

The Effect of Multiple Paternity on Genetic Diversity of Small Populations during and after Colonisation

Marina Rafajlović¹, Anders Eriksson^{2,3}, Anna Rimark¹, Sara Hintz-Saltin⁴, Grégory Charrier⁴, Marina Panova⁴, Carl André⁴, Kerstin Johannesson⁴, Bernhard Mehlig^{1*}

1 Department of Physics, University of Gothenburg, Gothenburg, Sweden, **2** Department of Zoology, University of Cambridge, Cambridge, United Kingdom, **3** Integrative Systems Biology Lab, King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia, **4** Department of Biological and Environmental Sciences-Tjärnö, University of Gothenburg, Strömstad, Sweden

Abstract

Genetic variation within and among populations is influenced by the genetic content of the founders and the migrants following establishment. This is particularly true if populations are small, migration rate low and habitats arranged in a stepping-stone fashion. Under these circumstances the level of multiple paternity is critical since multiply mated females bring more genetic variation into founder groups than single mated females. One such example is the marine snail *Littorina saxatilis* that during postglacial times has invaded mainland refuge areas and thereafter small islands emerging due to isostatic uplift by occasional rafting of multiply mated females. We modelled effects of varying degrees of multiple paternity on the genetic variation of island populations colonised by the founders spreading from the mainland, by quantifying the population heterozygosity during both the transient colonisation process, and after a steady state (with migration) has been reached. During colonisation, multiple mating by 2–10 males increased the heterozygosity by 10–300% in comparison with single paternity, while in the steady state the increase was 10–50% compared with single paternity. In the steady state the increase of heterozygosity due to multiple paternity is determined by a corresponding increase in effective population size. During colonisation, by contrast, the increase in heterozygosity is larger and it cannot be explained in terms of the effective population size alone. During the steady-state phase bursts of high genetic variation spread through the system, and far from the mainland this led to short periods of high diversity separated by long periods of low diversity. The size of these fluctuations was boosted by multiple paternity. We conclude that following glacial periods of extirpation, recolonization of isolated habitats by this species has been supported by its high level of multiple paternity.

Citation: Rafajlović M, Eriksson A, Rimark A, Hintz-Saltin S, Charrier G, et al. (2013) The Effect of Multiple Paternity on Genetic Diversity of Small Populations during and after Colonisation. PLoS ONE 8(10): e75587. doi:10.1371/journal.pone.0075587

Editor: Dirk Steinke, Biodiversity Institute of Ontario - University of Guelph, Canada

Received: April 9, 2013; **Accepted:** August 15, 2013; **Published:** October 28, 2013

Copyright: © 2013 Rafajlović et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors gratefully acknowledge financial support by Vetenskapsrådet, by the Goran Gustafsson Foundation for Research in Natural Sciences and Medicine, and by the platform “Centre for Theoretical Biology” at the University of Gothenburg. Further support was given from the Linnaeus Centre for Marine Evolutionary Biology (CeMEB, www.cemeb.science.gu.se). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Bernhard.Mehlig@physics.gu.se

Introduction

When local populations are established or re-established after local extirpation, genetic variation within the newly founded populations is initially governed by the genetic content of the founders, which, in the most extreme case, may be a single fertilised female. If new populations are established in a stepping-stone fashion at increasing distances from a large source population, the genetic content of founder groups is even more important, since loss of genetic variation by drift is expected to increase as the number of colonisation steps becomes larger. At a later stage, during continued input of variation through migration, the genetic composition of migrants may in a similar way contribute new variation and hence counteract loss by drift and selection.

Furthermore, the genetic variation carried by the founders is expected to be influenced by multiple paternity in brooding and sexual species. Females mating multiple males usually have broods of offspring sired by more than one male, although sperm competition or cryptic female choice may cause deviations from a

random distribution of paternity [1]. Multiple paternity, once believed to be rare in nature, is observed in a number of animal species including mammals, amphibians, fishes, reptiles and invertebrates [2–5]. In most species levels of multiple paternity are around two to four males, but in some species of fish and invertebrates it is rather six to ten [6–9]. The marine snail *Littorina saxatilis* is an extreme example with 15–23 males siring broods of single females [10].

In the near future, whole-genome sequences will become available for many species. In order to be able to deduce the factors driving the long-term evolution of natural populations from genome-wide patterns of genetic variation, it is necessary to quantify the effects of life history (e. g. mating patterns, multiple paternity) and geography (e. g. population structuring). The aim of this paper is to quantitatively understand how multiple paternity and geographic structure in the form of a stepping-stone model determine patterns of neutral genetic variation. The marine snail *Littorina saxatilis* is an example for which a stepping-stone colonisation model with a mainland as a source population

describes the establishment of new populations. This species is strictly intertidal and most abundant in rocky shores in the north Atlantic, with population densities of tens to hundreds of snails per square metre [11]. In contrast to many other marine snails, *L. saxatilis* does not have pelagic eggs or larvae, and therefore dispersal over a few tens of metres range is infrequent. However, snails occasionally migrate among islands, probably by rafting. It has been estimated that within an archipelago of small and large islands, 3% of the small islands receive a migrant snail each year [12]. In many areas, *L. saxatilis* forms local populations inhabiting discrete localities, such as islands of an archipelago, rocky outcrops and breakwaters intermingled by sandy substrates [13,14]. During the retreat of the ice sheet 12 000–15 000 years ago, *L. saxatilis* spread from refuge areas both in the northern Atlantic and south of the ice-cap [10]. Part of this postglacial expansion comprised colonisation of hundreds of islands in the archipelago along the Swedish west coast that successively became available by isostatic uplift, a process that is still ongoing. In this system, populations on the mainland and large islands are the oldest and largest, and these are likely to act as the ultimate sources of genetic variation during colonisation of emerging islands in a stepwise manner (Fig. 1 A). We have re-analysed genetic data from *L. saxatilis* populations in the archipelago of west Sweden (given in Table 1 in [13]) using a principal component analysis. The data consist of allelic frequencies at four polymorphic allozymes (we excluded the locus *Aat-1*) taken at thirteen skerries (skerry sizes $\approx 10m^2$), seven islands (island sizes $\approx 10^5m^2$), and five mainlands. We have found that the first principal component shows a largely linear relationship between population-genetic variation and size/age of the islands/populations, with mainland populations at the one end, skerry populations at the other end, and island populations in the middle (Fig. 1 B). This, together with the findings of [12] (see above), suggests that a section of the archipelago consisting of the neighbouring mainland-, island- and skerry-population can be modelled by a simple linear stepping-stone model with the mainland population acting as a source for colonisation of islands at successively younger age, and at increasing distance from the mainland (Fig. 1 C).

