INTRODUCTION

Gametophytic self-incompatibility (GSI) is a genetically controlled mating system, common in flowering plants. GSI prevents crossing between individuals sharing identical alleles at the self-incompatibility locus (S-locus), especially self-fertilization (Durand et al., 2020). Despite extensive empirical and theoretical studies from the seminal works by Emerson (1938) and Wright (1939), many properties of the evolutionary dynamics of GSI are not well understood yet. For instance, the stationary distribution of S-allele frequencies or the expected number of S-alleles in a finite population is still not clearly established. Yet, such predictions are necessary to interpret genetic data in GSI populations. In addition, the probability of invasion of a mutant, or an estimator of effective population size from genetic data, both classical quantities in population genetics, are still lacking. By using modern stochastic tools, we address these questions in this manuscript.

It was early observed that S-allele diversity in natural populations was surprisingly large: in a population of Oenothera organensis, at least 45 different S-alleles had been identified despite the population comprising at most one thousand individuals (Emerson, 1938; Lewis, 1949a). The large S-allele diversity is now known to be prevalent....
in GSI populations across various species and families (Castric & Vekemans, 2004). In the 1930s, explaining the origin, dynamics and maintenance of this diversity was a real challenge for the young fields of population and evolutionary genetics. Wright (1939) was the first to propose an approximation of the stationary distribution of S-allele frequencies in a finite population, one of the first applications of theoretical population genetics as introduced by Fisher (1930) and Wright (1937). Wright (1939) used his approximation of the stationary distribution to estimate the expected number of S-alleles in a given population of finite size with recurrent mutation. His prediction, applied to O. organensis populations, was much lower than the observed number of S-alleles (Emerson, 1938) under a reasonable mutation rate and population size. Wright (1939) then suggested that the discrepancy between theoretical prediction and empirical observation could be due to the population subdivision of the O. organensis population.

The failure of Wright’s attempt to apply a population genetic model to an empirical case generated a fierce and long debate about the goodness of stochastic models of population genetic dynamics. Fisher (1958) and Moran (1962) were the first to criticize Wright’s initial model, mostly for a lack of mathematical rigour. Wright further refined his GSI model (Wright, 1960, 1964) and provided, with others, computer simulations (Ewens & Ewens, 1966; Kimura & Crow, 1964; Mayo, 1966) that confirmed that his approximation of the stationary distribution of S-allele frequencies, and his prediction of the expected number of S-alleles, were correct.

Yet, to the best of our knowledge, there is still no theoretical derivation and justification of the allele frequency dynamics and stationary state in a GSI population, even though, more than 50 years ago, Moran (1962, p. 163) already suggested that the probabilistic model underlying GSI should be properly specified. Wright (1964) tentatively answered Moran’s criticism by framing it as ‘a basic difference in viewpoint’. In contrast, Wright’s approximations are far better than could reasonably be expected. An explicit derivation of the (macroscopic) properties of a GSI population from the individual (microscopic) level is necessary to determine the assumptions under which the approximation remains valid, and therefore can be used to interpret empirical observations and to conduct parameter inference. Indeed, paraphrasing Ewens and Ewens (1966): the validity of Wright’s approximation is ‘possibly fortuitous’, as suggested by some of the results by Mayo (1966) where the approximation can be incorrect (see Ewens (1964) and Ewens and Ewens (1966) for technical arguments).

In the end, one question remains unsolved: Wright’s approximation of the GSI stationary distribution seems to work well, but we do not know why. Here, we address this theoretical gap. We start by defining an individual-based model of gametophytic self-incompatibility. Taking the infinite population size limit, we first obtain the deterministic dynamics of gametophytic self-incompatibility, an equation that has already been derived and studied in great detail (Boucher, 1993; Fisher, 1958; Nagylaki, 1975). Second, we derive a diffusion approximation and compute the stationary distribution based on the central limit theorem for density-dependent Markov processes (Kurtz, 1971). We obtain an explicit expression for the stationary distribution that only depends on the population size and the number of S-alleles in the population, which formally confirms the validity of Wright’s approximation under general conditions. Finally, we also provide new results: (1) the invasion probability of a novel S-allele in a resident population, and (2) a method to estimate the effective population size from S-allele frequency data, based on our explicit result of the stationary distribution of S-allele frequencies.

2 | STOCHASTIC MODEL OF GAMETOPHYTIC SELF-INCOMPATIBILITY

We consider a diploid plant population with fixed size $N$ with overlapping generations (Moran-type model). The analogous model definition of the self-incompatibility dynamics with non-overlapping generations (Wright-Fisher-type model) is given in the Appendix S1. Section C. We assume a fixed number of S-alleles in the population and denote it by $M$. The number of plants with genotype $ij$ is denoted $A_{ij}$ with no allele order (i.e. $A_{ij} = A_{ji}$). Because of GSI, an $ij$ individual can only be fertilized by the pollen of type $k = i, j$ and then produces an offspring of type $jk$ or $kj$ with equal probability, consequently $A_{ii} = 0$ for all $i$. The dynamics of the number of $ij$ individuals are described in terms of coupled birth and death events to maintain the fixed population size $N$, as is common in Moran-type models. Then, the number of $ij$ individuals, $A_{ij}$, increases by one if

(b-i) a type $i$ pollen fertilizes a $[i*]$ individual that transmits allele $j$, or

(b-ii) a type $j$ pollen fertilizes a $[i*]$ individual that transmits allele $i$, and a non-$[ij]$ individual dies and is replaced.

The probability of fertilization of a $[jk]$ individual with an $i$ pollen is given by $p_i/(1 - p_i - p_j)$, where $p_i$ is the frequency of $i$-type pollen in the population $\left( 1 = \sum_j p_j \right)$. The form emerges from conditioning on successful fertilization, which reflects the assumption that the number of pollen is essentially infinite, or at least not limiting the (female) reproductive success of a plant, i.e. we assume no pollen limitation.

The birth rate, is then given by:

$$T_{ij} = \left( \frac{1}{2} \sum_{j \neq i} A_{ji} \frac{p_j}{1 - p_j - p_k} + \frac{1}{2} \sum_{i \neq j} A_{ij} \frac{p_i}{1 - p_i - p_k} \right) \frac{N - A_{ij}}{N} \text{non-} \langle ij \rangle \text{replacement.} \tag{1}$$

The reproduction terms between brackets were first derived by Fisher (1958).

Using the same arguments, the number of $ij$ individuals decreases by one if

(d-i) a non-$[ij]$ plant is fertilized (by any pollen), or

(d-ii) a $[i*]$ plant is fertilized, but the offspring is not $[ij]$. or
(d-iii) a (i•) plant is fertilized, but the offspring is not (i•), and a (i•) individual dies and is replaced.

