Hindawi
Complexity
Volume 2022, Article ID 6625709, 11 pages
https://doi.org/10.1155/2022/6625709

WILEY | Hindawi

Research Article

Dynamics of Positive Frequency Dependent Selection Triggers Selection for Silence

I. Hashem, V. De Buck, and J. Van Impe

KU Leuven, Chemical Engineering Department, BioTeC & OPTEC, Gebroeders De Smetstraat 1, Ghent 9000, Belgium

Correspondence should be addressed to J. Van Impe; jan.vanimpe@kuleuven.be

Received 10 December 2020; Revised 4 October 2021; Accepted 18 October 2021; Published 18 April 2022

Academic Editor: Saleh Mobayen

Positive frequency dependent selection is a natural selection regime where the fitness of a phenotype increases with its frequency in the population. Examples can be typically found in the spread of disease tolerance strategies in a population. A characterizing feature of PFDS is that the focal allele may experience favorable selection only when it becomes more frequent in the population, while being selected against when it is rare. In this paper, by applying a solution concept from evolutionary game theory, we hypothesize that the process of emergence of such alleles triggers selection for silence, where the emergence of alleles that can stay unexpressed for a period of time will be favored by selection forces. We illustrate our hypothesis using a mathematical model of a population of single locus diploid organisms with two alleles where a single mutant of an allele that experiences positive frequency dependent selection appears in a population where the other allele is a dominant. The model results show that the longer the silence interval of the new mutant before its expression, the better its chances of getting fixed in the population. However, this effect will be observable only to a certain limit after which further increase in the interval will not have an effect on the expected fixation rate. Two divergences from the simple scenario are also investigated: the nonsynchronous expression of the focal allele and its emergence in a spatial grid limited by a migration bottleneck. In all cases, it is shown that there is an evolutionary pressure inherent in the dynamics of positive frequency dependent selection on the genes to have a delayed expression. It is hypothesized that the regulation systems of such traits will be under selection for either internal silencing mechanisms or to be triggered by an external environmental condition.

1. Introduction

Genetic innovation happens via random mutations. A beneficial mutation, one which leads to an increase in the fitness of its carrier, will be favored by natural selection and its frequency will increase in the population. Yet, the relationship between the fitness of an allele and its frequency in the population goes in both directions. An allele whose fitness depends on its abundance in the population is said to be under frequency dependent selection. This can be further classified into two categories. In negative frequency dependent selection (NFDS), the allele's fitness increases if it is rarer in the population. Examples of NFDS are selection pressures acting on flu strains spreading in a population where resistance is more likely to be developed for the more common strains, giving rarer strains advantage (Williams et al. [1]) and also in rock-paper-scissor mating dynamics where rare mating strategies enjoy higher probability of getting advantageous interactions (Sinervo and Lively [2]). Moreover, such dynamics come to play in Fisher's principle of balancing selection on sex ratio; if one sex is rarer, it has higher chances of spreading its genes in the population (Fisher [3]). Emergence of new alleles via this mode of selection is relatively straightforward as the rare phenotype gets positively selected in population. Different dynamics however arise when considering positive frequency dependent selection (PFDS). PFDS is known to be a common mechanism in situations where there is selection for signaling and communication systems, as the more common the signal, the more efficient it is (Cartwright [4]). Examples can be found in the evolution of warning signals in predator-prey dynamics (Chouteau et al. [5]), the evolution of

quorum sensing communication system in bacterial communities (Eldar [6]), kin recognition systems to mediate altruistic social strategies among related animals (Sheehan et al. [7]), in models of assortative mating where most common genotypes enjoy selection advantage compared to rare alleles (Otto et al. [8]), the evolution of toxin resistance in situations of bacterial competition (Libberton et al. [9]), and selection dynamics of RhD alleles in a human population (Flegr et al. [10]). PFDS dynamics also arise in the control of mosquito-borne diseases via a biological agent. One strategy to control mosquito-borne human diseases is by introducing specific strains of *Wolbachia* bacteria to the population. *Wolbachia* is known then to undergo PFDS where its fitness increases as its frequency increases in the population (Hoffmann [11]), and hence, for *Wolbachia* to invade a population, it has to exist first at a proportion of the population higher than a critical "invasion" threshold, in order for the positive selection effect resulting from CI mechanism to outweigh the associated fitness cost (Turelli [12]; Reuter et al. [13]). An example of PFDS that we would like to highlight is the evolution of tolerance as a host defence response to pathogens. An allele inferring tolerance allows the host to cope with a disease by reducing the fitness offset associated with infection (Strauss and Agrawal [14]). Tolerance mechanisms were first observed in plants, which tolerate parasites and herbivores by increasing the number of their branches, as well as the size of their leaves and their chlorophyll concentration (Strauss and Agrawal [14]). In the animal kingdom, tolerance strategies have been first observed in mice to Plasmodium infection, the parasite causing malaria (Råberg et al. [15]). Additionally, sooty mangabey can tolerate simian immunodeficiency virus (SIV), consequently avoiding developing acquired immune deficiency syndrome (AIDS) (Chahroudi et al. [16]). Evolving tolerance is one of the two possible responses for a host against an infection. The other strategy is developing a resistance mechanism which limits or eliminates infection incidences. Both options require a reallocation of the host resources and hence come with a physiological cost. Nonetheless, in each case, the host's fitness increases through different evolutionary dynamics, as shown by Roy and Kirchner [17]. Host resistance mechanisms lead to a reduction in the pathogen prevalence in the population, hence reducing infection risk. As a result, the benefits associated with having resistance mechanisms diminish as the gene becomes more common in the population, while its cost remains the same. It follows then that a resistance gene will never get fixed in a population under selection forces alone. On the other hand, when a host develops tolerance, it can still survive and reproduce despite carrying the disease, consequently promoting the dissemination of the pathogen. Therefore, as a tolerance mechanism spreads in a population, pathogen prevalence also increases, leading to a subsequent fitness advantage for the organisms carrying the tolerance allele compared to those who do not. PFDS dynamics come to play in such situations. Hence, for a resistance allele, emergence is most feasible, fixation is most difficult, and the opposite is true for tolerance mechanisms. Furthermore, if the initial benefits from having a tolerance strategy does not offset the physiological cost, emergence will be not feasible by selection (Roy and Kirchner [17]).

To conclude, generally in PFDS, an allele's fitness improves as it becomes more common in the population. This means that the allele frequency is least likely to increase when it is rare, which poses the questions of how a focal allele that is getting selected against when it is rare can emerge in an evolving population?

