

On the Robustness of the Extension of the One-Third Law of Evolution to the Multi-Player Game

Sabin Lessard

Published online: 16 April 2011
© Springer Science+Business Media, LLC 2011

Abstract We show that the extension of the one-third law of evolution from the 2-player game to the d -player game is the same for all exchangeable models in the domain of application of the Kingman coalescent in the limit of a large neutral population. The extension relies on an approximation of the probability of fixation of a single mutant in terms of expected times in ancestral sample states that are calculated by induction. An interpretation based on the concept of projected average excess in payoff is discussed.

Keywords Evolutionary games · Multi-player game · One-third law of evolution · Fixation probability · Exchangeable models · Projected average excess

1 Introduction

The one-third law of evolution states that weak selection favors a single mutant strategy A replacing a resident strategy B in the limit of a large finite population with random pairwise interactions among individuals and payoffs as selection coefficients if

$$x^* = \frac{b_0 - a_0}{a_1 - a_0 - b_1 + b_0} < \frac{1}{3}. \quad (1)$$

Actually this condition is necessary if the degenerate case $x^* = 1/3$ is ignored. Here, a_k and b_k are the payoffs to A and B , respectively, in interaction with A if $k = 1$ or in interaction with B if $k = 0$. Note that x^* is the equilibrium frequency of A with respect to the deterministic evolutionary dynamics in an infinite population. This law was first established for the Moran model [14] and the Wright–Fisher model [5, 8]. It has been shown to hold more generally for exchangeable selection models extending the neutral Cannings model but only

Research supported in part by the Natural Sciences and Engineering Research Council of Canada.

S. Lessard (✉)
Département de mathématiques et de statistique, Université de Montréal, C.P. 6128 Succursale
Centre-ville, Montréal, Québec H3C 3J7, Canada
e-mail: lessards@dms.umontreal.ca

in the domain of application of the Kingman coalescent in the limit of a large population size [9, 10].

Kurokawa and Ihara [7] and Gokhale and Traulsen [4] have extended the one-third law of evolution for a d -player two-strategy game in a Moran population. One individual is replaced at a time but the payoff to an individual depends not only on its own strategy, but also of the strategies of $d - 1$ other interacting individuals chosen at random. This includes what is known as the public goods game with the payoff depending linearly on the number of cooperators among the interacting individuals in the same group. In the limit of a large population size, the condition for weak selection to favor a single strategy A replacing a strategy B is

$$\sum_{k=0}^{d-1} (d - k)a_k > \sum_{k=0}^{d-1} (d - k)b_k, \tag{2}$$

where a_k and b_k represent the payoffs to A and B , respectively, in interaction with k A players and $d - 1 - k$ B players. Moreover, condition (2) is necessary in generic cases and does not depend on the assumption that payoff differences are linear in abundance.

In this note, we establish the validity of condition (2) for the d -player game under exchangeable reproduction schemes that fall in the domain of the Kingman coalescent in the limit of a large neutral population. Such reproduction schemes go well beyond the Moran model and require a different approach.

2 The Model

Consider interactions between players within random groups of fixed size d in a population of fixed size N . There are two types of players, A and B , and the payoffs to A and B are a_k and b_k , respectively, if there are k players of type A and $d - 1 - k$ players of type B among the other $d - 1$ players in the same group.

If there are j players of type A in the whole population such that the frequency of A is $x = j/N$, then the expected payoffs to A and B are

$$\pi_A(x) = \sum_{k=0}^{d-1} \frac{\binom{j-1}{k} \binom{N-j}{d-1-k}}{\binom{N-1}{d-1}} a_k \tag{3}$$

and

$$\pi_B(x) = \sum_{k=0}^{d-1} \frac{\binom{j}{k} \binom{N-j-1}{d-1-k}}{\binom{N-1}{d-1}} b_k, \tag{4}$$

respectively. We have the approximations

$$\pi_A(x) = \sum_{k=0}^{d-1} \binom{d-1}{k} x^k (1-x)^{d-1-k} a_k + O(N^{-1}) \tag{5}$$

and

$$\pi_B(x) = \sum_{k=0}^{d-1} \binom{d-1}{k} x^k (1-x)^{d-1-k} b_k + O(N^{-1}), \tag{6}$$

where $O(N^{-1})$ designates a function of order N^{-1} uniformly with respect to x , that is, $|O(N^{-1})| \leq cN^{-1}$ for N large enough and some constant c for $0 \leq x \leq 1$.