The rate at which the number of (i•) individuals decreases by one, the death rate, is then given by (detailed derivation in the Appendix S1, Section A):

\[
\begin{align*}
T_{ij}^- &= \left( N - \frac{1}{2} \sum_{k \neq i} A_k \frac{p_i}{1 - p_i - p_k} - \frac{1}{2} \sum_{k \neq j} A_k \frac{p_i}{1 - p_i - p_k} \right) \left( \frac{A_{ij}}{N} \right) \\
&= \left( N - \frac{1}{2} \sum_{k \neq i} A_k \frac{p_i}{1 - p_i - p_k} - \frac{1}{2} \sum_{k \neq j} A_k \frac{p_i}{1 - p_i - p_k} \right) \left( \frac{A_{ij}}{N} \right). \\
&= (i\text{ replacement}) \\
\end{align*}
\]

Since all individuals are equally likely to reproduce (at rate 1), i.e. the overall rate of reproduction is N, the reproduction part of the death rate is given by N minus the reproduction rate of an (i•) individual (Equation 1), which explains the term between brackets in Equation (2).

Last, there is also the possibility that no change in the number of (i•) individuals occurs. This is the case if an (i•) individual is produced and replaces an (i•) individual or a non-(i•) individual is produced and replaces a non-(i•) individual. The rate for this to happen is \( N - T_{ij}^- - T_{ij}^+ \).

We now proceed with the theoretical analysis of this stochastic model.

3 | RESULTS

3.1 | Approximation of the dynamics in a finite population

First, we derive an approximation of the dynamics of genotypes and S-alleles in a large but finite GSI plant population. We assume that the number of S-alleles in the population is constant, i.e. we ignore the generation of new S-alleles through mutations and the loss of S-alleles through stochastic extinction events. To derive the genotype and S-allele dynamics, we apply diffusion-theoretical arguments (reviewed in the context of evolutionary applications in Czuppon & Traulsen, 2021). The mathematical details are stated in the Appendix S1, Section A. Briefly, if we denote by \( a_{ij}(t) = A_{ij}/N \) the proportion of (i•) individuals in the population at time \( t \), we find an equation of the form

\[
da_{ij}(t) = \mu_{ij}(t)dt + \sqrt{\frac{\sigma_{ij}^2(t)}{N}}dW_t, \tag{3}
\]

where \( \mu_{ij}(t) \) is the deterministic dynamics of the genotype frequencies (see below), \( \sigma_{ij}^2(t) \) is the infinitesimal variance of genotype frequencies (explicitly given in Equation (A.4) of Section A in Appendix S1) and \((W_t)_{t \geq 0}\) is a Brownian motion, i.e. \( W_t \) is normally distributed with mean zero and variance \( t \), related to the stochastic fluctuations of genotype frequencies, which models the inherent randomness of births and deaths in a finite population. Notably, the stochastic fluctuations scale with \( 1/\sqrt{N} \) and vanish in the limit of infinite populations (\( N \to \infty \)), where only the deterministic dynamics \( \mu_{ij} \) remain. The deterministic change of the genotype frequencies is given by

\[
\mu_{ij} = \frac{1}{2} \left( \sum_{k \neq i} A_k \frac{p_i}{1 - p_i - p_k} + \sum_{k \neq j} A_k \frac{p_i}{1 - p_i - p_k} \right) - a_{ij}. \tag{4}
\]

Solving \( \mu_{ij} = 0 \) in Equation (4) gives the steady state \( p^*_{ij} = 1/M \) and \( a_{ij} = 2/(M(M-1)) \) where all genotypes, on the one hand, and all S-alleles, on the other hand, have identical frequencies (Bouchier, 1993; Nagylaki, 1975).

The diffusion approximation of the dynamics given in Equation (3) describes genotype trajectories. The stochastic component of the equation, the second term on the right-hand side, quantifies how strongly the genotype frequency fluctuates around the deterministic trajectory, the first term on the right-hand side. These random fluctuations are of order \( 1/\sqrt{N} \), which means that they decrease with increasing population size \( N \) and vanish in the limit \( N \to \infty \) (we refer to Ethier & Kurtz, 1986 for a rigorous treatment). The precise scaling of these stochastic fluctuations, which are a result of the finite population size, is obtained by explicitly computing the infinitesimal variance \( \sigma_{ij}^2(t) \) (Appendix S1, Section A).

The derivation of the S-allele dynamics is obtained by similar arguments and using the relation between the frequency of S-allele \( i \) and the genotype (i•) frequencies: \( p_i = \sum a_{ij}^2 \) (details in Appendix S1, Section A). The deterministic change in S-allele frequency, i.e. the infinitesimal mean, is then given by

\[
\frac{dp_i}{dt} = \mu_i = \frac{1}{2} \sum_{j \neq i} a_{ij} = \frac{1}{2} \left( \frac{1}{2} \sum_j a_{ij} \frac{1}{1 - p_j - p_{ij}} - 1 \right). \tag{5}
\]

Equation (5) shows that, since we assume no pollen limitation, the deterministic dynamics of allele \( i \) are driven by pollination of non-\( i \) plants (first term in the brackets), rather than pollination of \( i \) plants by pollen of a different type (Bouchier, 1993; Fisher, 1958; Nagylaki, 1975).

The diffusion approximation of S-allele frequencies is

\[
\frac{dp_i(t)}{dt} = \mu_i(t)dt + \sqrt{\frac{\sigma_i^2(t)}{N}}dW_t, \tag{6}
\]

where \((W_t)_{t \geq 0}\) is a Brownian motion that reflects the stochastic fluctuations of the \( i \)-th S-allele frequency around the deterministic trajectory \( \mu_i \). The exact expression of the infinitesimal variance of the S-allele frequency change \( \sigma_i^2(t) \) is computed in Appendix S1, Section A, Equation (A.9). Note that the infinitesimal variance of S-alleles depends on the covariance between genotype frequencies because of our restriction of a fixed population size. Therefore, an increase in the frequency of one S-allele (birth event) inevitably entails a decrease in another S-allele (death event).

3.2 | Approximation of the stationary distribution

In an infinite plant population, the genotype and S-allele frequencies will converge to the steady states \( a_{ij}^* = 2/(M(M-1)) \) and \( p_{ij}^* = 1/M \),
respectively. This deterministic steady state is globally stable (Boucher, 1993). This means that in a finite population the stochastic trajectories of S-allele frequencies will fluctuate around the value 1/M, at least as long as the probability of extinction of an S-allele is negligibly small. This will typically be the case if M ≪ V/N. Using arguments motivated by the central limit theorem of density-dependent Markov processes (Ethier & Kurtz, 1986; van Kampen, 2007), we can then approximate the fluctuations around the deterministic steady state by a normal distribution (details in Appendix S1, Section B). We find that the stationary distribution of S-allele frequencies, denoted by \( \psi \), is distributed as (denoted by the ‘\( - \)’ sign)

\[
\psi \sim \mathcal{N}\left( \frac{1}{M}, \frac{(M-2)^3}{NM^2(2M-3)} \right),
\]

where \( \mathcal{N}(m,s^2) \) denotes a normal distribution with mean \( m \) and variance \( s^2 \).

