



HAL
open science

Effects of complex life cycles on genetic diversity: cyclical parthenogenesis

Romuald Rouger, Katja Reichel, Florent Malrieu, Jean-Pierre Masson, Solenn
Stoeckel

► **To cite this version:**

Romuald Rouger, Katja Reichel, Florent Malrieu, Jean-Pierre Masson, Solenn Stoeckel. Effects of complex life cycles on genetic diversity: cyclical parthenogenesis. *Heredity*, 2016, 117 (5), pp.336-347. 10.1038/hdy.2016.52 . hal-01462754

HAL Id: hal-01462754

<https://hal.science/hal-01462754v1>

Submitted on 4 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ORIGINAL ARTICLE

Effects of complex life cycles on genetic diversity: cyclical parthenogenesis

R Rouger¹, K Reichel¹, F Malrieu², JP Masson¹ and S Stoeckel¹

Neutral patterns of population genetic diversity in species with complex life cycles are difficult to anticipate. Cyclical parthenogenesis (CP), in which organisms undergo several rounds of clonal reproduction followed by a sexual event, is one such life cycle. Many species, including crop pests (aphids), human parasites (trematodes) or models used in evolutionary science (*Daphnia*), are cyclical parthenogens. It is therefore crucial to understand the impact of such a life cycle on neutral genetic diversity. In this paper, we describe distributions of genetic diversity under conditions of CP with various clonal phase lengths. Using a Markov chain model of CP for a single locus and individual-based simulations for two loci, our analysis first demonstrates that strong departures from full sexuality are observed after only a few generations of clonality. The convergence towards predictions made under conditions of full clonality during the clonal phase depends on the balance between mutations and genetic drift. Second, the sexual event of CP usually resets the genetic diversity at a single locus towards predictions made under full sexuality. However, this single recombination event is insufficient to reshuffle gametic phases towards full-sexuality predictions. Finally, for similar levels of clonality, CP and acyclic partial clonality (wherein a fixed proportion of individuals are clonally produced within each generation) differentially affect the distribution of genetic diversity. Overall, this work provides solid predictions of neutral genetic diversity that may serve as a null model in detecting the action of common evolutionary or demographic processes in cyclical parthenogens (for example, selection or bottlenecks).

Heredity (2016) **117**, 336–347; doi:10.1038/hdy.2016.52; published online 20 July 2016

INTRODUCTION

For decades, genetic models of idealised populations have described how genetic diversity would be distributed in natural populations according to a given range of assumptions (for example, the Hardy–Weinberg principle or the Wright–Fisher model). When comparing models with empirical data, such an approach has the great advantage of reducing the complexity of a biological system to a much more convenient approximation adhering to a known and fixed range of parameters (Hamilton, 2009). Predictions made under neutral conditions, that is, considering the evolution of molecular data under conditions of only mutation and genetic drift, are essential in evolutionary studies. These predictions provide a strong frame of reference for interpreting deviations due to other evolutionary processes such as natural selection (Kimura, 1983). If deviations are observed, models can be extended later by relaxing one or more assumptions until the models accurately describe the population genetics of the organism under scrutiny (for example, demographic history or mating pattern).

Strict sexuality is a very common feature of idealised model populations; however, clonality is also a widespread mode of reproduction across all types of organisms (De Meeùs *et al.*, 2007). Numerous field observations of partially clonal organisms have shown strong deviations from the predictions of intrapopulation genetic diversity resulting from strictly sexual models (Papura *et al.*, 2003;

Stoeckel *et al.*, 2006; Kanbe and Akimoto, 2009; Allen and Lynch, 2012). The main observed effects of clonal reproduction are (1) decreased genotypic diversity within populations, (2) an excess of heterozygotes resulting in strongly negative F_{IS} values and (3) an increased gametic phase disequilibrium due to the absence of recombination between loci (Halkett *et al.*, 2005b).

Different mathematical models and simulations have relaxed the assumption of strict sexuality to better understand the effects of clonality on the parameters of neutral genetic diversity (Marshall and Weir, 1979; Balloux *et al.*, 2003; de Meeùs and Balloux, 2004; Prugnolle *et al.*, 2005b; De Meeùs and Balloux, 2005). In addition to giving a precise description of the case of full clonality, these models agree that moderate to intermediate levels of clonality (that is, between 0 and 90% clonally produced offspring) have limited effects on the estimators of genetic diversity and structure (De Meeùs *et al.*, 2006). For example, mean F_{IS} values in moderately clonal populations are nearly the same as those under panmixia (Balloux *et al.*, 2003). Similarly, individual-based simulations to examine multilocus indices of genetic diversity have shown that intermediate levels of clonality have limited effects on \bar{r}_d , a measure of gametic phase disequilibrium (De Meeùs and Balloux, 2004; De Meeùs *et al.*, 2006; Navascués *et al.*, 2010). Furthermore, coalescent models that include partial clonality have demonstrated that genealogies differ from those with full sexuality when the rate of sex is on the order of $1/N$, where N is

¹INRA, UMR1349 Institute for Genetics, Environment and Plant Protection, Le Rheu, France and ²Laboratoire de Mathématiques et Physique Théorique (UMR CNRS 7350), Fédération Denis Poisson (FR CNRS 2964), Université François Rabelais, Tours, France

Correspondence: Dr S Stoeckel, INRA, UMR1349 Institute for Genetics, Environment and Plant Protection, Domaine de la Motte au Vicomte, BP 35327, 35653 Le Rheu Cedex, France.

E-mail: solenn.stoeckel@rennes.inra.fr

Received 29 September 2015; revised 2 June 2016; accepted 8 June 2016; published online 20 July 2016

the population size (Bengtsson, 2003; Ceplitis, 2003; Hartfield *et al.*, 2015). However, the effects of higher levels of sex in a coalescent context remain unclear.

Recently, Stoeckel and Masson (2014) proposed a forward-in-time stochastic model that exactly describes the full probabilistic distribution of genetic diversity in partially clonal organisms. This approach confirmed the existence of a small, nearly indistinguishable, difference in mean F_{IS} values between populations with moderate levels of clonality and those with strict sexuality. In addition, this new model highlighted the notable effects of moderate levels of clonality on both the probability of a positive F_{IS} and the dynamics of F_{IS} over generations. Our study builds on these previous results and extends them to address a particular form of partial clonality.

Most population genetics models of partial clonality have been designed to fit populations in which clonal and sexual reproduction co-occur in time. This type of life cycle, hereafter referred to as acyclic partial clonality (APC), is common in many plants. However, many species, especially in the animal kingdom, use a variation of this life cycle that is conventionally called cyclical parthenogenesis (CP). In CP, clonal and sexual reproductions alternate in time with one to many generations of clonality followed by one generation of sexual reproduction. Analysing how CP affects intrapopulation genetic diversity is particularly relevant given that crop pests (for example, aphids), human parasites (for example, trematodes) and classical biological models used in evolutionary science (for example, *Daphnia*) are among the organisms that possess this life cycle. To date, models studying CP have primarily focussed on detecting the intrapopulation genetic effects of both migration and variance on reproductive success (Prugnolle *et al.*, 2005a, b) or on the amount of genetic differentiation between populations of cyclical parthenogens (Berg and Lascoux, 2000). Simulations have shown that a parthenogenetic phase reduces clonal diversity within populations and can result in negative F_{IS} values when the number of parthenogenetic generations is sufficiently high (Pålson, 2001; Vanoverbeke and De Meester, 2010).

In this study, we investigated the effects of various levels of clonality on neutral genetic variation under CP. First, we describe how genetic diversity is affected by the number of clonal generations in a cycle. Second, we quantify how genetic diversity varies seasonally in CP, with the sexual event hypothesised to reset the entire population towards predictions made under the condition of full sexuality. Third, because APC model outputs are often used to discuss results obtained for cyclical parthenogens, we compare genetic diversity under APC and CP with similar levels of clonality. Genetic diversity is modelled at one locus using an adaptation of the stochastic model of Stoeckel and Masson (2014) and simulated at two unlinked neutral loci using classical individual-based simulations.

MATERIALS AND METHODS

Single-locus model

The model used in this study is an adaptation of the mathematical model developed by Stoeckel and Masson (2014). This model describes a single biallelic locus (A and a) in a population of N diploid individuals for which genotypic frequencies rather than allelic frequencies are computed. The numbers of individuals for each possible genotype are $r_{aa} \in \mathbb{N}$, $r_{Aa} \in \mathbb{N}$, and $r_{AA} \in \mathbb{N}$. Their frequencies at time t are thus $p_{ij}^t = \frac{r_{ij}^t}{N}$, where i and j are the alleles A or a . The genotypic composition of the population is given by its genotypic state (r_{aa}, r_{Aa}, r_{AA}) . At each time step, genotypic frequencies are only modified by the action of a reciprocal mutation between the two alleles which occurs at rate μ ($1 - \mu$ is the rate at which an allele remains the same from one generation to the next, genetic drift and reproduction). In APC, a proportion c of individuals in the population are clonally produced at each time step. In CP,

each time step either represents an event of strict clonal reproduction ($c = 1$) or an event of strict sexual reproduction ($c = 0$). The number of successive clonal generations in a cycle is n_{clonal} and the number of sexual generations, n_{sex} , is set to 1 as observed in most biological systems undergoing CP. Panmixia is assumed during sexuality.