Assume that the payoff has an additive effect on the relative success in reproduction such that the fitness of A can be expressed as

$$f_A(x) = 1 + s_N \pi_A(x) \tag{7}$$

and the fitness of B as

$$f_B(x) = 1 + s_N \pi_B(x). \tag{8}$$

The quantity s_N represents an intensity of selection. It will be assumed throughout that $s_N = \sigma N^{-1}$, where the population-scaled parameter σ is supposed to be small and positive.

Time is discrete and an expected fraction γ of the population is replaced from one time step to the next. The case $\gamma = 1$ corresponds to non-overlapping generations as in the Wright–Fisher model [3, 18]. At the other extreme, we have $\gamma = N^{-1}$ in the case of single birth-death events as in the Moran model [13]. Note that the fraction of the population that is replaced may be a random variable. However, only its expected value will matter.

If the frequency of A at a given time step is x , then this frequency at the next time step will have an expected value x in the fraction of the population that is not replaced, and an expected value $x f_A(x) / \bar{f}(x)$ in the fraction that is replaced, where

$$\bar{f}(x) = x f_A(x) + (1-x) f_B(x) \tag{9}$$

is the mean fitness in the whole population. Here, we assume that the individuals replaced are chosen at random. Moreover, each of them is replaced, not necessarily independently from the others, by a copy of any given individual at the previous time step with a probability proportional to the fitness of that individual. Actually, we assume that the contributions of players of the same type from one time step to the next are exchangeable random variables whose expected value depends linearly on the fitness of the type. In the absence of selection ($\sigma = 0$), the contributions of all individuals are exchangeable as in the neutral Cannings model [1, 10].

Then the change in the frequency of A , denoted by Δx , will have a conditional expected value given by

$$E_\sigma(\Delta x | x) = \frac{\gamma x (f_A(x) - \bar{f}(x))}{\bar{f}(x)}. \tag{10}$$

Note that the conditional distribution of Δx depends not only on the selection pressure but also on the reproduction scheme.

3 Fixation Probability

Suppose that A is a mutant strategy represented only once initially (time step $t = 0$) and let $x(t)$ be the frequency of A at time step $t = 0, 1, 2, \dots$. In particular, we have $x(0) = N^{-1}$. As

a result of the combined effects of selection and drift, the frequency of A will converge as t goes to ∞ to a limit random variable $x(\infty)$ which will take the value 1 with some probability $\rho_A(\sigma)$ and 0 with the complementary probability $1 - \rho_A(\sigma)$. In fact, $\rho_A(\sigma)$ represents the probability of fixation of A as a function of the scaled intensity of selection. Note that this probability is given under neutrality by the initial frequency of A , that is, $\rho_A(0) = x(0)$. As a matter of fact, one of the individuals at $t = 0$ will be the ancestor of the whole population in the long run. Moreover, if no selection takes place, this ancestor will be one individual chosen at random at $t = 0$ by symmetry.

Following Rousset [16] and Lessard and Ladret [10], we write the limit frequency of A in the population as

$$x(\infty) = x(0) + \sum_{t \geq 0} \Delta x(t), \tag{11}$$

where

$$\Delta x(t) = x(t + 1) - x(t) \tag{12}$$

is the change in the frequency of A from time step t to time step $t + 1$. Taking expectation on both sides of the above equality and assuming mild regularity conditions [10], we obtain that

$$E_\sigma(x(\infty)) = x(0) + \sum_{t \geq 0} E_\sigma(\Delta x(t)). \tag{13}$$

Moreover, the law of total expectation gives

$$E_\sigma(\Delta x(t)) = E_\sigma(E_\sigma(\Delta x(t)|x(t))). \tag{14}$$

For σ small enough, (10) leads to the approximation

$$E_\sigma(\Delta x(t)|x(t)) \approx \frac{\gamma\sigma}{N} x(t)(1 - x(t))(\pi_A(x(t)) - \pi_B(x(t))). \tag{15}$$

Plugging (5) and (6), this can be written as

$$\begin{aligned} E_\sigma(\Delta x(t)|x(t)) &\approx \frac{\gamma\sigma}{N} \sum_{k=0}^{d-1} \binom{d-1}{k} x(t)^{k+1} (1 - x(t))^{d-k} (a_k - b_k) \\ &\quad + \frac{\gamma\sigma}{N} x(t)(1 - x(t)) O(N^{-1}). \end{aligned} \tag{16}$$