2. Materials and Methods

2.1. Problem of Emergence of Altruism in Evolutionary Game Theory. A field that typically deals with the problem of how new phenotypes can emerge in a population is evolutionary game theory (EGT), founded by Smith and Price [18]. It defines a formal mathematical framework to study how the frequencies of different phenotypes/strategies change in an evolving population. One of the central problems of EGT is the evolution of altruistic behaviors. The problem is typically studied in the context of a prisoner's dilemma game. The game set-up, summarized in Table 1, is as follows: a player can be either a cooperator or a defector. If two cooperators interact together, each of them gets $g - c$. Two defectors on the other hand will get nothing when encountering each other. Hence, a homogeneous group of cooperators will fare better than an analogous population of defectors. However, when a defector plays with a cooperator, the cooperator is exploited, gets $-c$, and the defector reaps g. The primary question posed here is how cooperation can be evolutionary stable in face of cheaters. An answer was found by including a spatial component to the game (Nowak and May [19]). When the game is played by a population distributed over a grid, the cooperator phenotype can be evolutionary stable, if it was initially found in a large enough cluster of individuals. This is because in such case, the benefits resulting from the intracooperators interactions inside the cluster will outweigh the losses due to the exploitation at the boundaries (Nowak and May [19]; Fu et al. [20]; Lindgren and Nordahl [21]; Gang et al. [22]). While incorporating a spatial aspect to the game provides a reasonable solution to the problem of evolutionary stability of altruism, it still leaves another classical EGT question unanswered, the invasion question. Starting in a world of defectors, how can a lattice be invaded by rare mutations producing a cooperator phenotype? The challenge here is that if a mutation produced a single cooperative individual, it will get instantly heavily exploited by surrounding defectors and gets wiped out of the population. A recent solution to this problem has been provided by including a temporal component to the expression of the cooperative strategy. If the mutation stayed silent, unexpressed, long enough, it can spread via stochastic processes till reaching enough individuals to form a cluster exceeding the critical size, and hence the new trait becomes evolutionary stable once later expressed (Hashem et al. [23]). The problem of emergence of cooperation shares some similarities with the problem of emergence of a that is subjected to PFDS. In both cases, a mutation producing the new trait will face adverse negative selection uphill to climb, and in both cases, there is a certain prevalence after which the new

Table 1: Payoffs for two agents a and b engaged in a prisoner's dilemma game, $g > c > 0$.

	b	
a	C	D
C	$g - c$	$-c$
D	g	0

trait becomes favored by selection. There are few differences, nonetheless, in the spatial prisoner's dilemma settings, the switch in dynamics happens swiftly after a critical limit, this is not necessarily the case in PFDS dynamics. Nonetheless, the solution concept for the classical EGT problem seems to be able to get extended to the general problem of emergence of traits experiencing PFDS.

2.2. Concept. In this paper, a mathematical model is constructed for a population of single locus diploid organisms, where a mutation gives rise to an allele which experiences PFDS, and the fate of that allele is investigated. Our model relies on the following two ideas: the new mutation has to remain silent, unexpressed for an interval of time, and a stochastic process has to come to play to spread it till reaching a favorable frequency. The core concept is that if the new mutant allele remained silent after its emergence for a number of a generations, it could spread via stochastic processes till it reaches a frequency that favors its spread in the population. We discuss the two ideas afterwards in detail.

The expression of any phenotype is generally the result of a complex interplay between the genome of the organism and its environment (Nachtomy et al. [24]; Carlberg and Molnár [25]). Genetic factors can affect the time and level of expression as well as its sensitivity to environmental factors (Prasun et al. [26]; Golding et al. [27]). The optimal genotype, in the eyes of natural selection, is the one which gives rise to the expression parameters leading to highest fitness, given a certain environment (Kioukis and Pavlidis [28]). The parameter of interest in this work is the time of gene expression. In some situations, an allele could stay silent, unexpressed, yet remain inheritable for a number of generations (Carlberg [29]). The expression of a gene can rely on environmental conditions for its activity (Atkinson and Walden [30]; Gibson [31]). This process is called epigenetic activation. (Huang [32]; Russo [33]). In such case, an allele will not get expressed except when the organism is at the requisite environmental conditions for its activity. Such conditions can be physical, like the temperature of the surroundings (Sturtevant [34]) or chemical, such as the environment's acidity (Silverman [35]; Olson [36]). Hence, if the allele appeared in unfavorable environmental conditions, it will not get expressed, yet it will remain inheritable. If the environment came later to favorable conditions, all the organisms carrying said allele will start expressing it. A gene also could remain silent due to "internal mechanisms." In trans-generational gene silencing, RNA molecules act to suppress the expression of the gene for a number of generations (Qutob et al. [37]), which can reach 80 generations in some cases (Vastenhouw et al. [38]). Furthermore, a genetic timer can act so that the gene stays unexpressed for a tunable number of generations, acting like a fuse for a time bomb. Here again, if such gene appears, it will remain silent yet inheritable (Houri-Ze'evi et al. [39]; Waldron [40]).

An unexpressed gene does not have natural selection vouching for it, yet it can still spread in a population. Two possible mechanisms are genetic drift and genetic hitchhiking. Genetic hitchhiking refers to the process by which a neutral allele increases in frequency due to being associated with another gene on the same chromosome that is undergoing a selective sweep (Smith and Haigh [41]). Genetic drift on the other hand does not require the assumption of selection acting on an additional locus. It is the stochastic changes in the frequency of an allele due to random sampling effects, thought to play a key role in the fixation of neutral alleles (Kimura et al. [42]). Here, we would like to generalize the idea from spatial evolutionary game theory to the general case of the emergence of an allele in a population, while being under PFDS. The situation we would like to illustrate is how the dynamics of PFDS exert an evolutionary pressure on the novel mutations to stay silent for a number of generations after their inception, such that the rare alleles may increase in frequency via genetic drift till reaching a concentration where they are favored by the force of natural selection.

2.3. Model. We would like to investigate the fate of a mutation which produces an allele that experiences PFDS dynamics, such that the selection coefficient associated with the new allele is less than zero when it is less frequent in the population and larger than zero once it becomes the more established allele in the population. Hence, such mutations are expected to be wiped away from the population by selection forces. We are especially interested in the effect of the activation time of the allele on its chances to be fixed in the population. The allele will be assumed to be silent for a number of generations, during which the selection coefficient is zero, and then it will be expressed in all the organisms carrying the mutation synchronously. The trigger of the activation process will not be explicitly modeled; however, it can be interpreted as an environmental factor that is necessary for the activation of the gene or an internal, genetic, timing mechanism. The silence interval could help the mutation to spread via stochastic processes till it reaches favorable frequency for a subsequent spread via selection forces. Afterwards, two extensions of the model will be incorporated. First, the effect of a nonsynchronous activation of the focal allele will be investigated whether it is crucial for the allele's fixation. Additionally, a spatial extension of the model will be simulated to investigate how the migration bottleneck influences the outcomes of the model.

Consider a single locus diploid population with two alleles A_1 and A_2, whose frequencies in the population for a given generation i are denoted by p_i and q_i, respectively. The frequencies of the three possible genotypes, A_1A_1, A_1A_2, and A_2A_2, are assumed to be at Hardy-Weinberg equilibrium and are summarized in Table 2, with η_i as the frequency of the genotype, π_i as its fitness, h as the heterozygous effect,

Table 2: Frequencies and expected frequencies of possible genotypes, formed by two alleles A_1 and A_2 in a sexual population at a Hardy-Weinberg equilibrium.