The genotypic frequencies at $t+1$ as functions of the genotypic frequencies at t under clonal (p_{ij}^{t+1}) and sexual (q_{ij}^{t+1}) reproduction are respectively given by Equations (1) and (2) of Stoeckel and Masson (2014):

$$\begin{cases} p_{aa}^{t+1} = (1 - \mu)^2 p_{aa}^t + \mu(1 - \mu)p_{Aa}^t + \mu^2 p_{AA}^t \\ p_{Aa}^{t+1} = 2\mu(1 - \mu)p_{aa}^t + [\mu^2 + (1 - \mu)^2]p_{Aa}^t + 2\mu(1 - \mu)p_{AA}^t \\ p_{AA}^{t+1} = \mu^2 p_{aa}^t + \mu(1 - \mu)p_{Aa}^t + (1 - \mu)^2 p_{AA}^t \end{cases} \quad (1)$$

$$\begin{cases} q_{aa}^{t+1} = [(1 - \mu)p_{aa}^t + \frac{1}{2}p_{Aa}^t + \mu p_{AA}^t]^2 \\ q_{Aa}^{t+1} = 2[(1 - \mu)p_{aa}^t + \frac{1}{2}p_{Aa}^t + \mu p_{AA}^t][\mu p_{aa}^t + \frac{1}{2}p_{Aa}^t + (1 - \mu)p_{AA}^t] \\ q_{AA}^{t+1} = [\mu p_{aa}^t + \frac{1}{2}p_{Aa}^t + (1 - \mu)p_{AA}^t]^2 \end{cases} \quad (2)$$

In APC, the overall genotypic frequencies at $t+1$ (π_{ij}^{t+1}) are functions of p_{ij}^{t+1} and q_{ij}^{t+1} , given the clonality rate c as follows:

$$\pi_{ij}^{t+1} = c p_{ij}^{t+1} + (1 - c) q_{ij}^{t+1} \quad (3)$$

The transition probability from any current genotypic state (r_{aa}, r_{Aa}, r_{AA}) to each subsequent state (s_{aa}, s_{Aa}, s_{AA}) is calculated using the multinomial expression for every possible two-state combination:

$$p[(s_{aa}, s_{Aa}, s_{AA}) | (r_{aa}, r_{Aa}, r_{AA})] = \frac{N!}{s_{aa}! s_{Aa}! s_{AA}!} (\pi_{aa}^{t+1})^{s_{aa}} (\pi_{Aa}^{t+1})^{s_{Aa}} (\pi_{AA}^{t+1})^{s_{AA}} \quad (4)$$

Knowing the probability distribution of all states at a given time t , noted in a column vector $P^t (r_{aa}, r_{Aa}, r_{AA})$, the probability distribution of all states at the next time step $t+1$, $P^{t+1}(s_{aa}, s_{Aa}, s_{AA})$, is derived from the following recurrence relation:

$$P^{t+1}(s_{aa}, s_{Aa}, s_{AA}) = P_{APC} \cdot P^t(r_{aa}, r_{Aa}, r_{AA}), \quad (5)$$

where P_{APC} is the transition matrix among states at two consecutive time steps. Obviously, this matrix must be oriented such that the sum of each column is $\sum_{(s_{aa}, s_{Aa}, s_{AA})} p[i | (r_{aa}, r_{Aa}, r_{AA})] = 1$.

For CP, the probabilities of transition for a single time step were computed separately for sexual (p_{sex}) and clonal (p_{clonal}) modes of reproduction as follows:

$$\begin{aligned} p_{sex}[(s_{aa}, s_{Aa}, s_{AA}) | (r_{aa}, r_{Aa}, r_{AA})] &= \frac{N!}{s_{aa}! s_{Aa}! s_{AA}!} (q_{aa}^{t+1})^{s_{aa}} (q_{Aa}^{t+1})^{s_{Aa}} (q_{AA}^{t+1})^{s_{AA}} \\ p_{clonal}[(s_{aa}, s_{Aa}, s_{AA}) | (r_{aa}, r_{Aa}, r_{AA})] &= \frac{N!}{s_{aa}! s_{Aa}! s_{AA}!} (p_{aa}^{t+1})^{s_{aa}} (p_{Aa}^{t+1})^{s_{Aa}} (p_{AA}^{t+1})^{s_{AA}} \end{aligned} \quad (6)$$

Recurrence equations were also expressed for each reproductive mode:

$$\begin{aligned} P^{t+1}(s_{aa}, s_{Aa}, s_{AA}) &= P_{sex} \cdot P^t(r_{aa}, r_{Aa}, r_{AA}) \\ P^{t+1}(s_{aa}, s_{Aa}, s_{AA}) &= P_{clonal} \cdot P^t(r_{aa}, r_{Aa}, r_{AA}) \end{aligned} \quad (7)$$

Whereas only one sexual generation occurs during each cycle of CP, n_{clonal} parthenogenetic generations occur. The probability distribution of all states after n_{clonal} parthenogenetic generations $P^t + n_{clonal}(s_{aa}, s_{Aa}, s_{AA})$ can be inferred from the recurrence equation:

$$P^{t+n_{clonal}}(s_{aa}, s_{Aa}, s_{AA}) = (P_{clonal})^{n_{clonal}} \cdot P^t(r_{aa}, r_{Aa}, r_{AA}) \quad (8)$$

Consequently, the probability distribution of all states after n_{clonal} parthenogenetic generations followed by one sexual generation ($n_{sex} = 1$) is

$$P^{t+n_{clonal}+n_{sex}}(s_{aa}, s_{Aa}, s_{AA}) = [P_{sex} \cdot (P_{clonal})^{n_{clonal}}] \cdot P^t(r_{aa}, r_{Aa}, r_{AA}) \quad (9)$$

Similarly, the probability distribution of all states after one sexual generation followed by n_{clonal} parthenogenetic generations is

$$P^{t+n_{sex}+n_{clonal}}(s_{aa}, s_{Aa}, s_{AA}) = [(P_{clonal})^{n_{clonal}} \cdot P_{sex}] \cdot P^t(r_{aa}, r_{Aa}, r_{AA}) \quad (10)$$

A graphical explanation for the ordering of the matrix multiplication in Equations (9) and (10) is given in Supplementary Information 1.

The transition matrix in Equation (9), $P_{CP \text{ after sex}} = [P_{sex} \cdot (P_{clonal})^{n_{clonal}}]$, thus permits a study of the genotypic frequencies of populations just after the sexual phase of CP (*CP after sex*). Similarly, the transition matrix in Equation (10),

$P_{CP \text{ before sex}} = [(P_{clonal})^{n_{clonal}} \cdot P_{sex}]$, permits a study of the genotypic frequencies of populations at the end of their clonal phase (that is, before the sexual event, *CP before sex*).

Through generations (that is, when $t \rightarrow \infty$), the probability distribution of the next states $P^i(s_{aa}, s_{Aa}, s_{AA})$ will converge towards a stationary probability distribution (hereafter referred to as s.p.d.) of the genotypic states. Because all the transition matrices generated are irreducible and ergodic Markov chains, the s.p.d. of the genotypic states is given by the eigenvector corresponding to the largest eigenvalue as stated by the Perron–Frobenius theorem. The transition matrices and resulting s.p.d. were calculated using Python 2.7 (<http://www.python.org/>).

We selected the parameter sets for our CP model based on descriptions of the life cycles of common cyclical parthenogens (Hebert, 1987; Moran, 1992; Strathdee *et al.*, 1993; Kanbe and Akimoto, 2009). The number of clonal generations in each cycle was set to $n_{clonal} = \{1; 9; 99; 999\}$, permitting us to study various cases, from those involving short cycles comprising only a few clonal events (for example, Cynipids, some aphid lineages) to those involving long cycles comprising a large number of clonal events (for example, those assumed to occur in permanent populations of cladocerans or monogonot rotifers). Comparisons of CP with APC were based on APC transition matrices with a constant rate of clonality $c = \frac{n_{clonal}}{n_{sex} + n_{clonal}} = \{0.5; 0.9; 0.99; 0.999\}$. In addition, transition matrices for full sexuality ($c=0$) and full clonality ($c=1$) were computed. The time required to compute these matrices strongly depends upon the population size. The parameters that regulate the balance between mutation and genetic drift were therefore chosen to allow the computation, within a reasonable timeframe, of a system in which the dominant force is

either genetic drift ($N\mu = 2 \times 10^{-4}$, $N=200$ and $\mu = 10^{-6}$) or mutation ($N\mu = 2$, where $N=200$ and $\mu = 10^{-2}$).

Single-locus genetic diversity

The probability of fixation (that is, the probability that all individuals in the population possess the genotype aa or AA) and the probability of heterozygote fixation (that is, the probability that all individuals in the population possess the genotype Aa) were obtained from the s.p.d. of each scenario. Because monomorphic markers are considered noninformative in empirical population genetics, all states in which one allele was fixed in the population were removed from the s.p.d. in subsequent analyses. The stationary probability of each genotypic state was rescaled such that the vector of stationary probabilities across all states summed to 1.