In summary, (13) yields

$$\rho_A(\sigma) \approx \frac{1}{N} + \frac{\gamma\sigma}{N} \sum_{k=0}^{d-1} d_k (a_k - b_k) + \frac{\gamma\sigma}{N} d_0 O(N^{-1}) \tag{17}$$

as a first-order approximation for the probability of fixation of A with respect to the intensity of selection, where

$$d_k = \binom{d-1}{k} \sum_{t \geq 0} E_0(x(t)^{k+1} (1 - x(t))^{d-k}). \tag{18}$$

Here, E_0 is used for the expected value in the neutral model, which differs only by terms of order σ from the expected value E_σ in the selection model.

In order to go further, we will assume that the following condition holds.

Assumption A The probability under neutrality for two individuals chosen at random without replacement at any given time step to have the same parent one time step back, represented by c_N , and the corresponding probability for three individuals, represented by b_N , satisfy

$$\lim_{N \rightarrow \infty} c_N = 0 \tag{19}$$

and

$$\lim_{N \rightarrow \infty} b_N c_N^{-1} = 0. \tag{20}$$

Assumption A means that c_N and b_N tend to 0 as N tends to infinity but b_N does it more quickly so that $b_N = o(c_N)$. This ensures that the probability of a 3-merger, and then the probability of a k -merger for any $k \geq 3$ by monotonicity, can be neglected compared to the probability of a 2-merger in the absence of selection when the population size is large enough. Actually this assumption guarantees that the ancestral neutral process in the limit of a large population size with c_N^{-1} time steps as unit of time will be the Kingman [6] coalescent [12]. This assumption holds unless the contributions of individuals from one time step to the next are highly skewed [2, 15, 17], which is unlikely in most real situations.

Under Assumption A (see Lemma 1 below), we obtain that

$$\lim_{N \rightarrow \infty} N c_N \sum_{t \geq 0} E_0(x(t)^{k+1}(1-x(t))^{d-k}) = \frac{2k!(d-k)!}{(d+1)!}. \tag{21}$$

In the neutral Wright–Fisher model with the numbers of descendants left by the N individuals of a given generation having a multi-nomial distribution, we have $c_N = N^{-1}$. On the other hand, $c_N = 2(N(N-1))^{-1}$ in the Moran model with one individual chosen at random to be replaced by an offspring of another individual chosen at random at every time step.

Using (21) in (18), we find that

$$d_k = \frac{2(d-k)}{d(d+1)N c_N} (1 + o(1)), \tag{22}$$

where $o(1) \rightarrow 0$ as $N \rightarrow \infty$. Therefore, we conclude the following.

Result 1 Under Assumption A, a first-order approximation for the probability of fixation of A represented once initially is given by

$$\rho_A(\sigma) \approx \frac{1}{N} + \frac{2\gamma\sigma}{N^2 c_N d(d+1)} \sum_{k=0}^{d-1} (d-k)(a_k - b_k), \tag{23}$$

for σ small enough and N large enough.

This approximation agrees with the result given in [7] and [4] in the case of a Moran model in continuous time and extends it to a much wider class of models. Moreover, it leads to the condition given in (2) to have $\rho_A(\sigma) > \rho_A(0)$ for σ small enough and N large enough.

4 Key Lemma

Let us define

$$F_N(n, m) = N c_N \sum_{t \geq 0} E_0(x(t)^n (1 - x(t))^m), \tag{24}$$

where $x(t)$ denotes the frequency of A at time step $t \geq 0$ with $x(0) = N^{-1}$. In this section, we prove the following result.

Lemma 1 *Under Assumption A, we have*

$$\lim_{N \rightarrow \infty} F_N(n, m) = \frac{2(n - 1)!m!}{(n + m)!}, \tag{25}$$

for $n, m \geq 1$.

