nature ecology & evolution
SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41559-018-0580-9

The rate of facultative sex governs the number of expected mating types in isogamous species

George W. A. Constable[1]* and Hanna Kokko[2]

[1]Department of Mathematical Sciences, University of Bath, Bath, UK. [2]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland. *e-mail: g.w.a.constable@bath.ac.uk

Supplementary Information: The rate of facultative sex governs the number of expected mating types in isogamous species

George W. A. Constable*
Department of Mathematical Sciences
The University of Bath
g.w.a.constable@bath.ac.uk

Hanna Kokko
Department of Evolutionary Biology and Environmental Studies
University of Zurich

Contents

S1 Background mathematical theory

We envisage a continuous-time Markov process describing the time-evolution of being in a state $\boldsymbol{n}$ at a time t, $P(\boldsymbol{n}, t)$. Here $\boldsymbol{n}$ is a vector of integers of length $M_{\max}$ with each element representing the number of individuals in the population of a given type. Suppose that $T(\boldsymbol{n}'|\boldsymbol{n})$ is a function that gives the probability per unit time of transitioning from a state $\boldsymbol{n}$ to a state $\boldsymbol{n}'$. The probabilistic dynamics of the system are then described by the master equation;

$$\frac{\mathrm{d}P(\boldsymbol{n},t)}{\mathrm{d}t} = \sum_{\boldsymbol{n}'\neq\boldsymbol{n}} \left[T(\boldsymbol{n}|\boldsymbol{n}')P(\boldsymbol{n}',t) - T(\boldsymbol{n}'|\boldsymbol{n})P(\boldsymbol{n},t)\right]. \tag{S1}$$

This is relatively straightforward to understand. The first term on the right-hand side is made up of the probability of being in state $\boldsymbol{n}'$ multiplied by the probability of being in that state and making a transition to state $\boldsymbol{n}$. It therefore represents the probability of starting in state $\boldsymbol{n}'$ and making at transition to state $\boldsymbol{n}$. In the same way the second term on the right-hand side represents the probability of starting in state $\boldsymbol{n}$ and making at transition to state $\boldsymbol{n}'$. Their difference, summed over all states $\boldsymbol{n}'$, different to $\boldsymbol{n}$, gives the rate of increase of $P(\boldsymbol{n}, t)$ with time. A realization of a stochastic process with probability distribution $P(\boldsymbol{n}, t)$ can be efficiently and accurately simulated using the Gillespie simulation [4].

We aim to identify stationary distribution of the process, $P^{\mathrm{st}}(\boldsymbol{n})$. This is analogous to finding the fixed points in a deterministic system, and essentially gives us the probability of finding the system in a state $\boldsymbol{n}$ at infinitely long times. Recognizing that the stationary distribution is time-independent, we find that it can be obtained as the solution to the following set of difference equations;

$$\sum_{\boldsymbol{n}'\neq\boldsymbol{n}} \left[T(\boldsymbol{n}|\boldsymbol{n}')P^{\mathrm{st}}(\boldsymbol{n}') - T(\boldsymbol{n}'|\boldsymbol{n})P^{\mathrm{st}}(\boldsymbol{n})\right] = 0\,. \tag{S2}$$

In general, obtaining an analytic solution for Eq. (S2) is difficult. However, we will show that an analytic solution is obtainable if we can express each and every transition in the system by

$$T(\boldsymbol{n} + \boldsymbol{\Delta}^{(ij)}|\boldsymbol{n}) = b(n_i)d(n_j)\,, \qquad \forall\, i,\, j \tag{S3}$$

where

$$\begin{aligned} \Delta_k^{(ij)} &= +1 \quad \text{if} \quad k = i\,, \\ \Delta_k^{(ij)} &= -1 \quad \text{if} \quad k = j\,, \\ \Delta_k^{(ij)} &= 0 \quad \text{otherwise.} \end{aligned} \tag{S4}$$

Under these conditions, we propose that the stationary distribution is proportional to

$$P^{\mathrm{st}}(\boldsymbol{n}) \propto \prod_{i=1}^{M-1} \prod_{k=0}^{n_i^{\downarrow}-1} \frac{b(k)d\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow} - 1\right)d(k+1)}\,, \tag{S5}$$

where $\boldsymbol{n}^{\downarrow}$ is similar to the vector $\boldsymbol{n}$ but reordered with its largest elements first and M is the number of non-zero entries of $\boldsymbol{n}$. The validity of this solution can be demonstrated by substitution of Eq. (S5) into Eq. (S2). Here we have essentially shown that we can solve for the stationary distribution if the system is a one-step process in which every transition involves the birth of one individual and the death of another (see Eq. (S4)) and if additionally the probability of each-birth death event is dependent only on the frequencies of the type reproducing and the type dying (see Eq. (S3)). This solution is a result of the fact that under the above conditions, there is symmetry under the re-labeling of types, i.e.

$$P^{\mathrm{st}}(\boldsymbol{n}) = P^{\mathrm{st}}(\boldsymbol{n}^{\downarrow})\,. \tag{S6}$$

The symmetry in the solution can be appreciated more fully by comparing the stationary solution for a birth-death process with two types [6] to Eq. (S5), which we rewrite for emphasis as;

$$P^{\mathrm{st}}(\boldsymbol{n}) \propto \prod_{i=1}^{M-1} \prod_{k=0}^{n_i^{\downarrow}-1} \frac{b(k)d\left(N_i - k\right)}{b\left(N_i - k - 1\right)d(k+1)}\,, \tag{S7}$$

where

$$N_i = N - \sum_{j=1}^{i-1} n_j^{\downarrow}\,. \tag{S8}$$

Eq. (S7) can be re-expressed in the following way, which we will find slightly more useful;

$$\begin{aligned}
P^{\mathrm{st}}(\boldsymbol{n}) &\propto \prod_{i=1}^{M-1}\left[\prod_{k=0}^{n_i^{\downarrow}-1}\frac{b(k)}{d(k+1)}\right]\left[\prod_{k=0}^{n_i^{\downarrow}-1}\frac{d\left(N-k-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}{b\left(N-(k+1)-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}\right]\\
P^{\mathrm{st}}(\boldsymbol{n}) &\propto \prod_{i=1}^{M-1}\left[\frac{b(0)}{d(n_i)}\prod_{k=1}^{n_i^{\downarrow}-1}\frac{b(k)}{d(k)}\right]\left[\frac{d\left(N-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}{b\left(N-n_i-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}\prod_{k=1}^{n_i^{\downarrow}-1}\frac{d\left(N-k-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}{b\left(N-k-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}\right]\\
P^{\mathrm{st}}(\boldsymbol{n}) &\propto \prod_{i=1}^{M-1}\left[\frac{b(0)}{d(n_i)}\prod_{k=1}^{n_i^{\downarrow}-1}\frac{b(k)}{d(k)}\right]\left[\frac{d\left(N-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}{b\left(N-\sum_{j=1}^{i}n_j^{\downarrow}\right)}\prod_{k=1}^{n_i^{\downarrow}-1}\frac{d\left(N-k-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}{b\left(N-k-\sum_{j=1}^{i-1}n_j^{\downarrow}\right)}\right]\,.
\end{aligned} \tag{S9}$$

Systems such as Eq. (S1) with transitions obeying Eq. (S4) can also be analyzed in the deterministic limit, which is valid when the population has infinite size. In order to obtain a description of the dynamics in this limit, we make a diffusion approximation [2] (see also Kramers-Moyal expansion [7]); assuming N is large, since transitions between states occur only locally in state space (see Eq. (S4)), we can treat the vector of variables $\boldsymbol{x} = \boldsymbol{n}/N$ as approximately continuous and Taylor expand Eq. (S1) in a series in $1/N$. Taking only the leading order (deterministic) term that survives when we take the limit $N \to \infty$ we find that the frequencies of the mating types $\boldsymbol{x}$ evolve according to

$$\begin{aligned}
\frac{\mathrm{d}x_i}{\mathrm{d}\tau} &= \sum_{j=1}^{M_{\max}}\left[T(N\boldsymbol{x}+\boldsymbol{\Delta}^{(ij)}|N\boldsymbol{x}) - T(N\boldsymbol{x}+\boldsymbol{\Delta}^{(ji)}|N\boldsymbol{x})\right]\Big|\lim_{N\to\infty}\,,\\
&= \sum_{j=1}^{M_{\max}}\left[b\left(Nx_i\right)d\left(Nx_j\right) - b\left(Nx_j\right)d\left(Nx_i\right)\right]\Big|\lim_{N\to\infty}\,.
\end{aligned} \tag{S10}$$

where $\tau = t/N$ [5] and $M_{\max}$ is the length of vectors $\boldsymbol{n}$ and $\boldsymbol{x}$.

In comparing this notation with that used in the main text, note that

$$T(\boldsymbol{n}+\boldsymbol{\Delta}^{(ij)}|\boldsymbol{n}) = T\left[(\ldots,n_i+1,\ldots,n_j-1,\ldots)\,|\,(\ldots,n_i,\ldots,n_j,\ldots)\right] \equiv \mathcal{T}_{ij} \tag{S11}$$

S2 Deterministic model analysis

For simplicity, we begin by assuming that m is zero. Using Eq. (S10) with Eqs. (6-7) in the main text it is straightforward to calculate the deterministic dynamics;

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \sum_{j=1}^{M_{\max}}\left\{\left[cD_jx_ix_j + \frac{(1-c)}{2}D_jx_i(1-x_i)x_j\right] - \left[cD_ix_jx_i + \frac{(1-c)}{2}D_ix_j(1-x_j)x_i\right]\right\}\,. \tag{S12}$$

Recall that the terms D_i are chosen from a normal distribution with mean 1 and variance σ. To simplify the above equation, as well as to more clearly delineate the effect of the frequency dependent selection terms D_i, we make a parameter change $D_i = 1 + \sigma d_i$, where d_i are chosen from a normal distribution with mean zero and variance 1. The equation can then be rewritten as Eq. (8) in the main text.