An important well-known characteristic of *L. saxatilis* is multiple paternity [10]. In the following, we investigate how multiple paternity affects the spatial and temporal structure of neutral genetic diversity of subpopulations arranged in a linear stepping-stone model. With the extreme level of multiple paternity of the marine snail in mind, we construct a mating model for this species and analyse how multiple paternity affects population-genetic variation and structure in a population distributed over both mainland sites and adjacent small island habitats. In our analysis, we consider established local populations in a global steady state, but also populations under establishment (during initial colonisation of previously empty habitats). We first derive an expression for the effective genetic population size [15] resulting from the mating behaviour observed in earlier empirical studies. This further allows us to derive simple approximations describing the genetic diversity of subpopulations during colonisation, and in the steady state with migration. We use simulations to assess the temporal effects of migration on genetic diversity as new mutations from the mainland spread to distant islands.

Description of the Model

Stepping-stone colonisation model

We construct a stepping-stone colonisation model with the following basic assumptions mimicking how *L. saxatilis* colonised the post-glacial archipelago of western Sweden (and continues to

colonise new emerging islands). First, colonisation was occasional and rapid, as rafted fertilised females release a few hundred offspring already in the first generation [12]. Second, small skerries are likely to be colonised within a few years after emergence, and hence all newly established populations are limited by the small size of the habitat, resulting in census sizes of $\approx 10^2 - 10^3$ [12]. Third, colonisation is likely to take place in a stepping-stone manner with smaller and more distantly related islands being colonised from their closest already colonised islands. For simplicity, we consider a system consisting of a mainland and of linearly arranged islands of equal carrying capacities (substantially smaller than that of the mainland).

In our model, islands are linearly arranged and numbered from 1 to k , with k being furthest away from the mainland (see Fig. 1 C). We include high values of k in our model (such as $k=10$) in order to be able to assess saturation effects. The mainland is labelled by 0. Generations are assumed to be non-overlapping. Before the process of colonisation starts, the mainland is the only populated habitat, and the population heterozygosity on the mainland is stationary (that is, the mainland population is assumed to be old). We assume that mutations accumulate according to the infinite-alleles model [16].

Within our model, an empty island becomes fully colonised in a single generation after the arrival of one or more founder females from the nearest neighbouring island. This is motivated by the very large capacity for population growth of *L. saxatilis* in a suitable habitat [12,13]. In our model the founder females give rise to $2N$ offspring in total, with equal sex ratio, where $2N$ denotes the carrying capacity of the island. Upon establishing a given island population, its population size remains constant over time. In our model, mating takes place before migration, which allows us to trace only the movement of females (males also migrate, but since they will not mate after migration, they do not contribute to the progeny on the island they migrated to). Individuals are equally likely to migrate to each of their closest neighbours (but the mainland and the last island have only a single neighbour, Fig. 1 C). On average, M females migrate per generation from one island to a neighbouring island, except for empty islands that only receive migrants.

In addition to the above, we assume that the population size on the mainland is much larger than the population size of a colonised island. In our computer simulations we set the mainland heterozygosity to unity. This simplifies the simulations, since the dynamics of the mainland does not need to be simulated explicitly.

An important aspect of our model is the mainland, providing a source of genetic variation through repeated founder events. We remark that our model differs essentially from the mainland-island model analysed in [17] (see also references cited in this work). In [17] it is assumed that all island populations (called sink populations in [17]) either receive migrants from a source simultaneously, or are equally likely to receive migrants. In this case the degree of genetic variation in the steady state turns out to be the same for all islands. This is also true for stepping-stone models without a mainland (see for example Refs. [18,19]).

Mating model

In order to study the consequences of multiple paternity for genetic diversity, we introduce a mating model to describe different levels of multiple paternity in mating systems.

Based on known life-history traits [13], we assume that the duration of the reproductive cycle of females is the same for all females. Each female carries beneath her shell juveniles of varying degrees of maturity, and juveniles are released at an approximately steady rate. We also assume that after a successful mating, the

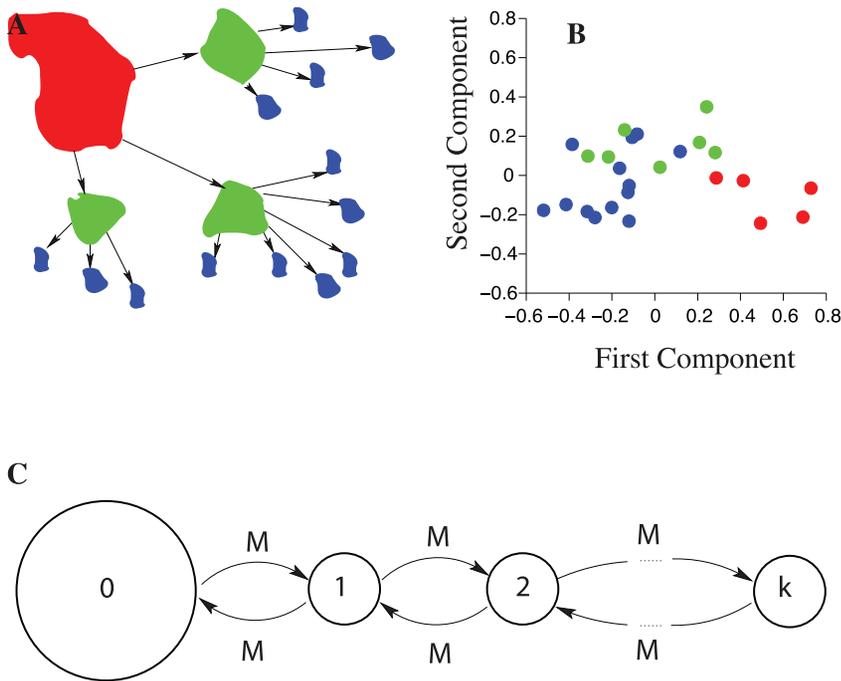


Figure 1. Spatial structure of the marine habitats of *Littorina saxatilis*. (A) A schematic illustration of the geographical structure of populations with mainland (red), islands (green), and skerries (blue). (B) Principal components of allozyme population differentiation in *L. saxatilis* (data from [13], the presumably selected locus *Aat-1* is excluded). Populations are classified as mainlands (red), islands (green), or skerries (blue). (C) Stepping-stone model of a section of the archipelago, with the mainland (labelled by 0) acting as a source for establishing the island populations (1 to *k*). On average, each island sends *M* females per generation to its right and to its left neighbouring island (but, note that the mainland and the island furthest from the mainland have only one neighbour). doi:10.1371/journal.pone.0075587.g001

mated female obtains a sperm package able to fertilise female eggs during its persistence time. Our observations show that sperm can be stored and used up to a year after mating. The number of eggs fertilised by a single sperm package is assumed to be the same for all sperm packages that a female receives during the reproductive cycle. The total number of active sperm packages received by each female during the reproductive cycle is denoted by *l*, and *l* is assumed not to depend on time (it has the same value in every generation).