Genotype	A_1A_1	A_1A_2	A_2A_2
η_i	p_i^2	$2p_i$ qi	q_i^2
π_i	$1+s$	$1+hs$	1
η_{i+1}	$p_i^2(1+s)/\overline{\omega_i}$	$2p_i\,q_i(1+hs)/\overline{\omega_i}$	$q_i^2/\overline{\omega_i}$

assumed to be equal to 0.5, s as the selection coefficient of the allele A_1, and $\overline{\omega_i}$ as the average fitness of a given generation, expressed by the following equation:

$$\overline{\omega_i} = p_i^2(1+s) + 2p_iq_i(1+\text{hs}) + q_i^2. \tag{1}$$

Hence, the expected frequency of the focal allele A_1 in the next generation, p', is calculated as follows:

$$p' = \frac{p_i^2(1+s)}{\overline{\omega_i}} + \frac{p_iq_i(1+hs)}{\overline{\omega_i}}. \tag{2}$$

It follows that the frequency of A_1 is expected to increase whenever $s > 0$ and declines if $s < 0$. However, in a frequency dependent selection, s is a function of the frequency of A_1 itself. The scenario of interest is the emergence of the focal allele A_1 in a population consisting exclusively of the allele A_2, when the selection dynamics of A_1 is positive frequency dependent. The relationship between the selection coefficient against the allele frequency is shown in Figure 1(b). In this work, the focal allele does not become expressed, except when the number of generations t elapses an interval of time t', termed as a silence interval. Therefore, the selection coefficient of the emerging allele can be modeled as

$$\begin{aligned} s_i &= 0, \\ &= 2^{2p_i-1} - 1, \\ t &< t', \\ t &\geq t'. \end{aligned} \tag{3}$$

An overview of the model's parameters is provided in Table 3. The simulation model is illustrated by the flowchart depicted in Figure 1(a), a single mutation happens in a population consisting of A_2 alleles. The expected change in frequency due to selection is calculated via equation (2). After that, an individual-based simulation is performed to generate the next generation. (i) In a loop, each A_1 allele gets passed to the next generation with a transition probability, P_{trans}, which is set to be equal to p'. (ii) After that, the rest of the population gets filled with A_2 alleles. Hence, while the expected fraction of A_1 alleles in the next generation should be equal to p', the stochastic nature of this step simulates the drift effect. Finally, (iv) after the number of generations exceeds a certain t', both the selection coefficient and p_{trans} become dependent on the fraction of A_1 in the population.

Furthermore, the model is later extended to a spatial model in which the population is distributed over a grid with periodic boundary conditions. Migration takes place between a focal cell and the cells in its von Neumann neighborhood, as shown in Figure 1(c). Consequently, this cellular automata model can be used to investigate the effect of migration in a structured population (van Dijk et al. [43] and Sherratt [44]). Hence, in this set-up, the P_{trans} in a certain patch will be affected by p' at the focal patch as well as neighboring patches, depending on the migration coefficient, as expressed in the following equation:

$$P_{\text{trans}} = (1-m)p'_{i,j} + \frac{m}{4}\left(p'_{i+1,j} + p'_{i-1,j} + p'_{i,j+1} + p'_{i,j-1}\right). \tag{4}$$

Hence, for the new generation, in a certain patch, there will be a probability $1-m$ that the A_1 allele is descendant from an allele of the same patch and a probability m that it is a result of a migration event. The selection coefficient here also becomes a local property which depends on the allele frequency in a patch. The simulation is initialized by a single mutation in a random patch for the first generation. Since the boundary conditions are periodic, all patches are mathematically equivalent.

All simulations have been run using MATLAB R2020b on a 64-bit Windows 10 workstation with an Intel(R) Core(TM) i7-7820HQ CPU @ 2.90GHz processor and 32.0 GB RAM. For the cellular automata model, the time complexity of the simulations is linearly dependent on both the number of patches, N_p, and the number of experiments, N_e. The time complexity is found to be $O(N_p \times N_e)$, which approaches $O(N_e)$, as $N_e \gg N_p$ (Panahi and Navimipour [45]; Neghabi et al. [46]). Finally, for every set of conditions, all the simulations have been repeated for 2×10^7 times and the mean of the normalized fixation rates obtained has been plotted, with a standard deviation <1% for all results.

3. Results and Discussion

Genetic drift refers to the process by which allele frequencies in a population are subjected to random fluctuations due to the effect of chance in natural selection. Its role is known to be more significant in small populations (Ewens [47]). In our model, the population size is equal to 1000 individuals and the update process includes a stochastic element through which genetic drift can operate. For a neutral allele, in the absence of any selection forces, the probability that it will get fixed in the population due to genetic drift, its fixation rate, is equal to its initial frequency in the population, here $1/2N_p$. For this model, the fixation rate is calculated as the ratio of the number of experiments that ended with the fixation of the mutation to the total number of experiments. An allele under PFDS on the other hand experiences negative selection when it is rare and positive selection once it gets established in the population. Hence, its emergence via random drift is not straightforward. In Figure 2, we investigate the effect of the length of the silence interval for an allele under PFDS on its normalized fixation rate in the population, relative to the fixation rate of a neutral allele. It can be seen that for an instantly expressed allele, when the simulation is run for $2 * 10^7$ times, the fixation rate of an instantly expressed allele is equal to 0.0%. The selection against the rare allele prohibits its increase in frequency till it can enjoy positive selection and hence it goes extinct. Yet, as the length of the silence interval increases, the fixation rate of