The same derivation is possible for a model with non-overlapping generations. In this case, we find that the stationary distribution differs only by a factor 2 in the denominator of the variance (Equation [C.9], Section C, Appendix S1), in line with the generally found difference between the variances of the Moran and Wright-Fisher model (Czuppon & Traulsen, 2021).

Figure 1 shows that the approximation of the stationary distribution in Equation (7) predicts well the stationary distributions obtained by stochastic simulations, especially if the number of S-alleles (M) is small. The approximation of the stationary distribution is slightly worse for large numbers of S-alleles because the probability that an allele is lost by chance increases with the number of S-alleles in the population, as shown by the bar located at zero for \( M = 15 \) in Figure 1. As a consequence of an allele extinction, the mean allele frequency is displaced from 1/M to 1/(M−1) in some simulations, or even lower values of M if multiple extinction events occurred. This explains the slightly worse fit of the approximated stationary distribution for \( M = 15 \), compared with lower numbers of S-alleles.

### 3.3 The number of S-alleles maintained in a finite population

We now allow the number of S-alleles to vary over time. In particular, at each reproduction event, a mutation may occur and give rise to a new S-allele. The mutation probability is denoted by \( u \). We assume an infinitely many alleles model, which means that every mutation will generate a new S-allele. Then, the rate at which new S-alleles arise is simply given by the mutation probability \( u \) (because the overall rescaled reproduction rate in the population is 1 due to the conditioning on successful reproduction).

To determine the number of S-alleles that are stably maintained in a population of finite size, we follow the partially heuristic reasoning by Wright (1939) and fill in the analytical gaps using the stationary distribution and the individual-based model. The basic idea is as follows (Wright, 1939): First, given an approximation of the stationary distribution of the allele frequencies for a given number of S-alleles \( M \) in a finite population of size \( N \) (Equation (7)), the loss rate of a focal S-allele is computed by multiplying the probability for this S-allele to be at frequency 1/2N with its death rate \( T^- \) at frequency 1/2N (computation in Appendix S1, Section D). Second, the loss rate is balanced with the gain rate of new S-alleles, which is just the mutation probability \( u \). This yields the following equality

\[
u = \frac{M}{N} \times \int_{1/2N}^{0} \psi(x) \, dx = \frac{1}{NM} \int_{T^- \times M}^{1/2N} \psi(x) \, dx,
\]

which can be solved numerically for \( M \). Figure 2 shows a comparison between the prediction obtained by solving Equation (8) for different values of \( M \) and stochastic simulations. The number of S-alleles maintained in a finite population is reasonably well predicted by our model, even though it is overestimated as population size increases. The reason for this is that we underestimate the extinction risk of an S-allele in large populations. This is a consequence of our approximation of the stationary allele frequency distribution being a less accurate description of the dynamical system for large population sizes and a large number of S-alleles. The more S-alleles are found in a population, the less strong is the deterministic force due to self-incompatibility and the system dynamics appear more neutral. This explains why our derivation of the stationary allele frequency distribution, which is based on the assumption of a (strongly) attracting fixed point, becomes less accurate. This in turn leads to an overestimation of the number of stably maintained S-alleles in large populations.

The overall good fit confirms the previous conclusion from numerical simulations (Ewens & Ewens, 1966; Mayo, 1966) that Wright’s (1939) general methodology indeed correctly predicts...
3.4 Invasion probability of a new S-allele

We now derive new results on the invasion behaviour of a novel S-allele. In this section, we again ignore mutations and study the situation where at time 0, there is a single S-allele that is present in exactly one individual, i.e. this S-allele is at frequency 1/2N. All the other S-alleles are assumed to be in equilibrium, a reasonable assumption in view of the stability of the steady state (Boucher, 1993). The invasion probability can then be computed from the diffusion approximation (Equation (6)) by applying results from stochastic diffusion theory (mathematical details in Appendix S1, Section E). We find that a new S-allele, appearing at frequency 1/2N by mutation (or immigration), establishes in a population with M resident S-alleles with probability

$$\varphi = \frac{2}{2M-3} \quad (M \geq 3).$$

(9)

This approximation agrees well with simulation results (Figure 3). Surprisingly though, it fits the simulation results better than the non-approximated formula from which Equation (9) is derived; more details in Appendix S1, Section E.

The invasion probability decreases with M because the rare allele advantage is smaller when the number of resident S-alleles M increases. This is explained by an increase in the number of plants that can be fertilized by a single resident S-allele for larger numbers of resident alleles. The invading allele is initially present in a single copy, which means that the number of individuals it can fertilize is always N−1. All the resident S-alleles, assuming that genotype frequencies are in equilibrium at 2/M(M−1) (except for one genotype that is at frequency 2/M(M−1) − 1/N because of the invading allele), can fertilize

$$\frac{(M-1)(M-2)}{2} \times \frac{2}{M(M-1)} = \frac{M-2}{M} N. \quad (M \geq 3)$$

(10)

individuals. This shows that as M increases, the difference in the rate of successful pollen fertilization between the invading S-allele, i.e. N-1, and the resident alleles, i.e. N(M−2)/M, decrease. In contrast, since we assume no pollen limitation, the probability for a plant to be fertilized only depends on its genotype frequency in the population.

3.5 Estimation of the number of S-alleles and the population size from data

Our goal in this section is to provide a joint estimator of both the population size and the number of S-alleles in a sampled plant population genotyped at the S-locus. The main idea is to use the mean and the variance of the stationary distribution (Equation (7)) to estimate the two population parameters of interest, M and N, given the sample size n and the number of S-alleles observed in the sample. We have developed two different estimators. The first estimator...
accounts for the sampling process and uses this correction to estimate the number of $S$-alleles and the population size. The second estimator computes the first two moments of the sample allele frequency distribution. These estimated moments are then used to estimate the mean and variance of the theoretical allele frequency distribution, which is then translated to the number of $S$-alleles and the population size.

The first estimator is an extension of Paxman’s (1963) urn model and the ideas presented in Yokoyama and Hetherington (1982) (details are given in Appendix S1, Section F). We derived a least-squares estimator for the number of $S$-alleles and the population size from the mean and variance of the number of occurrences $Y$ of a given $S$-allele observed in the sample (Appendix S1, Section G). We show that the mean $E[Y]$ and the variance $V[Y]$ can be used to estimate the number of $S$-alleles and the population size. However, the minimization of the least-squares function showed convergence instability, presumably because the least-squares function is very flat along the axis of the population size (Appendix S1, Section G, Figure C). Even though the number of $S$-alleles was correctly estimated, the population size estimator performed rather poorly (Figure 4a).