The probability *p* that two eggs are fertilised by the same sperm package is $p = l^{-1}$ assuming that all sperm packages a female received during her reproductive cycle persist until the end of the reproductive cycle, that all eggs are fertilised after all sperm packages have been collected, and that sperm packages are chosen with replacement to fertilise eggs.

The scheme presented above models the process of mating at an individual level. We assume that individuals belong to a well mixed diploid population of N_m males and N_f females, and we take $N_m \gg 1$ and $N_f \gg 1$. In our model, each female encounters $s \leq N_m$ different males during the reproductive cycle (thus, *s* is assumed to be the same for all females, and it does not depend on time). Having $s < N_m$ reflects the limited movement of snails during the reproductive cycle. Unlike under random mating, where all males are treated as on average equally successful mates, it is possible that females exhibit cryptic choice, or that some of the available mates of each female are more dominant than others. We include this in our model by assigning different levels of dominance to the available mates: we assume that one of the *s* males (randomly chosen) has on average a higher success in mating, whereas the remaining *s* - 1 males have lower success. The dominant male mates the female with probability α , whereas the remaining *s* - 1

males mate the female with probability $(1 - \alpha)/(s - 1)$. We also assume that all females are on average equally successful mothers.

In our model the probability that two offspring come from the same female, P_f , is given by

$$P_f = \frac{1}{N_f} \tag{1}$$

Note that this implies that the brood sizes of females are multinomially distributed, similarly to the commonly used Wright-Fisher population [20,21]. Furthermore, the probability that two offspring come from the same male, P_m , is

$$P_m = \frac{1}{N_f} \kappa + \left(1 - \frac{1}{N_f}\right) \frac{1}{N_m} \tag{2}$$

In Eq. (2), the first term corresponds to the probability that two offspring share both a mother and a father, and the second term is the probability that two offspring come from different mothers but they share a father. The factor κ stands for the probability that two offspring having the same mother share a father, and it is given by

$$\kappa = p + (1 - p) \left(\alpha^2 + \frac{(1 - \alpha)^2}{s - 1} \right) \tag{3}$$

Here, the first term is the probability that two eggs are fertilised by the same sperm package, and the second term stands for the probability that two eggs are fertilised by two distinct sperm packages, both coming from the same male. The probability that two packages come from the same male has two contributions.

The first contribution, α^2 , stands for the probability that both packages come from the dominant male, and the second contribution, $(1-\alpha)^2/(s-1)$, is the probability that two packages come from one of the remaining $s-1$ males. Note that, according to Eq. (3), κ decreases as the number of available mates, s , of a female increases (keeping the values of p and of α fixed). Therefore, we take κ^{-1} as a measure of the degree of multiple paternity.

Using Eqs. (1)–(3) we find that the effective size of a freely mixing population under the model described is given by:

$$N_e = 4 \left(\frac{1+\kappa}{N_f} + \frac{1}{N_m} \right)^{-1}. \quad (4)$$

The mathematical details of this derivation are given in **Appendix S1**. Our model reduces to the model in [22] in the case a female encounters all males present in the population, $s=N_m$, upon assuming that all males are on average equally successful mates, $\alpha=1/N_m$, and upon identifying the number of matings in [22] with p^{-1} . Furthermore, if $s=N_m$, $\alpha=1/N_m$, and the probability that two eggs are fertilised by the same sperm package is $p=0$, our model reduces to random mating. We also note that in two particular cases, our model reduces to two mating models analysed in [23]. First, for $p=1/l$, $\alpha=1/s$, and $s=N_m$ (or $l=1$), our model is analogous to the model of lottery polygyny in [23] in the case all females mate with certainty, and assuming non-overlapping generations (that is, upon setting $\alpha_f=1$, and $v=0$ in [23]). Second, for $s \ll N_m$, $\alpha=1/s$, and $p=0$, our model reduces to the model of dominance polygyny in [23] in the case of non-overlapping generations, and upon assuming that all males are dominant in this model (that is, for $v=0$, and $\alpha_m=1$ in [23]).

We show in Fig. 2 how the effective population size under our model depends upon p and s . As expected, by increasing the degree of multiple paternity, the effective population size becomes larger (as found in [24] for a different mating model). For the parameters set in Fig. 2, the maximum value of N_e is equal to $N_m + N_f$, which corresponds to the effective population size under

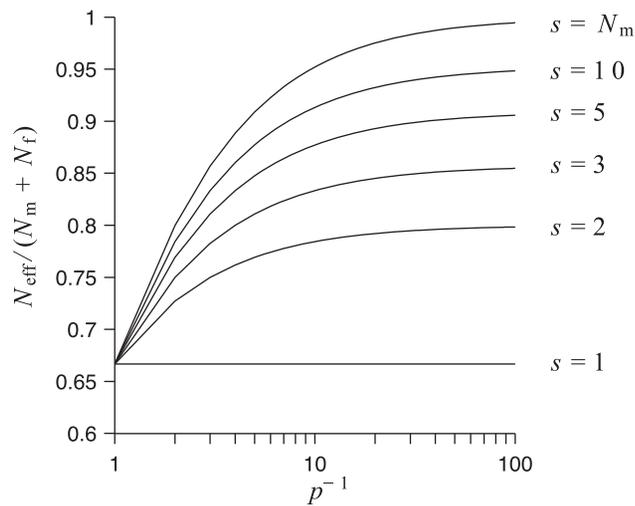


Figure 2. Effective population size. For different number of available mates, s , and probability that two eggs are fertilised by the same sperm package, p , we show the ratio of the effective population size (N_{eff}) to the total population size ($N_m + N_f$). Parameters: $\alpha=1/s$, $N_m=N_f=10^3$. doi:10.1371/journal.pone.0075587.g002

random mating. The increasing trend of N_e saturates at $s \approx 10$ for a given value of p and α .