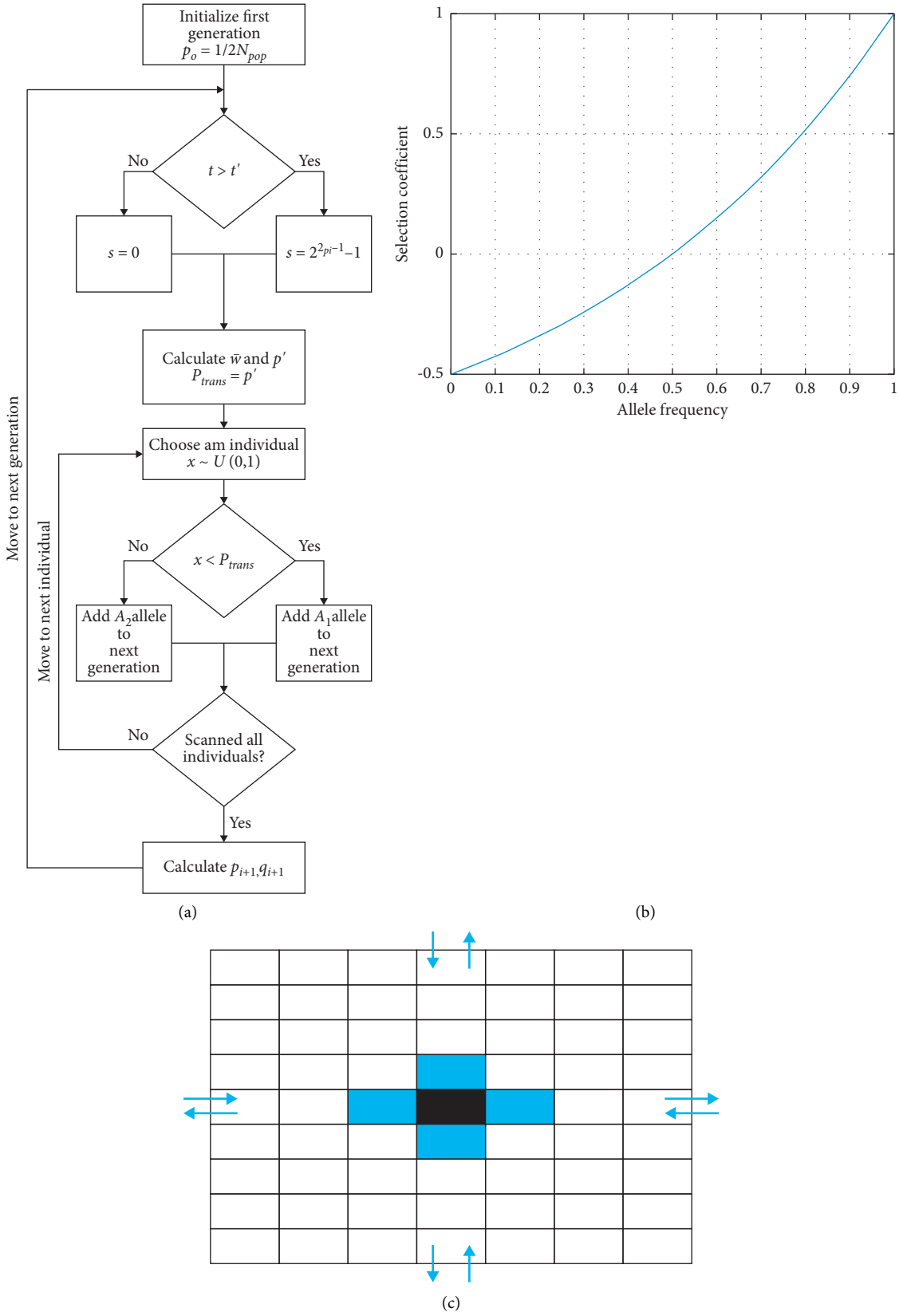

Figure 1: (a) An individual-based model for the emergence of a positive frequency dependent allele. (b) Selection coefficient of the said allele is negative when it is rare in the population, and is positive when the allele becomes more frequent within the population. (c) An extension of the model to a spatial grid with periodic boundary conditions where migration happens between residents of any focal patch and the surrounding patches in its von Neumann neighborhood.

Table 3: Model parameters.

Parameter	Definition
p_i	The frequency of the allele A_1 in a given generation i
q_i	The frequency of the allele A_2 in a given generation i
p'	The frequency of the allele A_1 in the next generation $i+1$
ω_i	The average fitness of a given generation i
s	The selection coefficient of the allele A_1
h	The heterozygous effect

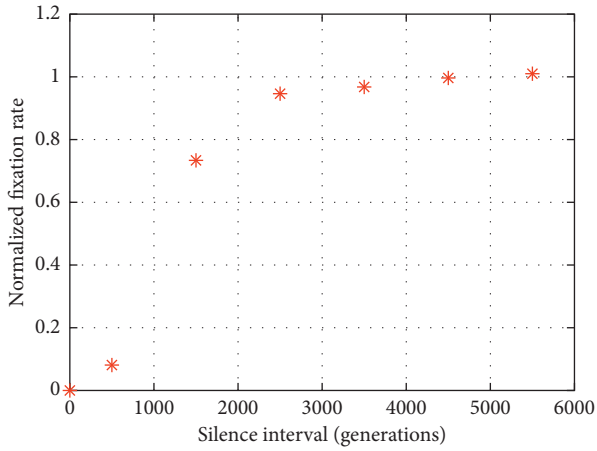

Figure 2: The relationship between the normalized fixation rate of the A_1 allele and the duration of its silence interval in case of a fully synchronous activation.

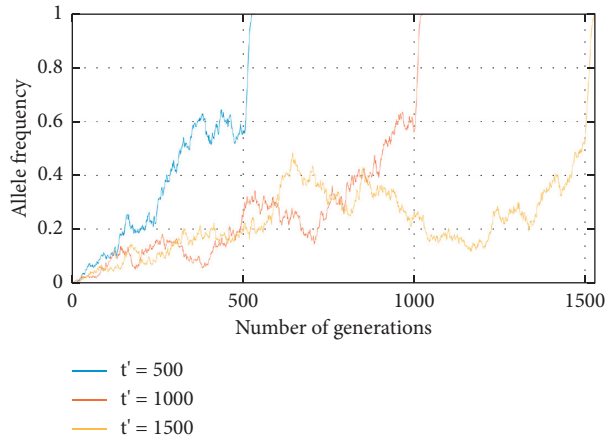

Figure 3: Evolution of the frequency of the focal allele A_1 in case of reaching fixation, starting from a single mutant with a silence time of 500, 1000, and 1500 generations.

the allele increases till it becomes equivalent to the fixation rate of a neutral allele.

Examples of successful fixations of a rare allele are examined in Figure 3. Here, the relationship between the number of generations and allele frequency is plotted for a silence interval equal to 500, 1000, and 1500 generations. It is seen that for fixation to happen, the allele frequency has to increase in the population via genetic drift till reaching a favorable frequency by the end of its silence interval. Hence, once the allele gets activated in all individuals carrying it, it becomes frequency dependent and experiences positive selection. Positive feedback loop dynamics arise here as the allele's fitness is an increasing function of its frequency in the population, hence the higher the frequency, the higher the fitness, which in turn leads to a subsequent increase in the frequency of the allele due to selection and so on. The result is that it gets rapidly fixed in the population, which is the reason for the sharp increase in frequency at the end of each silence interval. Furthermore, two aspects are to be highlighted about this figure. First, this figure displays examples of successful invasions for different silence intervals, and thus, for a successful invasion to be likely to happen, the allele frequency must be at a favorable frequency by the genetic random drift at the time of the activation. Hence, shorter silence intervals will require the allele to reach favorable frequency faster, for a successful invasion to occur. The second aspect is that one can notice the rapid increase in frequency that happens once the allele gets activated in a majority of the population; this is due to the dynamics of positive frequency dependent selection, which makes the fitness of the allele higher as its frequency in the population increases, leading to a positive feedback loop and rapid fixation within the population. It must be noted that one of the critical features of the model that allowed this effect to happen is that the allele gets activated at the same time in all the organisms carrying it, and hence the frequency of the expressed allele jumps above the limit after which it gets positively selected via evolution. It is reasonable to think that such perfectly synchronous expression of the allele would not happen in nature, whatever the underlying trigger of its activity. For example, if the allele needs specific environmental conditions to get expressed, the environment is unlikely to change homogeneously for all the population. Hence, we next investigate the effects of nonsynchronous activation.