The second estimator is a moment estimator and ignores the sampling process associated with data collection. We fit the theoretically derived normal distribution of the stationary $S$-allele frequency distribution (Equation (7)) to the empirical $S$-allele frequency distribution. The empirical mean, $\bar{Y}$, and variance, $\bar{Y}^2$, are used to fit a normal distribution. From this fitted normal distribution, we can directly compute the estimates for the number of $S$-alleles, $\hat{M}$, and the population size, $\hat{N}$, using Equation (7):

$$\hat{M} = \frac{1}{\bar{Y}} \quad \text{and} \quad \hat{N} = \frac{(\hat{M} - 2)^3}{\bar{Y}^2 \hat{M} (2\hat{M} - 3)} \quad (11)$$

Note that by definition the estimated number of $S$-alleles, $\hat{M}$, is exactly the observed number of $S$-alleles in the sample (the expectation of the allele frequency spectrum equals the number of observed $S$-alleles). This estimator therefore underestimates the true number of $S$-alleles in the population. Our least-squares estimator, which takes into account the sampling process, corrects for this bias and therefore does not encounter this problem.

We also applied this second estimator to a simulated plant population of size 3000 with different numbers of $S$-alleles. We find that the estimated population sizes mostly underestimate the true population size. This effect is most pronounced for small sample sizes (blue violin plots in Figure 4b). This is explained by smaller samples showing a larger variance in the allele frequency distribution, which translates to smaller values of the estimated population size $\hat{N}$ (Equation (11)). The larger the sample size, the better the population size approximated, yet there is still considerable variation in the estimates.

Lastly, we applied our estimator to an empirical data set from Stoeckel et al. (2012), where a population of wild cherry ($Prunus avium$) was quasi-exhaustively sampled ($n = 249$). We estimated $\hat{M} = 18$ and $\hat{N} = 120$, which is of the same order as, though under-estimates, the census population size $n$. From the 95% confidence interval (CI) of the sample variance, we computed the 95% CI for the population size, which spans from $\hat{N} = 53$ to $\hat{N} = 213$, still not covering the sample size. The discrepancy between the estimation and the census population size can be due to a bias of our estimator. Another possible reason for this underestimation is a violation of our modelling assumptions, e.g., spatial structure of the population or fecundity differences between individuals due to their size (Stoeckel et al., 2012). Overall, our results may serve as a starting point to develop new statistical estimators that can be used to infer ecological parameters like the effective population size from the diversity observed at an $S$-locus. In particular, as a perspective, it might be

![Figure 4](image-url)
possible to combine Paxman’s model with a fitted normal distribution to jointly estimate the number of \( S \)-alleles and the population size, similar to our first attempt in Appendix S1, Section G.

4 | DISCUSSION

4.1 | Self-incompatibility as an historical illustration of probabilistic thinking

As emphasized by Wright (1937), one of the central goals of population genetics is to find the frequency distribution of genetic variants under various conditions, in particular, for convenience, under stationarity. GSI was one of the first genetic systems used to validate theoretical predictions from the young field of population genetics in the 1930-40s. Before detailing how GSI was used to bring together theory and empirical observations, we first briefly summarize the general theory as stated by Fisher and Wright in the 1930s (see Ishida & Rosales, 2020, for a detailed and general historical review).

The approach taken by Wright was to split the frequency changes into two parts: a deterministic part (describing mutation, migration or selection) and a stochastic part (describing random sampling errors). Wright generally first states how the frequency of a given allele \( p_i \) changes in expectation from one generation to the next (by a difference equation), and second derives the form of the stochastic fluctuations, here denoted \( \Delta p_i^2 \). To derive the stochastic fluctuations, he usually assumes that the gene frequency dynamics are a succession of binomial draws, where the probability for a specific allele to be transmitted is equal to its frequency in the gamete offspring pool, assumed to be infinite. These allele frequencies in the offspring pool could have changed, compared with the parental generation, because of the processes of selection, migration or mutation (e.g. Wright, 1937). In this model, the stochastic fluctuations from one generation to the next are given by the variance of a binomial distribution \( \frac{\bar{p}(1-\bar{p})}{2N} \). The deterministic change of the allele frequency \( \Delta p_i \) should appear in the binomial distribution parameters, but Wright explicitly assumes that it is negligible and thus that stochastic fluctuations only depend on the population size and the actual allele frequency, generally referred to as ‘genetic drift’.

Using a continuous approximation of the discrete stochastic process, Wright (1937, 1938) calculated the change of the mean and variance of the allele frequencies, giving a possible approximation of the distribution of allele frequencies for various cases regarding selection, migration and mutation, and its explicit stationary distribution. In his papers, Wright (1937, 1938) acknowledged that Fisher (1930) found similar results, despite Fisher’s approach being different. Thanks to a letter from the mathematician Kolmogorov, Wright (1945) realized that the form he had found for the allele distribution was the solution of the more general Fokker-Planck equation used in physics, which describes a stochastic process combining deterministic directed change, on the one hand, and random motion, on the other hand. At the same time, Malécot considered the change of the allele frequencies as a Markov process (reviewed in Nagylaki, 1989). In particular, Malécot rigorously demonstrated that assuming weak selection and a large population size, the form conjectured by Wright for the distribution of the allele frequencies was correct, probably the first application of the diffusion approximation to population genetics (Malécot, 1945). After the establishment of theoretical population genetics in the 1930s, it then remained to apply this general framework to a real scenario. GSI was an ideal problem since it combines the selection, mutation and migration, with a parameter easily estimated from natural populations, the number of \( S \)-alleles.

Very rapidly Wright (1939) aimed at using the theory to resolve an apparent paradox from observations in one population of the flowering plant \( O. \) organensis. In this population, 45 self-incompatibility alleles were detected by Emerson (1938, 1939) even though the population size was thought to be no larger than one thousand individuals. Wright (1939) followed the general approach developed before (Fisher, 1930; Wright, 1938): he first provided the deterministic change of the frequency of a focal \( S \)-allele given the frequency of all the other \( S \)-alleles in the population, summarized in the parameter \( R \), which measures the relative success of fertilization of the focal \( S \)-allele compared with the other \( S \)-alleles: \( \Delta p_i = \frac{R_i(1-R_p(R-3R))}{2(1+p_i(1-R))} \). Wright remained elusive on the derivation and biological interpretation of \( R \). We suspect that it can be seen as a selection coefficient, which depends on the genotype structure of the population.

Second, as in his previous works (Wright, 1937, 1938), Wright assumed that the stochastic fluctuations of the \( i \)-th \( S \)-allele frequency in a generation are given by the variance of a binomial distribution, from which he derived an approximation of the \( S \)-allele frequency distribution. As there was no explicit solution for the stationary distribution, Wright (1939) provided numerical results. He finally concluded that the large number of \( S \)-alleles observed in the \( O. \) organensis population would be possible if the mutation rate was as large as \( 10^{-3} - 10^{-4} \) per generation, or if the observed population was a small part of a much larger subdivided population with a low dispersal rate and a mutation rate of \( 10^{-5} - 10^{-6} \).