We used empirical data to test our mating model from two experiments. In one experiment, six virgin females of similar sizes were placed in separate aquaria, and each was accompanied by ten equally sized males. The snails were allowed to mate freely during eight weeks, after which the males were removed, and offspring were collected from each female over a one-year period following the mating period. The sire of each offspring was determined by genotyping the males, females and juveniles at eight microsatellite loci and performing parentage analysis following an exclusion method (Saltin, S. H. *et al*, unpublished data). Only the offspring that could uniquely be assigned to a given sire were kept. Hence, 12 out of 176 offspring were removed from the analysis.

In the second experiment, genotype data from five microsatellite loci were collected in four wild females and their offspring [10]. The brood sizes of the four wild females were 79, 77, 71, and 53 [10]. Since the potential sires were unknown, COLONY [25] was used to estimate the sibships of each female’s offspring. In order to test the goodness of our model, we fitted our mating model to the distribution of male family sizes in the empirical data, using a chi-square test to assess goodness of fit (details of binning and degrees of freedom are given in the results section). Because the circumstances in the controlled mating experiments differ from that of females mated in the wild, we fitted the two datasets separately.

The brood sizes of the six females mated under experimental conditions in the lab was 59, 32, 25, 24, 19, and 17 (Saltin, S. H. *et al*, unpublished data). For the offspring of each female it was determined how many came from the same father, that is, male family sizes. For the snails mated in the lab, the maximum male family size was 25 (1 male). Remaining male families were of sizes 14 (1 male), 11 (1 male), 10 (1 male), 9 (1 male), 7 (3 males), 6 (1 male), 5 (4 males), 4 (5 males), 3 (3 males), 2 (3 males), and 1 (13 males). To avoid ambiguity in paternity, 12 offspring were excluded from male family assignment, yielding broods of sizes 59, 27, 22 (two broods), 19, and 15 for further analysis. Similarly, for the snails mated in the wild, the male families were of sizes 10 (2 males), 8 (2 males), 7 (2 males), 6 (4 males), 5 (12 males), 4 (21 males), 3 (10 males), 2 (8 males), and 1 (16 males). For the details on parental assignment in this case, see [10].

Results

Distribution of male family sizes

In Fig. 3 **A** we show the histogram of male family sizes obtained experimentally and in Fig. 3 **B** we show similar data collected from wild females in natural habitats. For both sets of data we use computer simulations in order to find the parameters in our model resulting in male family sizes that are closest to those empirically observed. In the computer simulations we vary the number of available mates s , the probability p , and the level of dominance of a dominant male, α . For the females mated in the lab, we vary the parameter s from 6 to 10, and for the wild females, we vary s from 6 to 30. In both cases, we vary the parameter α from 0.05 to 0.9 in steps of 0.05, and the inverse of parameter p from 6 to 30 in steps of 2, and we test $p=0$ as well. For each set of test values of s , α , and p , we simulate mating of six females (four for wild snails) and we generate broods of sizes corresponding to that in the empirical data. By simulating the process of mating 1000 times for each set of test parameters, we compute the average probability of obtaining a given male family size. Using these data, we apply a chi-square test to quantify the difference between the empirical and simulated data for each set of test parameters. The best fits

obtained are shown in Fig. 3 **A–B** by circles, and they correspond to $s=9, p=1/30, \alpha=0.35$ (Fig. 3 **A**), and to $s=20, p=0, \alpha=0.1$ (Fig. 3 **B**). We discuss these fits in Section **Discussion and conclusions**. The results shown in Fig. 3 suggest that our mating model describes empirical data well, but the agreement is not perfect. Possible improvements of our model are discussed in Section **Discussion and conclusions**.

To address the question of how multiple paternity affects genetic variation and structure in our subdivided population, we analyse genetic diversity under the colonisation model described above. We analyse separately two phases of population dynamics on each island: initial colonisation, and the steady state that develops once the colonisation phase is over. For a given island, we compute the expected heterozygosity in the generation when the island is colonised (colonisation phase), as well as the expected heterozygosity in the steady state. The corresponding analytical computations are described in detail in **Appendix S2** and **Appendix S3**. We also study temporal changes of heterozygosity under our model by computer simulations. In the following two subsections we present separately the results for the colonisation phase and for the steady state.

Colonisation phase

We compute the heterozygosity in the colonisation phase analytically using a coalescent approach [26]. Recall that at a given island, the colonisation-phase heterozygosity corresponds to the island heterozygosity in the generation when the island is colonised. Note that the population-size history of the population on island i at the time when this island is colonised can be represented as a sequence of i bottlenecks. In our model, each of the bottlenecks lasted one generation, and the most recent bottleneck occurred when the first colonisers (founder females) came to island i . Note that there are i bottlenecks, because i is the number of colonisation events that the population ancestral to that on island i went through before the island was colonised. We derive an analytical result upon assuming that the migration rate is small, $M \ll 1$. Under this assumption, colonisation of empty islands occurs rarely, but when it does, an island typically receives a single founder female. The resulting heterozygosity in island i in the colonisation phase, $H_c^{(i)}$, is (see **Appendix S2**):

$$H_c^{(i)} = P(0|i)H^{(0)}. \quad (5)$$

Here, $H^{(0)}$ is the mainland heterozygosity, and $P(0|i)$ is the probability that the most recent common ancestor of two alleles sampled from the newly established population in island $i > 0$ stems from an allele that was born on the mainland (see **Appendix S2**):

$$P(0|i) = \left(1 - \frac{1}{8}(1 + \kappa)\right)^i \left(\frac{2MN_e}{2MN_e + 1}\right)^{i-1}. \quad (6)$$

It follows from Eqs. (5) and (6) that the dependence of the heterozygosity in the colonisation phase on the degree of multiple paternity κ^{-1} can not be expressed only in terms of the effective population size. This suggests that the degree of multiple paternity has possibly a larger effect on the colonisation-phase heterozygosity in subpopulations, than on the heterozygosity of freely mixing populations.

In Fig. 4 **A** we show how the colonisation-phase heterozygosity depends on distance from the mainland. We observe two expected features. First, the colonisation-phase heterozygosity decays as distance from the mainland increases. Second, at any distance from the mainland, multiple paternity results in higher heterozygosity than single paternity. We find that mating two males ($s=2$) increases the values of single-paternity heterozygosity by 10–100% for the parameters used in Fig. 4 **A**, and mating ten males ($s=10$) increases the values of single-paternity heterozygosity by 10–300%, depending on distance from the mainland (Fig. 4 **B**). The largest increase is observed at the island furthest from the mainland. It is interesting to note that mating more than about 10 males only marginally increases the heterozygosity (results not shown), as found in the case of freely mixing populations (see our description of the mating model).

