by $N\rho^\circ$ and ρ' approaches a triangle ([Fig 13D](#)), the sides of which correspond to three scalings of ϵ with n: constant, inverse ($\epsilon = kn^{-1}$), and inverse square ($\epsilon = kn^{-2}$). The maximum value of ρ' is achieved when $n \to \infty$ with $\epsilon = (m - 1)/(m\sqrt{n})$; in this limit, ρ' approaches $(m + 1)/(2m)$.

For uniform initialization, $\rho^\circ = 1/N$ and $\rho' = \text{num}/\text{denom}$, where

$$\text{num} = m^2 n^2 \epsilon (m^3(m + 1)n^3 \epsilon + 2m^2 n^2 (\epsilon - 1)\epsilon$$
$$+ mn(3m - 2\epsilon + 1)(m + \epsilon - 1) + 4(m + \epsilon - 1)^2), \tag{23a}$$

$$\text{denom} = 2(mn + 1)(m^2 n^2 \epsilon + m + \epsilon - 1)$$
$$\times (m^3 n^2 \epsilon + mn(\epsilon + 1)(m + \epsilon - 1) + (m + \epsilon - 1)^2). \tag{23b}$$

We show in [S1 Text](#) that the Fan with uniform initialization is a suppressor of weak selection for $0 < \epsilon < (m - 1)/(mn - 1)$ and an amplifier of weak selection for $\epsilon > (m - 1)/(mn - 1)$. As

$n \to \infty$ with m and ϵ held constant, ρ' converges to $(m + 1)/(2m)$, matching the maximal limiting value of ρ' that was found with temperature initialization.

Discussion

The problem of identifying spatial structures that amplify or suppress selection has been actively investigated for the past decade and a half [1–13]. Our work introduces new methods for studying this problem in the weak selection regime. These methods allow for the effects of spatial structure on fixation probability (for $r \approx 1$) to be computed in polynomial time, enabling analytical and computational investigations that would previously have been infeasible. However, this weak-selection approach does not tell us the behavior of $\rho(r)$ away from $r = 1$, which may include complex phenomena such as multiple transitions between amplification and suppression [4, 9, 57]. We also have not explored the question of fixation time [10, 11, 26], which can significantly affect the overall rate of evolution [58].

Our work underscores the critical role that initialization schemes play in the determining how spatial structure affects selection [5, 11]. Different initialization schemes arise from different assumptions on how mutations occur. Uniform initialization, which is taken as the default in most previous works, corresponds to an assumption that mutation strikes all individuals equally regardless of age. Temperature initialization arises from instead considering a constant probability of mutation in each new offspring [14].

With temperature initialization comes new considerations, since the neutral fixation probability ρ° is affected in addition to the weak-selection coefficient ρ'. This leads to a distinction between relative versus absolute amplification and suppression. The absolute notions refer to the rate of increase fixation probability with respect to mutant fitness, while the relative notions refer to the expected ratio of beneficial to neutral mutations that fix over time. Of these, the relative notions are arguably more empirically relevant since they are directly related to observable genetic change.

A caveat with these definitions is that, for temperature initialization, absolute amplifiers of weak selection do not necessarily have larger fixation probability than the well-mixed population for any particular value of r. The minimal absolute amplifier (Fig 12) shows why: under temperature initialization, the neutral fixation probability ρ° is typically reduced from the well-mixed value, which may outweigh any increase in ρ'. On the other hand, it is certainly possible for a graph to have larger fixation probability than a well-mixed population under temperature initialization; the Cartwheel family provides examples of this.

Other initialization schemes aside from temperature and uniform may be considered. Since mutations can strike both existing organisms as well as new offspring, the relevant scheme for a given population may be a blend of uniform and temperature. Alternatively, for unicellular populations, mutations may occur in the parent cell upon reproduction; this would lead to initialization schemes that depend on the birth rate as well as the death rate. Fortunately, our method can be applied to any initialization scheme, using Eq (9) for coalescence times.

We have found that, for temperature initialization, the vast majority of small (unweighted, undirected) graphs are relative amplifiers but absolute suppressors of weak selection. This means that, for a spatially structured population satisfying the relevant modeling assumptions, one would typically see a greater ratio of beneficial-to-neutral mutations, but a slower overall rate of genetic change, compared to a well-mixed population. For uniform initialization, we find that the vast majority of graphs are amplifiers of weak selection, which accords with findings from other works using different methods [6, 10, 11].