Given the predicted response of the inbreeding coefficient to clonality (Halkett *et al.*, 2005b), F_{IS} was calculated for each genotypic state according to Rousset (2004) as follows:

$$F_{IS} = \frac{F - \theta}{1 - \theta}$$

where F is the average allelic identity within individuals in the population and θ is the average allelic identity between a pair of individuals in the population. The exact F_{IS} distribution based on the s.p.d. was inferred for each scenario. The mean and variance of each exact F_{IS} distribution were calculated together with the probability of obtaining a positive value for F_{IS} (Table 1). For the purpose of visualisation, the exact F_{IS} distribution was approximated by a weighted kernel density estimation in which the s.p.d. was used to weight the F_{IS} value of each genotypic state. This weighted kernel density estimation was

Table 1 Descriptive statistics of F_{IS} distribution depending on levels, modes of clonality and the balance between genetic drift and mutation ($N\mu$)

$N\mu$	c	Mode	$p(F_{IS} > 0)$	$p(\text{fixation})$	$p(\text{het})$	Mean	Variance
2×10^{-4}	0	Full sexuality	0.2391	0.9947	0	-0.0025	0.0037
		APC	0.2244	0.9947	0	-0.0043	0.0043
	0.5	<i>CP after sex</i>	0.2391	0.9947	0	-0.0025	0.0037
		<i>CP before sex</i>	0.2136	0.9947	0	-0.0044	0.0063
		APC	0.1674	0.9946	0	-0.0152	0.0108
		<i>CP after sex</i>	0.2399	0.9946	0	-0.0025	0.0037
	0.9	<i>CP before sex</i>	0.1526	0.9946	0	-0.0158	0.0201
		APC	0.0751	0.9942	0	-0.0889	0.0507
		<i>CP after sex</i>	0.2499	0.9942	0	-0.0025	0.0038
	0.99	<i>CP before sex</i>	0.0632	0.9942	0.0001	-0.0914	0.0870
		APC	0.0111	0.9905	0.0015	-0.4407	0.1850
		<i>CP after sex</i>	0.3264	0.9916	0	-0.0025	0.0042
		<i>CP before sex</i>	0.0014	0.9916	0.0036	-0.4927	0.2151
	1	Full clonality	<0.0001	0.4976	0.4976	-0.9947	0.0045
		Full sexuality	0.4813	<0.0001	<0.0001	-0.0025	0.0050
	2	0	Full sexuality	0.4813	<0.0001	<0.0001	-0.0025
APC			0.4715	<0.0001	<0.0001	-0.0048	0.0065
0.5		<i>CP after sex</i>	0.4813	<0.0001	<0.0001	-0.0025	0.0050
		<i>CP before sex</i>	0.4754	<0.0001	<0.0001	-0.0049	0.0096
		APC	0.4422	<0.0001	<0.0001	-0.0178	0.0190
		<i>CP after sex</i>	0.4814	<0.0001	<0.0001	-0.0025	0.0050
0.9		<i>CP before sex</i>	0.4491	<0.0001	<0.0001	-0.0204	0.0339
		APC	0.4055	<0.0001	<0.0001	-0.0470	0.0464
		<i>CP after sex</i>	0.4813	<0.0001	<0.0001	-0.0025	0.0050
0.99		<i>CP before sex</i>	0.3965	<0.0001	<0.0001	-0.0568	0.0560
		APC	0.3964	<0.0001	<0.0001	-0.0563	0.0549
		<i>CP after sex</i>	0.4813	<0.0001	<0.0001	-0.0025	0.0050
		<i>CP before sex</i>	0.3952	<0.0001	<0.0001	-0.0576	0.0560
1		Full clonality	0.3952	<0.0001	<0.0001	-0.0576	0.0560

Abbreviations: APC, acyclic partial clonality; CP, cyclical parthenogenesis.

In APC, c is the fraction of individuals in the population reproducing clonally, in CP, c is the proportion of clonal generation in a cycle ($c = n_{clonal} / (n_{clonal} + n_{sex})$); $p(F_{IS} > 0)$ is the probability of observing positive F_{IS} ; $p(\text{fixation})$ is the probability of fixing one of the allele in the population; and $p(\text{het})$ is the probability of fixing heterozygotes in the population.

made using a Gaussian kernel function with a bandwidth of 0.05. The F_{IS} density distribution was then estimated at 512 equally spaced points across the F_{IS} range (that is, $[-1,1]$) using the R function *density* (R Core Team, 2013).

Synthetic parameters such as F_{IS} are almost always used in population genetics studies to describe a genetic data set. Although convenient, summarising genetic information using such indices is associated with the loss of some of the information encompassed in a full data set. In our case, the s.p.d. of genotypic states comprises the total genetic information available. The s.p.d. of genotypic states was visualised using 'de Finetti diagrams' (De Finetti, 1926, Supplementary Information 2). A colour scale was used to represent the stationary probability of each genotypic state using the R package *ggtern* (Hamilton, 2015). Pairwise divergences between scenarios were quantified using the Jensen–Shannon measure of divergence between s.p.d. (D_{JS}) wherein each genotypic state represented a discrete class (all comparisons can be found in Supplementary Information 3).

Two-loci simulations

Extending the model of Stoeckel and Masson for multiple loci is theoretically possible. Unfortunately, the number of genotypic states increases exponentially with the number of possible genotypes. Consequently, the calculation of the exact s.p.d. becomes rapidly intractable for current computational resources (Reichel *et al.*, 2015).

Individual-based simulations following the same conceptual model were therefore used to approximate the distribution of genetic diversity encountered at two unlinked loci. Each individual in these simulations possessed two biallelic loci. Clonal reproduction was simulated by a random draw with replacement of individuals from the parental generation. Sexual reproduction was simulated by a random draw with replacement of two parents, each randomly transmitting one allele per locus to their offspring. The scenarios analysed by simulations were similar to those analysed in the single-locus model, with $n_{clonal} = \{1; 9; 99; 999\}$ in CP, $c = \{0.5; 0.9; 0.99; 0.999\}$ in APC, $c = 0$ in full sexuality and $c = 1$ in full clonality. Similar to the single-locus model, the mutation drift balance was set to cases where the dominant force is either genetic drift ($N\mu = 0.2$, $N = 200$ and $\mu = 10^{-3}$) or mutation ($N\mu = 2$, $N = 200$ and $\mu = 10^{-2}$). Each simulation was repeated 10 000 times and ran for 10 000 generations.

Gametic phase disequilibrium

We used the two-loci simulations to approximate the distribution of gametic phase disequilibrium between two loci. In APC, full sexuality and full clonality, gametic phase disequilibrium was measured for each repetition at generation 10 001. In CP, disequilibrium was measured before and after the next sexual event (for example, generations 10 009 and 10 010 if $n_{clonal} = 9$). We calculated the gametic phase disequilibrium using Pearson's r correlation coefficient:

$$r = \frac{|D|}{\sqrt{p_A p_a p_B p_b}}$$

where D is the gametic phase disequilibrium parameter, p_A and p_a are the frequencies of the alleles A and a at the first locus, and p_B and p_b are the frequencies of the alleles B and b at the second locus. For each scenario, repetitions in which at least one of the two loci reached fixation were removed from the analysis. Kernel density estimations using all repetitions (at most 10 000 gametic phase disequilibrium observations) were used to visualise the approximate gametic phase disequilibrium distribution for each scenario. The distribution of gametic phase disequilibrium was also calculated using Lewontin's D . Qualitatively, both Pearson's r and Lewontin's D lead to the same distributions of gametic phase disequilibrium. In the present paper, only the distributions of Pearson's r will be discussed. The distributions of Lewontin's D are available for comparison in Supplementary Information 4.

RESULTS

Single-locus model: full sexuality/clonality

When $N\mu < 1$ in full clonality, genetic drift is the main force acting on genotypic frequencies. Genetic drift leads to the fixation of a genotype (Aa , AA or aa ; Figure 1 and Table 1). States in which an allele is fixed in the population were removed for the weighted kernel density

estimation of the exact F_{IS} distribution. As a result, the F_{IS} distribution is tightly grouped around $F_{IS} = -1$ (Figure 2, dotted area). In full sexuality, only drift towards the fixation of homozygotes is possible because drift towards the fixation of heterozygotes is impeded by sexual reproduction (Figure 1). Observed deviations from Hardy–Weinberg equilibrium (Table 1, negative mean F_{IS} and probability of positive $F_{IS} < 0.5$; Figure 2, grey shaded area) are common expectations in populations of finite size because of the difference in allelic frequencies between the two gametic pools participating in sexual reproduction (Wang, 1996).

When $N\mu > 1$ in full clonality, mutation is the only force acting against genetic drift. Random mutation impedes the fixation of a genotype by decreasing the frequency of the most common genotype (Figure 3 and Table 1). In our biallelic model, such mutations tend to relocate the distribution of genetic diversity towards Hardy–Weinberg equilibrium (Reichel *et al.*, 2016). This process is, however, slower than under sexual reproduction, implying that strongly positive or negative values are likely to be observed (Figure 4, dotted area, and Table 1). In full sexuality, both random mutation and sexual reproduction act against genetic drift. The genotypic states with the highest stationary probabilities are located at the summit of the Hardy–Weinberg parabola (Figure 3). The resulting F_{IS} distribution is nearly symmetrical around its negative mean, and the probability of observing a positive F_{IS} is close to 0.5 (Figure 4, grey shaded area, and Table 1).

Single-locus model: CP

When $N\mu > 1$, the divergence between predictions of genetic diversity under conditions of CP *before sex* and full sexuality increases with the number of clonal generations until reaching predictions made under a condition of full clonality (Figure 3 and Supplementary Information 3). Regarding the F_{IS} distribution, this increasing divergence translates into a progressive shift of the mean towards negative values and an increased variance (Figure 4, dashed-line distributions, and Table 1). In contrast, for CP *after sex*, only one generation of sexual reproduction is sufficient to reset the distribution of genetic diversity within populations towards the full-sexuality predictions (Figures 3 and 4).