We note that the results of our computer simulations (see Fig. S1 **A–C**) agree well with the analytical results for low migration rates. For large migration rates ($M=0.5$, i. e. 0.5 females on average per generation) by contrast, the simulations assume somewhat higher values than the theory. The reason for this deviation is that for $M=0.5$ it is probable that more than one founder female come to

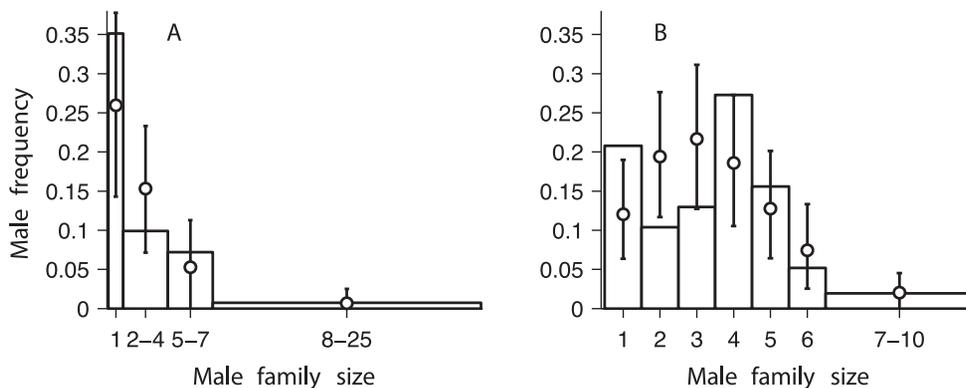


Figure 3. Distribution of number of progeny attributed to each sire. Bars in panel **A** show the empirical data obtained under experimental conditions; the data correspond to six broods, two of which have size 22, and the remaining four have sizes 15, 19, 27, and 59. Bars in panel **B** show the results from females mated in the field, data taken from [10]; the data correspond to four broods of sizes 79, 77, 71, and 53. The width of the bins are chosen so that the expected number of counts in each bin is not smaller than 5. The probability assigned to each bar is proportional to the bar area. Symbols and error bars show the result of the best fit to the experimental data, together with their 95% confidence intervals: $s=9, p=1/30, \alpha=0.35$ in **(A)**, and $s=20, p=0, \alpha=0.1$ in **(B)**. We simulated 10^3 independent realisations of the mating process. doi:10.1371/journal.pone.0075587.g003

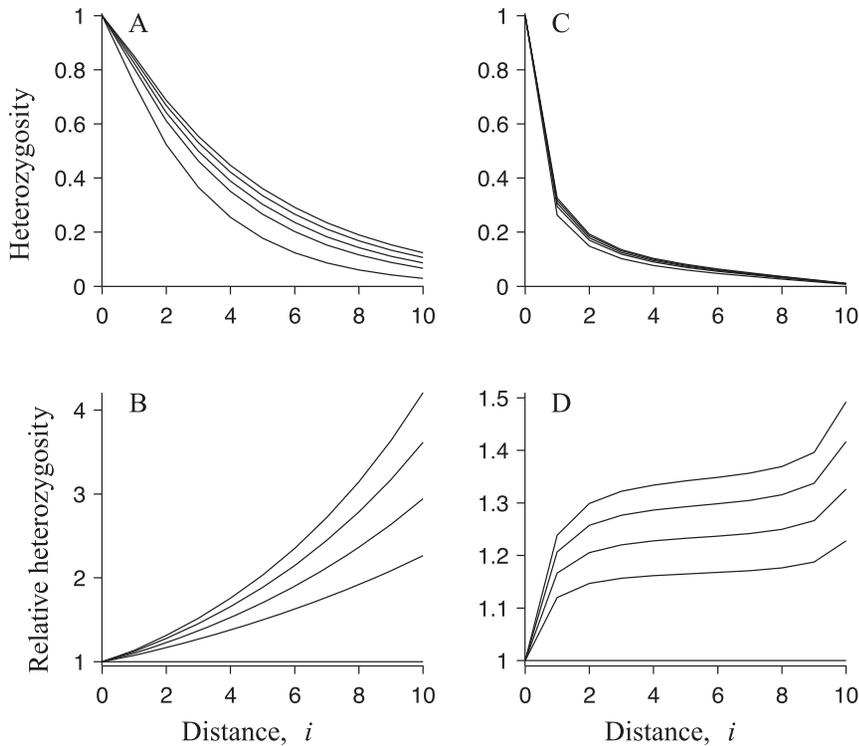


Figure 4. Analytically computed heterozygosity during colonisation and in the steady state. (A) Heterozygosity during colonisation as a function of distance from the mainland. The lines shown from top to bottom correspond to: $s=10$, $s=5$, $s=3$, $s=2$, and $s=1$. (B) Heterozygosity during colonisation relative to its corresponding value for $s=1$ as a function of distance from the mainland. The order of lines corresponds to that in A. (C) and (D) Same as in (A) and (B), respectively, but for the steady state. Remaining parameters: all available males are on average equally successful, $\alpha=1/s$, the mainland heterozygosity is set to unity, the scaled female migration rate is $M=0.05$, the number of females in each populated island is $N=100$, the probability that two eggs are fertilised by the same sperm package is $p=0.1$, and the number of islands is $k=10$. doi:10.1371/journal.pone.0075587.g004

an empty island to establish the population, and, consequently, will contribute with more genetic variation than just one founder female. For natural populations of *L. saxatilis* it has been estimated that 3% of empty islands receive a migrant each year [12]. This estimate is close to the lower value of M used in our simulations ($M=0.05$).

Steady state

We show in **Appendix S3** how to compute the heterozygosity in the steady state at distance i from the mainland using recursion relations (this is a standard procedure, see for example [19,21]). Note that this derivation does not require M to be small. In Fig. 4 C, we show how the resulting steady-state heterozygosity depends on distance from the mainland. As in the colonisation phase, the steady-state heterozygosity decreases as distance from the mainland increases. Also, by increasing the degree of multiple paternity, the heterozygosity at a given island increases (this effect saturates at $s \approx 10$, results not shown). In contrast to the strong effect of multiple paternity during colonisation, the effect is substantially smaller in the steady state. We find that the single-paternity heterozygosity in the steady state increases by 10–20% for $s=2$, and by 20–50% for $s=10$ (Fig. 4 D). As in the colonisation phase, the largest increase is observed at the island furthest from the mainland.

In addition, we examined the variation in heterozygosity over consecutive generations in a particular realisation of our model. We find that the heterozygosity shows strong temporal fluctuations (Fig. 5 A). Notably, the fluctuations are strongest furthest from the

mainland, with periods of high diversity separated by long periods of near or complete fixation. Hence the distribution of heterozygosity at large distances from the mainland is bimodal. The heterozygosity is expected to have a bimodal distribution if genetic variation is introduced at a very small rate (see, for example, [27] and references cited therein).