For almost 20 years, no important advances were made in the theory of stochastic processes applied to self-incompatibility systems. After experimental estimations of the mutation rate, found to be much lower than \( 10^{-6} \) in \( O. \) organensis by Emerson (1939) and Lewis (1948, 1949b), there was a revival of interest in having the best approximation of the frequency distribution of \( S \)-alleles in a population, launched in particular by a reexamination of the theory by Fisher (1958). Fisher (1958) challenged Wright’s (1939) model, called for developing the best stochastic approximation of the GSI system and used it to identify the mechanisms underlying the pattern first highlighted by Emerson (1938). Fisher adopted the same approach as Wright (1939): first he derived the deterministic frequency change and then he assumed a form for the variance of the allelic change from a binomial distribution. In contrast to Wright, however, Fisher (1958) motivated his deterministic change in allele frequency by an individual-based reasoning on the level of the plants. He was thus the first to give an explicit expression for the deterministic
change of a given $S$-allele in the form given in Equation (5) above. Moreover, he also proposed a different form, compared with Wright (1939), for the variance of the $S$-allele frequency change, $\Delta \rho^2 = \frac{p(1-2p)}{2N}$. Similar to Wright, Fisher did not provide an explicit derivation of the infinitesimal variance from a probabilistic model. Finally, after deriving an allele frequency distribution under some technical simplification, Fisher came back to the plausible mechanisms explaining the large number of $S$-alleles observed in the O. organensis population (Emerson, 1938). He concluded that the population had most likely suffered from a strong population bottleneck, resulting in the population not being at drift-mutation equilibrium. Fisher discarded Wright’s (1939) hypothesis that the population could be strongly subdivided, with no clear reasons since Fisher did not consider any population structure or dispersal in his model. Fisher also took the opportunity to criticize that Wright (1939) only relied on numerical computations to analyse his model and was not able to derive an explicit formula.

Wright (1960) answered Fisher’s (1958) criticisms mostly by comparing his 1939’s and new 1960’s models, and Fisher’s model through numerical simulations. Wright found no notable difference between the different models. Still, he pointed out that Fisher’s approximation gave the worst estimation of $S$-allele frequency change among the three models. Overall, Wright (1960) and Fisher (1958), compared with Wright (1939), brought no important conceptual progress to the theory of GSI. Yet, these two papers illustrate that finding the correct form of approximation of the genetic dynamics of a GSI system was a major challenge for population genetics that led to another dispute between Wright and Fisher, besides the famous controversy about genetic dominance (Bagheri, 2006; Billiard & Castric, 2011).

A few years after this dispute, many papers were at least partly devoted to address the question of the correct approximation of the stationary distribution of $S$-allele frequencies in a finite population (Ewens, 1964; Ewens & Ewens, 1966; Kimura & Crow, 1964; Mayo, 1966; Moran, 1962; Wright, 1964, 1966). Yet again, no theoretical advances were made; most results relied on computer simulations that confirmed that Wright’s (1964) approximation was good enough. Later, Yokoyama and Ni (1979) and Yokoyama and Hetherington (1982) obtained new results by novel diffusion theory methods: instead of computing the allele frequency of a single allele, they computed the entire allele frequency distribution to directly infer the number of stably maintained $S$-alleles in a population. Yet again, similar to Wright’s different models (1939, 1960, 1964), they relied on a numerical approximation of the stationary distribution. Finally, and more recently, Muirhead and Wakeley (2009) derived recursion formulae to approximate the $S$-allele frequency distribution. This latest approach does not take a continuum limit (infinite population size limit) but approaches the problem in the discrete state space. While giving good results when compared to simulation results, this approach makes the formulae intractable and complicated to evaluate.

It is remarkable that the population genetics of self-incompatibility systems generated a lot of theoretical progress motivated by the aim to have theoretical predictions compatible with empirical observations. There was clearly a challenge for theoretical population geneticists, as if the stochastic theory of population genetics would be valuable only if it could help to explain the genetic diversity observed in natural populations, or at least if it could give correct predictions about parameter values like the mutation rate. In short, the main question that theoreticians tried to address was: what is the correct way to model the stochastic dynamics of allele frequencies in finite populations?

Two points of view opposed: the need for rigorous vs. practical derivations. Moran (1962) particularly criticized the lack of rigour of Wright’s approach, while Wright (1964) claimed that he only aimed for a good rather than a justified approximation. Such an opposition raised two difficulties: Is it possible to evaluate the validity of a model independently of data obtained from natural populations? How can the robustness and precision of the model’s approximation be evaluated? As emphasized by Moran (1962) and Malécot (reviewed in Nagylaki, 1989), these goals can be achieved by using a probabilistic way of thinking, i.e. by clearly justifying a stochastic model, and by properly deriving its approximation with probabilistic, mathematical tools. Here, we applied this theoretical framework to GSI.

4.2 Reexamination of the number of $S$-alleles in a finite population

One aim of our study was to analytically revisit the results on the number of $S$-alleles under GSI, as derived by Wright (1939, 1964). Even though Wright’s results were numerically confirmed (see the previous section), a theoretical confirmation was still lacking (to the best of our knowledge). Our results on the stationary distribution fit the numerically derived normal distribution obtained by Wright (1964) almost perfectly, e.g. Figure 5, thus confirming Wright’s results also from a theoretical perspective.

Following Wright’s (1939) strategy, we computed the number of $S$-alleles that can be maintained in a finite population of fixed size and for a fixed mutation rate (Figure 2). Our analytical prediction slightly overestimates the simulated values. This is explained by the idealized assumption that all $S$-allele frequencies are centred around the steady state. Therefore, deviations from this stationary distribution, i.e. the rate of extinction of an $S$-allele is underestimated, which results in the overestimation of the simulated number of $S$-alleles. If we were to compute the number of $S$-alleles without the normal approximation of the stationary distribution, we would instead slightly underestimate the number of $S$-alleles as found in Czuppon and Rogers (2019) in the related haploid self-incompatibility (HSI) system, a mating system found for example in some fungi and ciliates.

Qualitatively, HSI is very similar to GSI because individuals can reproduce successfully only by mating with an individual with a different $S$-allele. This, as in GSI, produces negative frequency-dependent selection on the frequency of a focal $S$-allele, which
results in a stable equilibrium frequency at an intermediate value. Recently, quantities like the number of mating types maintained in a population, the invasion probability, extinction dynamics, the effect of clonal reproduction and of fitness differences between different mating types have been addressed in this simpler genetic system (Berrios-Caro et al., 2021; Constable & Kokko, 2018; Czuppon & Constable, 2019; Czuppon & Rogers, 2019; Krumbeck et al., 2020). Even though simpler to analyse, we still suspect the qualitative results to carry over from the HSI to the GSI setting because the overall behaviour of the two dynamical systems is similar.

For example, here our results were derived assuming neither fitness differences between different \(S\)-alleles, nor phenomena such as partial self-compatibility. Accounting for these mechanisms would alter the prediction of the number of \(S\)-alleles, e.g. it was shown that variation in zygote viability reduces the diversity at the \(S\)- locus (Uyenoyama, 2003), which is similar to results found in HSI systems (Krumbeck et al., 2020). In addition, we can speculate that the effect of self-fertilization on \(S\)-allele diversity maintained under GSI would be similar to that under HSI. We would therefore expect that the number of \(S\)-alleles is reduced for increasing rates of self-fertilization (Berrios-Caro et al., 2021; Constable & Kokko, 2018; Czuppon & Constable, 2019).