Let us first consider the case where there are no fitness differences between the mating types (i.e. $\sigma = 0$). Eq. (8) then simplifies to

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{(1-c)}{2}\sum_{j=1}^{M_{\max}} x_ix_j(x_j-x_i)\,. \tag{S13}$$

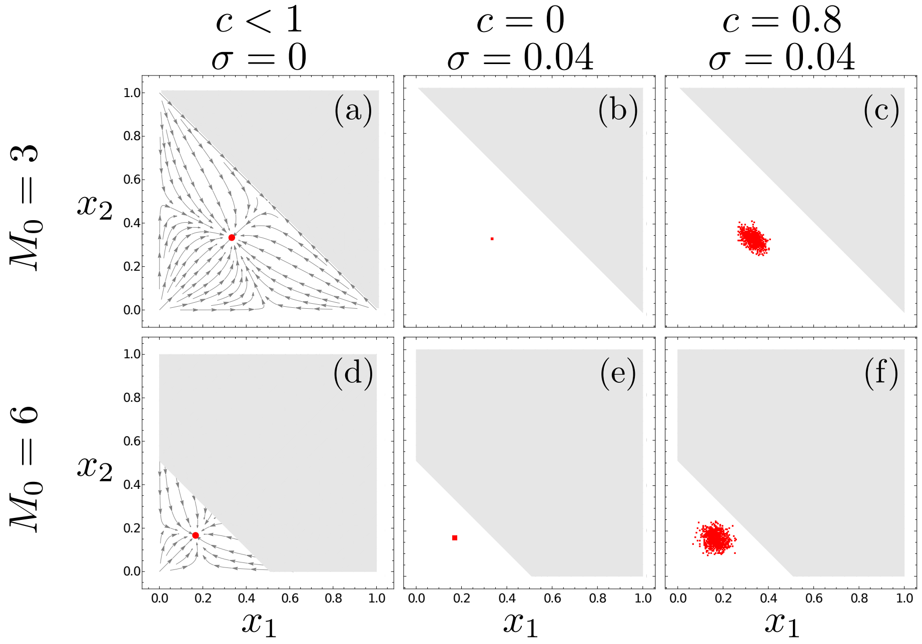

Figure S1: The dynamics of mating type frequencies in the zero mutation, infinite population size (deterministic) limit. Plots (a) and (d) show phase diagrams of the dynamics when there are no selective differences between the mating types, so that a stable fixed point exists (red dot) at $x_i = 1/M$ (unplotted variables, $i > 3$ are evaluated at this fixed point value). These dynamics hold so long as $c < 1$. When selection is introduced ($\sigma > 0$) the fixed point value is a function of the stochastically chosen parameters d_i; plots (b), (c), (e) and (f) show histograms of the fixed point locations for random values of d_i. When sex is obligate (plots (b) and (e)), the fixed points are near even sex ratios. When sex is rarer (plots (c) and (f)) these fixed points can be more biased. Increasing the number of mating types (plots (e) and (f)) also makes the distribution of fixed points more spread out, as well as increasing their proximity to extinction boundaries.

To this system we periodically add mutations at a rate m. Suppose that we have arbitrary initial conditions but with exactly M_0 mating types initially present (that is, M_0 non-zero entries of $\boldsymbol{x}$). After some transient dynamics, the system relaxes to a stable fixed point at $x_i = 1/(M_0)$ for each i^{th} mating type that was initially present in the population ($x_i = 0$ otherwise, as we are ignoring mutation here). This is illustrated graphically in Figure S1 for the cases $M = 3$ and $M = 6$ (see (a) and (d)). The eigenvalues of the system at this fixed point are

$$\lambda_i = -\frac{1-c}{2M_0}\,. \tag{S14}$$

in the absence of mutation. Note that the stability of the fixed point relies on $1 > c$; in other words, there must be at least some sexual reproduction in the population. Without sexual reproduction (that is, when $c = 1$), the model essentially reverts to the neutral Moran model (see Eqn. (6), main text). The stability of the fixed point is also dependent on mutation being zero; the fixed point becomes unstable if a new mating type arises, so that the number of mating types in now $M_0 + 1$. In this scenario a new stable fixed point emerges at $x_i = 1/(M_0 + 1)$

$$\lambda_i = -\frac{1-c}{2(M_0+1)} \tag{S15}$$

to which the system quickly approaches. Therefore the picture that develops is that each new mating type that arises invades the population with certainty and drives the system to a fixed point in which all mating types are present at equal frequency. Given that new mating types arise on average at a rate m, the deterministic prediction is that the number of mating types in the population at a given time should follow the equation

$$\begin{aligned} M(t) &= M_0 + mt\,, \\ M(\tau) &= M_0 + mN\tau\,. \end{aligned} \tag{S16}$$

This is shown graphically in Fig. 3a in the main text.

Now let us instead assume that there are fitness differences between the mating types (i.e. $\sigma > 0$). The fixed point solutions are now given by Eq. (11) in the main text. Clearly these depend on the values of the stochastically determined parameters d_i. In general however, we can say that as c increases (and thus selection for even sex ratios decreases) smaller values of σ are sufficient to generate skewed sex ratios, and even complete extinctions of a given mating type allele. Further, as M increases, the probability that a randomly generated value of d_i will lead to an unstable system increases, as: (i) the fixed point moves closer to extinction boundaries; (ii) selection for even sex ratios weakens (see Eq. (S14)). These points are illustrated graphically in Fig. S1.

Aside on the interpretation of τ and generation times. The model that we have formulated features continuous birth and death processes in the spirit of the Moran model. It does not feature discrete generations (as in Wright-Fisher models). However for developing biological intuition, it is useful to consider the timescale on which we might expect the population to go through N birth-death events, which can be viewed as analogous to a generation. First let's look at the case $c = 1$. In this case the population reproduces entirely asexually, and our model reduces to the Moran model. In this case, the average time between each birth-death events is one, leading to a generation time (i.e. time for N birth-death events) of N. This is a well known result for mapping between the haploid Wright-Fisher and Moran models [3]. Now let's look at the case $c = 0$; here the time for each reproductive event to occur is dependent on the population composition. However, intuition can be gained by considering the birth-death rate when the population resides at a fixed point in which $\boldsymbol{x} = \boldsymbol{n}/N \approx N/M$. When there are two mating types, $M = 2$, there will be a birth-death event approximately once every two time-steps. However this number decreases as the number of mating types increases; when M and N are very large, the term $N - n_i/N$ in Eqn. (6) (main text) tends to one. In this case, the rate of reproduction becomes the same as that in the limit $c = 1$; that is the average time between each birth-death events is one. We therefore expect that for general c, the generation time should be somewhere between $t = N$ and $t = 2N$. Since both are of order N, we simplify our analysis and assume that the generation time can be approximated by N in all scenarios. Under this assumption, the variable τ can be considered as measuring time in units of generations. We can also then define $m_g = mN$ as the mutation rate per generation. Eq. (S16) can then be re-written as

$$M(\tau) = M_0 + m_g \tau \, .$$

S3 Explicit stationary probability distribution for the neutral $\sigma = 0$ model

In this section we re-express Eq. (S9) for the neutral described in the main text (see Eq. (6-7)) under three sexual scenarios; asexuality, obligate sexuality and facultative sexuality. We will employ the general result expressed in Eq. (S5). In order for Eq. (6-7) in the main text to obey the functional form stated in Eq. (S3), we must assume $\sigma = 0$; that is we can only obtain analytical results when there are no fitness differences between the mating type alleles (see Eq. (13)). Additionally, we note that the mutation term given by Eq. (13) in the main text also does not strictly obey the conditions described in Section S1 that are required to calculate the stationary distribution for the process; $b(0)$ is a function of M (the number of non-zero entries of $\boldsymbol{n}$) and therefore the birth rate of a given type is not independent of the numbers of the remaining types. However, if we take the limit of infinite M_{max} (as we will implement later in the analysis, see Section S4.2) the dependence of $b(0)$ on M drops out of the equations.

S3.1 Entirely clonal model ($c = 1$): Recapitulation of an infinite allele Moran model

Taking Eq. (S9) and substituting $b(k)$ and $d(k)$ from Eq. (13) (see main text) with $c = 1$, we can obtain an expression for the stationary distribution in the case of purely asexual reproduction. Let us start by simplifying the individual terms in Eq. (S9);

$$
\begin{aligned}
\frac{b(0)}{d(n_i^{\downarrow})} &= \left(\frac{m}{M_{\max} - M}\right)\frac{N}{n_i^{\downarrow}}\,, \\
\frac{b(k)}{d(k)} &= 1\,, \\
\frac{d\left(N - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - \sum_{j=1}^{i} n_j^{\downarrow}\right)} &= \frac{N - \sum_{j=1}^{i-1} n_j^{\downarrow}}{N - \sum_{j=1}^{i} n_j^{\downarrow}}\,, \\
\frac{d\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)} &= 1\,.
\end{aligned}
\tag{S17}
$$

Conducting the product over i from 1 to $M - 1$ in Eq. (S9) using the expression in Eq. (S17), we find

$$
\begin{aligned}
\prod_{i=1}^{M-1} \frac{b(0)}{d(n_i)} &= \left(\frac{mN}{M_{\max} - M}\right)^{(M-1)} \prod_{i=1}^{M-1} \frac{1}{n_i^{\downarrow}}\,, \\
\prod_{i=1}^{M-1} \frac{d\left(N - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - \sum_{j=1}^{i} n_j^{\downarrow}\right)} &= \prod_{i=1}^{M-1} \frac{N - \sum_{j=1}^{i-1} n_j^{\downarrow}}{N - \sum_{j=1}^{i} n_j^{\downarrow}}\,, \\
&= \frac{N}{N - \sum_{i=1}^{M-1} n_i^{\downarrow}} \prod_{i=1}^{M-1} 1\,, \\
&= \frac{N}{n_M^{\downarrow}} \prod_{i=1}^{M-1} 1\,,
\end{aligned}
\tag{S18}
$$

Putting these expression back together, we find that for $c = 1$ Eq. (S9) takes the form

$$
P^{\mathrm{st}}(\boldsymbol{n}) \propto \left(\frac{mN}{M_{\max} - M}\right)^{(M-1)} \prod_{i=1}^{M} \frac{1}{n_i^{\downarrow}}\,,
\tag{S19}
$$

where we note that we have omitted a constant factor N since this will drop out in any normalization of $P^{\mathrm{st}}(\boldsymbol{n})$. This result can be seen to perfectly capture the results of simulations, as illustrated in Figure S2.

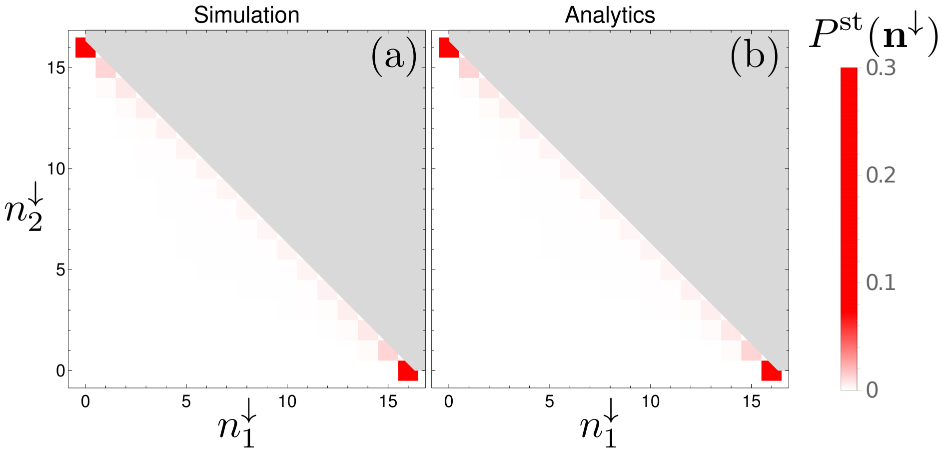

Figure S2: Density plots of the stationary probability distribution of $\boldsymbol{n}^{\downarrow}$) in the $(n_1^{\downarrow}, n_2^{\downarrow})$ plane for the case of purely asexual reproduction ($c = 1$). The distribution is strongly peaked in states in which there is only a single mating type. Analytical results are obtained using Eq. (S19) and can be seen to capture the behavior of simulations extremely well. Low population numbers, $N = 16$, are here utilized to facilitate evaluation of the analytic expression. Remaining parameters are $\sigma = 0$ and $m = 0.01$. The simulation is conducted with $M_0 = 5$ and given $10,000$ generations to equilibrate before the distribution is time-averaged over $400,000$ generations.