In what follows, we analyse how the durations of the phases of low and high heterozygosity are affected by multiple paternity. We consider values of the heterozygosity smaller than 0.1 to be low, and values of the heterozygosity larger than 0.4 to be high. The minimum value for the high phase (0.4) is chosen since the typical maximum value that the heterozygosity has at the island furthest from the mainland is ≈ 0.5 for the parameters chosen in Fig. 5 A. Note that the maximum value of the heterozygosity at a locus with only 2 allelic types is equal to 0.5.

Using the results of our computer simulations, we compute the average durations of low- and high-heterozygosity phases at the island furthest from the mainland. We also derive corresponding analytical results under the assumption that the scaled migration rate M is small (see **Appendix S4**). For island $i=10$ we show in Fig. 5 B (see also Fig. S2 B, D, F) the durations of low- and high-heterozygosity phases relative to their corresponding values for a single mate ($s=1$). Fig. 5 B shows that multiple paternity prolongs the duration of the high-heterozygosity phase, and decreases the duration of the low-heterozygosity phase. For example, the high-heterozygosity phase for the highest level of multiple paternity shown ($s=10$) is prolonged by around 40% compared to its value under single paternity ($s=1$). The low-heterozygosity phase is

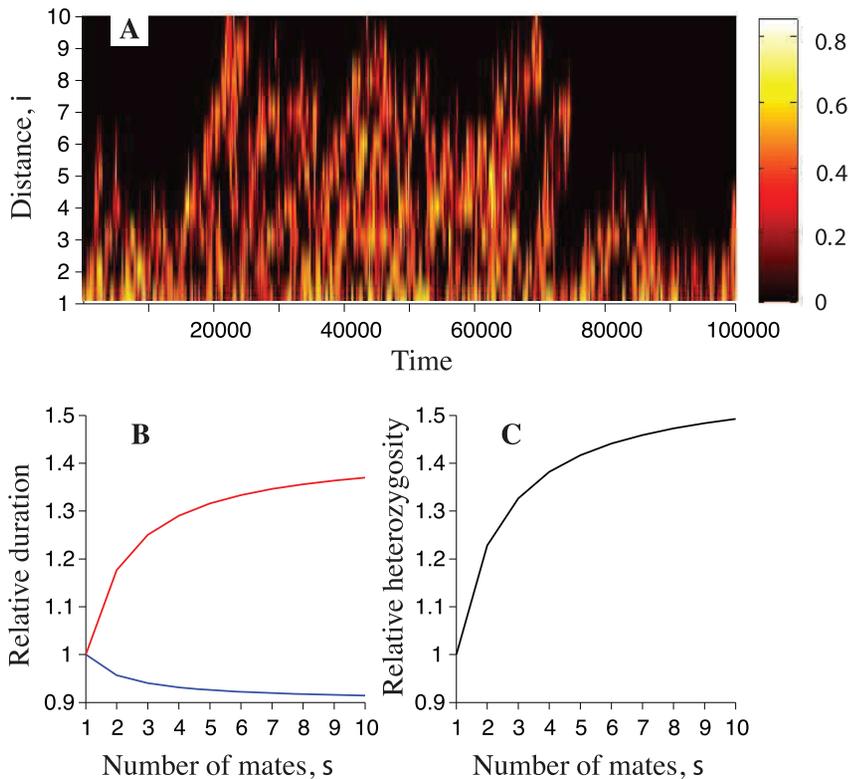


Figure 5. Temporal fluctuations of heterozygosity. (A) Heterozygosity as a function of distance from the mainland and of time (single realisation of the model described). Mainland is not shown. The data correspond to 10^5 generations after the initial $7 \cdot 10^6$ generations. The number of available mates is $s = 10$. (B) Analytically computed durations of low- and high-heterozygosity phases (blue, and red) relative to their corresponding values for $s = 1$. (C) Analytically computed steady-state heterozygosity (black) relative to its corresponding value for $s = 1$. Remaining parameters are as in Fig. 4 C.

doi:10.1371/journal.pone.0075587.g005

shortened by only around 10% for $s = 10$ (Fig. 5 B). For comparison, Fig. 5 C shows the steady-state heterozygosity relative to its corresponding value for a single mate ($s = 1$). In conclusion, multiple paternity promotes heterozygosity by prolonging the duration of peaks of variation that reach the most distant islands.

Discussion and Conclusions

In this study we analysed the effect of multiple paternity on genetic diversity over spatial and temporal scales in a population living in islands distributed in one dimension. We quantified the effect of multiple paternity during the colonisation of the semi-isolated islands and in the steady state developed after the colonisation phase. Our conclusions given below can be generalised to a population inhabiting relatively small patches that are partly isolated from each other and approximately linearly distributed, such as, rocky outcrops along a sandy coast, sandy beaches along a rocky coast, distinct patchy habitats along a stream or a river, patchy habitats found at a specific mountain altitude, small lakes at an increasing distance from a larger lake, etc..

We introduced a mating model which allows for different levels of multiple paternity in a population. As expected, we found that by increasing the degree of multiple paternity within our mating model, the effective population size increases, and thus the heterozygosity increases. Yet, the consequences of multiple paternity on heterozygosity in relatively small and semi-isolated populations are substantial, which is summarised in the following.

At a given distance from the mainland, populations with high degree of multiple paternity establish higher heterozygosity than populations with low degree of multiple paternity. In the steady state, this effect is expected, since the heterozygosity in the steady state depends on the effective population size and the rate of income of new genetic material to a given habitat. It is also expected that this effect saturates at the same degree of multiple paternity as the effective population size of a freely mixing population, which is consistent with our findings. By comparing the effect of multiple paternity on the heterozygosity in the steady state to that on the heterozygosity in the colonisation phase, we find that the latter is substantial. This difference between the steady state and the colonisation phase is explained in the following. Upon the arrival of founder females to an empty island, the carrying capacity of an island is reached within a single generation. Therefore, such a newly established population receives genetic material of most males that the founder females were inseminated by. By contrast, in the steady state, all mothers present in an island contribute to the population in the next generation, and hence the impact of immigrant females to the next generation is rather small. From this reasoning, we find that it is possible that the effect of multiple paternity upon heterozygosity during colonisation might decrease if the growth rate of the island populations up to the carrying capacity were less than infinite (as assumed in our model).