What would happen if the silent alleles in the population do not get synchronously expressed after the elapse of the silence interval? To test that, the activation time of the alleles in the population is allowed to be a random variable drawn from a normal distribution, such that $t' = N(\mu, \sigma)$. This way, the focal allele in the population will not get expressed at all the individuals of the same generation carrying it. This makes it more difficult for the allele to survive as the fraction of the alleles that will become active first will experience negative selection, and it will continue to do so till the proportion of active alleles exceeds 50%. When including a variance of 4 generations, as shown in Figure 4, the relationship between the silence interval length and the fixation

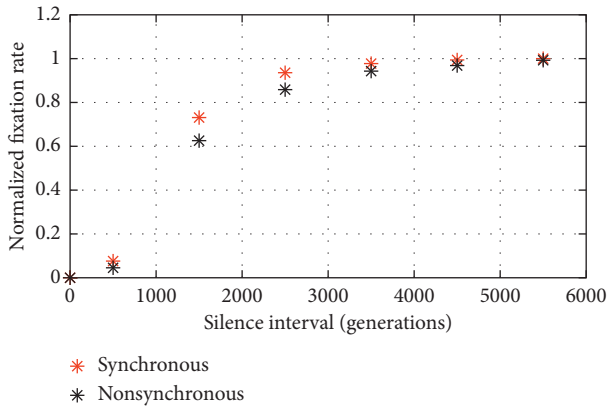

Figure 4: The discrepancies between the fixation rates of the A_1 allele in case of a synchronous activation and a nonsynchronous activation with a standard deviation of two generations.

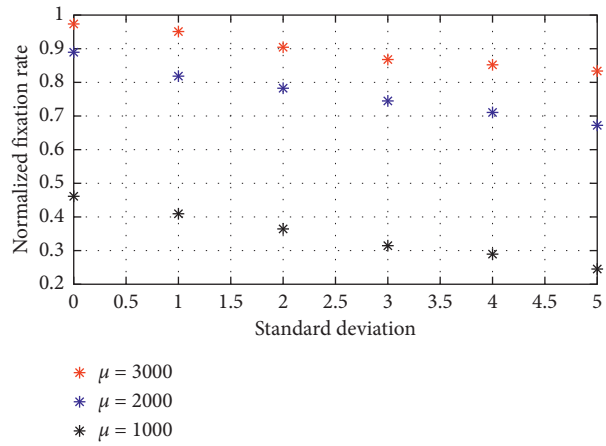

Figure 5: The relationship between the fixation rate of the focal allele A_1 and the value of the standard deviation of the expression time t', in case of nonsynchronous expression for different expression times.

rate had a similar overall behavior to the synchronous activation case while generally having significantly less fixation rate for the same silence interval length. Additionally, it could be observed that while there are notable differences between the fixation rates for synchronous versus nonsynchronous activation, at shorter silence intervals, this gab progressively diminishes for longer silence intervals. The reason for that is that the longer the silence interval, the more time an allele has to increase its frequency in the population. Hence, eventually, the allele will be expressed at high frequency within the population, experiencing strong positive selection and thus outweighing the negative selection that will occur at the beginning to the alleles which will be activated earlier than the rest, and this will result in fixation rates comparable to the ones observed in case of perfectly synchronized expression, as noted in the figure.

The effect of the variance of the activation time on the fixation rate is investigated further in Figure 5. For a constant activation mean, the standard deviation is increased progressively. The fixation rate is observed to decrease linearly with increasing variance. For the same activation mean, the higher the variance, the higher the proportion of alleles getting "prematurely" expressed. These alleles are likely to face adverse circumstance, being frequency dependent, their fitness will be lying on low end of the spectrum, and they will get selected against until the rest of the silent alleles also get expressed. So, while a synchronous activation of the silent alleles tend to maximize the fixation rate, introducing the variance reduces but does not diminish positive effect of silence. While the alleles that get activated early face negative selection, they do not get wiped out completely from the population before the activation of the rest of the silent alleles. However, once the rest of the alleles get activated, a positive selection force is generated and fixation can occur. Another idealized aspect of the model so far is that the population is fully mixed with smooth gene flow between different segments of the population. When considering a real population that is spatially distributed, gene flow will be restricted between different segments, depending on the extent of the migration forces. This in turn will put a limit on genetic drift's action of spreading the allele to a sufficient extent. We consider the spatial case next.

The model is extended to simulate the emergence of an allele under PFDS in a population distributed spatially. Here, we look into the fate of a single focal allele one patch of the grid. A single successful run is illustrated in Figure 6, where the activation time has been set to 1000 generations. We start by a grid full of A_2 alleles, which is seeded with a single A_1 allele at the center of the grid. Since selection forces are local, depending on the local frequency of the allele at each patch, the allele needs to propagate in silence to enough patches in the grid for subsequent activation to cause fixation. The two limit cases here are that if the migration rate is close to zero, it will act as a bottleneck for the action of drift, the focal allele will not spread, and when it gets activated, its action will remain locally limited. On the other hand, a migration rate of one brings us back to the fully mixed state. Here also, it is noticed in Figure 7 that for different lengths of the silence interval, the allele frequency in successful runs has increased via genetic drift till it exceeded 50% of the population. After that, activation occurs and selection forces took over.

Genetic drift is expected to perform slower on a spatial grid as low migration rates act to preserve the heterogeneity of different patches. In Figure 8, it is observed that also in the spatial case, the instantly expressed alleles cannot make it to fixation. Additionally, the evolution of the fixation rate with increasing silence time is similar to the mixed case, reaching a plateau later, while being with significantly less fixation rate than the fully mixed case for the same activation time. Figure 9 shows that these features are general regardless of the value of the migration coefficient. Also, for a fixed interval length, the fixation rate will improve with increasing migration coefficient. As migration increases, genetic drift will act quicker to spread the focal allele to a suitable frequency such that it gets picked up by natural selection. It is seen from the two model extensions that while the fixation rate will decrease compared to the ideal case, the general relationship between the ability of allele to undergo delayed expression and its emergence in the population remains valid.