S3.2 Entirely sexual model ($c = 0$)

Taking Eq. (S9) and substituting $b(k)$ and $d(k)$ from Eq. (13) (see main text) with $c = 1$, we can obtain an expression for the stationary distribution in the case of purely sexual reproduction. Again, let us start by simplifying the individual terms in Eq. (S9);

$$\begin{aligned}
\frac{b(0)}{d(n_i^{\downarrow})} &= \left(\frac{m}{M_{\max} - M}\right)\frac{N}{n_i^{\downarrow}}\,, \\
\frac{b(k)}{d(k)} &= \frac{1}{2}\left(\frac{N-k}{N}\right)\,, \\
\frac{d\left(N - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - \sum_{j=1}^{i} n_j^{\downarrow}\right)} &= \frac{\left(\frac{N-\phi_i}{N}\right)}{\frac{1}{2}\left(\frac{N-\phi_{i+1}}{N}\right)\frac{\phi_{i+1}}{N}}\,, \\
\frac{d\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N - k - \sum_{j=1}^{i-1} n_j^{\downarrow}\right)} &= \frac{2N}{k+\phi_i}\,,
\end{aligned} \tag{S20}$$

where

$$\phi_i = \sum_{j=1}^{i-1} n_j^{\downarrow}\,. \tag{S21}$$

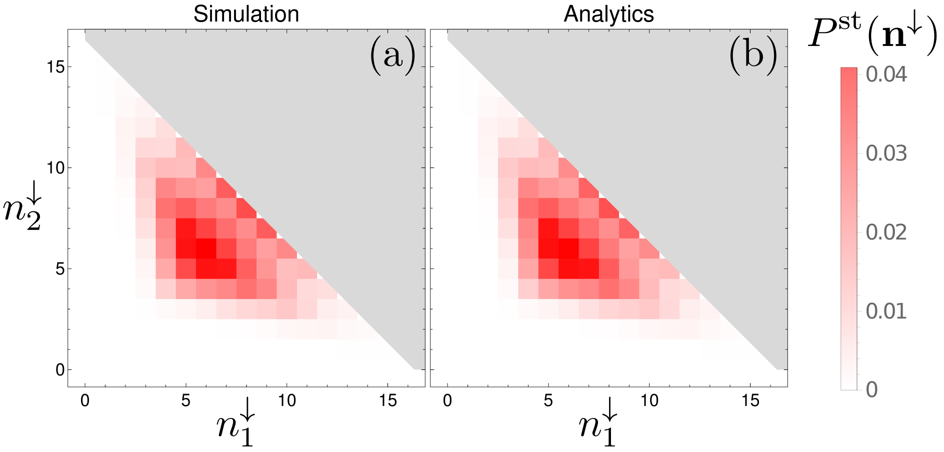

Figure S3: Density plots of the stationary probability distribution of $\boldsymbol{n}^{\downarrow}$) in the $(n_1^{\downarrow}, n_2^{\downarrow})$ plane for the case of obligately sexual reproduction ($c = 0$). The distribution is strongly peaked in states containing approximately three mating types. Analytical results are obtained using Eq. (S25) and can be seen to capture the behavior of simulations extremely well. Low population numbers, $N = 16$, are here utilized to facilitate evaluation of the analytic expression. Remaining parameters are $\sigma = 0$ and $m = 0.01$. The simulation is conducted with $M_0 = 5$ and given $10,000$ generations to equilibrate before the distribution is time-averaged over $400,000$ generations.

Conducting the products over k in Eq. (S9) using the expression in Eq. (S20), we find

$$\begin{aligned}
\prod_{k=1}^{n_i^{\downarrow}-1} \frac{b(k)}{d(k)} &= \left(\frac{1}{2N}\right)^{(n_i^{\downarrow}-1)} \left(\frac{(N-1)!}{(N-n_i^{\downarrow})!}\right), \\
&= \frac{1}{N}\left(\frac{1}{2N}\right)^{(n_i-1)} \left(\frac{N!}{(N-n_i)!}\right), \\
\prod_{k=1}^{n_i^{\downarrow}-1} \frac{d\left(N-k-\sum_{j=1}^{i-1} n_j^{\downarrow}\right)}{b\left(N-k-\sum_{j=1}^{i-1} n_j^{\downarrow}\right)} &= (2N)^{(n_i^{\downarrow}-1)} \left(\frac{(\phi_i)!}{(\phi_{i+1}-1)!}\right), \\
&= (2N)^{(n_i^{\downarrow}-1)} \left(\frac{\phi_{i+1}(\phi_i)!}{(\phi_{i+1})!}\right). \qquad \text{(S22)}
\end{aligned}$$

Putting these terms back together, we find that for $c = 0$ Eq. (S9) takes the form

$$\begin{aligned}
P^{\mathrm{st}}(\boldsymbol{n}) &\propto \prod_{i=1}^{M-1} \left[\left(\frac{m}{M_{\max}-M}\right)\frac{N}{n_i^{\downarrow}}\right]\left[\frac{1}{N}\left(\frac{1}{2N}\right)^{(n_i^{\downarrow}-1)}\left(\frac{N!}{(N-n_i^{\downarrow})!}\right)\right]\left[\frac{2N\,(N-\phi_i)}{(N-\phi_{i+1})\,\phi_{i+1}}\right]\left[(2N)^{(n_i^{\downarrow}-1)}\left(\frac{\phi_{i+1}(\phi_i)!}{(\phi_{i+1})!}\right)\right] \\
P^{\mathrm{st}}(\boldsymbol{n}) &\propto \prod_{i=1}^{M-1} \left[\left(\frac{m}{M_{\max}-M}\right)\frac{1}{n_i^{\downarrow}}\right]\left[\left(\frac{N!}{(N-n_i^{\downarrow})!}\right)\right]\left[\frac{2N\,(N-\phi_i)}{(N-\phi_{i+1})}\right]\left[\left(\frac{(\phi_i)!}{(\phi_{i+1})!}\right)\right] \qquad \text{(S23)}
\end{aligned}$$

Eq. (S23) can be simplified further. Conducting the products over i, in Eq. (S23) we find

$$\begin{aligned}
\prod_{i=1}^{M-1}\left[\left(\frac{m}{M_{\max}-M}\right)\frac{1}{n_i^{\downarrow}}\right] &= \left(\frac{m}{M_{\max}-M}\right)^{(M-1)}\prod_{i=1}^{M-1}\frac{1}{n_i^{\downarrow}}\,,\\
\prod_{i=1}^{M-1}\left[\left(\frac{N!}{(N-n_i^{\downarrow})!}\right)\right] &= (N!)^{(M-1)}\prod_{i=1}^{M-1}\left(\frac{1}{(N-n_i^{\downarrow})!}\right)\,,\\
\prod_{i=1}^{M-1}\left[\frac{2N\,(N-\phi_i)}{(N-\phi_{i+1})}\right] &= (2N)^{(M-1)}\frac{N-\phi_0}{N-\phi_M}\,,\\
&= (2N)^{(M-1)}\frac{N}{n_M^{\downarrow}}\,,\\
\prod_{i=1}^{M-1}\left[\left(\frac{(\phi_i)!}{(\phi_{i+1})!}\right)\right] &= \frac{(\phi_0)!}{(\phi_M)!}\\
&= \frac{1}{(N-n_M^{\downarrow})!}\,. \qquad \text{(S24)}
\end{aligned}$$

This allows us to finally express the stationary distributions as

$$P^{\mathrm{st}}(\boldsymbol{n}) \propto \left(\frac{2mNN!}{M_{\max}-M}\right)^{(M-1)}\prod_{i=1}^{M}\frac{1}{n_i^{\downarrow}}\frac{1}{(N-n_i^{\downarrow})!},\,. \qquad \text{(S25)}$$

This result can be seen to perfectly capture the results of simulations, as illustrated in Figure S3.

S3.3 Facultative model ($1 > c > 0$)

Taking Eq. (S9) and substituting $b(k)$ and $d(k)$ from Eq. (13) (see main text) with c now between 0 and 1, we can obtain an expression for the stationary distribution in the case of facultatively reproduction. Again, let us start by simplifying the individual terms in Eq. (S9);

$$\begin{aligned}
\frac{b(0)}{d(n_i^{\downarrow})} &= \left(\frac{m}{M_{\max}-M}\right)\frac{N}{n_i^{\downarrow}}\,,\\
\frac{b(k)}{d(k)} &= c+\frac{1}{2}(1-c)\left(\frac{N-k}{N}\right)\,,\\
\frac{d\,(N-\phi_i)}{b\,(N-\phi_{i+1})} &= \frac{N-\phi_i}{c(N-\phi_{i+1})+(1-c)(N-\phi_{i+1})\phi_{i+1}/(2N)}\,,\\
\frac{d\,(N-k-\phi_i)}{b\,(N-k-\phi_i)} &= \frac{1}{c+(1-c)(k+\phi_i)/(2N)}\,, \qquad \text{(S26)}
\end{aligned}$$

where again,

$$\phi_i = \sum_{j=1}^{i-1} n_j\,. \qquad \text{(S27)}$$

Clearly now the sums over k in Eq. (S26) are slightly more complicated than in the earlier cases where $c = 1$ and $c = 0$. However, let's begin by rewriting the terms involving k in Eq. (S26);

$$\begin{aligned}
\frac{b(k)}{d(k)} &= c+\frac{1}{2}(1-c)\left(\frac{N-k}{N}\right)\,,\\
&= \frac{(1+c)N-(1-c)k}{2N}\,,\\
&= \frac{(1-c)}{2N}\left[\left(\frac{1+c}{1-c}\right)N-k\right]\,. \qquad \text{(S28)}
\end{aligned}$$

From this equation we have

$$\begin{aligned}\prod_{k=1}^{n_i^{\downarrow}-1} \frac{b(k)}{d(k)} &= \left(\frac{1-c}{2N}\right)^{(n_i^{\downarrow}-1)} \frac{\left[\left(\frac{1+c}{1-c}\right)N-1\right]!}{\left[\left(\frac{1+c}{1-c}\right)N-(n_i^{\downarrow}-1)-1\right]!}, \\ &= \frac{1}{\left[\left(\frac{1+c}{1-c}\right)N\right]} \left(\frac{1-c}{2N}\right)^{(n_i^{\downarrow}-1)} \frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{\left[\left(\frac{1+c}{1-c}\right)N-n_i^{\downarrow}\right]!}.\end{aligned} \tag{S29}$$