Our findings reveal new families of graphs with interesting effects on selection, and shed new light on known families. In particular, our work has identified Cartwheel graphs (Fig 9) as strong absolute amplifiers for temperature initialization—with the "spider" case, $n = h$ and

$m = 2$, having especially interesting properties. In one limit, the fixation probability jumps discontinuously, from zero for all deleterious mutations, to more than one-third for all beneficial mutations (Fig 10). Stronger amplifiers for temperature initialization were found by Pavlogiannis et al. [8], but only with self-loops, meaning that new offspring can displace their own parents. Among graphs with no self-loops, Cartwheels appear to be the strongest family of amplifiers for temperature initialization that have been discovered thus far.

Detour graphs, in contrast, are shown to strongly suppress the ρ'/ρ° ratio for temperature initialization. This family of graphs was previously identified by Möller et al. [10] for their extreme properties under uniform initialization. In that context, Möller et al. found that Detours are powerful suppressors of selection for fitness values of r close to 1, but transition into amplifiers as r increases beyond a threshold value $r^* > 1$. Overall, Detours appear to be a particularly interesting example for evolutionary graph theory. We have not come up with closed-form solutions for ρ° or ρ' for Detours; doing so appears difficult although perhaps not impossible.

Our analysis of the Fan, meanwhile, highlights the determinative role that edge weight can play. By increasing the weight of a single type of edge, one can change the graph from a relative suppressor, to an absolute amplifier, to an absolute suppressor but relative amplifier. Interestingly, both ρ° and ρ' are maximized for intermediate values of this edge weight. Attention to edge weight will therefore be crucial in connecting this theory to applications.

Although our general results apply to arbitrary (strongly connected) weighted digraphs, we focused our numerical exploration and examples on undirected graphs without self-loops. Extending this analysis to directed graphs would allow for many more possible structures, and may reveal new phenomena with regard to the amplification or suppression of selection.

Conclusion

Our work and other recent contributions [5, 6, 9, 11–13, 27, 28] make clear that whether a population structure amplifies for suppresses selection depends not only on the graph, but also on the update rule, the initialization scheme, and the regime of fitness values being considered, and whether absolute or relative genetic change is of interest. This combinatorial explosion of modeling choices may seem daunting, but it also suggests an increased opportunity for applications of the theory. Different modeling choices will be relevant to animal, plant, and microbial populations, to somatic tissue [18–20], and to infectious diseases [15, 21]. Given recent advances in theory [8–13, 26, 33], the toolkit now exists to apply evolutionary graph theory in a wide range of biological settings.

Supporting information

S1 Text. Mathematical derivations and proofs. Contains derivations and proofs of our overall method, as well as mathematical analysis of particular graph families.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Benjamin Allen.

References

1. Lieberman E, Hauert C, Nowak MA. Evolutionary dynamics on graphs. Nature. 2005; 433(7023):312–316. https://doi.org/10.1038/nature03204

2. Antal T, Redner S, Sood V. Evolutionary dynamics on degree-heterogeneous graphs. Physical Review Letters. 2006; 96(18):188104. https://doi.org/10.1103/PhysRevLett.96.188104

3. Shakarian P, Roos P, Johnson A. A review of evolutionary graph theory with applications to game theory. Biosystems. 2012; 107(2):66–80. https://doi.org/10.1016/j.biosystems.2011.09.006

4. Voorhees B, Murray A. Fixation probabilities for simple digraphs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 2013; 469(2154):20120676. https://doi.org/10.1098/rspa.2012.0676

5. Adlam B, Chatterjee K, Nowak M; The Royal Society. Amplifiers of selection. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 2015;471(2181):20150114.

6. Hindersin L, Traulsen A. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. PLOS Computational Biology. 2015; 11(11):e1004437. https://doi.org/10.1371/journal.pcbi.1004437

7. Cuesta FA, Sequeiros PG,Rojo ÁL. Suppressors of selection. PLOS One. 2017; 12(7):e0180549. https://doi.org/10.1371/journal.pone.0180549

8. Pavlogiannis A, Tkadlec J, Chatterjee K, Nowak MA. Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory. Communications Biology. 2018; 1(1):71. https://doi.org/10.1038/s42003-018-0078-7

9. Cuesta FA, Sequeiros PG,Rojo ÁL. Evolutionary regime transitions in structured populations. PLOS One. 2018; 13(11):e0200670. https://doi.org/10.1371/journal.pone.0200670

10. Möller M, Hindersin L, Traulsen A. Exploring and mapping the universe of evolutionary graphs identifies structural properties affecting fixation probability and time. Communications Biology. 2019; 2(1):137. https://doi.org/10.1038/s42003-019-0374-x

11. Tkadlec J, Pavlogiannis A, Chatterjee K, Nowak MA. Population structure determines the tradeoff between fixation probability and fixation time. Communications Biology. 2019; 2.