The heterozygosity at distances far from the mainland fluctuates significantly. Long periods of almost complete loss of genetic variation are interrupted by bursts of high heterozygosity, and we

quantified how this effect is boosted by multiple paternity. The durations of high- and low-heterozygosity phases could be an important survival factor in natural populations. For example, the low-heterozygosity phase could be disadvantageous if a malignant disease appears in the population, assuming that only a particular mutation (not present in the population, or being rare) can survive the disease. Moreover, in some species evidence was found for selection acting against homozygous individuals [28]. The wave-like nature of the spread of new alleles from the mainland population is also seen in the correlation of genetic diversity at neighbouring islands. We find that the correlation between heterozygosities at a pair of nearest-neighbouring islands increases as distance from the mainland increases (results not shown). These results suggest intermittent bursts of genetic diversity in remote islands, an effect which becomes stronger as the degree of multiple paternity increases. We note that a similar trend is expected in our model even when the process of mutation on islands is explicitly included in simulations, as long as the migration rate is not too small. This is because for a given value of the scaled mutation rate on the mainland, the scaled mutation rate on islands inevitably approaches zero (the mainland is infinitely larger than islands).

The conclusions given above are confirmed by our computer simulations. In order to minimise computing time during simulations, we assumed that the mainland heterozygosity is equal to unity, which guarantees that whenever a migrant from the mainland comes to the first island, it carries genetic material that previously did not exist on any of islands (and thus the population dynamics on the mainland does not need to be simulated explicitly). However, we emphasise that the conclusions given above qualitatively do not depend on the value of heterozygosity on the mainland.

We note that, unlike in our model, it is possible that the rate of successful colonisation in natural habitats is smaller than the rate of migration. For example, if an immigrant female carries a small number of progeny, her progeny alone may not be enough to colonise an empty island successfully. By allowing for the rate of successful colonisation to be smaller than the rate of migration in the colonisation model, the steady-state values of heterozygosity remain equal to those obtained under the assumption that the colonisation and migration rates are the same (as in our model). The values of heterozygosity during colonisation, by contrast, are expected to differ from those found here.

In order to determine how realistic our mating model is, we compared the male family sizes of female broods obtained within our model, to empirically observed family sizes in populations of *L. saxatilis* from natural habitats and under experimental conditions. By computing chi-square values, we found that empirical male family size distributions were in agreement with our mating model, that is our mating model could not be rejected. The best-fitting parameters for the data obtained under experimental conditions indicated that the number of available mates for a female is on average less than the number of males that a female was surrounded by in her aquarium. This might be due to the success of mates is variable, so that some of the available mates did not mate at all during the time of the experiments (males were removed after eight weeks). Moreover, many males are likely to have mated multiple times during the experiment, while some may have mated just a few times, and for those males it is possible that they did not sire any offspring at all. In addition, females may exhibit cryptic choice of sperm. We also found that the best-fitting parameters for the data obtained under experimental conditions indicated fewer matings than brood sizes, suggesting that some of the eggs were fertilised by sperm retained between matings. By contrast, the corresponding empirical distribution in natural

populations was best fitted by assuming an unlimited number of matings (i. e. no sperm retention). This is consistent with the high density of snails observed in the wild. However, both empirical distributions showed an excess of males with a single progeny compared to the mating model with the best-fitting parameters. This discrepancy could be because the success of available mates in natural conditions is highly variable, or a female exhibits cryptic choice of sperm.

We note that for the females mated in the lab, only the offspring that could uniquely be assigned to a given sire were used in the analysis. This resulted in removing 12 out of 176 offspring from the parentage assignment. While this could in principle bias offspring towards males with more unique alleles, we estimated that the likelihood to inherit the same set of alleles from two males is very small for all pairs of males (the majority of pairs of males have at least two loci with no shared alleles). This also means that it is unlikely that genotyping errors would lead to misassignments. By contrast, for the data corresponding to wild females, genotypes of possible fathers were unknown. They were estimated using a maximum likelihood approach [10]. Since it is possible that there were a number of fathers with similar genotypes, the maximum likelihood method in [10] could have underestimated the true number of fathers. In particular, some of the large male families found might have consisted of a number of smaller families, as a given large male family could have belonged to multiple alike fathers. As a consequence, we expect that the assignment in this case was biased so that not only the total number of fathers was underestimated, but also the number of small male families. However, it was argued in [10] that, since the estimated number of fathers of the four broods analysed was large, it is likely that these estimates are very close to the corresponding true number of fathers. For this reason, we expect that the effect of genotyping errors in this case would also be small.

Since mating is considered to be costly [10], we raise the question whether or not mating multiple males is an evolutionary strategy of *L. saxatilis* to increase the heterozygosity. Recall that we estimated that in natural populations of *L. saxatilis* the probability that two eggs are fertilised by the same sperm package is likely to be very small. Under our model, this probability is equal to zero if each sperm package fertilises one egg, or if the actual number of sperm packages a female receives during her reproductive cycle is very large. If the latter applies, we find that it is unlikely that the heterozygosity increase is the main reason for the extreme multiple paternity in this species. As earlier suggested, it seems likely that the cost of rejecting an intercourse is higher than the cost of accepting it, and a consequence of convenience polyandry [10].

In summary, this study can be used to quantify the gene flow between partly isolated natural populations using allelic frequencies at a number of neutral loci. In particular, since the spatial patterns of heterozygosity during the colonisation phase differ from those in the steady state, this study can be used for determining whether or not the colonisation process started recently.

Furthermore, since our results show that the heterozygosity exhibits extreme fluctuations in populations founded through repeated founder events, we raise the question of whether similar fluctuations can be observed at any given time at neutral loci sampled genome-wide. In order to answer this and related questions, the effect of recombination needs to be analysed. Since island populations in our model experience at least one severe bottleneck, we expect that the degree of linkage disequilibrium in the colonisation phase is constant over a range of genetic distances, as shown in [29]. However, how multiple paternity affects linkage disequilibrium during colonisation and in the steady state is yet to

be investigated. It would also be interesting to analyse how selection combined with recombination affects genetic diversity in a subdivided population. Such results would provide an advance in the endeavor of identifying genes under selection, and especially, the genes underlying speciation [30–32].

All simulations in this study were written in Matlab. The simulations as well as their raw results used in this study are freely available upon request.