When $N\mu < 1$, the divergence between full sexuality and CP *before sex* also increases with the number of clonal generations in the cycle (Figure 1 and Supplementary Information 3). However, for $N\mu = 10^{-4}$, even 999 generations of clonality are not sufficient to reach the predictions made under an assumption of full clonality (Figures 1 and 2, dashed-line distribution). One generation of sexual reproduction (CP *after sex*) does not reset the distribution of genetic diversity within populations towards the full-sexuality predictions if the number of preceding clonal generations is large (Figure 1). In this case, the probability of fixing the heterozygote genotype during the clonal phase is high (Table 1, CP *before sex*, $c = 0.999$, $N\mu = 10^{-4}$). This probability is then reported in CP *after sex* as a state in which both alleles are equipotent at the summit of the Hardy–Weinberg parabola. Genetic drift additionally 'spreads' this probability to neighbouring states (Figure 1, CP *after sex*, $c = 0.999$). As a result, the variance of F_{IS} slightly increases, and the probability of obtaining a positive F_{IS} value approaches 0.5 (Figure 2, $c = 0.999$, solid-line distribution, and Table 1).

Single-locus model: APC

Disregarding the balance between mutations and genetic drift ($N\mu$), the divergence between predictions of genetic diversity under conditions of APC and full sexuality increases with the rate of clonality (Figures 1 and 3 and Supplementary Information 3).

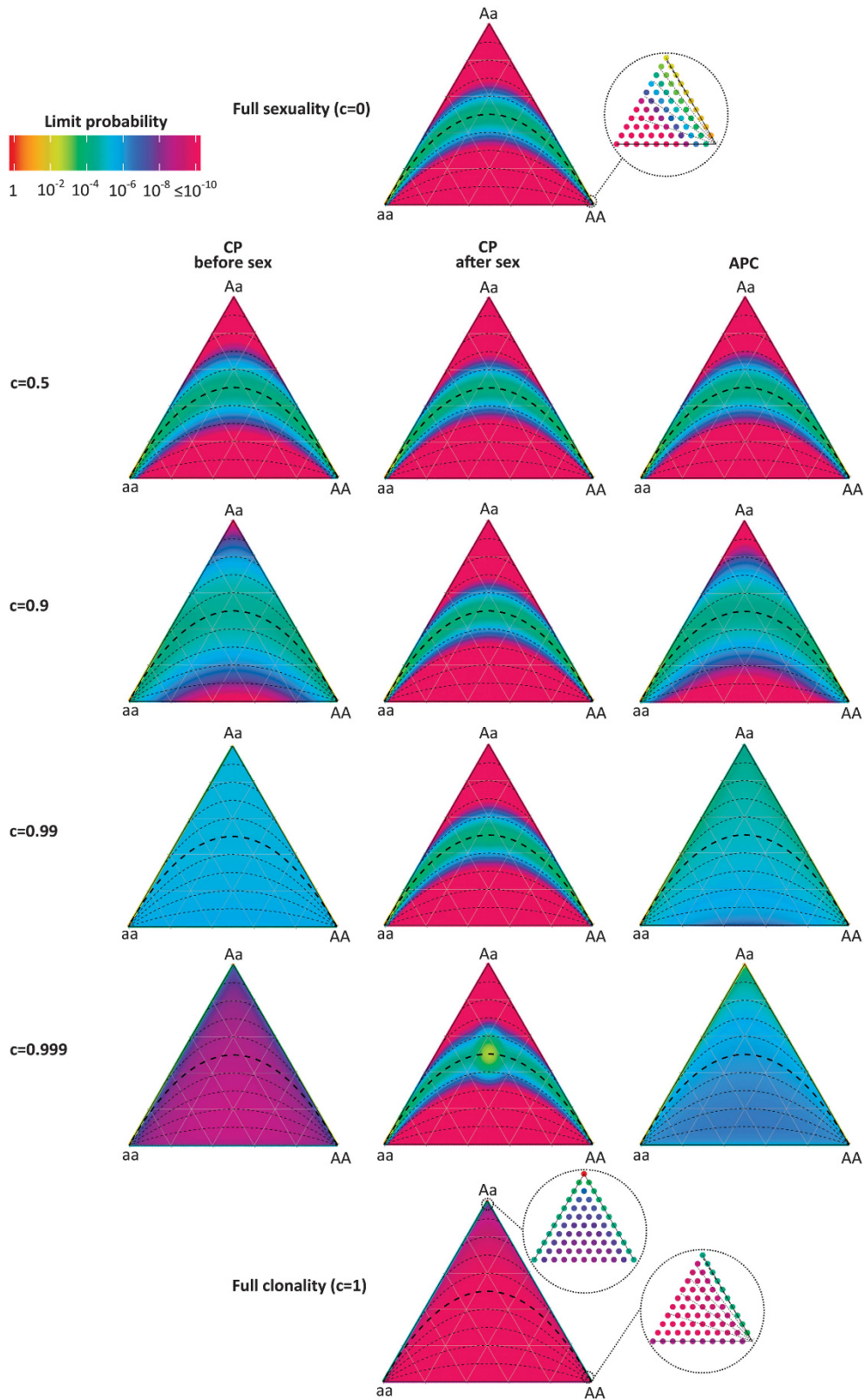


Figure 1 De Finetti diagrams illustrating stationary probability distributions of genotypic states for distinct modes and levels of clonality at $N\mu < 1$ ($N\mu = 2 \times 10^{-4}$; $N = 200$, $\mu = 10^{-6}$). In APC, c is the fraction of individuals clonally produced in the population; in CP, c is the proportion of clonal generations in a cycle ($c = n_{clonal} / (n_{clonal} + n_{sex})$). De Finetti diagrams are ternary diagrams displaying the frequency of each genotype in the population. Any point in the diagram corresponds to a given genotypic state (r_{AA} , r_{Aa} , r_{aa}). Each genotypic state in the diagram has a given stationary probability. Black dotted lines are F_{IS} isoclines, with the thickest located at $F_{IS} = 0$. More details on how to read de Finetti diagrams can be found in Supplementary Information 2.

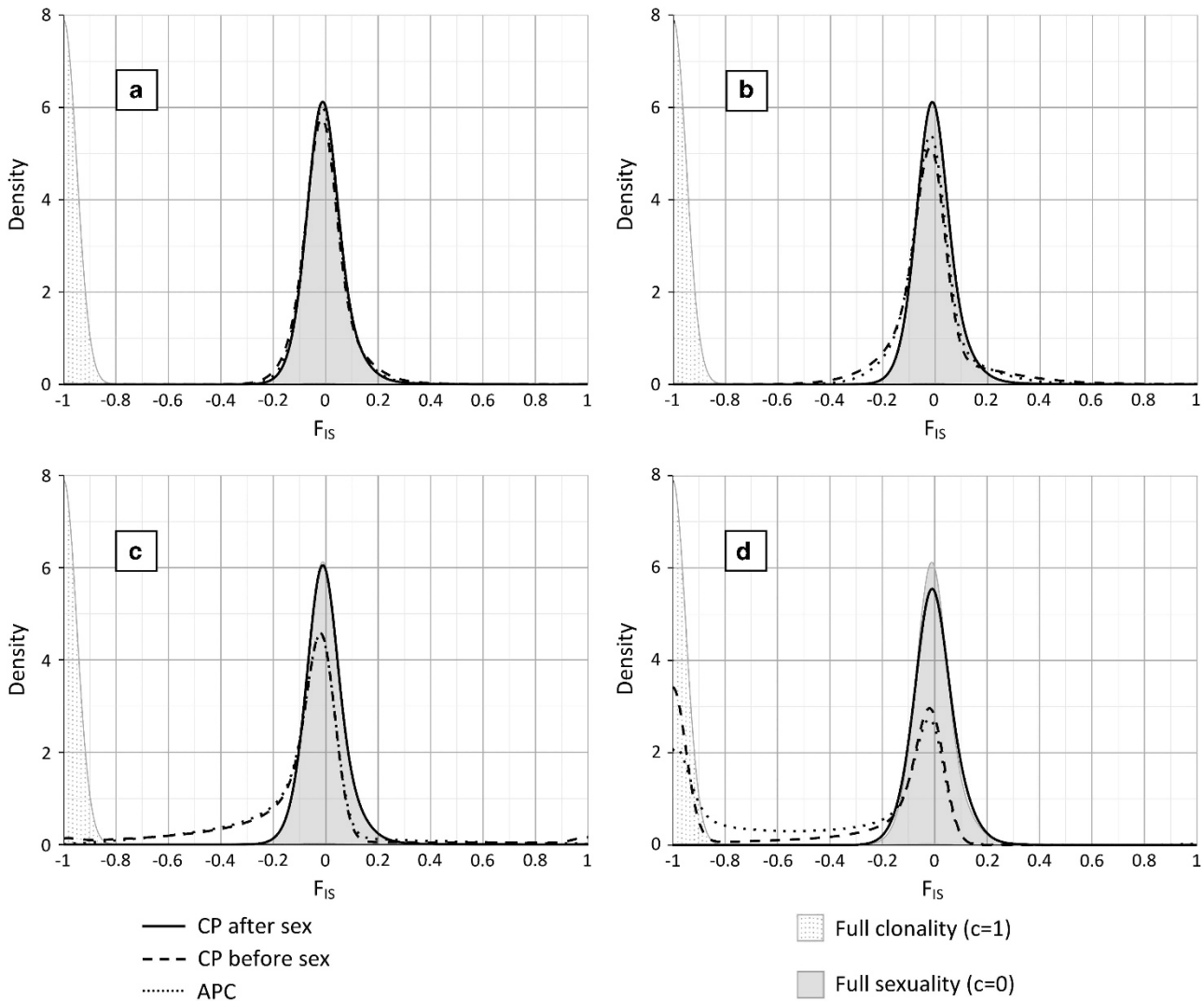


Figure 2 Weighted kernel density estimation of the full F_{IS} probability density function for distinct levels and modes of clonality at $N\mu < 1$ ($N\mu = 2 \times 10^{-4}$, $N = 200$, $\mu = 10^{-6}$). (a) The $n_{clonal} = 1$ in CP and $c = 0.5$ in APC; (b) $n_{clonal} = 9$ in CP and $c = 0.9$ in APC; (c) $n_{clonal} = 99$ in CP and $c = 0.99$ in APC; and (d) $n_{clonal} = 999$ in CP and $c = 0.999$ in APC.