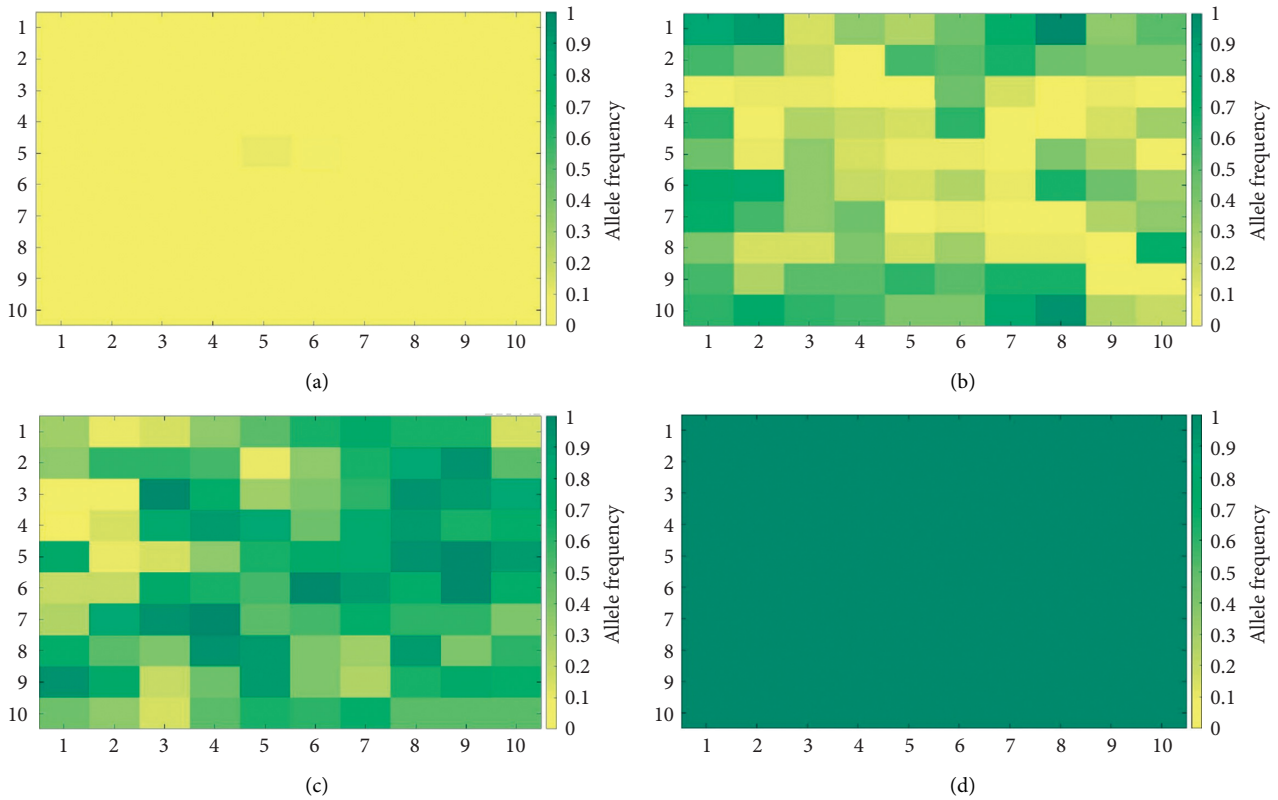

FIGURE 6: Evolution of the frequency of the focal allele A_1 distributed over a spatial grid. The allele frequency increases by genetic drift and it spreads in the population via migration till getting activated at $t' = 1000$ generations, resulting in subsequent fixation of the allele. (a) $t = 0$. (b) $t = 500$. (c) $t = 1000$. (d) End of the solution.

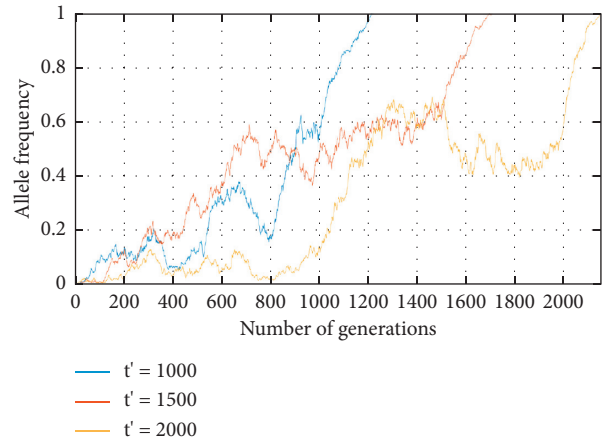

FIGURE 7: Evolution of the frequency of the focal allele A_1 in the population, starting from a single mutant in a spatial grid with migration between neighboring patches, with a silence interval of 1000, 1500, and 2000 generations.

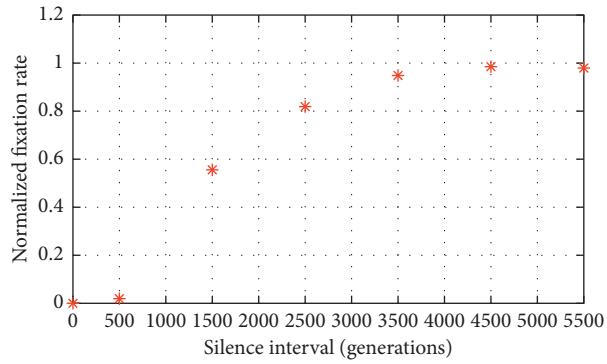

Figure 8: The relationship between the normalized fixation rate of the A1 allele and the duration of its silence interval in case of a population distributed over a spatial grid with $m = 0.1$.

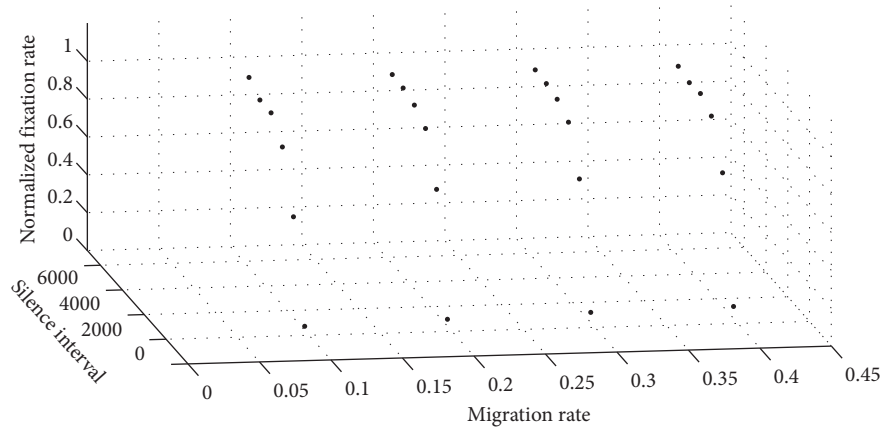

Figure 9: The relationship between the normalized fixation rate of the A_1 allele, the duration of its silence interval, and different values of the migration rate parameter in a spatial grid.

4. Conclusion

A recently proposed solution to the problem of emergence of cooperation, one of the central questions of evolutionary game theory, relied on the time of expression of the strategy. The background of the problem dealt with here is that, for a cooperative strategy to be evolutionary stable on a spatial grid, it has to be played by a large enough group of players distributed over clusters with size that exceeds a certain critical limit. Hence, the natural following question is how could a cooperative strategy perform an invasion to a population starting from a rare mutant. A proposed solution relied on the time of expression of the mutant strategy, if it remained silent yet inheritable, for a long enough interval. It could spread in the population by the virtue of being associated with another advantageous trait, in analogous with the genetic hitchhiking phenomenon. This problem shares some features with the emergence of a trait under PFDS, where the mutant allele will be at a fitness minimum in the moment of its inception in the population, as its fitness decreases with decreasing frequency, and if selection works against the novel allele when it is rare, a critical frequency will be essential for it to get fixed as well. PFDS is known to play a key role in the evolution of signaling and communication systems in both the animal and the microbial worlds, in strategies which rely on introducing biological agents to manage mosquito-borne diseases and in the spread of tolerance-inducing alleles in a population facing a disease (Hamilton [48]). Hence, understanding how such mutations can emerge in a population can help in both deciphering and manipulating such phenomena. Hence, in this paper, we apply and generalize the core concepts from the EGT solution to the general problem of emergence of biological innovation via mutations under PFDS.