Proof Let us introduce the notation

$$F_\infty(n, m) = \lim_{N \rightarrow \infty} F_N(n, m), \tag{26}$$

for $n, m \geq 1$. In the case $n + m = 2$, that is, $n = m = 1$, we have

$$F_N(1, 1) = N c_N \sum_{t \geq 0} (1 - N^{-1}) E_0(\xi_1(t) (1 - \xi_2(t))), \tag{27}$$

where $(1 - N^{-1})$ is the probability for two individuals arbitrarily labeled by 1 and 2 that are chosen at random at time step t not to be the same, and $\xi_i(t) = 1$ if individual i is of type A and 0 otherwise, for $i = 1, 2$. On the other hand,

$$E_0(\xi_1(t) (1 - \xi_2(t))) = N^{-1} P_0(t_2 > t), \tag{28}$$

where N^{-1} is the probability for the ancestor of individual 1 at time step 0 to be of type A , and t_2 stands for the number of time steps back for the ancestors of individuals 1 and 2 to be the same. Moreover, P_0 denotes probability under neutrality. Actually,

$$P_0(t_2 > t) = (1 - c_N)^t, \tag{29}$$

and, therefore,

$$\sum_{t \geq 0} E_0(\xi_1(t) (1 - \xi_2(t))) = N^{-1} c_N^{-1}. \tag{30}$$

We conclude that

$$F_N(1, 1) = 1 - N^{-1}, \tag{31}$$

from which

$$F_\infty(1, 1) = 1. \tag{32}$$

This establishes (25) for $n + m = 2$.

Suppose now that (25) holds for $m + n = 2, \dots, r - 1$ and consider the case where $m + n = r$. The induction hypothesis and the fact that $n + m$ individuals chosen at random at the same time step t will not be all different with probability $O(N^{-1})$ guarantee that

$$F_N(n, m) = G_N(n, m) + O(N^{-1}), \tag{33}$$

for $n + m \leq r$, where

$$G_N(n, m) = N c_N \sum_{t \geq 0} E_0 \left(\left(\prod_{i=1}^n \xi_i(t) \right) \left(\prod_{i=n+1}^{n+m} (1 - \xi_i(t)) \right) \right), \tag{34}$$

with $\xi_i(t)$ defined as previously for $n + m$ individuals arbitrarily labeled by $i = 1, \dots, n + m$ that are chosen at random without replacement at time step t . This can be obtained by expressing the frequency of A at time step t in the form $x(t) = N^{-1} \sum_{i=1}^N \xi_i(t)$ with $\xi_1(t), \dots, \xi_N(t)$ being exchangeable indicator random variables for the frequency of A in the N individuals of the population. In particular, in the limit of a large population size, we have

$$G_\infty(n, m) = \lim_{N \rightarrow \infty} G_N(n, m) = F_\infty(n, m), \tag{35}$$

for $n + m \leq r$.

Let t_{n+m} be the time back before the occurrence of the first merger in the ancestry of $n + m$ individuals chosen at random without replacement at time t . Under neutrality, this time will take the value $s \geq 1$ with probability

$$p_{n+m}(s) = (1 - q_{n+m})^{s-1} q_{n+m}, \tag{36}$$

where q_{n+m} is the probability of at least one merger among $n + m$ lineages under neutrality. Under our assumptions, this probability takes the form

$$q_{n+m} = \binom{n+m}{2} c_N + o(c_N) = \frac{(n+m)(n+m-1)c_N}{2} + o(c_N). \tag{37}$$

Therefore,

$$E_0(t_{n+m}) = q_{n+m}^{-1} = \frac{2}{(n+m)(n+m-1)c_N} (1 + o(1)). \tag{38}$$

Moreover, given that $t_{n+m} = s$, the individuals labeled by $1, \dots, n$ at time step $t < s$ have a single ancestor at time step 0 with probability 1 in the case $n \geq 1$ but probability 0 in the case $n \geq 2$. In this case, they have $n - 1$ ancestors at time step $r = t - s \geq 0$ with probability

$$u_{n-1,m} = \frac{\binom{n}{2} c_N + o(c_N)}{\binom{n+m}{2} c_N + o(c_N)} = \frac{n(n-1)}{(n+m)(n+m-1)} + o(c_N), \tag{39}$$

and $n - l$ ancestors with probability $u_{n-l,m} = o(c_N)$ for $l \geq 2$. Similarly, given that $t_{n+m} = s$ and $m \geq 2$, the individuals labeled by $n + 1, \dots, n + m$ will have $m - 1$ ancestors at time

step $r = t - s \geq 0$ with probability

$$v_{n,m-1} = \frac{\binom{m}{2} c_N + o(c_N)}{\binom{n+m}{2} c_N + o(c_N)} = \frac{m(m-1)}{(n+m)(n+m-1)} + o(c_N), \tag{40}$$

and $n - l$ ancestors with probability $v_{n,m-l} = o(c_N)$ for $l \geq 2$.