However, convergence towards the full-clonality predictions is faster in *CP before sex* than in APC (Figures 1 and 3). This observation can also be made by observing the F_{IS} distributions (Figures 2 and 4).

Two-loci simulations: full sexuality/clonality

Regarding the gametic phase disequilibrium, strong genetic drift ($N\mu < 1$) in a fully clonal population leads to three possible outcomes. In the first, the population consists entirely of a single genotype that is homozygous at both loci; r is undefined. In the second outcome, the population consists entirely of a single genotype that is heterozygous at one locus and homozygous at the other; r is also undefined. In the example of $N\mu = 0.2$, mutation prevents fixation such that drift towards the two above-mentioned scenarios results in a high probability density for low, rather than undefined, r values (Figure 5, dotted area). In the third possible outcome, the population consists of a single genotype that is heterozygous at both loci, resulting in $r = 1$. Again, at $N\mu = 0.2$, trajectories reaching $r = 1$ are rare because of the limited dominance of genetic drift over mutation. High probability densities are therefore located around states with intermediate r values (Figure 5, dotted area). In full sexuality, drift only

leads towards homozygote fixation at both loci. For $N\mu = 0.2$, high probability densities at low r values (Figure 5, grey shaded distribution) are produced by the combined effects of mutations that prevent fixation and sexuality that recombines gametic phases.

When $N\mu > 1$ in full clonality, mutation impedes gametic phase disequilibrium by decreasing the frequency of the most common gametic phase (Figure 6, dotted area). Compared with sexual reproduction, mutation is a slow process by which to relocate r towards 0. Thus, we observe a large amount of variation around the gametic phase equilibrium, and the r distribution is logically right skewed (Figure 6, dotted area). In full sexuality, constant recombination yields a probability distribution tightly grouped around small values of r , with the right tail occurring only because of rare stochastic events (Figure 6, grey shaded area).

Two-loci simulations: CP

When $N\mu < 1$, the divergence between r distributions obtained under conditions of full sexuality and *CP before sex* is visible even for the smallest cycle (Figure 5, $n_{clonal} = 1$, dashed-line distribution). As the number of clonal generations increases in the cycle, the distribution

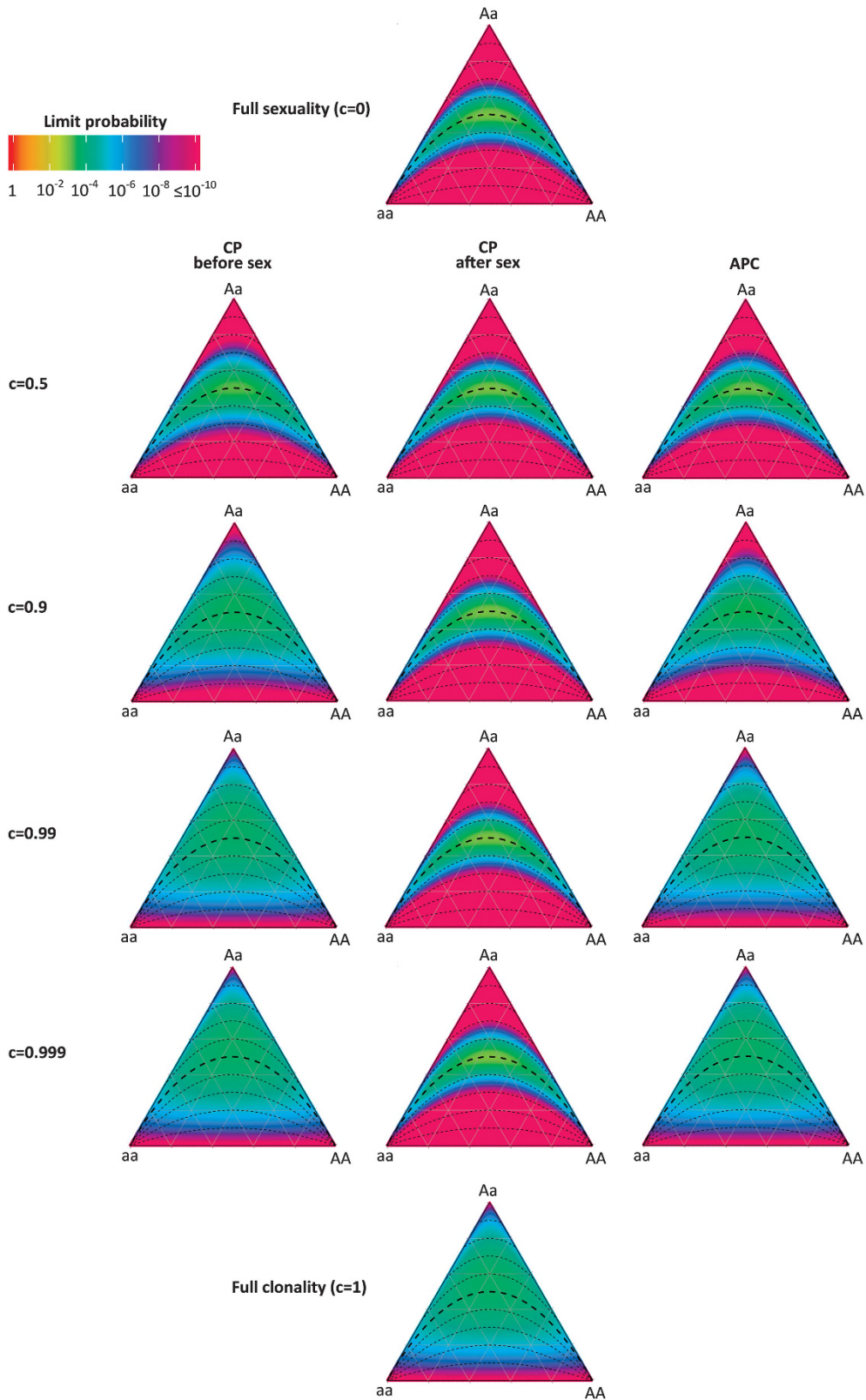


Figure 3 De Finetti diagrams illustrating stationary probability distributions of genotypic states for distinct modes and levels of clonality at $N\mu > 1$ ($N\mu = 2$; $N = 200$, $\mu = 10^{-2}$). In APC, c is the fraction of individuals clonally produced in the population; in CP, c is the proportion of clonal generation in a cycle ($c = n_{clonal} / (n_{clonal} + n_{sex})$). De Finetti diagrams are ternary diagrams displaying the frequency of each genotype in the population. Any point in the diagram corresponds to a given genotypic state (r_{AA} , r_{Aa} , r_{aa}). Each genotypic state in the diagram has a given stationary probability. Black dotted lines are F_{IS} isoclines, with the thickest located at $F_{IS} = 0$. More details on how to read de Finetti diagrams can be found in Supplementary Information 2.