12. Allen B, Sample C, Jencks R, Withers J, Steinhagen P, Brizuela L, et al. Transient amplifiers of selection and reducers of fixation for death-Birth updating on graphs. PLoS Computational Biology. 2020; 16(1): e1007529. https://doi.org/10.1371/journal.pcbi.1007529 PMID: 31951612

13. Tkadlec J, Pavlogiannis A, Chatterjee K, Nowak MA. Limits on amplifiers of natural selection under death-Birth updating. PLoS Computational Biology. 2020; 16(1):e1007494. https://doi.org/10.1371/journal.pcbi.1007494

14. Allen B, Sample C, Dementieva Y, Medeiros RC, Paoletti C, Nowak MA. The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure. PLOS Computational Biology. 2015; 11(2):e1004108. https://doi.org/10.1371/journal.pcbi.1004108

15. Krieger MS, Denison CE, Anderson TL, Nowak MA, Hill AL. Population structure across scales facilitates coexistence and spatial heterogeneity of antibiotic-resistant infections. PLOS Computational Biology. 2020; 16(7):e1008010. https://doi.org/10.1371/journal.pcbi.1008010

16. Nowak MA, Michor F, Iwasa Y. The linear process of somatic evolution. Proceedings of the National Academy of Sciences. 2003; 100(25):14966–14969. https://doi.org/10.1073/pnas.2535419100

17. Vermeulen L, Morrissey E, van der Heijden M, Nicholson AM, Sottoriva A, Buczacki S, et al. Defining stem cell dynamics in models of intestinal tumor initiation. Science. 2013; 342(6161):995–998. https://doi.org/10.1126/science.1243148 PMID: 24264992

18. Hindersin L, Werner B, Dingli D, Traulsen A. Should tissue structure suppress or amplify selection to minimize cancer risk? Biology Direct. 2016; 11(1):41.

19. Cannataro VL, McKinley SA, St Mary CM. The implications of small stem cell niche sizes and the distribution of fitness effects of new mutations in aging and tumorigenesis. Evolutionary Applications. 2016; 9(4):565–582. https://doi.org/10.1111/eva.12361

20. Cannataro VL, McKinley SA, St Mary CM. The evolutionary trade-off between stem cell niche size, aging, and tumorigenesis. Evolutionary Applications. 2017; 10(6):590–602. https://doi.org/10.1111/eva.12476

21. Ottino-Loffler B, Scott JG, Strogatz SH. Evolutionary dynamics of incubation periods. eLife. 2017; 6.

22. Broom M, Rychtář J. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 2008; 464(2098):2609–2627. https://doi.org/10.1098/rspa.2008.0058

23. Monk T, Green P, Paulin M. Martingales and fixation probabilities of evolutionary graphs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 2014; 470(2165):20130730. https://doi.org/10.1098/rspa.2013.0730

24. Jamieson-Lane A, Hauert C. Fixation probabilities on superstars, revisited and revised. Journal of Theoretical Biology. 2015; 382:44–56. https://doi.org/10.1016/j.jtbi.2015.06.029

25. Pavlogiannis A, Tkadlec J, Chatterjee K, Nowak MA. Amplification on undirected population structures: comets beat stars. Scientific Reports. 2017; 7(1):82. https://doi.org/10.1038/s41598-017-00107-w

26. Hathcock D, Strogatz SH. Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs. Physical Review E. 2019; 100(1):012408. https://doi.org/10.1103/PhysRevE.100.012408

27. Frean M, Baxter G. Death-birth ordering and suppression of fitness in networks; 2008. Available from: http://homepages.mcs.vuw.ac.nz/marcus/manuscripts/FreanBaxterJTB.pdf.

28. Kaveh K, Komarova NL, Kohandel M. The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. Royal Society Open Science. 2015; 2(4):140465. https://doi.org/10.1098/rsos.140465

29. Pattni K, Broom M, Rychtář J, Silvers LJ; The Royal Society. Evolutionary graph theory revisited: when is an evolutionary process equivaent to the Moran process? Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science. 2015;471(2182):20150334.

30. Cuesta FA, Sequeiros PG,Rojo ÁL. Fast and asymptotic computation of the fixation probability for Moran processes on graphs. Biosystems. 2015; 129:25–35. https://doi.org/10.1016/j.biosystems.2015.01.007

31. Hindersin L, Möller M, Traulsen A, Bauer B. Exact numerical calculation of fixation probability and time on graphs. Biosystems. 2016; 150:87–91. https://doi.org/10.1016/j.biosystems.2016.08.010

32. Monk T. Martingales and the fixation probability of high-dimensional evolutionary graphs. Journal of Theoretical Biology. 2018; 451:10–18. https://doi.org/10.1016/j.jtbi.2018.04.039