For the next term involving k in Eq. (S26) we have

$$\frac{d\left(N-k-\phi_i\right)}{b\left(N-k-\phi_i\right)} = \frac{2N}{(1-c)}\left[\frac{1}{2\left(\frac{c}{1-c}\right)N+\phi_i+k}\right]. \tag{S30}$$

From this we have

$$\begin{aligned}\prod_{k=1}^{n_i^{\downarrow}-1} \frac{d\left(N-k-\phi_i\right)}{b\left(N-k-\phi_i\right)} &= \left[\frac{2N}{(1-c)}\right]^{(n_i^{\downarrow}-1)} \left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_i+n_i^{\downarrow}-1\right)!}\right], \\ &= \left(2\left(\frac{c}{1-c}\right)N+\phi_i+n_i^{\downarrow}\right)\left[\frac{2N}{(1-c)}\right]^{(n_i^{\downarrow}-1)} \left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_i+n_i^{\downarrow}\right)!}\right], \\ &= \left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)\left[\frac{2N}{(1-c)}\right]^{(n_i^{\downarrow}-1)} \left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)!}\right].\end{aligned} \tag{S31}$$

Putting these terms together, we see that Eq. (S9) in the case of facultative sex can be simplified to

$$\begin{aligned}P^{\mathrm{st}}(\boldsymbol{n}) \propto\ & \prod_{i=1}^{M-1}\left\{\left(\frac{m}{M_{\max}-M}\right)\frac{N}{n_i^{\downarrow}}\right\}\left\{\frac{1}{\left[\left(\frac{1+c}{1-c}\right)N\right]}\left(\frac{1-c}{2N}\right)^{(n_i^{\downarrow}-1)}\frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{\left[\left(\frac{1+c}{1-c}\right)N-n_i^{\downarrow}\right]!}\right\}\times \\ & \left\{\frac{N-\phi_i}{c(N-\phi_{i+1})+(1-c)(N-\phi_{i+1})\phi_{i+1}/(2N)}\right\}\left\{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)\left[\frac{2N}{(1-c)}\right]^{(n_i^{\downarrow}-1)}\times\right. \\ & \left.\left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)!}\right]\right\} \\ \propto\ & \left\{\left(\frac{mN}{M_{\max}-M}\right)\frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{\left[\left(\frac{1+c}{1-c}\right)N\right]}\right\}^{(M-1)}\prod_{i=1}^{M-1}\left\{\frac{1}{n_i^{\downarrow}}\right\}\left\{\frac{1}{\left[\left(\frac{1+c}{1-c}\right)N-n_i^{\downarrow}\right]!}\right\}\times \\ & \left\{\left(\frac{2N}{1-c}\right)\frac{(N-\phi_i)}{(N-\phi_{i+1})(2\left(\frac{c}{1-c}\right)N+\phi_{i+1})}\right\}\left\{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)\left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)!}\right]\right\}, \\ \propto\ & \left\{\left(\frac{mN}{M_{\max}-M}\right)\frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{\left[\left(\frac{1+c}{1-c}\right)N\right]}\right\}^{(M-1)}\prod_{i=1}^{M-1}\left\{\frac{1}{n_i^{\downarrow}}\right\}\left\{\frac{1}{\left[\left(\frac{1+c}{1-c}\right)N-n_i^{\downarrow}\right]!}\right\}\times \\ & \left\{\left(\frac{2N}{1-c}\right)\frac{(N-\phi_i)}{(N-\phi_{i+1})}\right\}\left\{\left[\frac{\left(2\left(\frac{c}{1-c}\right)N+\phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N+\phi_{i+1}\right)!}\right]\right\}\end{aligned} \tag{S32}$$

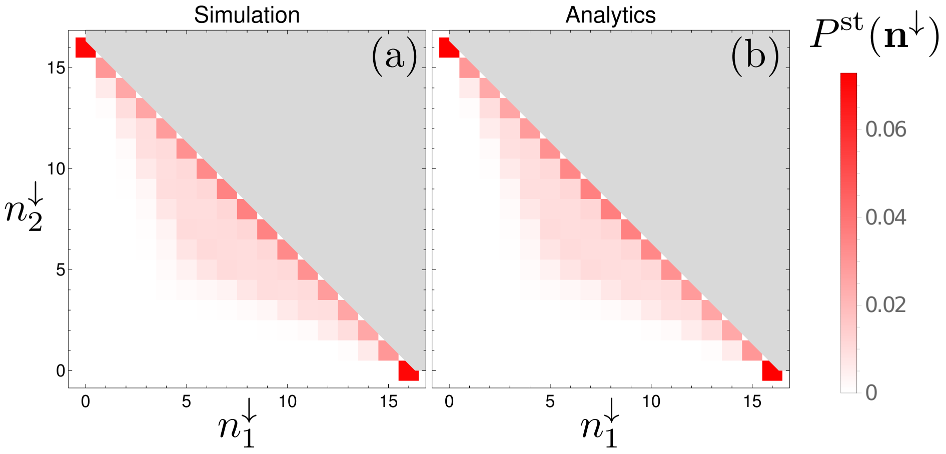

Figure S4: Density plots of the stationary probability distribution of $\boldsymbol{n}^{\downarrow}$) in the $(n_1^{\downarrow}, n_2^{\downarrow})$ plane for the case of facultative sexual reproduction ($c = 0.4$). The distribution is strongly peaked in states containing a single mating type, though states with two mating types are more likely than in the asexual scenario (see Figure S2). Note also that in comparison with Figure S3, having less asexual reproduction has here shifted the typical number of mating types away from three. Analytical results are obtained using Eq. (S35) and can be seen to capture the behavior of simulations extremely well. Low population numbers, $N = 16$, are here utilized to facilitate evaluation of the analytic expression. Remaining parameters are $\sigma = 0$ and $m = 0.01$. The simulation is conducted with $M_0 = 5$ and given 10,000 generations to equilibrate before the distribution is time-averaged over 400,000 generations.

Now we can take the product over i in the remaining relevant terms in Eq. (S32);

$$\prod_{i=1}^{M-1} \frac{N-\phi_i}{N-\phi_{i+1}} = \frac{N}{n_M^{\downarrow}}, \tag{S33}$$

and

$$\begin{aligned} \prod_{i=1}^{M-1} \left[\frac{\left(2\left(\frac{c}{1-c}\right)N + \phi_i\right)!}{\left(2\left(\frac{c}{1-c}\right)N + \phi_{i+1}\right)!} \right] &= \left[\frac{\left(2\left(\frac{c}{1-c}\right)N\right)!}{\left(2\left(\frac{c}{1-c}\right)N + (N - n_m^{\downarrow})\right)!} \right], \\ &= \left[\frac{\left(2\left(\frac{c}{1-c}\right)N\right)!}{\left(\left(\frac{1+c}{1-c}\right)N - n_m^{\downarrow}\right)!} \right], \end{aligned} \tag{S34}$$

Putting all of this back together, we have a final expression for the stationary probability distribution;

$$\begin{aligned} P^{\mathrm{st}}(\boldsymbol{n}) &\propto \left\{ \left(\frac{mN}{M_{\max}-M}\right) \frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{\left[\left(\frac{1+c}{1-c}\right)N\right]} \left(\frac{2N}{1-c}\right) \right\}^{(M-1)} \left[2\left(\frac{c}{1-c}\right)N\right]! \prod_{i=1}^{M} \left\{\frac{1}{n_i^{\downarrow}}\right\} \left\{ \frac{1}{\left[\left(\frac{1+c}{1-c}\right)N - n_i^{\downarrow}\right]!} \right\}, \\ &\propto \left\{ \left(\frac{2mN}{M_{\max}-M}\right) \frac{\left[\left(\frac{1+c}{1-c}\right)N\right]!}{(1+c)} \right\}^{(M-1)} \left[2\left(\frac{c}{1-c}\right)N\right]! \prod_{i=1}^{M} \left\{\frac{1}{n_i^{\downarrow}}\right\} \left\{ \frac{1}{\left[\left(\frac{1+c}{1-c}\right)N - n_i^{\downarrow}\right]!} \right\}. \end{aligned} \tag{S35}$$

The expression can be seen to perfectly capture the results of simulations, as illustrated in Figure S4.

S4 Utilizing the stationary distribution to estimate the number of mating types in the system when $\sigma = 0$

In Section S3, expressions for the stationary distribution $P^{\mathrm{st}}(\boldsymbol{n})$ are obtained. However, the normalization factor required to evaluate $P^{\mathrm{st}}(\boldsymbol{n})$ has not been included. While this factor is clearly calculable, it involves conducting the sum of the expressions for $P^{\mathrm{st}}(\boldsymbol{n})$ over all possible states. Since the number of possible states is very large, this sum becomes computationally expensive for large N. In addition, such a numeric calculation arguably removes the primary advantage of tackling the problem analytically in the first place, that of obtaining intuition as to the system's behavior. In this section therefore, we develop approximations that will gives us insight into how the mode of the distribution of the number of mating types varies with the parameters N, m and c.

S4.1 Estimating the mode of $P^{\mathrm{st}}(\boldsymbol{n})$

In this section we will show that it is relatively straightforward to calculate the mode of $P^{\mathrm{st}}(\boldsymbol{n})$ under relatively few, physically well-motivated assumptions. Let us begin by stating these assumptions:
Assumption 1a. **The mode of $P^{\mathrm{st}}(\boldsymbol{n})$ will contain a set of local maxima that lie in close proximity to the fixed points of the deterministic analogue of the probabilistic dynamics.** Simply put, we assume that the distribution consists of a series of peaks around the fixed points of the deterministic dynamics. This makes intuitive sense. Say we start with a system with M mating types. Over short times we might expect the distribution to be centered around a state with M mating types, each one consisting of roughly N/M individuals. After some time a mating type is lost, such that we now have $M-1$ mating types. We would now expect the system to relax to a state in which each of those $M-1$ mating types consists of approximately $N/(M-1)$ individuals. In time, a new mating type may arise in the population. Assuming that it successfully invades (rather than being lost quickly at low frequency) we would again expect to see individuals equally distributed across the M mating types. This gives rise to a series of peaks in $P^{\mathrm{st}}(\boldsymbol{n})$, each centered around states with M mating types, each of which is approximately at frequency $1/M$, for M between zero and M_{max}.
Assumption 1b. **The arrival rate of new mutations is much less than once per generation**. We note that for the picture in *Assumption 1a.* to hold, another assumption must be true; that the time between new mutations arriving is long enough that upon losing a mating type, the system has time to relax to its new quasi-stationary state (at approximately $N/(M-1)$) before another mutation arises. To enforce this, we must assume that $mN \ll 1$ or equivalently $m_g \ll 1$ (see Section S2).
Assumption 2. **There is in general only one set of local maxima that constitute the global maxima of the system.** For each M, we have assumed that there are a series of local maxima centered on states containing M mating types, each at a frequency of $1/M$. For a given M, each of these local maxima must have an identical height, since by symmetry it does not matter how we label each of the M existing mating types. The number of these local maxima is given by the binomial distribution, $\mathrm{Binomial(M_{max}, M)}$. Let's now sum each of these degenerate peaks together, so that we now have the total probability of being in a state with M mating types, each with frequency N/M, for every M. This series of probabilities should be unimodal; there should be at most a single peak. As an example, if we are more likely to find the system in a state with four mating types (each at a frequency 1/4) than in a state with five mating types (each at a frequency 1/5), then we should not be more likely to find the system in a state with six mating types rather than five.