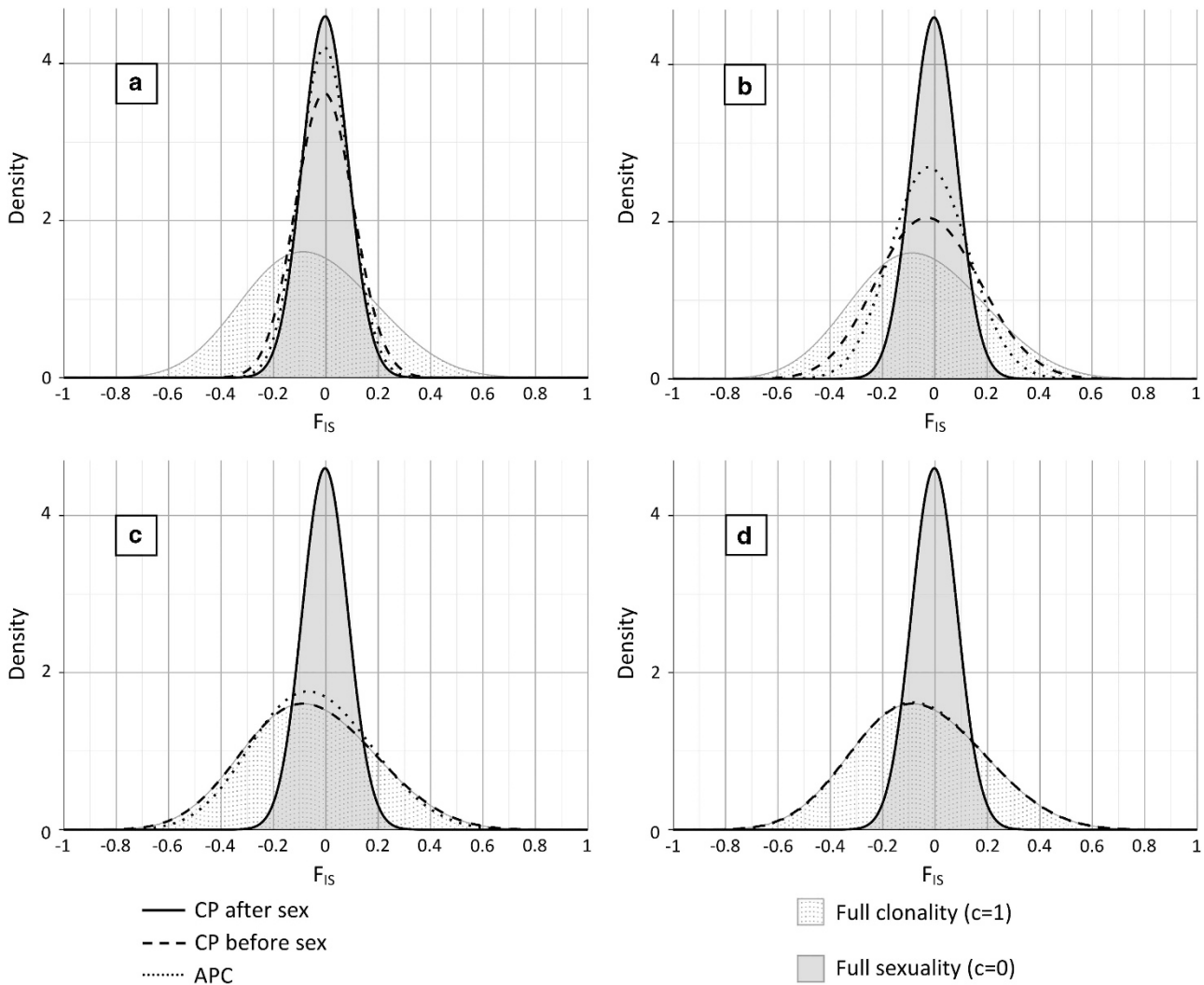


Figure 4 Weighted kernel density estimation of the full F_{IS} probability density function for distinct levels and modes of clonality at $N\mu > 1$ ($N\mu=2$, $N=200$, $\mu=10^{-2}$). (a) The $n_{clonal}=1$ in cyclical CP and $c=0.5$ in APC; (b) $n_{clonal}=9$ in CP and $c=0.9$ in APC; (c) $n_{clonal}=99$ in CP and $c=0.99$ in APC; and (d) $n_{clonal}=999$ in CP and $c=0.999$ in APC.

for *CP before sex* converges towards the full-clonality prediction, but the difference between the two distributions remains even after numerous clonal generations (Figure 5, $n_{clonal}=999$, dashed-line distribution). We predicted that this difference would increase if genetic drift becomes stronger compared with mutation (that is, $N\mu$ approaches 0). In theory, a single recombination event between two loci in gametic disequilibrium is not sufficient to produce gametic phase equilibrium in the next generation. The differences observed between *CP after sex* and full sexuality confirmed this prediction at every level of clonality considered (Figure 5, solid-line distributions).

The pattern observed for CP when mutation dominates over genetic drift ($N\mu > 1$) is very similar to that described when genetic drift dominates over mutation ($N\mu < 1$). The distribution of r diverges from the full-sexuality predictions even for the smallest cycle (Figure 6, $n_{clonal}=1$, dashed-line distribution). As the number of clonal generations increases in the cycle, the distribution for *CP before sex* converges towards predictions made under a condition of full clonality, and the two distributions are very similar at $n_{clonal}=99$. Finally, a single

event of recombination fails to reset the distribution of r towards predictions made under full-sexuality conditions (Figure 6, solid-line distributions).

Two-loci simulations: APC

Similar to the observations made at a single locus, the divergence between predictions of genetic diversity for APC and full sexuality continues to increase with the rate of clonality (Figures 5 and 6). Moreover, predictions of linkage disequilibrium with *CP before sex* converge faster than those with APC towards predictions made under conditions of full clonality (Figures 5 and 6).

DISCUSSION

Several studies have quantified and formalised the effect of clonality on genetic diversity in species in which sexual and clonal reproduction co-occur in time (Marshall and Weir, 1979; Orive, 1993; Balloux *et al.*, 2003; Bengtsson, 2003; De Meeûs and Balloux, 2004, 2005; Stoeckel and Masson, 2014). Despite many examples of CP in the wild, analytical models or simulations studying the effect of this

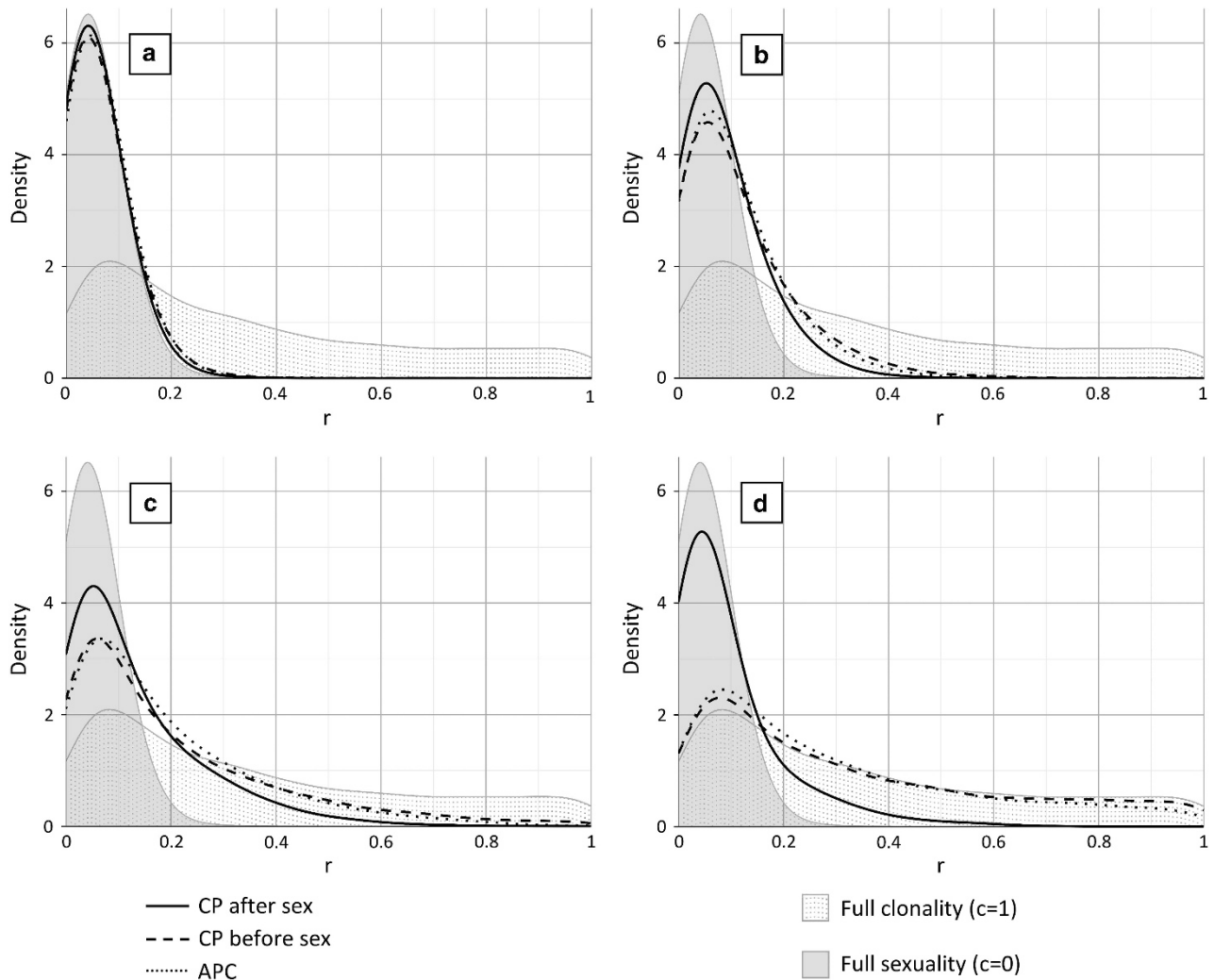


Figure 5 Probability density function of the gametic phase disequilibrium (r) for distinct levels and modes of clonality at $N\mu < 1$ ($N\mu = 0.2$, $N = 200$, $\mu = 10^{-3}$). Distributions were obtained by a kernel density estimation based on 10 000 simulation results. (a) The $n_{clonal} = 1$ in cyclical parthenogenesis and $c = 0.5$ in APC; (b) $n_{clonal} = 9$ in CP and $c = 0.9$ in APC; (c) $n_{clonal} = 99$ in CP and $c = 0.99$ in APC; and (d) $n_{clonal} = 999$ in CP and $c = 0.999$ in APC.

reproductive system on genetic diversity are rare (Lynch and Deng, 1994; Berg and Lascoux, 2000; Pålson, 2001; Prugnolle *et al.*, 2005a, b; Vanoverbeke and De Meester, 2010). The model presented here is the first to provide a probability distribution of neutral genetic diversity at a single locus and at two unlinked loci, depending on the level of clonality (length of the clonal phase) in CP. The two theoretical cases investigated in this paper can be used to predict how neutral genetic diversity would be distributed in a large range of situations, from cases in which genetic drift dominates ($N\mu < 1$, typically in small populations such as those used in experimental evolutionary studies and mesocosms) to cases in which mutation plays an important role in maintaining genetic diversity ($N\mu > 1$, typically in large populations such as those found in the wild).

A few generations of clonality cause departures from the assumptions of full sexuality

Our analysis confirms earlier studies showing that intermediate levels of clonality ($0 < c < 0.9$), or a limited number of clonal generations, can substantially impact the probability distribution of F_{IS} (*CP before sex*) (Stoekel and Masson, 2014) despite having a limited effect on the

mean F_{IS} (Balloux *et al.*, 2003). This observed pattern is even stronger in cases where mutation dominates ($N\mu > 1$). However, analysing only the F_{IS} distribution provides little information about the processes causing such deviation. Moreover, the analysis of the entire distribution of genetic diversity indicated that such processes are very distinct depending on whether mutation or genetic drift is the dominant force in the population (Figures 1 and 3). This result stresses the importance of not restricting analyses to synthetic parameters, such as the mean F_{IS} , to describe the effects of complex life cycles on genetic diversity and structure.