33. McAvoy A, Allen B. Fixation probabilities in evolutionary dynamics under weak selection. Journal of Mathematical Biology. 2021; 82:14. Available from: https://doi.org/10.1007/s00285-021-01568-4

34. Nowak MA, Sasaki A, Taylor C, Fudenberg D. Emergence of cooperation and evolutionary stability in finite populations. Nature. 2004; 428(6983):646–650. https://doi.org/10.1038/nature02414

35. Rousset F. Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. Theoretical Population Biology. 2006; 69(2):165–179. https://doi.org/10.1016/j.tpb.2005.08.008

36. Chen YT. Sharp benefit-to-cost rules for the evolution of cooperation on regular graphs. The Annals of Applied Probability. 2013; 23(2):637–664. https://doi.org/10.1214/12-AAP849

37. Lessard S, Ladret V. The probability of fixation of a single mutant in an exchangeable selection model. Journal of Mathematical Biology. 2007; 54(5):721–744. https://doi.org/10.1007/s00285-007-0069-7

38. Kingman JFC. The coalescent. Stochastic Processes and Their Applications. 1982; 13(3):235–248. https://doi.org/10.1016/0304-4149(82)90011-4

39. Cox JT. Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. Annals of Probability. 1989; 17(4):1333–1366. https://doi.org/10.1214/aop/1176991158

40. Liggett TM. Interacting Particle Systems. Springer Science & Business Media; 2006.

41. Wakeley J. Coalescent Theory: An Introduction. Greenwood Village, CO: Roberts & Company Publishers; 2009.

42. Allen B, Tarnita CE. Measures of success in a class of evolutionary models with fixed population size and structure. Journal of Mathematical Biology. 2014; 68(1-2):109–143. https://doi.org/10.1007/s00285-012-0622-x

43. Moran PAP. The Statistical Process of Evolutionary Theory. Clarendon Press; 1962.

44. Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. Trends in Ecology & Evolution. 2000; 15(12):496–503. https://doi.org/10.1016/S0169-5347(00)01994-7

45. Allen B, McAvoy A. A mathematical formalism for natural selection with arbitrary spatial and genetic structure. Journal of Mathematical Biology. 2018;. PMID: 30430219

46. Maciejewski W. Reproductive value in graph-structured populations. Journal of Theoretical Biology. 2014; 340:285–293. https://doi.org/10.1016/j.jtbi.2013.09.032

47. Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, et al. Evolutionary dynamics on any population structure. Nature. 2017; 544(7649):227–230. https://doi.org/10.1038/nature21723 PMID: 28355181

48. McKay B. Graphs; [cited 18 December 2020]. Available from: http://users.cecs.anu.edu.au/~bdm/data/graphs.html.

49. Sample C, Shapiro J, King M, Goncalves M. Computing fixation probabilities in graph-structured populations under weak selection; 2020. Zenodo. Available from: http://doi.org/10.5281/zenodo.4338335.

50. Broom M, Rychtář J, Stadler B. Evolutionary dynamics on small-order graphs. Journal of Interdisciplinary Mathematics. 2009; 12(2):129–140. https://doi.org/10.1080/09720502.2009.10700618

51. Brightwell G, Winkler P. Maximum hitting time for random walks on graphs. Random Structures and Algorithms. 1990; 1(3):263–276. https://doi.org/10.1002/rsa.3240010303

52. Feige U. A tight upper bound on the cover time for random walks on graphs. Random Structures and Algorithms. 1995; 6(1):51–54. https://doi.org/10.1002/rsa.3240060106

53. Jonasson J. Lollipop graphs are extremal for commute times. Random Structures & Algorithms. 2000; 16(2):131–142. https://doi.org/10.1002/(SICI)1098-2418(200003)16:2%3C131::AID-RSA1%3E3.0.CO;2-3

54. Hadjichrysanthou C, Broom M, Rychtář J. Evolutionary games on star graphs under various updating rules. Dynamic Games and Applications. 2011; 1(3):386. https://doi.org/10.1007/s13235-011-0022-7

55. Moon JW, Moser L. On cliques in graphs. Israel journal of Mathematics. 1965; 3(1):23–28. https://doi.org/10.1007/BF02760024

56. Hsu DF. Harmonious labelings of windmill graphs and related graphs. Journal of Graph Theory. 1982; 6(1):85–87. https://doi.org/10.1002/jgt.3190060110

57. Choi JO, Yu U. Fixation probability on clique-based graphs. Physica A: Statistical Mechanics and its Applications. 2018; 492:2129–2135. https://doi.org/10.1016/j.physa.2017.11.131

58. Frean M, Rainey PB, Traulsen A. The effect of population structure on the rate of evolution. Proceedings of the Royal Society B: Biological Sciences. 2013; 280(1762):20130211. https://doi.org/10.1098/rspb.2013.0211