If the above assumptions hold, then the mode of $P^{\mathrm{st}}(\boldsymbol{n})$ can be efficiently found by comparing the total height of $P^{\mathrm{st}}(\boldsymbol{n})$ at a given local maxima (by *Assumption 1a.*, located only near the deterministic fixed points) with the total height of $P^{\mathrm{st}}(\boldsymbol{n})$ at adjacent maxima (by *Assumption 2.*, if the distribution is unimodal, we need only consider the local gradient of the distribution). Further, if we consider the ratio of these heights, then we do not need to evaluate the normalization of the distribution, significantly simplifying the calculation.

Let us now mathematically formalize the above discussion. Let us define $r_{\mathrm{LB}}(M)$ as the ratio of the probability being in any of $\mathrm{Binomial(M_{max}, M)}$ states with M mating types each at frequency $1/M$ and the probability being in any of $\mathrm{Binomial(M_{max}, M-1)}$ states with $M-1$ mating types each at frequency $1/(M-1)$;

$$r_{\mathrm{LB}}(M) = \frac{\mathrm{Binomial}(M_{\mathrm{max}}, M)}{\mathrm{Binomial}(M_{\mathrm{max}}, M-1)} \frac{P^{(\mathrm{st})}(\boldsymbol{\eta}^{(M)})}{P^{(\mathrm{st})}(\boldsymbol{\eta}^{(M-1)})} \tag{S36}$$

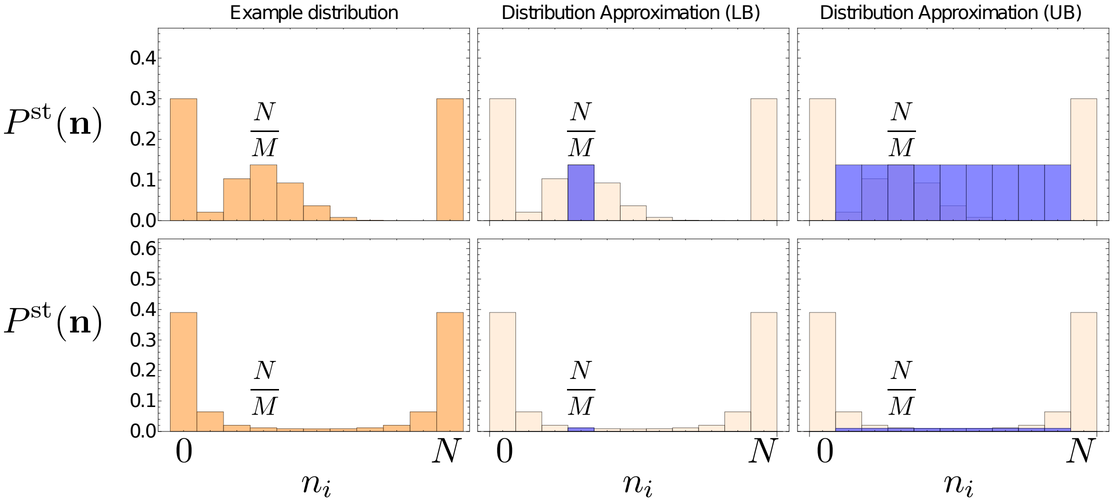

Figure S5: Figure describing the physical logic underlying the approximation which estimates the upper and lower bound of the mode number of mating types in the stationary distribution. Two different scenarios are illustrated in the top row and bottom row; $m_g \ll 1$ (top row) and $m_g < 1$ (bottom row). **Top row**: The peak of the actual distribution (orange) when there are M mating types present in the population (i.e. $N > n_i > 0$) lies near a deterministic fixed point at N/M. We wish to obtain an estimate for the probability of being in any of the states in the range $N > n_i > 0$ (i.e. the probability of there being M mating types). However, we wish to do this in such a way that we do not need to evaluate the probability of being in each state. In the center panel, we assume that the probability of being in this range is entirely concentrated at the state near the deterministic fixed point (blue bar). This underestimates the probability of $N > n_i > 0$ but does provide a lower bound estimate for the probability. In the right panel we assume that each state in this range is equally as probable as the probability of being in the state near the deterministic fixed point (blue bars). This overestimates the probability of $N > n_i > 0$ but does provide an upper bound estimate for the probability. **Bottom row**: The peak of the actual distribution (orange) when there are M mating types present in the population does not lie near the deterministic fixed point at N/M. Instead, the probability is gathered at the boundaries. This situation can arise when the condition $1 \gg mN$ is violated. In this case the system does not have enough time to relax to a deterministic fixed point in the period between an extinction event, leading to the distribution on the range $N > n_i > 0$ being peaked at the edges. Note that in this scenario, the 'upper bound' approximation in the bottom right panel actually underestimates the probability of the system being in the interval $N > n_i > 0$. This leads to a breakdown of our approximations for the mode number of mating types if the condition $mN \ll 1$ is not met.

where

$$\begin{aligned}\boldsymbol{\eta}_i^{(M)} &= \frac{N}{M}, \quad \text{if} \quad i \leq M\,, \\ \boldsymbol{\eta}_i^{(M)} &= 0, \quad \text{otherwise}\,. \end{aligned} \tag{S37}$$

Note that formally the argument of $P^{\rm st}(\boldsymbol{n})$ can only take integer vectors, however there is nothing in our analytic expressions that prevents us from using non-integer values of $\boldsymbol{n}$ to obtain approximations for $P^{\rm st}(\boldsymbol{n})$ in the region of a deterministic fixed point. Let us now consider how $r_{\rm LB}(M)$ varies with M. If it is more likely that we find the system at a fixed point $1/M$ than a fixed point at $1/(M-1)$, then $r_{\rm LB}(M)$ is greater than one. If it is less likely, then $r_{\rm LB}(M)$ is less than one. Therefore, if *Assumption 2.* holds (that is, if the distribution of probabilities of finding the system at each fixed point is unimodal), then finding the mode of the entire distribution of $P^{\rm st}(\boldsymbol{n})$ amounts to cycling through $r_{\rm LB}(M)$ for $M = 2, \dots M_{max}$, and finding the last M for which $r_{\rm LB}(M)$ is positive. This then tells us that the mode of $P^{\rm st}(\boldsymbol{n})$ exists in a state with M mating types.

S4.2 Estimating the mode number of mating types

We first begin by noting that the distribution we calculated is for the probability of being in a state $\boldsymbol{n}$, $P^{\mathrm{st}}(\boldsymbol{n})$. In reality the quantities that we are most interested in are related to the probability that the system has a given number of mating types M, $\mathcal{P}^{\mathrm{st}}(M)$. These distributions are of course related. We can obtain the distribution of the number of mating types as follows;

$$\mathcal{P}^{\mathrm{st}}(M) \propto \sum_{\boldsymbol{n}\in S^M} P^{\mathrm{st}}(\boldsymbol{n}) \,, \tag{S38}$$

where S^M is the set of all vectors $\boldsymbol{n}$ that contain M non-zero elements. Note that the number of vectors in S^M is given by

$$\frac{\mathrm{Binomial}\left(M_{\mathrm{max},M}\right)}{(M-1)!} \frac{\prod_{k=0}^{M-1}(N-k)}{N} \,. \tag{S39}$$

In a similar manner to the issue of calculating the normalization factor, while it is entirely possible to evaluate Eq. (S38) numerically, one can see that it is both computationally expensive and lacking in intuition. Ideally, we would like the calculate some summary statistics for the distribution $\mathcal{P}^{\mathrm{st}}(M)$. We will focus on its mode, however we will see empirically from simulations that for large N the mode and median of $\mathcal{P}^{\mathrm{st}}(M)$ are nearly identical.

Let us begin by considering a quantity that we have shown how to obtain already in Section S4.1; the mode of the distribution $P^{\mathrm{st}}(\boldsymbol{n})$. Recall that we obtained this by comparing the probability of sitting at each of the fixed-points of the problem. Note then that the number of mating types present in the mode state of $P^{\mathrm{st}}(\boldsymbol{n})$ is not the same as the mode of $\mathcal{P}^{\mathrm{st}}(M)$, as this latter distribution involves also accounting for the probability of the system residing in states far from the deterministic fixed points of the problem (see Eq. (S38)). However, if the distribution of $P^{\mathrm{st}}(\boldsymbol{n})$ consisted of a series of delta peaks, each centered on the deterministic fixed points, then the mode of each distribution would converge. While this scenario is clearly implausible in general, it is likely to be close to the real situation when m and c are small and N is very large. Outside this regime, the number of mating types present in the mode state of $P^{\mathrm{st}}(\boldsymbol{n})$ provides an approximate lower-bound for the mode of $\mathcal{P}^{\mathrm{st}}(M)$. Thus we have calculated an approximation for the lower bound of the mode of $\mathcal{P}^{\mathrm{st}}(M)$, the greatest value of M for which $r_{\mathrm{LB}}(M)$ is greater than one (see Eq. (S36)).