The simulations of the evolution of two loci also indicated that the effects of clonality are more evident when examining the distribution of the gametic phase disequilibrium rather than its mean (De Meeüs and Balloux, 2004). Similar to the single-locus model, the relationship between the level of clonality and the distribution of gametic phase disequilibrium is largely dependent upon the balance between genetic drift and mutation. Our results indicate that the distribution of gametic phase disequilibrium, alone or in combination with the F_{IS} distribution, represents an interesting tool with which to infer the level of clonality in a population.

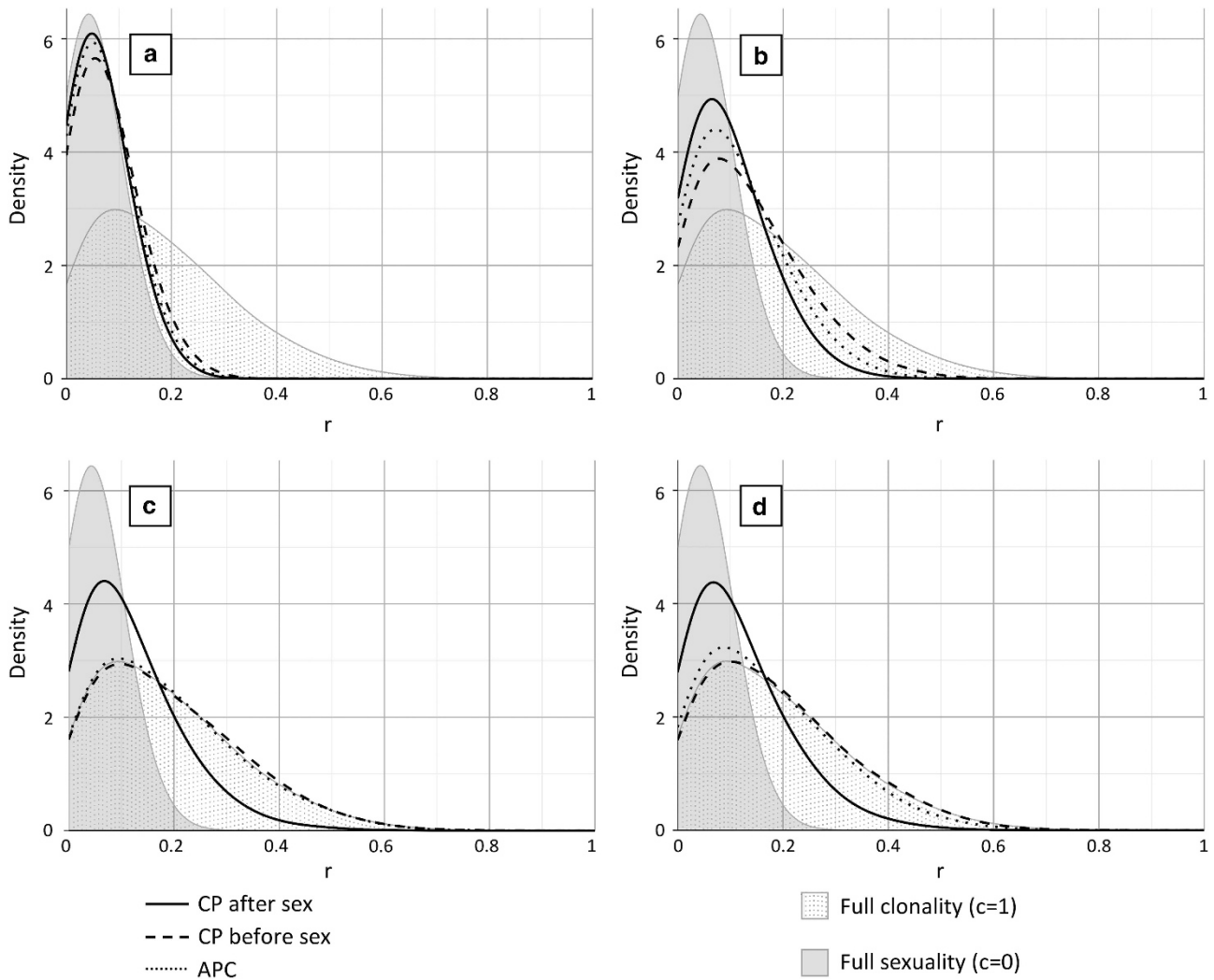


Figure 6 Probability density function of the gametic phase disequilibrium (r) for distinct levels and modes of clonality at $N\mu > 1$ ($N\mu = 2$, $N = 200$, $\mu = 10^{-2}$). Distributions were obtained by a kernel density estimation based on 10 000 simulation results. (a) The $n_{clonal} = 1$ in CP and $c = 0.5$ in APC; (b) $n_{clonal} = 9$ in CP and $c = 0.9$ in APC; (c) $n_{clonal} = 99$ in CP and $c = 0.99$ in APC; and (d) $n_{clonal} = 999$ in CP and $c = 0.999$ in APC.

Sexuality in CA does not reset genetic diversity towards predictions made under full-sexuality conditions

As anticipated in previous studies (Lynch and Deng, 1994; Pfrender and Lynch, 2000; De Meester *et al.*, 2006; Allen and Lynch, 2012), our model confirms that a single generation of sexual reproduction following a low number of clonal generations rearranges genetic diversity at a single locus towards predictions made under a condition of full sexuality (*CP after sex*). Notably, this is not true in particular cases in which genetic drift dominates ($N\mu < 1$, for example, in very small populations) and the number of clonal generations preceding the sexual event is large enough to increase the probability of obtaining a high frequency of heterozygotes in the population.

In contrast, the simulations at two loci show that the distribution of gametic phase disequilibrium after a sexual event never relocates towards the full-sexuality predictions. This result is theoretically predictable because the decay in gametic phase disequilibrium is a function of the rate of recombination between two loci (Lynch and Deng, 1994; Rice, 2004). In our simulations, free recombination was allowed between loci. Therefore, we predict this effect to be more pronounced when two loci are physically linked. Overall, these results

stress the importance of taking the time of sampling into account when analysing population genetic data sets of cyclical parthenogens (Berg and Lascoux, 2000). Using our model/simulation results as a reference for the interpretation of data from empirical studies will help to distinguish between neutral scenarios and the actions of other evolutionary processes (for example, epistatic interactions causing high gametic phase disequilibrium; Lehmann *et al.*, 2004).

Long periods of clonality are not always equivalent to full clonality

The number of clonal generations required to reach the predictions of genetic diversity under full-clonality conditions also depends on the balance between genetic drift and mutation. A long clonal phase, that is, more than 999 generations of clonality, is required before reaching the predictions made under a condition of full clonality when genetic drift largely dominates ($N\mu = 2 \times 10^{-4}$). In contrast, when mutation dominates, a shorter period of clonality is sufficient for the distribution of genetic diversity to fully converge towards the full-clonality predictions. As mentioned previously, an approach combining multiple summary statistics, including F_{IS} and gametic phase disequilibrium, may be used to infer intermediate levels of clonality in natural

populations. However, this method may be of limited use in cases where the distribution of genetic diversity is identical to the predictions made under full-clonality conditions. For this purpose, a recently developed coalescent approach may provide a good alternative to inference methods based on distributions of multiple parameters (Hartfield *et al.*, 2015). Such a method would be particularly relevant for the detection of very low levels of sex in natural or experimental populations of putatively ancient asexuals (Danchin *et al.*, 2011).

CA and APC yield different distributions of genetic diversity

The probability distributions of both F_{IS} and gametic phase disequilibrium under CP *after* or *before* sex are different from those observed under APC. Furthermore, the probability distributions obtained under APC conditions do not represent an average situation between the two seasonal distributions of CP. For similar levels of clonality, convergence towards the predictions made under full-clonality conditions is always stronger in CP *before* sex than in APC. However, the two scenarios yield very similar distributions when the level of clonality is high.

From an evolutionary perspective, this result suggests that genetic diversity tends to be lost more quickly in CP than in APC. By extension, we predict that selection would also be more rapid in CP to fix advantageous alleles or advantageous epistatic associations. In APC, however, such advantageous epistatic association is rarely broken by sexual reproduction. This is not the case in CP, in which the synchronous massive sexual event of the cycle largely recombines genotypes that not only produces many new recombined associations (ensuring adaptive potential in new environmental conditions) but also maintains the previously advantageous epistatic associations at high frequencies in the population (ensuring adaptive potential if environmental conditions remain unchanged from one cycle to the next).

Because of the absence of an explicit model describing the effect of clonality on the distribution of neutral genetic diversity in CP, many studies have relied on models describing predictions of genetic diversity under APC conditions in the discussion of their CP data (see, for example, Halkett *et al.*, 2005a; Vorburger, 2006; Kanbe and Akimoto, 2009; Allen and Lynch, 2012). Specially designed to fit the CP lifecycle, our model refines these predictions and is particularly suited to testing precise hypotheses regarding genetic diversity in cyclical parthenogens.

CONCLUSION AND PERSPECTIVES

Our model yields detailed predictions of neutral genetic diversity under conditions of CP. The model demonstrates that CP has specific effects on the neutral evolution of molecular data that are different from other reproductive modes (that is, full sexuality, full clonality and APC).