Considering a single panmictic population is another important simplification of our model. The subdivision of population and the migration rate between demes strongly affect the number of \(S\)-alleles maintained in a population (Schierup et al., 2000; Stoeckel et al., 2008; Wright, 1939). In particular, Schierup et al. (2000) showed that the total number of \(S\)-alleles at the scale of the whole metapopulation varies non-monotonically with the migration rate (the number of \(S\)-alleles is minimal for an intermediate value of the migration rate). This striking result was obtained with numerical simulations. A formal derivation of the \(S\)-diversity in a structured population was suggested by Wright (1939), yet surprisingly, no analytical progress has been made since then (to the best of our knowledge). We suspect that the framework developed in this manuscript could help to make analytical progress, e.g. by studying the edge cases of (fully) isolated populations, i.e. very low migration rates, and a well-mixed population without population structure, i.e. very high migration rates (essentially the model of this manuscript). Importantly, the invasion dynamics of new \(S\)-alleles into a local population will need to be taken into account.

### 4.3 Invasion behaviour under GSI and HSI

The invasion probability of a new \(S\)-allele under GSI, denoted by \(\varphi\) (Equation (9)), can also be compared with the case of HSI. Czuppon and Rogers (2019) showed that a good approximation for the invasion probability of a single \(S\)-allele in an HSI population with \(M\) resident \(S\)-alleles is \(1/M\). Here, for GSI, we find \(\varphi > 1/M\), i.e. the invasion probability of a new \(S\)-allele is always larger for GSI than for HSI. This can be understood by comparing the ratio between the number of individuals that are compatible with the invading and with resident \(S\)-alleles in both situations. Applying the same reasoning that led to Equation (10), we find

\[
\frac{(N-1)\times \frac{M}{N(M-2)}}{\frac{M}{N-1}} \approx \frac{M}{M-1} \text{ for GSI,}
\]

\[
\frac{(N-1)\times \frac{M}{N(M-2)}}{\frac{M}{N(M-1)}} \approx \frac{M}{M-1} \text{ for HSI.}
\]

(12)

This shows that an invading (or rare) \(S\)-allele always has a larger fertilization advantage relative to a resident \(S\)-allele in GSI than in HSI. As a consequence, for similar population sizes, we would expect a higher \(S\)-allele diversity under GSI than under HSI dynamics because the rare allele advantage is stronger in the GSI mating system. The equations also show that as \(M\) increases, the difference between GSI and HSI vanishes as new \(S\)-alleles are getting closer to being neutral. Hence, as \(M\) increases, GSI and HSI tend to have similar invasion dynamics.

### 4.4 \(S\)-haplotype diversification dynamics: Does the underlying molecular architecture matter?

The number of \(S\)-alleles maintained in the population emerges from the balance between the \(S\)-allele loss rate by chance (genetic drift), the appearance of a new \(S\)-allele in the population (mutation rate) and its probability of invasion, which only depends on the number of resident \(S\)-alleles. As with most stochastic models of GSI dynamics after Wright (1939), we assumed a one-step mutation model where mutations instantly generate new, fully functional \(S\)-alleles. If we further assume that mutations are rare
enough so that the $S$-allele frequencies equilibrate between successive mutation events, we can quantify the diversification rate. That is, using (9) and a mutation rate $u$, we find that the diversification dynamics, i.e. the rate at which the number of $S$-alleles $n(t)$ accumulates if the initial number of $S$-alleles is three, follows a square-root function of time: $n(t) = 2 + \sqrt{2tu + 1}$ (mathematical details in Section H in the Appendix S1).

However, assuming a one-step mutation model is (very likely) an oversimplification as shown by the description of the genetic and molecular architectures underlying the SI phenotypes in different families (reviewed in Durand et al., 2020). Generally, a 'superlocus' composed of several genes is involved, which makes the appearance of new $S$-alleles more complicated as it requires several compensatory mutations allowing the self or nonself SI recognition. Using stochastic computer simulations, Gervais et al. (2011), Sakai (2016), Bod'ová et al. (2018) and Harkness et al. (2021) studied the diversification dynamics of $S$-alleles taking explicitly into account the succession of compensatory mutations invading a population. These multi-step mutation models were studied to better reflect the biological complexity associated with the appearance of a new $S$-allele. Gervais et al. (2011) and Bod'ová et al. (2018) showed that the diversification dynamics are due to complex interactions between the selection regimes of the different mutations present in the population (functional SI alleles or self-compatible alleles). Starting with a low number of $S$-alleles, they also showed that the diversification rate was first very fast, and rapidly reached either a plateau, where the number of $S$-alleles was stationary or a regime where the diversification rate was low.

Comparing the diversification dynamics expected under our one-step mutation model to the dynamics observed in the multi-step mutation models studied by Gervais et al. (2011), Sakai (2016), Bod'ová et al. (2018) and Harkness et al. (2021), our results suggest that the genetic architecture underlying SI seems to have only small effects on the diversification dynamics. Indeed, two types of diversification dynamics are observed in multi-step mutation models (Bod'ová et al., 2018; Gervais et al., 2011; Harkness et al., 2021; Sakai, 2016): either the number of $S$-alleles asymptotically approaches its stationary state, or the stationary number of $S$-alleles is rapidly reached after a diversification burst. These two diversification dynamics can be described by our one-step mutation model (Figure 6), respectively, when the mutation rate is low vs. high. When the mutation rate is low, the diversification dynamics follow a square-root function of time (Figure 6a). When the mutation rate is high, the diversification dynamics follow a square-root function of time truncated at the stationary number of $S$-alleles. The number of $S$-alleles accumulates and reaches the stationary phase quickly, which leads to an abrupt halt of an increase in the number of $S$-alleles (Figure 6b). In stochastic simulations, the number of $S$-alleles then fluctuates around this stably maintained number of $S$-alleles. Overall, our results suggest that the diversification dynamics are mainly affected by an effective mutation rate, i.e. the rate of appearance of new $S$-alleles in multi-step models, rather than by the molecular and genetic details of the mechanism underlying SI. However, how this effective mutation rate is related to these mechanisms is an open issue that would need to be addressed by a specific model.

The diversification dynamics derived in our one-step mutation model can also be used as a null model of $S$-allele diversification. One possible application could be to compare the distribution of times separating the appearance of $S$-alleles in phylogenies and to compare it with a square-root function of time. If a substantial discrepancy between empirical vs. null diversification dynamics is observed, this would suggest that, contrarily to what our model suggests, the molecular and genetic mechanisms underlying SI significantly affect the evolution of SI.