Supporting Information

Appendix S1
(PDF)

Appendix S2
(PDF)

Appendix S3
(PDF)

Appendix S4
(PDF)

Figure S1 Heterozygosity in the colonisation phase, and in the steady state. (A)–(C) Colonisation phase. The analytical results are shown as solid lines, and the results of computer simulations are shown as symbols. The number of available mates is $s = 1$ (blue), $s = 2$ (red), $s = 3$ (green), $s = 5$ (magenta), and $s = 10$ (black). Averages are over $2 \cdot 10^4$ realizations of the process of colonisation of empty islands. (D)–(F) Same as in (A)–(C) but for the steady state. In panel D, averaging is done over $1.5 \cdot 10^7$ generations (the initial 10^7 generations being discarded), and over three independent realisations of our model. In panel E, averages are over $4 \cdot 10^7$ generations (the initial $5.5 \cdot 10^7$ generations being

discarded) and over four independent realisations of our model. In panel F, averaging is done over $5 \cdot 10^7$ generations (the initial $5 \cdot 10^6$ generations being discarded) and over five independent realisations of our model. Remaining parameters used: all available males are on average equally successful, $\alpha = 1/s$, mainland heterozygosity $H^{(0)} = 1$, number of islands $k = 10$. (EPS)

Figure S2 Low- and high-heterozygosity phases. (A) Illustration of the method used to determine the duration of low- and high-heterozygosity phases. The panel shows the heterozygosity of the population on island 10 as a function of time. The heterozygosity represented in terms of the low and high phases is shown by the magenta line. The black line depicts the result of computer simulation. The data shown correspond to those in Fig. 5 A in the main text. (B), (D), and (F) Durations of low- and high heterozygosity phases relative to their corresponding values for $s = 1$ (blue and red, respectively) in island 10. (C), (E), and (G) Steady-state heterozygosity in island 10 relative to its corresponding value for $s = 1$. The results of computer simulations are shown as symbols, and the analytical results are shown as solid lines. The parameters in B, and C correspond to those in Fig. S1 D. The parameters in D, and E correspond to those in Fig. S1 E. The parameters in F, and G correspond to those in Fig. S1 F. (EPS)

Author Contributions

Conceived and designed the experiments: KJ MP CA. Performed the experiments: SHS GC. Analyzed the data: MR AE MP CA KJ SHS GC. Wrote the paper: MR AE KJ BM CA MP SHS GC. Performed computer simulations: MR AR. Formulated the question: KJ BM. Derived the theoretical results: MR AE BM.

References

- Birkhead TR, Pizzari T (2002) Postcopulatory sexual selection. *Nature Reviews Genetics* 3: 262–273.
- Birkhead TR (2000) *Promiscuity: an evolutionary history of sperm competition*. Cambridge (MA): Harvard University Press.
- Jennions MD, Petrie M (2000) Why do females mate multiply? A review of the genetic benefits. *Biol Rev Camb Philos Soc* 75: 21–64.
- Uller T, Olsson M (2008) Multiple paternity in reptiles: patterns and processes. *Molecular Ecology* 17: 2566–2580.
- Avise JC, Tatarenkov A, Liu JX (2011) Multiple mating and clutch size in invertebrate brooders versus pregnant vertebrates. *Proceedings of the National Academy of Sciences USA* 108: 11512–11517.
- Paterson IG, Partridge V, Buckland-Nicks J (2001) Multiple paternity in *Littorina obtusata* (Gastropoda, Littorinidae) revealed by microsatellite analyses. *The Biological Bulletin* 200: 261–267.
- Trontú K, Thurin N, Sundström L, Aron S (2007) Mating for convenience or genetic diversity? Mating patterns in the polygynous ant *Plagiolepis pygmaea*. *Behavioral Ecology* 18: 298–303.
- Brante A, Fernandez M, Viard F (2011) Microsatellite evidence for sperm storage and multiple paternity in the marine gastropod *Crepidula coquimbensis*. *J Exp Mar Biol Ecol* 396: 83–88.
- Coleman SW, Jones AG (2011) Patterns of multiple paternity and maternity in fishes. *Biological Journal of the Linnean Society* 103: 735–760.
- Panova M, Boström J, Hofving T, Areskoug T, Eriksson A, et al. (2010) Extreme female promiscuity in a non-social invertebrate species. *PLoS ONE* 5: e9640.
- Reid DG (1996) *Systematics and Evolution of Littorina*. London: The Ray Society.
- Johannesson K, Johannesson B (1995) Dispersal and population expansion in a direct developing marine snail (*Littorina saxatilis*) following a severe population bottleneck. *Hydrobiologia* 309: 173–180.
- Janson K (1987) Genetic drift in small and recently founded populations of the marine snail *Littorina saxatilis*. *Heredity* 58: 31–37.
- Johannesson K, Warmoes T (1990) Rapid colonization of Belgian breakwaters by the direct developer *Littorina saxatilis* Oliv. *Hydrobiologia* 193: 99–108.
- Ewens W (1982) On the concept of the effective population size. *Theoretical Population Biology* 21: 373–378.
- Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49: 725–738.
- Gaggiotti OE (1996) Population genetic models of source-sink metapopulations. *Theoretical population biology* 50: 178–208.
- Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49: 561–576.
- Austerlitz F, Jung-Muller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology* 51: 148–164.
- Fisher RA (1930) *The genetical theory of natural selection*. Oxford: Clarendon.
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16: 97–159.
- Balloux F, Lehmann L (2003) Random mating with a finite number of matings. *Genetics* 165: 2313–2315.
- Nunney L (1993) The influence of mating system and overlapping generations on effective population size. *Evolution* 47: 1329–1341.
- Pearse DE, Anderson EC (2009) Multiple paternity increases effective population size. *Molecular Ecology* 18: 3124–3127.
- Wang J (2004) Sibship reconstruction from genetic data with typing errors. *Genetics* 166: 1963–1979.
- Kingman JFC (1982) The coalescent. *Stoch Proc Appl* 13: 235–248.
- Eriksson A, Haubold B, Mehlig B (2002) Statistics of selectively neutral genetic variation. *Phys Rev E* 65: 040901.
- Bensch S, Andrén H, Hansson B, Pedersen HC, Sand H, et al. (2006) Selection for heterozygosity gives hope to a wild population of inbred wolves. *PLoS ONE* 1: e72.
- Schaper E, Eriksson A, Rafajlovic M, Sagitov S, Mehlig B (2012) Linkage disequilibrium under recurrent bottlenecks. *Genetics* 190: 217–229.
- Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology* 17: 2103–2106.
- Johannesson K, Panova M, Kemppainen P, André C, Rolán-Alvarez E, et al. (2010) Repeated evolution of reproductive isolation in a marine snail: unveiling mechanisms of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 1735–1747.
- Butlin RK, Debelle A, Kerth C, Snook RR, Beukeboom LW, et al. (2012) What do we need to know about speciation? *Trends Ecol Evol* 27: 27–39.