To date, many studies have linked summary statistics such as F_{IS} or linkage disequilibrium to the relative importance of sexual vs clonal reproduction within species displaying CP (Allen and Lynch, 2012; Delmotte *et al.*, 2002; Halkett *et al.*, 2005a; Ivens *et al.*, 2012; Kanbe and Akimoto, 2009; Papura *et al.*, 2003; Vorburger, 2006; Vorwerk and Forneck, 2006). Based on previous models and observations, a negative F_{IS} and strong linkage disequilibrium have been interpreted as signatures of strong clonality. Our analysis shows that strongly positive F_{IS} values can also be observed for highly clonal species if mutation dominates over genetic drift ($N\mu > 1$). Such positive F_{IS} values were often observed but were difficult to explain until now (Ivens *et al.*, 2012). Overall, this model provides a theoretical framework that

permits improved estimations of the contribution of each mode of reproduction within a particular population.

The detection in natural systems of departures from our model predictions will also permit the identification of the actions of other evolutionary forces such as natural selection. In this strictly neutral approach, the influence of selection was voluntarily overlooked. However, background selection has been shown to have strong effects on the distribution of genetic diversity at linked neutral markers in cyclical parthenogens (Pålson, 2001). We hope that our results will also lead to further investigations aiming to refine predictions of genetic diversity in populations of cyclical parthenogens while accounting for other classical aspects of their life cycles (for example, population bottlenecks following sexual reproduction).

DATA ARCHIVING

All data are included in the article or can be reproduced by the program codes available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.r08vm>.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Jurgen Angst, Pierre Nouhaud and Jean-Christophe Simon for helpful discussions, the Région Bretagne and the Plant Health and Environment (SPE) division of the French National Institute for Agricultural Research (INRA) for financing Katja Reichel as well as the French National Research Agency (ANR) for funding within the CLONIX project (ANR-11-BSV7-0007).

- Allen DE, Lynch M (2012). The effect of variable frequency of sexual reproduction on the genetic structure of natural populations of a cyclical parthenogen. *Evolution* **66**: 919–926.
- Balloux F, Lehmann L, De Meeüs T (2003). The population genetics of clonal and partially clonal diploids. *Genetics* **164**: 1635–1644.
- Bengtsson BO (2003). Genetic variation in organisms with sexual and asexual reproduction. *J Evol Biol* **16**: 189–199.
- Berg LM, Lascoux M (2000). Neutral genetic differentiation in an island model with cyclical parthenogenesis. *J Evol Biol* **13**: 488–494.
- Ceplitis A (2003). Coalescence times and the Meselson effect in asexual eukaryotes. *Genet Res* **82**: 183–190.
- Danchin EGJ, Flot JF, Perfus-Barbeoch L, van Doninck K (2011). Genomic perspectives on the long-term absence of sexual reproduction in animals. In: Pontarotti P (ed) *Evolutionary Biology—Concepts, Biodiversity, Macroevolution and Genome Evolution*. Springer: Verlag Berlin Heidelberg. pp 223–242.
- De Finetti B (1926). Considerazioni matematiche sull'ereditarietà mendeliana. *Metron* **6**: 3–41.
- De Meester L, Vanoverbeke J, De Gelas K, Ortells R, Spaak P (2006). Genetic structure of cyclic parthenogenetic zooplankton populations—a conceptual framework. *Arch Für Hydrobiol* **167**: 217–244.
- De Meeüs T, Balloux F (2004). Clonal reproduction and linkage disequilibrium in diploids: a simulation study. *Infect Genet Evol* **4**: 345–351.
- De Meeüs T, Balloux F (2005). F-statistics of clonal diploids structured in numerous demes. *Mol Ecol* **14**: 2695–2702.
- De Meeüs T, Lehmann L, Balloux F (2006). Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. *Infect Genet Evol* **6**: 163–170.
- De Meeüs T, Prugnolle F, Agnew P (2007). Asexual reproduction: genetics and evolutionary aspects. *Cell Mol Life Sci* **64**: 1355–1372.
- Delmotte F, Leterme N, Gauthier J-P, Rispé C, Simon J-C (2002). Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Mol Ecol* **11**: 711–723.
- Halkett F, Plantegenest M, Prunier-Leterme N, Mieuze L, Delmotte F, Simon JC (2005a). Admixed sexual and facultatively asexual aphid lineages at mating sites. *Mol Ecol* **14**: 325–336.
- Halkett F, Simon J-C, Balloux F (2005b). Tackling the population genetics of clonal and partially clonal organisms. *Trends Ecol Evol* **20**: 194–201.
- Hamilton M (2009). *Population Genetics*. Wiley-Blackwell: New Jersey, USA.
- Hamilton N (2015). ggtern: an extension to 'ggplot2', for the creation of ternary diagrams. R package version 1.0.6.0.

- Hartfield M, Wright SI, Agrawal AF (2015). Coalescent times and patterns of genetic diversity in species with facultative sex: effects of gene conversion, population structure, and heterogeneity. *Genetics* **202**: 297–312.
- Hebert PD (1987). Genotypic characteristics of cyclic parthenogens and their obligately asexual derivatives. In: Stearns SC (ed), *The Evolution of Sex and its Consequences*. Birkhäuser Verlag: Basel, Switzerland. pp 175–195.
- Ivens IBF, Kronauer DJC, Pen I, Weissing FJ, Boomsma JJ (2012). Reproduction and dispersal in an ant-associated root aphid community. *Mol Ecol* **21**: 4257–4269.
- Kanbe T, Akimoto S (2009). Allelic and genotypic diversity in long-term asexual populations of the pea aphid, *Acyrtosiphon pisum* in comparison with sexual populations. *Mol Ecol* **18**: 801–816.
- Kimura M (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press: New York, NY, USA.
- Lehmann T, Graham DH, Dahl ER, Bahia-Oliveira LMG, Gennari SM, Dubey JP (2004). Variation in the structure of *Toxoplasma gondii* and the roles of selfing, drift, and epistatic selection in maintaining linkage disequilibria. *Infect Genet Evol* **4**: 107–114.
- Lynch M, Deng H-W (1994). Genetic slippage in response to sex. *Am Nat* **144**: 242–261.
- Marshall DR, Weir BS (1979). Maintenance of genetic variation in apomictic plant populations. *Heredity (Edinb)* **42**: 159–172.
- Moran NA (1992). The evolution of aphid life cycles. *Annu Rev Entomol* **37**: 321–348.
- Navascués M, Stoeckel S, Mariette S (2010). Genetic diversity and fitness in small populations of partially asexual, self-incompatible plants. *Heredity (Edinb)* **104**: 482–492.
- Orive ME (1993). Effective population size in organisms with complex life-histories. *Theor Popul Biol* **44**: 316–340.
- Papura D, Simon J-C, Halkett F, Delmotte F, Le Gallic J-F, Dedyver C-A (2003). Predominance of sexual reproduction in Romanian populations of the aphid *Sitobion avenae* inferred from phenotypic and genetic structure. *Heredity (Edinb)* **90**: 397–404.
- Pfrender ME, Lynch M (2000). Quantitative genetic variation in *Daphnia*: temporal changes in genetic architecture. *Evolution (NY)* **54**: 1502–1509.
- Prugnolle F, Liu H, De Meeüs T, Balloux F (2005a). Population genetics of complex life-cycle parasites: an illustration with trematodes. *Int J Parasitol* **35**: 255–263.
- Prugnolle F, Roze D, Théron A, de Meeüs T (2005b). F-statistics under alternation of sexual and asexual reproduction: a model and data from schistosomes (platyhelminth parasites). *Mol Ecol* **14**: 1355–1365.
- Pålsson S (2001). The effects of deleterious mutations in cyclically parthenogenetic organisms. *J Theor Biol* **208**: 201–214.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
- Reichel K, Bahier V, Midoux C, Parisey N, Masson J-P, Stoeckel S (2015). Interpretation and approximation tools for big, dense Markov chain transition matrices in population genetics. *Algorithms Mol Biol* **10**: 31.
- Reichel K, Masson J-P, Malrieu F, Arnaud-Haond S, Stoeckel S (2016). Rare sex or out of reach equilibrium? The dynamics of FIS in partially clonal organisms. *BMC Genet* **17**: 76.
- Rice SH (2004). *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates: Sunderland, MA, USA.
- Rousset F (2004). *Genetic Structure and Selection in Subdivided Populations (MPB-40)*. Princeton University Press: Princeton, NJ, USA.
- Stoeckel S, Grange J, Fernández-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species – *Prunus avium* L. *Mol Ecol* **15**: 2109–2118.
- Stoeckel S, Masson J-P (2014). The exact distributions of Fis under partial asexuality in small finite populations with mutation. *PLoS One* **9**: e85228.
- Strathdee AT, Bale JS, Block WC, Webb NR, Hodgkinson ID, Coulson SJ (1993). Extreme adaptive life-cycle in a high arctic aphid, *Acyrtosiphon svalbardicum*. *Ecol Entomol* **18**: 254–258.
- Vanoverbeke J, De Meester L (2010). Clonal erosion and genetic drift in cyclical parthenogens—the interplay between neutral and selective processes. *J Evol Biol* **23**: 997–1012.
- Vorburger C (2006). Temporal dynamics of genotypic diversity reveal strong clonal selection in the aphid *Myzus persicae*. *J Evol Biol* **19**: 97–107.
- Vorwerk S, Forneck A (2006). Reproductive mode of grape phylloxera (*Daktulosphaira vitifoliae*, Homoptera: Phylloxeridae) in Europe: molecular evidence for predominantly asexual populations and a lack of gene flow between them. *Genome* **49**: 678–687.
- Wang J (1996). Deviations from Hardy-Weinberg proportions in finite populations. *Genet Res* **68**: 249–257.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)