### 4.5 Connecting theoretical results to data

Returning to the question about the observed diversity at the $S$-locus in O. organensis, Wright (1960) concluded that the empirically estimated low mutation rate obtained by Lewis (1948) would not be

---

**FIGURE 6** $S$-allele diversification dynamics in our one-step mutation model. The black solid line is the mean obtained from 1000 stochastic simulations of a population with size $N = 5000$ and started with three $S$-alleles in stationarity. The black dashed lines correspond to the 25 and 75 percentiles of the simulations. The blue line shows the theoretical prediction $n(t) = 2 + \sqrt{2tu + 1}$, which fits well with the simulation results. The consistent underestimation is a result of the underestimation of the invasion probability (Figure 3). The mutation rate is set to $u = 1/50N$ in (a) and $u = 1/10N$ in (b). The theoretically predicted number of stably maintained $S$-alleles for the high mutation rate is 37, which, as explained above, is slightly overestimating the number of stably maintained $S$-alleles in simulations.
The allelic frequency spectrum is expected not to be Gaussian anymore at intermediate migration rate; Muirhead, 2001; Schierup, 1998; Schierup et al., 2000; Stoeckel et al., 2008; Wright, 1939). A better understanding of the limitations of this new estimator is therefore needed to reliably apply it to data sets.

5 | CONCLUSION

To summarize, we have revisited the question of the possible number of S-alleles maintained in a finite population from a theoretical point of view. Our explicit expression of the stationary allele frequency distribution confirms the numerically obtained results by Wright (1964). Additionally, we provide an approximation for the invasion probability of a new S-allele and find that it is similar to the corresponding expression in populations with haploid self-incompatibility. Lastly, we define an estimator for the population size of a GSI population, which provides a starting point to assess the state of the sampled population. For example, a larger estimated diversity of S-alleles than predicted by the estimated population size in Figure 2 might be indicative of a recent population decline. Lastly, our model might additionally be useful for conservation purposes. The reproductive output of plants can be reduced by pollen limitation, which may occur especially with low S-allele diversity in SI species (Leducq et al., 2010; Olivieri et al., 2016). Hence, if the population size is known, the prediction of the number of S-alleles for that size can help evaluate the extinction risk.

AUTHOR CONTRIBUTIONS

PC and SB designed the study; PC conducted the formal analysis; PC and SB discussed the results and wrote the manuscript.

ACKNOWLEDGMENTS

We warmly thank Solenn Stoeckel and Stéphanie Mariette who kindly shared their Prunus avium dataset. We also thank Vincent Castric and Xavier Vekemans for fruitful discussions. We thank two anonymous reviewers for thorough comments that resulted in the discussion of diversification dynamics. We are grateful to the High Performance Computing service of the University of Münster for providing computational resources on the computer cluster PALMA II. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The C++ codes, data files and Python scripts used to generate the figures are available at Dryad https://doi.org/10.5061/dryad.0zpc86712.

PEER REVIEW

The peer review history for this article is available at https://publon.com/publon/10.1111/jeb.14061.

ORCID

Peter Czuppon https://orcid.org/0000-0003-1462-7237

Sylvain Billiard https://orcid.org/0000-0003-2135-3246
REFERENCES
Bagheri, H. (2006). Unresolved boundaries of evolutionary theory and the question of how inheritance systems evolve: 75 years of debate on the evolution of dominance. Journal of Experimental Zoology Part B, 306, 329–359. https://doi.org/10.1002/jez.b.21069
Berrios-caro, E., Galla, T., & Constable, G. W. (2021). Switching environments, synchronous sex, and the evolution of mating types. Theoretical Population Biology, 138, 28–42. https://doi.org/10.1016/j.tpb.2021.02.001
Billiard, S., & Castric, V. (2011). Evidence for Fisher’s dominance theory: How many ‘special cases’? Trends in Genetics, 27, 441–445. https://doi.org/10.1016/j.tig.2011.06.005
Bod’ořová, K., Prikořil, T., Field, D. L., Barton, N. H., & Pickup, M. (2018). Evolutionary pathways for the generation of new self-incompatibility haplotypes in a non-self-recognition system. Genetics, 209(3), 861–883.
Boucher, W. (1993). A deterministic analysis of self-incompatibility alleles. Journal of Mathematical Biology, 31(2), 149–155. https://doi.org/10.1007/bf00171223
Castric, V., & Vekemans, X. (2004). Plant self-incompatibility in natural populations: A critical assessment of recent theoretical and empirical advances. Molecular Ecology, 13, 2873–2889. https://doi.org/10.10111/j.1365-294X.2004.02267x
Constable, G. W. A., & Kokko, H. (2018). The rate of facultative sex governs the number of expected mating types in isogamous species. Nature Ecology & Evolution, 2(7), 1168–1175. https://doi.org/10.1038/s41559-018-0580-9
Czuppon, P., & Constable, G. W. A. (2019). Invasion and extinction dynamics of mating types under facultative sexual reproduction. Genetics, 213(2), 567–580. https://doi.org/10.1534/genetics.119.302306
Czuppon, P., & Rogers, D. W. (2019). Evolution of mating types in finite populations: The precarious advantage of being rare. Journal of Evolutionary Biology, 32(11), 1290–1299. https://doi.org/10.1111/jeb.13528
Czuppon, P., & Traulsen, A. (2021). Understanding evolutionary and ecological dynamics using a continuum limit. Ecology and Evolution, 11(11), 5857–5873. https://doi.org/10.1002/ece3.7205
Durand, E., Chantreau, M., Le Veve, A., Stetsenko, R., Dubin, M., Genete, M., Laurens, V., Poux, C., Roux, C., Billiard, S., Vekemans, X., & Castric, V. (2020). Evolution of self-incompatibility in the Brassicaceae: Lessons from a textbook example of natural selection. Evolutionary Applications, 13, 1279–1297. https://doi.org/10.1111/eva.12933
Emerson, S. (1938). The genetics of self-incompatibility in Oenothera organensis. Genetics, 23(2), 190–202. https://doi.org/10.1093/genet/23.2.190
Emerson, S. (1939). A preliminary survey of the Oenothera organensis population. Genetics, 24(4), 524–537. https://doi.org/10.1093/genetics/24.4.524
Ethier, S., & Kurtz, T. (1986). Markov processes: Characterization and convergence. In Wiley series in probability and mathematical statistics. J. Wiley & Sons. https://doi.org/10.1002/9780470316658
Evans, W. (1964). On the problem of self-sterility alleles. Genetics, 50(6), 1433–1438. https://doi.org/10.1093/genetics/50.6.1433
Evans, W. J., & Evans, P. M. (1966). The maintenance of alleles by mutation — Monte Carlo results for normal and self-sterility populations. Heredity, 21(3), 371–378. https://doi.org/10.1038/hdy.1966.38
Fisher, R. (1930). The genetical theory of natural selection. Oxford University Press.
Fisher, R. (1958). The genetical theory of natural selection - A complete, Variorum edition. Oxford University Press.
Gervais, C. E., Castric, V., Ressayre, A., & Billiard, S. (2011). Origin and diversification dynamics of self-incompatibility haplotypes. Genetics, 188(3), 625–636.
Harkness, A., Goldberg, E. E., & Brandvain, Y. (2021). Diversification or collapse of self-incompatibility haplotypes as a rescue process. The American Naturalist, 197(3), E89–E109.
Ishida, Y., & Rosales, A. (2020). The origins of the stochastic theory of population genetics: The Wright-Fisher model. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 79, 101226.
Kimura, M., & Crow, J. (1964). The number of alleles that can be maintained in a finite population. Genetics, 49(4), 725–738. https://doi.org/10.1093/genetics/49.4.725
Krumbeek, Y., Constable, G. W. A., & Rogers, T. (2020). Fitness differences suppress the number of mating types in evolving isogamous species. Royal Society Open Science, 7(2), 192126. https://doi.org/10.1098/rsos.192126
Kurtz, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. Journal of Applied Probability, 8(2), 344–356. https://doi.org/10.2307/3211904
Leduca, J.-B., Gosset, C. C., Poirot, M., Hendoux, F., Vekemans, X., & Billiard, S. (2010). An experimental study of the S-Allee effect in the self-incompatible plant Biscutella neustriaca. Conservation Genetics, 11(2), 497–508.
Levin, D. A., Ritter, K., & Elstrand, N. C. (1979). Protein polymorphism in the narrow endemic Oenothera organensis. Evolution, 33(2), 534–542. https://doi.org/10.1111/1558-5646.1979.tb04708.x
Lewis, D. (1948). Structure of the incompatibility gene. I spontaneous mutation rate. Heredity, 2(2), 219–236. https://doi.org/10.1038/hdy.1948.12
Lewis, D. (1949a). Incompatibility in flowering plants. Biological Reviews, 24(4), 472–496. https://doi.org/10.1111/j.1469-185x.1949.tb00584.x
Lewis, D. (1949b). Structure of the incompatibility gene. II induced mutation rate. Heredity, 3(3), 339–355. https://doi.org/10.1038/hdy.1949.25
Malécot, G. (1945). La diffusion des gènes dans une population mendélienne. Comptes Rendus de l'Académie des Sciences, 221, 340–342.
Mayo, O. (1966). On the problem of self-incompatibility alleles. Biometrika, 22(1), 111. https://doi.org/10.2307/2528218
Moran, P. (1962). The statistical process of evolution theory. Oxford University Press.
Muirhead, C., & Wakeley, J. (2009). Modeling multiallelic selection using a Moran model. Genetics, 3182, 1141–1157. https://doi.org/10.1534/genetics.108.089474
Muirhead, C. A. (2001). Consequences of population structure on genes under balancing selection. Evolution, 55(8), 1532–1541. https://doi.org/10.1534/evol.1.0014-3820.2001.tb00473.x
Nagylaki, T. (1975). The deterministic behavior of self-incompatibility alleles. Genetics, 79(3), 545–550. https://doi.org/10.1093/genet/79.3.545
Nagylaki, T. (1989). Gustave Malécot and the transition from classical to modern population genetics. Genetics, 122, 253–268. https://doi.org/10.1093/genetics/122.2.253
Oliveri, I., Tonnabel, J., Ronce, O., & Mignot, A. (2016). Why evolution matters for species conservation: Perspectives from three case studies of plant metapopulations. Evolutionary Applications, 9(1), 196–211.
Paxman, G. J. (1963). The maximum likelihood estimation of the number of self-sterility alleles in a population. Genetics, 48(8), 1029–1032.
Sakai, S. (2016). How have self-incompatibility haplotypes diversified? Generation of new haplotypes during the evolution of self-incompatibility from self-compatibility. The American Naturalist, 188(2), 163–174.
Schierup, M. H. (1998). The number of self-incompatibility alleles in a finite, subdivided population. Genetics, 149(2), 1153–1162. https://doi.org/10.1093/genetics/149.2.1153
Schiereup, M. H., Vekemans, X., & Charlesworth, D. (2000). The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genetical Research, 76*(1), 51–62. https://doi.org/10.1017/s0016672300004535