We now seek an estimate for the upper bound. While for the lower bound we assumed that the distribution of states around the fixed point in $P^{\mathrm{st}}(\boldsymbol{n})$ was tightly peaked, we will here assume that it is very flat. We assume that each state containing M mating types is as probable as highest probability state containing M types, the state near the fixed point. We pause to emphasize that as with *Assumption 1.*, this approximation for the upper bound will only be accurate when $mN \ll 1$, such that the system has time to relax to a quasi-stationary distribution centered on the deterministic fixed point. These arguments are illustrated graphically in Figure S5. Recalling that the total number of states containing M mating types is given by Eq. (S39), in a similar manner to $r_{\mathrm{LB}}(M)$ (see Eq. (S36)), we define $r_{\mathrm{UB}}(M)$ as

$$r_{UB}(M) = \frac{(M-2)!}{(M-1)!} \frac{\mathrm{Binomial}(M_{\mathrm{max}}, M)}{\mathrm{Binomial}(M_{\mathrm{max}}, M-1)} \frac{\prod_{k=0}^{M-1}(N-k)}{\prod_{k=0}^{M-2}(N-k)} \frac{P^{(\mathrm{st})}(\boldsymbol{\eta}^{(M)})}{P^{(\mathrm{st})}(\boldsymbol{\eta}^{(M-1)})} \,. \tag{S40}$$

In the same way r_{LB} was utilized, r_{UB} can be used to find an approximate upper bound for the mode number of mating types, by cycling through r_{UB} for increasing values of M and identifying the last M for which r_{UB} is greater than one. Noting the similarity in form between r_{UB} and r_{LB} in Eqs. (S40) and (S36), we can in fact map between them in a straightforward way. We first note that

$$\frac{(M-2)!}{(M-1)!} = \frac{1}{M-1} \,, \qquad \text{and} \qquad \frac{\prod_{\mathrm{k=0}}^{\mathrm{M-1}}(\mathrm{N-k})}{\prod_{\mathrm{k=0}}^{\mathrm{M-2}}(\mathrm{N-k})} = \mathrm{N-(M-1)} \,. \tag{S41}$$

From this we find that

$$r_{\mathrm{UB}} = \left(\frac{N}{M-1} - 1\right) r_{\mathrm{LB}} \,. \tag{S42}$$

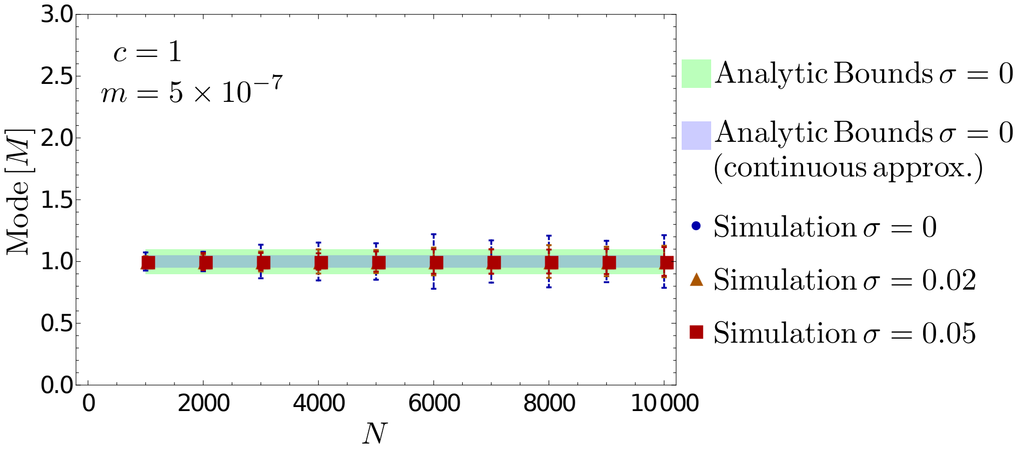

Figure S6: The expected number of mating types under purely asexual reproduction ($c = 1$). The analytic results obtained by solving Eqs. (S44-S45) (filled green bounds) which has an approximate analytic solution given by Eq. (S47) (filled blue bounds) can be seen to compare well with the results of simulations. Note that as all of our analytic results are dependent on $mN \ll 1$, for fixed m we expect the above results to breakdown as N is increased. The simulations are run with $M_0 = 3$ and given 10^4 generations to equilibrate before the distribution is time-averaged over 10^6 generations, sampled every 200 generations. Error bars give the standard deviation of simulation samples.

We now proceed to calculate r_{UB} and r_{LB} for each of the sexual scenarios (obligately sexual, facultatively sexual, asexual). However before proceeding we note that each of the stationary distributions for $\boldsymbol{n}$ contain a factor $[1/(M_{\max} - M)]^{M-1}$ (see Eqs. (S19), (S25) and (S35)), while both the functions r_{UB} and r_{LB} contain a factor Binomial$(M_{\max}, M)$/Binomial$(M_{\max}, M-1)$. Pulling these terms out of the expressions for $P^{(\text{st})}(\boldsymbol{n})$, we find that all terms in r_{UB} and r_{LB} that involve $M_{\max}$ disappear when we take the limit $M_{\max}$ to infinity;

$$\frac{\text{Binomial}(M_{\max}, M)}{\text{Binomial}(M_{\max}, M-1)} \frac{[M_{\max} - (M-1)]^{M-2}}{[M_{\max} - M]^{M-1}} \Bigg| \lim_{M_{\max} \to \infty} = \frac{1}{M} \, . \tag{S43}$$

In this way we can take the appropriate limit of an infinite number of theoretically possible mating types, while obtaining a finite, normalizable distribution, without having to actually calculate the normalization factor.

S4.2.1 Asexual model ($c = 1$)

We begin by evaluating and simplifying r_{LB} from Eq. (S40) for $c = 1$ using Eq. (S19). Taking the limit of $M_{\max} \to \infty$, using Eq. (S43) we find

$$\begin{aligned} r_{\text{LB}}(M) &= \frac{mN}{M} \frac{\left(\frac{N}{M-1}\right)^{M-1}}{\left(\frac{N}{M}\right)^{M}} \, , \\ &= m \left(\frac{M}{M-1}\right)^{M-1} \, . \end{aligned} \tag{S44}$$

The function r_{UB} therefore can be straightforwardly obtained from Eq. (S42);

$$r_{\text{UB}}(M) = m \left(\frac{N}{M-1} - 1\right) \left(\frac{M}{M-1}\right)^{M-1} \, . \tag{S45}$$

Recall that, as discussed in Section S4.1 and Section S4.2, we estimate upper and lower bounds for the mode of $\mathcal{P}^{\text{st}}(M)$ to coincide with the last integer value of M for which $r_{\text{LB}}(M) \geq 1$ and $r_{\text{UB}}(M) \geq 1$. The function

$r_{\mathrm{LB}}(M)$ is independent of N. However, in the limit $M \to 2$, $r_{\mathrm{LB}}(M) = 2m$. Therefore, if $m \leq 0.5$, which is very biologically reasonable, the lower bound for the most common number of mating types in $\mathcal{P}^{\mathrm{st}}(M)$ is one. The function for $r_{\mathrm{UB}}(M)$ does features N dependence. However, when N is large we find

$$r_{\mathrm{UB}}(M) \approx mN \frac{M^{M-1}}{(M-1)^M} \, . \tag{S46}$$

This function equals $2mN$ in the limit $M \to 2$. Therefore, if $mN \leq 0.5$, then the upper bound for the most common number of mating types in $\mathcal{P}^{\mathrm{st}}(M)$ is also one. Since we have already assumed $mN \ll 1$ (or equivalently $m_g \ll 1$) in our derivation of r_{LB} and r_{UB}, we can then say

$$\mathrm{Mode}\left[\mathcal{P}^{\mathrm{st}}(\mathrm{M})\right] \approx 1 \tag{S47}$$

in an entirely asexual population. The theory agrees well with simulations, as illustrated in Figure S6. Essentially this says that if mutations that yield a new mating type are arriving at a rate much less than once per generation, then genetic drift purges mutations (which are neutral in the asexual $c = 1$ limit) at a faster rate than they arrive.

S4.2.2 Obligately sexual model ($c = 0$)

Let us begin by evaluating and simplifying r_{LB} from Eq. (S40) for $c = 0$ using Eq. (S25). Taking the limit of $M_{\max} \to \infty$, using Eq. (S43) we find

$$\begin{aligned} r_{\mathrm{LB}}(M) &= \frac{2mNN!}{M} \frac{\left[\frac{N}{M-1}\left(N - \frac{N}{M-1}\right)!\right]^{M-1}}{\left[\frac{N}{M}\left(N - \frac{N}{M}\right)!\right]^{M}} \, , \\ &= 2mN! \left(\frac{M}{M-1}\right)^{(M-1)} \left\{\left[N\left(\frac{M-2}{M-1}\right)\right]!\right\}^{(M-1)} \left\{\left[N\left(\frac{M}{M-1}\right)\right]!\right\}^{-M} \, . \end{aligned} \tag{S48}$$

The function r_{UB} therefore can be straightforwardly obtained from Eq. (S42);

$$r_{\mathrm{UB}}(M) = 2mN! \left(\frac{N}{M-1} - 1\right) \left(\frac{M}{M-1}\right)^{(M-1)} \left\{\left[N\left(\frac{M-2}{M-1}\right)\right]!\right\}^{(M-1)} \left\{\left[N\left(\frac{M}{M-1}\right)\right]!\right\}^{-M} \, . \tag{S49}$$

As we have discussed in Section S4.1 and Section S4.2, identifying the range over which we expect to find the mode number of mating types is now simply a matter of finding when the functions $r_{\mathrm{LB}}(M)$ and $r_{\mathrm{UB}}(M)$ cross one. This is straightforward (and computationally inexpensive) to achieve numerically, and yields accurate results (see Figure S7, green bounds). We stress however that for the upper bound to be a good approximation, we require mN to be small. In practical terms, $mN \approx 0.01$ (one mutation every hundred generations) is sufficient for the approximation to work. As mN is lowered further, the mode and mean of the simulation move towards the curve predicted by the lower bound.

Although numerically obtaining the analytic approximation for the mode M in the system is now possible with Eqs. (S48-S49), we can progress toward an analytic solution with further approximations. The first is Stirling's approximation for large factorials. This states that if the argument of a factorial k is large, then

$$k! \approx \sqrt{2\pi k} \left(\frac{k}{\mathrm{e}}\right)^{k} \, . \tag{S50}$$

Applying this to the terms in Eq. (S48), we find

$$N! \approx \sqrt{2\pi N} \left(\frac{N}{\mathrm{e}}\right)^{N} \, , \tag{S51}$$

$$\left\{\left[N\left(\frac{M-1}{M-2}\right)\right]!\right\}^{(M-1)} \approx \left\{\sqrt{2\pi\left(\frac{M-1}{M-2}\right)N} \left[\left(\frac{M-1}{M-2}\right)\frac{N}{\mathrm{e}}\right]^{(M-1)N/(M-2)}\right\}^{(M-1)} \, . \tag{S52}$$

$$\left\{\left[N\left(\frac{M}{M-1}\right)\right]!\right\}^{-M} \approx \left\{\sqrt{2\pi\left(\frac{M}{M-1}\right)N} \left[\left(\frac{M}{M-1}\right)\frac{N}{\mathrm{e}}\right]^{MN/(M-1)}\right\}^{-M} \, . \tag{S53}$$

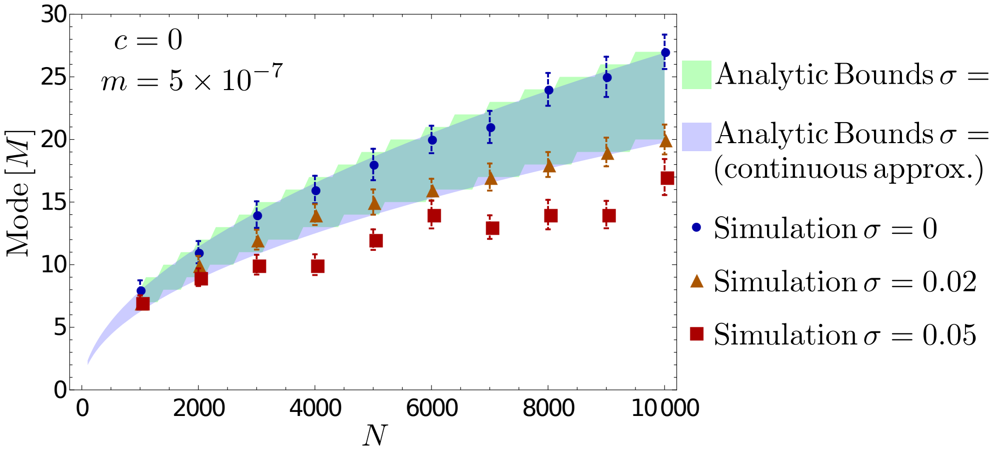

Figure S7: The expected number of mating types under obligately sexual reproduction ($c = 0$). The analytic results obtained by numerically solving Eqs. (S48-S49) (filled green bounds) which has an approximate analytic solution given by Eq. (S58) (filled blue bounds) can be seen to compare well with the results of neutral simulations ($\sigma = 0$). Note that as all of our analytic results are dependent on $mN \ll 1$, for fixed m we expect the above results to breakdown as N is increased. The simulations are run with $M_0 = 30$ and given 10^4 generations to equilibrate before the distribution is time-averaged over 10^6 generations, sampled every 200 generations. Error bars give the standard deviation of simulation samples.