Previous research has been directed to the impact of PFDS on alleles that are already established in the population and hence positively selected (Chouteau et al. [5]; Otto et al. [8]; Lehtinen et al. [49]; Trotter and Spencer [50]; Huang et al. [51]). Here, the focus is directed on the emergence of such alleles in a population. In terms of the implemented methods, a cellular automata spatial model has been developed to investigate the effect of immigration on the results following van Dijk et al. [43] and Sherratt [44]. The framework developed in this paper is general for populations under PFDS dynamics. The main idea is that if a mutation remained silent for long enough time, it could spread in the population via stochastic processes, till reaching a favorable frequency for selection forces. A mutation could remain silent either due to the nonsuitability of the environmental conditions for its expression or due to a built-in generational genetic timer. Whenever a mutation is silent, it is considered to be neutral in the eyes of natural selection. In a

single locus diploid population, a neutral mutation can still spread in the population via genetic drift and stochastic fluctuations in alleles' frequencies due to random sampling effects. A model was built to investigate the rate of fixation of a trait under PFDS in an ideal fully mixed population where the allele experiences activation synchronously in all organisms carrying it after the elapse of a silence interval of varying length. It is found that the rate of fixation improves with increasing silence interval till it reaches a plateau, approaching fixation rates by genetic drift for a neutral allele.

With respect to the population size, the spread of an allele to a high frequency in a population is expected to be faster, the smaller the population is. Hence, while shorter silence intervals could be sufficient for the focal allele to get fixed in a small population, larger populations will require the allele to have significantly longer silence intervals so that it does not get expressed before reaching a high enough frequency to get positively selected within the population. Genetic drift is well known to be most observable in smaller populations (LaBar and Adami [52]), and while no minimum population size is theoretically required for this model, the population size was nevertheless chosen to be 1000 individuals in accordance with previous genetic drifts models in literature (Bataillon et al. [53]; Saunders et al. [54]). One must note as well that the emergence of a mutation under PFDS will become more difficult as the population size increases, assuming perfect mixing, since this will increase the probability that the activation of the allele happens before its spread to a sufficient extent within the population.

Subsequently, two nonidealistic deviations of the model were investigated: the case of nonsynchronous activation and the case of a population which is distributed spatially, not ideally mixed. In both extensions, the fixation rate decreases, compared to the idealized case, however a silence interval still correlates with an improved fixation rate compared to an instantly expressed mutation and the general features of the invasion dynamics still remained similar to the fully mixed case. The expression of a trait is usually the result of an orchestrated action of numerous factors, either internal regulation mechanisms or external environmental inputs, and natural selection is the primary audience which makes sure that the gene expression process operates optimally. This paper shows, based on a general mathematical model, that in case of emergence of traits under PFDS, factors that are related to the expression time will be under an evolutionary pressure to keep the novel allele silent for a number of generations, as this will increase the chances of the new trait to emerge in the population. Silence could be a virtue in the backstage of the evolutionary theater.

Data Availability

All data are available on request from the corresponding author.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgments

This work was supported by the KU Leuven Research Council (OPTEC Center-of-Excellence Optimization in Engineering OPTEC and Project C24/18/046), by the ERA-NET FACCESurPlus FLEXIBI Project, cofunded by VLAIO Project HBC.2017.0176, by the Fund for Scientific Research-Flanders (Projects G.0863.18 and G.0B41.21N), and by the European Union's Horizon 2020 Research and Innovation Programme (Marie Sklodowska-Curie Grant Agreement Numbers 813329 and 956126). VDB was supported by FWO-SB Grant 1SC0920N.

References

[1] E. S. C. P. Williams, N. M. Morales, B. R. Wasik, V. Brusic, S. P. J. Whelan, and P. E. Turner, "Repeatable population dynamics among vesicular stomatitis virus lineages evolved under high co-infection," *Frontiers in Microbiology*, vol. 7, p. 370, 2016.

[2] B. Sinervo and C. M. Lively, "The rock-paper-scissors game and the evolution of alternative male strategies," *Nature*, vol. 380, no. 6571, pp. 240–243, 1996.

[3] R. Fisher, *The Genetical Theory of Natural Selection*, Oxford University Press, Oxford, UK, 1930.

[4] R. A. Cartwright, "Bards, poets, and cliques: frequency-dependent selection and the evolution of language genes," *Bulletin of Mathematical Biology*, vol. 73, no. 9, pp. 2201–2212, 2011.

[5] M. Chouteau, M. Arias, and M. Joron, "Warning signals are under positive frequency-dependent selection in nature," *Proceedings of the National Academy of Sciences*, vol. 113, no. 8, pp. 2164–2169, 2016.

[6] A. Eldar, "Social conflict drives the evolutionary divergence of quorum sensing," *Proceedings of the National Academy of Sciences*, vol. 108, no. 33, pp. 13635–13640, 2011.

[7] M. J. Sheehan, C. Miller, and H. K. Reeve, "Identity signaling and patterns of cooperative behavior," *Integrative and Comparative Biology*, vol. 57, no. 3, pp. 580–588, 2017.

[8] S. P. Otto, M. R. Servedio, and S. L. Nuismer, "Frequency-dependent selection and the evolution of assortative mating," *Genetics*, vol. 179, no. 4, pp. 2091–2112, 2008.

[9] B. Libberton, M. J. Horsburgh, and M. A. Brockhurst, "The effects of spatial structure, frequency dependence and resistance evolution on the dynamics of toxin-mediated microbial invasions," *Evolutionary applications*, vol. 8, no. 7, pp. 738–750, 2015.

[10] J. Flegr, R. Hoffmann, and M. Dammann, "Worse health status and higher incidence of health disorders in rhesus negative subjects," *PLoS One*, vol. 10, no. 10, Article ID e0141362, 2015.

[11] A. A. Hoffmann, "Cytoplasmic incompatibility in insects," *Influential Passenger: Inherited Microorganisms and Arthropod Reproduction*, pp. 42–80, 1997.

[12] M. Turelli, "Evolution of incompatibility-inducing microbes and their hosts," *Evolution*, vol. 48, no. 5, pp. 1500–1513, 1994.

[13] M. Reuter, L. Lehmann, and F. Guillaume, "The spread of incompatibility-inducing parasites in sub-divided host populations," *BMC Evolutionary Biology*, vol. 8, no. 1, p. 134, 2008.

[14] S. Y. Strauss and A. A. Agrawal, "The ecology and evolution of plant tolerance to herbivory," *Trends in Ecology & Evolution*, vol. 14, no. 5, pp. 179–185, 1999.

[15] L. Råberg, D. Sim, and A. F. Read, "Disentangling genetic variation for resistance and tolerance to infectious diseases in animals," *Science (New York, N.Y.)*, vol. 318, no. 5851, pp. 812–814, 2007.

[16] A. Chahroudi, S. E. Bosinger, T. H. Vanderford, M. Paiardini, and G. Silvestri, "Natural siv hosts: showing aids the door," *Science*, vol. 335, no. 6073, pp. 1188–1193, 2012.

[17] B. A. Roy and J. W. Kirchner, "Evolutionary dynamics of pathogen resistance and tolerance," *Evolution*, vol. 54, no. 1, pp. 51–63, 2000.