Therefore, conditioning on the time step t_{n+m} and the numbers of ancestor at this time step, we obtain

$$N c_N \sum_{s \geq 1} p_{n+m}(s) \sum_{t=0}^{s-1} E_0 \left(\xi_1(0) \left(\prod_{i=2}^{m+1} (1 - \xi_i(0)) \right) \right) + N c_N \sum_{s \geq 1} p_{n+m}(s) \sum_{l=1}^{m-1} v_{1,m-l} \sum_{t \geq s} E_0 \left(\xi_1(t-s) \left(\prod_{i=2}^{m-l+1} (1 - \xi_i(t-s)) \right) \right) \tag{41}$$

for $G_N(1, m)$, and

$$N c_N \sum_{s \geq 1} p_{n+m}(s) \sum_{l=1}^{n-1} u_{n-l,m} \sum_{t \geq s} E_0 \left(\left(\prod_{i=1}^{n-l} \xi_i(t-s) \right) \left(\prod_{i=n-l+1}^{n+m-l} (1 - \xi_i(t-s)) \right) \right) + N c_N \sum_{s \geq 1} p_{n+m}(s) \sum_{l=1}^{m-1} v_{n,m-l} \sum_{t \geq s} E_0 \left(\left(\prod_{i=1}^n \xi_i(t-s) \right) \left(\prod_{i=n+1}^{n+m-l} (1 - \xi_i(t-s)) \right) \right)$$

for $G_N(n, m)$ in the case $n \geq 2$. This leads to

$$G_N(1, m) = c_N E_0(t_{n+m}) + \sum_{l=1}^{m-1} v_{n,m-l} G_N(1, m - l) \tag{42}$$

and

$$G_N(n, m) = \sum_{l=1}^{n-1} u_{n-l,m} G_N(n - l, m) + \sum_{l=1}^{m-1} v_{n,m-l} G_N(n, m - l), \tag{43}$$

respectively.

Then the induction hypothesis yields

$$G_\infty(1, m) = \frac{2}{m(m+1)} + \frac{2(m-1)(m-1)!}{(m+1)!} = \frac{2m!}{(m+1)!}, \tag{44}$$

and

$$G_\infty(n, m) = \frac{2n!m!}{(n+m-1)(n+m)!} + \frac{2(m-1)(n-1)!m!}{(n+m-1)(n+m)!} = \frac{2(n-1)!m!}{(n+m+1)!}, \tag{45}$$

in the case $n \geq 2$, which completes the proof. □

5 Discussion

By analogy with Lessard and Lahaie [11] in the case where the intensity of selection is σN^{-1} , (17) means that

$$a_A(0) = \gamma N \sum_{k=0}^{d-1} d_k (a_k - b_k) \tag{46}$$

is the projected average excess in payoff of the mutant strategy A . This is the difference between the marginal payoff to an A player and the mean payoff in the whole population not only at the initial time step but also at all future time steps.

Note that

$$d_k = \binom{d-1}{k} \sum_{t \geq 0} E_0(x(t)^{k+1}(1-x(t))^{d-k}) \tag{47}$$

is the expected time back in number of time steps under neutrality that a focal individual and k other individuals out of $d - 1$ will have the same common ancestor of a given label, different from the ancestors of $d - k - 1$ other individuals and different from the ancestor of another individual in competition with the focal individual. Here, the individuals are chosen at random with replacement and an ancestor at any time step back has a label from 1 to N . Lemma 1 entails that

$$\lim_{N \rightarrow \infty} N c_N d_k = \frac{2(d-k)}{d(d+1)}. \tag{48}$$

Therefore, the condition for weak selection to favor A replacing B in the limit of a large population size, given by (2), is equivalent to $a_A(0) > 0$. This holds in the case where A is represented only once initially.

In general, the quantity

$$H_N(n, m) = \binom{n+m}{n} \sum_{t \geq 0} E_0(x(t)^n(1-x(t))^m), \tag{49}$$

represents the expected time back in number of time steps under neutrality that n individuals out of $n + m$ individuals chosen at random with replacement will have the same ancestor of a given label, different from the ancestors of the others. Lemma 1 ascertains that

$$\lim_{N \rightarrow \infty} N c_N H_N(n, m) = \frac{2}{n}. \tag{50}$$

It is rather surprising that this limit does not depend on m .