Substituting these expressions into Eq. (S48) and simplifying, we obtain

$$r_{\mathrm{LB}}(M) \approx 2mM^{-1+\frac{3}{2}M+(M-1)N}(M-1)^{\frac{3}{2}-2M+(3-2M)N}(M-2)^{\frac{1}{2}(M-1)+(M-2)N}\,. \tag{S54}$$

Obtaining a similar expression for the upper bound condition requires taking account of the form of r_{UB} (see Eq. (S42)). We find

$$r_{\mathrm{UB}}(M) \approx 2m\left(\frac{N}{M-1}-1\right)M^{-1+\frac{3}{2}M+(M-1)N}(M-1)^{\frac{3}{2}-2M+(3-2M)N}(M-2)^{\frac{1}{2}(M-1)+(M-2)N}\,. \tag{S55}$$

These expression are valid when N is large.

Simplifying these terms further is not entirely straightforward. We are aiming to calculate the value of M that satisfies $r_{\mathrm{LB}}(M) = 1$ and $r_{\mathrm{UB}}(M) = 1$. Since this M value grows with N, we cannot simply take the limit of large N or M independently. However, empirically we see that the M that satisfies these equations is approximately proportional to $\sqrt{N}$. We utilize this empirical observation, and expand $r_{\mathrm{LB}}(M)$ and $r_{\mathrm{UB}}(M)$ in a series in $1/N$, assuming $M \approx \mathcal{O}(\sqrt{N})$. We obtain;

$$r_{\mathrm{LB}}(M) \approx 2m\exp\left(1+\frac{N}{2M^2}\right)+\mathcal{O}\left(N^{-1}\right)\,, \tag{S56}$$

$$r_{\mathrm{UB}}(M) \approx 2m\frac{N}{M}\exp\left(1+\frac{N}{2M^2}\right)+\mathcal{O}\left(N^{-1}\right)\,. \tag{S57}$$

Solving the equations $r_{\mathrm{LB}}(M) = 1$ and $r_{\mathrm{UB}}(M) = 1$ for M using the above approximations then allows us to approximately bound the mode of $\mathcal{P}^{\mathrm{st}}(M)$;

$$\sqrt{\frac{N}{\mathrm{W}\left[\frac{1}{4\mathrm{e}^2\mathrm{m}^2\mathrm{N}}\right]}} > \mathrm{Mode}\left[\mathcal{P}^{\mathrm{st}}(\mathrm{M})\right] > \mathrm{i}\sqrt{\frac{\mathrm{N}}{2\left\{1+\log\left[2\mathrm{m}\right]\right\}}}\,, \tag{S58}$$

where i is an imaginary number and $W[z]$ is the Lambert W function [1]. This equation can now be used to obtain an estimate for the most frequently occurring number of mating types in an obligately sexual system. The approximations can be seen to compare well against both simulations and numerical evaluation of the points at which the functions $r_{\mathrm{LB}}(M)$ and $r_{\mathrm{UB}}(M)$ cross one, as illustrated in Figure S7.

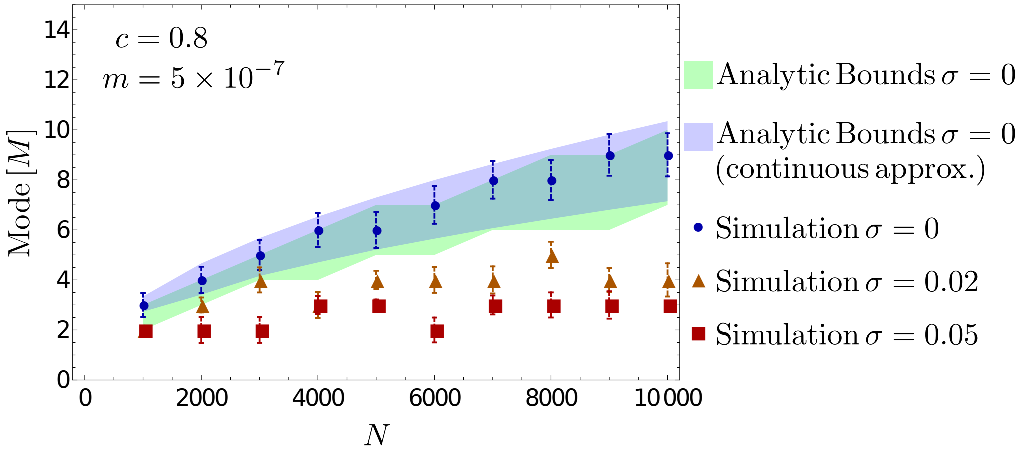

Figure S8: The expected number of mating types under facultatively sexual reproduction ($1 > c > 0$). The analytic results obtained by solving Eqs. (S59-S60) (filled green bounds) and solving the approximate Eq. (S61) (filled blue bounds) can be seen to compare well with the results of neutral simulations ($\sigma = 0$). Note that as all of our analytic results are dependent on $mN \ll 1$, for fixed m we expect the above results to breakdown as N is increased. The simulations are run with $M_0 = 10$ and given 10^4 generations to equilibrate before the distribution is time-averaged over 10^6 generations, sampled every 200 generations. Error bars give the standard deviation of simulation samples.

S4.2.3 Facultatively sexual model ($1 > c > 0$)

As in Sections S4.2.1 and S4.2.2, we begin by simplifying r_{LB} from Eq. (S40) for general c using Eq. (S35). Taking the limit of $M_{\max} \to \infty$, using Eq. (S43) we find;

$$\begin{aligned} r_{\mathrm{LB}}(M) &= \frac{2mN}{M} \frac{\left[\frac{1+c}{1-c}N\right]!}{1+c} \frac{\left[\frac{N}{M-1}\left(\frac{1+c}{1-c}N - \frac{N}{M-1}\right)!\right]^{M-1}}{\left[\frac{N}{M}\left(\frac{1+c}{1-c}N - \frac{N}{M}\right)!\right]^{M}}, \\ &= 2m\frac{\left[\frac{1+c}{1-c}N\right]!}{1+c}\left(\frac{M}{M-1}\right)^{M-1} \frac{\left[\left(\frac{1+c}{1-c}N - \frac{N}{M-1}\right)!\right]^{M-1}}{\left[\left(\frac{1+c}{1-c}N - \frac{N}{M}\right)!\right]^{M}}. \end{aligned} \tag{S59}$$

Again, the function r_{UB} therefore can be straightforwardly obtained from Eq. (S42);

$$r_{\mathrm{UB}}(M) = 2m\frac{\left[\frac{1+c}{1-c}N\right]!}{1+c}\left(\frac{N}{M-1} - 1\right)\left(\frac{M}{M-1}\right)^{M-1} \frac{\left[\left(\frac{1+c}{1-c}N - \frac{N}{M-1}\right)!\right]^{M-1}}{\left[\left(\frac{1+c}{1-c}N - \frac{N}{M}\right)!\right]^{M}}. \tag{S60}$$

Applying Stirling's approximation (see Eq. (S50)) allows us to simplify the expressions a little further when the terms in the factorials are large;

$$\begin{aligned} r_{\mathrm{LB}}(M) &\approx 2m(1+c)^{-\frac{1}{2}+\frac{c+1}{1-c}N} M^{-1+\frac{3}{2}M+\frac{1+c}{1-c}MN-N}(M-1)^{-\frac{3}{2}(M-1)+\frac{2-cM-M}{1-c}N} \times \\ &\quad (cM+M-2)^{\frac{1}{2}(M-1)-\frac{2-cM-M}{1-c}N}(cM+c+M-1)^{-\frac{M}{2}-\frac{1+c}{1-c}MN+N}, \\ r_{\mathrm{UB}}(M) &\approx \left(\frac{N}{M-1} - 1\right) r_{\mathrm{LB}}(M). \end{aligned} \tag{S61}$$

While we cannot simplify these expressions any further without making additional assumptions on the value of c, upper and lower bounds for Mode $[\mathcal{P}^{\mathrm{st}}(\mathrm{M})]$ can now be efficiently by found using root-finding algorithms to identify the intersection of $r_{\mathrm{LB}}(M)$ and $r_{\mathrm{UB}}(M)$ with one. These methods are found to well-predict the values obtained from simulations (see Figure S8). In the main text, we rewrite the function $r_{\mathrm{LB}}(M)$ having applied the Stirling approximation as

$$\begin{aligned} R(M) \quad &= \quad 2m(1+c)^{-\frac{1}{2}+\frac{c+1}{1-c}N} M^{-1+\frac{3}{2}M+\frac{1+c}{1-c}MN-N} (M-1)^{-\frac{3}{2}(M-1)+\frac{2-cM-M}{1-c}N} \times \\ &\quad\;\; (cM+M-2)^{\frac{1}{2}(M-1)-\frac{2-cM-M}{1-c}N} (cM+c+M-1)^{-\frac{M}{2}-\frac{1+c}{1-c}MN+N} . \end{aligned} \tag{S62}$$

S5 Simulation results on the expected number of mating types when $\sigma > 0$

When $\sigma > 0$ we can not obtain analytic estimates on the expected number of mating types in the population. Instead we use simulations to obtain a more qualatative description of the dynamics in this regime. In general, it can be seen that increasing variance in the fitness of mating type alleles leads to fewer mating types (see Figures S7 and S8). As the rate of sex is decreased (increasing c) this effect becomes larger (compare Figure S7 with S8); this is because as the rate of sex is reduced, so too is selection for even sex ratios, leading to more frequent extinction of mating types (see also Figure S1).