Stoeckel, S., Castric, V., Mariette, S., & Vekemans, X. (2008). Unequal allelic frequencies at the self-incompatibility locus within local populations of Prunus avium L.: An effect of population structure? *Journal of Evolutionary Biology, 21*(3), 889–899. https://doi.org/10.1111/j.1420-9101.2008.01504.x

Stoeckel, S., Klein, E., Oddou-Muratorio, S., Musch, B., & Mariette, S. (2012). Microevolution of S-allele frequencies in wild cherry populations: Respective impacts of negative frequency dependent selection and genetic drift. *Evolution, 66*, 486–504. https://doi.org/10.1111/j.1558-5646.2011.01457.x

Uyenoyama, M. K. (2003). Genealogy-dependent variation in viability among self-incompatibility genotypes. *Theoretical Population Biology, 63*(4), 281–293. https://doi.org/10.1016/s0040-5809(03)00020-0

van Kampen, N. (2007). *Stochastic processes in physics and chemistry.* North Holland. https://doi.org/10.1016/B978-0-444-52965-7.00004-4

Wright, S. (1937). The distribution of gene frequencies in populations. *Proceedings of the National Academy of Sciences of the United States of America, 23*, 307–320. https://doi.org/10.1073/pnas.23.6.307

Wright, S. (1938). The distribution of gene frequencies under irreversible mutation. *Proceedings of the National Academy of Sciences of the United States of America, 24*, 253–259. https://doi.org/10.1073/pnas.24.7.253

Wright, S. (1939). The distribution of self-sterility alleles in populations. *Genetics, 24*(4), 538–552. https://doi.org/10.1093/genetics/24.4.538

Wright, S. (1945). The differential equation of the distribution of gene frequencies. *Proceedings of the National Academy of Sciences of the United States of America, 31*, 382–389. https://doi.org/10.1073/pnas.31.12.382

Wright, S. (1960). On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of given size - A reexamination. *Biometrics, 16*(1), 61–85. https://doi.org/10.2307/2527956

Wright, S. (1964). The distribution of self-incompatibility alleles in populations. *Evolution, 18*(4), 609–619. https://doi.org/10.1111/j.1558-5646.1964.tb01675.x

Wright, S. (1966). Polyallelic random drift in relation to evolution. *Proceedings of the National Academy of Sciences of the United States of America, 55*(5), 1074–1081. https://doi.org/10.1073/pnas.55.5.1074

Yokoyama, S., & Hetherington, L. E. (1982). The expected number of self-incompatibility alleles in finite plant populations. *Heredity, 48*(2), 299–303. https://doi.org/10.1038/hdy.1982.35

Yokoyama, S., & Nei, M. (1979). Population-dynamics of a sex-determining alleles in honey bees and self-incompatibility alleles in plants. *Genetics, 91*(3), 609–626. https://doi.org/10.1093/genetics/91.3.609

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Czuppon, P., & Billiard, S. (2022). Revisiting the number of self-incompatibility alleles in finite populations: From old models to new results. *Journal of Evolutionary Biology, 35*, 1296–1308. https://doi.org/10.1111/jeb.14061