On the other hand, the binomial theorem allows us to develop the quantity $F_N(n, m)$ in Lemma 1 in the form

$$F_N(n, m) = N c_N \sum_{t \geq 0} \sum_{l=0}^m (-1)^l \binom{m}{l} E_0(x(t)^{n+l}). \tag{51}$$

Note that

$$E_0(x(t)^{n+l}) = N^{-1} (1 - P_0(S_{n+l} > t)), \tag{52}$$

where S_{n+l} represents the number of time steps back before the occurrence of the last merger in the ancestry of $n + l$ individuals sampled at random with replacement at any given time step. Moreover, we have

$$P_0(S_{n+l} > t) \asymp P_0(T_{n+l} > t), \tag{53}$$

where T_{n+l} represents the corresponding time in the case of sampling without replacement. Here, we use the notation $f(N) \asymp g(N)$ to mean that $f(N)$ is asymptotic to $g(N)$ in the sense that $f(N)/g(N) \rightarrow 1$ as $N \rightarrow \infty$. Convergence to the Kingman coalescent ensures that

$$c_N \sum_{t \geq 0} P_0(T_{n+l} > t) = c_N E_0(T_{n+l}) \asymp \sum_{j=2}^{n+l} \frac{2}{j(j-1)} = 2 \left(1 - \frac{1}{n+l} \right). \tag{54}$$

Using this in the above expression for $F_N(n, m)$ and applying twice the binomial theorem yield

$$\lim_{N \rightarrow \infty} F_N(n, m) = 2 \sum_{l=0}^m (-1)^l \binom{m}{l} \frac{1}{n+l}. \tag{55}$$

Therefore, Lemma 1 states that

$$\sum_{l=0}^m (-1)^l \binom{n+m}{m} \binom{m}{l} \frac{n}{n+l} = 1. \tag{56}$$

Acknowledgements The author is grateful to Arne Traulsen who raised the question of the extension of the one-third law of evolution to the d -player game in exchangeable models during the BIRS workshop on Evolutionary Games held in Banff from June 13 to June 18, 2010.

References

1. Cannings C (1974) The latent roots of certain Markov chains arising in genetics: a new approach. I. Haploid models. *Adv Appl Probab* 6:260–290
2. Eldon B, Wakeley J (2006) Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* 172:2621–2633
3. Fisher RA (1930) *The genetical theory of natural selection*. Clarendon, Oxford
4. Gokhale CS, Traulsen A (2010) Evolutionary games in the multiverse. *Proc Natl Acad Sci USA* 107:5500–5504
5. Imhof LA, Nowak MA (2006) Evolutionary game dynamics in a Wright–Fisher process. *J Math Biol* 52:667–681
6. Kingman JFC (1982) The coalescent. *Stoch Process Appl* 13:235–248
7. Kurokawa S, Ihara Y (2009) Emergence of cooperation in public goods games. *Proc R Soc B* 276:1379–1384
8. Lessard S (2005) Long-term stability from fixation probabilities in finite populations: new perspectives for ESS theory. *Theor Popul Biol* 68:19–27
9. Lessard S (2007) Cooperation is less likely to evolve in a finite population with a highly skewed distribution of family size. *Proc R Soc B* 274:1861–1865
10. Lessard S, Ladret V (2007) The probability of fixation of a single mutant in an exchangeable selection model. *J Math Biol* 54:721–744
11. Lessard S, Lahaie P (2009) Fixation probability with multiple alleles and projected average allelic effect on selection. *Theor Popul Biol* 75:266–277
12. Möhle M (2000) Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Adv Appl Probab* 32:983–993

13. Moran PAP (1958) Random processes in genetics. *Proc Camb Philos Soc* 54:60–71
14. Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428:646–650
15. Pitman J (1999) Coalescents with multiple collisions. *Ann Probab* 27:1870–1902
16. Rousset F (2003) A minimal derivation of convergence stability measures. *J Theor Biol* 221:665–668
17. Sagitov S (1999) The general coalescent with asynchronous mergers of ancestral lines. *J Appl Probab* 36:1116–1125
18. Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159