Obtaining estimates on the number of mating types expected in large populations with very low mutation rates is hampered by the prohibitively large simulation times required to obtain a good estimate of $\mathcal{P}^{\mathrm{st}}(M)$. Indeed, the strength of the anayltic results in the neutral $\sigma = 0$ limit is that such simulations can be circumvented. However we can still demonstrate that when sex is rare only weak selective differences between the mating types are sufficient to reduce their expected number. In Figures S9-S11 we see that less than a two per cent fitness variance ($\sigma = 0.02$) is required to reduce the expected number of mating types below that which is predicted neutrally. In particular, as the rate of asexual reproduction increases (even as the population size increases) the fitness variance required to reduce the number of mating types decreases; this can be seen sequentially through Figures S9-S11. Figure S11, in which $N = 10^6$ and sex occurs a thousand times less than asex, shows an especially startling result; while a fitness variance of just $\sigma = 10^{-4}$ (0.01%) is required to reduce the expected number of mating types to two (left panels), a fitness variance of merely $\sigma = 5 \times 10^{-3}$ (0.5%) is capable of reducing the expected number of mating types to one. Together these results show that when sex is rare, even in large populations, only minor fitness differences are required to reduce the number of expected mating types.

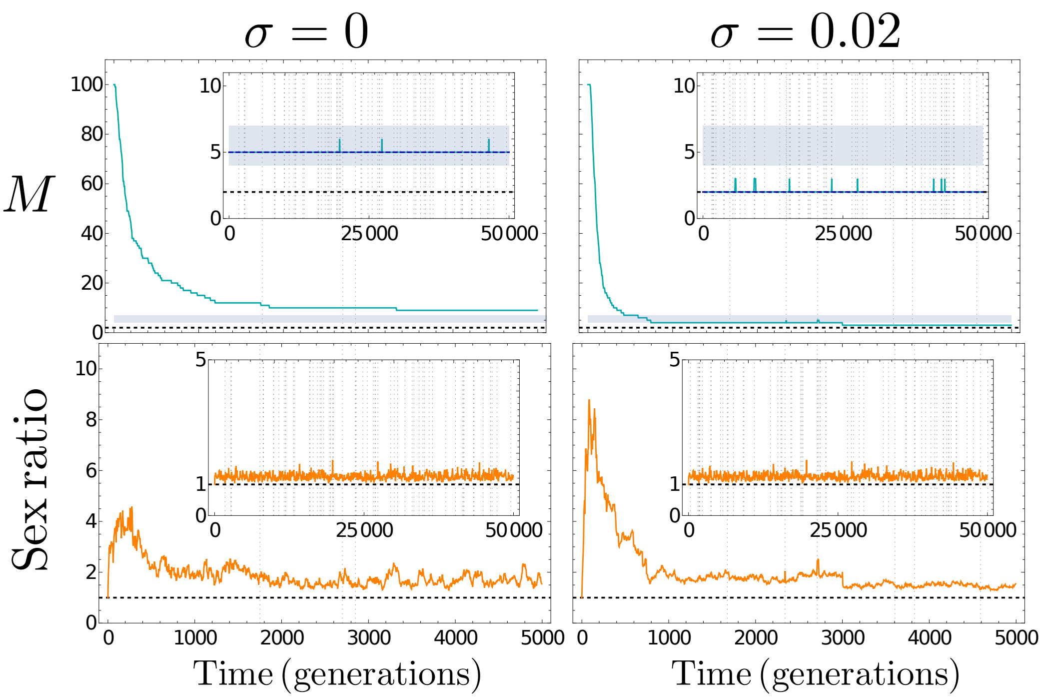

Figure S9: Individual stochastic trajectories for M (top panels) and scaled sex ratio of the most frequent type $M \times \mathrm{Max}[\boldsymbol{n}/N]$ (bottom panels) in a population with $N = 10^4$, $m_g = 0.01$ and $c = 0.9$. **Top panels:** When $\sigma = 0$ (left panel) the number of mating types in the population decreases when M_0 is large and rests for a long time in the region predicted by the neutral theory described analytically in Section S4.2.3 (blue shaded band, inset). When $\sigma > 0$ (right panel), the number of mating types decreases more quickly and rests for a long time in a state in which there are only two mating types (dashed line), below the neutral prediction (blue shaded band, see inset). This is in spite of novel mating types being regularly introduced to the population (vertical dashed lines). **Bottom panels:** When the number of mating type is large, demographic fluctuations cause large deviations in the sex ratio ($M \times \mathrm{Max}[\boldsymbol{n}/N] \approx 1$ under even sex ratios, see horizontal black dashed line). These deviations are more prominent when there are selective differences between the mating type alleles (right panel) but decrease with increasing M.

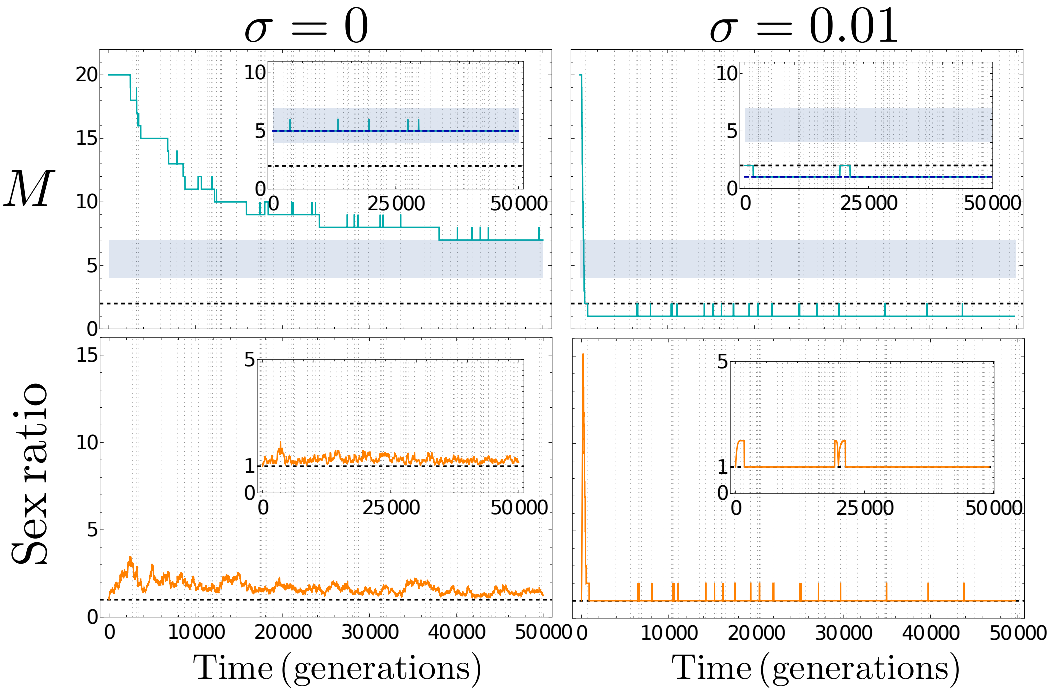

Figure S10: Individual stochastic trajectories for M (top panels) and scaled sex ratio of the most frequent type $M \times \mathrm{Max}[\boldsymbol{n}/N]$ (bottom panels) in a population with $N = 10^5$, $m_g = 0.01$ and $c = 0.99$. **Top panels:** As in Figure S9, the long-time neutral dynamics (left panel) are well predicted by the neutral theory (blue shaded band), while selective differences between the mating types reduces the expected number (right panel). Notice however since sex is rarer here, smaller fitness differences between the mating types have led to a larger decrease in the expected number of types at long times. **Bottom panels:** When the number of mating type is larger, as in the left panels, demographic fluctuations drive deviations away from an even sex ratio (black dashed line). When only one mating type is present (see right panels), deviations are only present during the invasion of a new type.

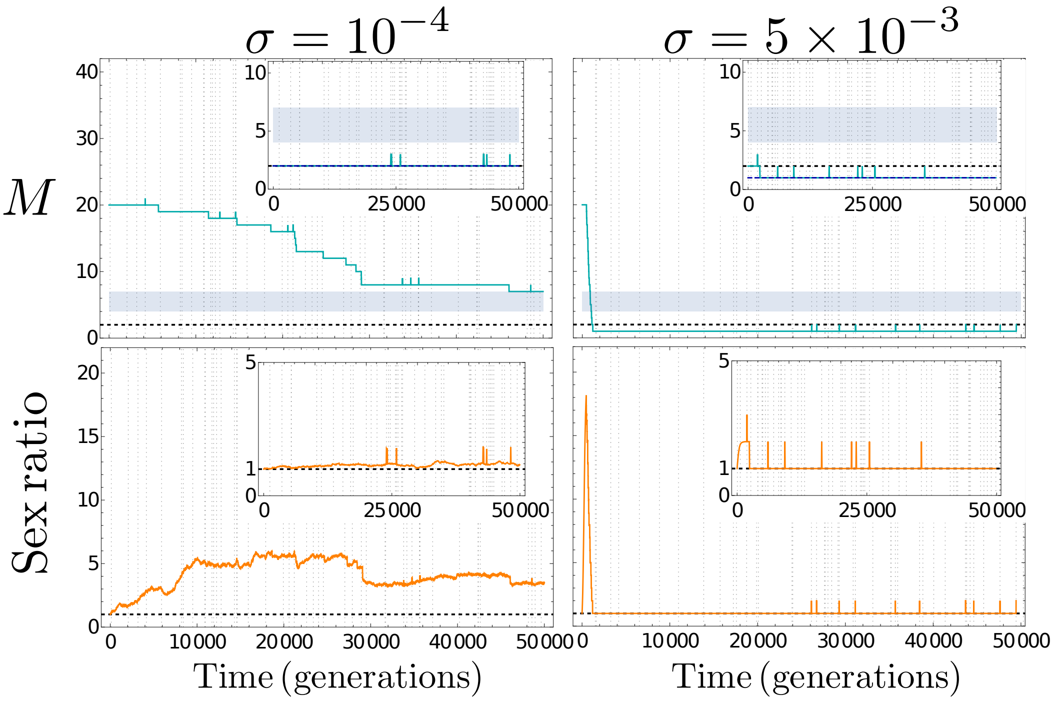

Figure S11: Individual stochastic trajectories for M (top panels) and scaled sex ratio of the most frequent type $M \times \text{Max}[\boldsymbol{n}/N]$ (bottom panels) in a population with $N = 10^6$, $m_g = 0.01$ and $c = 0.999$. **Top panels:** When sex is rare even very small fitness differences between the mating types reduce their expected number compared to the neutral prediction (blue shaded band). From the insets plots we see that the system remains for long periods at $M = 2$ when $\sigma = 10^{-4}$ (left panels), while increasing σ to just 5×10^{-3} reduces the number of mating types to one (right panels). **Bottom panels:** Sex ratios deviating strongly from one (horizontal black dashed line) are indicitive of the system approaching an extinction boundary (see left panel, main plot), while sex ratios remaining close to one (left panel, inset) indicate an even sex ratio is maintained, far from the extinction boundaries.

References

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions.* Dover Publications, New York, 1965.

[2] J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory.* The Blackburn Press, New Jersey, 1970.

[3] J. Felsenstein. Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics*, 68:581–597, 1971.

[4] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[5] A. J. McKane, T. Biancalani, and T. Rogers. Stochastic pattern formation and spontaneous polarisation: the linear noise approximation and beyond. *Bull. Math. Biol.*, 76:895–921, 2014.

[6] M. A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life.* Harvard University Press, Cambridge, Massachusetts, 2006.

[7] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier, Amsterdam, 2007.