[18] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.

[19] M. A. Nowak and R. M. May, "The spatial dilemmas of evolution," *International Journal of bifurcation and chaos*, vol. 03, no. 01, pp. 35–78, 1993.

[20] F. Fu, M. A. Nowak, and C. Hauert, "Invasion and expansion of cooperators in lattice populations: p," *Journal of Theoretical Biology*, vol. 266, no. 3, pp. 358–366, 2010.

[21] K. Lindgren and M. G. Nordahl, "Evolutionary dynamics of spatial games," *Physica D: Nonlinear Phenomena*, vol. 75, no. 1, pp. 292–309, 1994.

[22] W. Gang, G. Kun, Y. Han-Xin, and W. Bing-Hong, "Role of clustering coefficient on cooperation dynamics in homogeneous networks," *Chinese Physics Letters*, vol. 25, no. 6, pp. 2307–2310, 2008.

[23] I. Hashem, D. Telen, P. Nimmegeers, and J. Van Impe, "The silent cooperator: an epigenetic model for emergence of altruistic traits in biological systems," *Complexity*, vol. 2018, Article ID 2082037, 16 pages, 2018.

[24] O. Nachtomy, A. Shavit, and Z. Yakhini, "Gene expression and the concept of the phenotype," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, vol. 38, no. 1, pp. 238–254, 2007.

[25] C. Carlberg and F. Molnár, *Mechanisms of Gene Regulation*, Springer, New York, NY, USA, 2016.

[26] P. Prasun, M. Pradhan, S. Agarwal et al., "One gene, many phenotypes," *Journal of Postgraduate Medicine*, vol. 53, no. 4, pp. 257–61, 2007.

[27] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, "Real-time kinetics of gene activity in individual bacteria," *Cell*, vol. 123, no. 6, pp. 1025–1036, 2005.

[28] A. Kioukis and P. Pavlidis, "Evolution of gene regulatory networks by means of selection and random genetic drift," 2018, https://www.biorxiv.org/content/10.1101/449645v1, Article ID 449645.

[29] C. Carlberg, *Mechanisms of Gene Regulation*, Springer, Dordrecht, 2014.

[30] B. Atkinson and D. Walden, *Changes in Eukaryotic Gene Expression in Response to Environmental Stress*, Cell Biology - Academic Press, 1985.

[31] G. Gibson, "The environmental contribution to gene expression profiles," *Nature Reviews Genetics*, vol. 9, no. 8, pp. 575–581, 2008.

[32] S. Huang, *Epigenetic Gene Expression and Regulation*, Academic Press is an imprint of Elsevier, London, UK, 2015.

[33] V. E. A. Russo, *Epigenetic Mechanisms of Gene Regulation*, Cold Spring Harbor Laboratory Press, Plainview, N.Y, 1996.

[34] A. H. Sturtevant, "The himalayan rabbit case, with some considerations on multiple allelomorphs," *The American Naturalist*, vol. 47, no. 556, pp. 234–239, 1913.

[35] W. A. Silverman, "A cautionary tale about supplemental oxygen: the albatross of neonatal medicine," *Pediatrics*, vol. 113, no. 2, pp. 394–396, 2004.

[36] E. R. Olson, "Influence of ph on bacterial gene expression," *Molecular Microbiology*, vol. 8, no. 1, pp. 5–14, 1993.

[37] D. Qutob, B. Patrick Chapman, and M. Gijzen, "Transgenerational gene silencing causes gain of virulence in a plant pathogen," *Nature Communications*, vol. 4, no. 1, p. 1349, 2013.

[38] N. L. Vastenhouw, K. Brunschwig, K. L. Okihara, F. Müller, M. Tijsterman, and R. H. A. Plasterk, "Long-term gene silencing by RNAi," *Nature*, vol. 442, no. 7105, p. 882, 2006.

[39] L. Houri-Ze'evi, Y. Korem, H. Sheftel et al., "A tunable mechanism determines the duration of the transgenerational small \{RNA\} inheritance in c. elegans," *Cell*, vol. 165, no. 1, pp. 88–99, 2016.

[40] D. Waldron, "Regulating transgenerational epigenetics," *Nature Reviews Genetics*, vol. 17, no. 6, p. 315, 2016.

[41] J. M. Smith and J. Haigh, "The hitch-hiking effect of a favourable gene," *Genetical Research*, vol. 23, no. 1, pp. 23–35, 1974.

[42] M. Kimura et al., "Evolutionary rate at the molecular level," *Nature*, vol. 217, no. 5129, pp. 624–626, 1968.

[43] B. van Dijk, P. Hogeweg, H. Doekes, and N. Takeuchi, "Bacteria maintain slightly beneficial genes and selfish genetic elements through the evolution of horizontal gene transfer," 2020, https://www.biorxiv.org/content/10.1101/2020.02.13.947077v1.

[44] T. N. Sherratt, "Spatial mosaic formation through frequency-dependent selection in Müllerian mimicry complexes," *Journal of Theoretical Biology*, vol. 240, no. 2, pp. 165–174, 2006.

[45] V. Panahi and N. J. Navimipour, "Join query optimization in the distributed database system using an artificial bee colony algorithm and genetic operators," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 17, Article ID e5218, 2019.

[46] A. A. Neghabi, N. J. Navimipour, M. Hosseinzadeh, and A. Rezaee, "Energy-aware dynamic-link load balancing method for a software-defined network using a multi-objective artificial bee colony algorithm and genetic operators," *IET Communications*, vol. 14, no. 18, pp. 3284–3293, 2020.

[47] W. Ewens, *Mathematical Population Genetics 1: Theoretical Introduction. Interdisciplinary Applied Mathematics*, Springer, New York, NY, USA, 2012.

[48] M. B. Hamilton, *Population Genetics*, John Wiley & Sons, New Jersey, US, 2021.

[49] S. Lehtinen, J. S. Huisman, and S. Bonhoeffer, "Evolutionary mechanisms that determine which bacterial genes are carried on plasmids," *Evol. Lett.*vol. 18, 2020.

[50] M. V. Trotter and H. G. Spencer, "Models of frequency-dependent selection with mutation from parental alleles," *Genetics*, vol. 195, no. 1, pp. 231–242, 2013.

[51] W. Huang, B. Werner, and A. Traulsen, "The impact of random frequency-dependent mutations on the average population fitness," *BMC Evolutionary Biology*, vol. 12, no. 1, pp. 1–10, 2012.

[52] T. LaBar and C. Adami, "Evolution of drift robustness in small populations," *Nature Communications*, vol. 8, no. 1, pp. 1012–12, 2017.

[53] T. M. Bataillon, J. L. David, and D. J. Schoen, "Neutral genetic markers and conservation genetics: simulated germplasm collections," *Genetics*, vol. 144, no. 1, pp. 409–417, 1996.

[54] P. A. Saunders, S. Neuenschwander, and N. Perrin, "Sex chromosome turnovers and genetic drift: a simulation study," *Journal of Evolutionary Biology*, vol. 31, no. 9, pp. 1413–1419, 2018.