Habitat geometry, population viscosity and the rate of genetic drift

J.M.J. Travis\textsuperscript{a,*}, T.H.G Ezard\textsuperscript{a,b,1}

\textsuperscript{a}NERC Centre for Ecology and Hydrology, Hill of Brathens, Banchory, Aberdeenshire, AB31 4BW, United Kingdom
\textsuperscript{b}Imperial College at Silwood Park, Ascot, Berkshire, SL5 7PY, United Kingdom

\textbf{ARTICLE INFO}

Article history:
Received 22 November 2005
Received in revised form 8 March 2006
Accepted 12 March 2006

Keywords:
Genetic drift
Metapopulation viscosity
Habitat geometry

\textbf{ABSTRACT}

In all natural populations, individuals located close to one another tend to interact more than those further apart. The extent of population viscosity can have important implications for ecological and evolutionary processes. Here we develop a spatially explicit population model to examine how the rate of genetic drift depends upon both spatial population structure and habitat geometry. The results show that the time to fixation for a new and selectively neutral mutation is dramatically increased in viscous populations. Furthermore, in viscous populations the time to fixation depends critically on habitat geometry. Fixation time for populations of identical size increases markedly as landscape width decreases and length increases. We suggest that similar effects will also be important in metapopulations, with the spatial arrangement of subpopulations and their connectivity likely to determine the rate of drift. We argue that the recent increases in computer power should facilitate major advances in our understanding of evolutionary landscape ecology over the next few years, and suggest that the time is ripe for a unification of spatial population dynamics theory, landscape ecology and population genetics.

© 2006 Elsevier B.V. All rights reserved.

\textbf{1. Introduction}

The spatial structure of a population can have important implications for ecological and evolutionary dynamics. This is true both when we consider a single isolated population and when we consider a series of subpopulations linked by dispersal forming a metapopulation. Within a single isolated population individuals are more likely to interact with neighbours than with distant individuals, and within a metapopulation, interaction (mediated through dispersal) is more likely to be between neighbouring subpopulations. In this paper we are concerned with understanding how this spatial structure influences the process of genetic drift: in particular we ask how the time to fixation of a new neutral mutation differs according to the spatial scale of interaction between individuals. We are also interested in establishing how the timescale of evolution is influenced by habitat shape: how much does time to fixation vary depending upon whether habitat is long and thin, or relatively square?

When Kimura first derived analytical solutions for the time to fixation of a new mutation, he validated the maths through the use of simulations (Kimura and Ohta, 1969). Using state-of-the-art computers he was able to simulate a freely mixing population of 10 individuals. Rapid increases in computer power meant that by 1981, it was possible to simulate 10,000 individuals in which the freely mixing assumption was relaxed, albeit for a relatively small number of generations (Turner et al., 1982), and certainly not for long enough to determine the mean time to fixation of a new mutation. Turner et al. (1982) demonstrated that even when genotypes were initially distributed randomly across the landscape, genetic drift rapidly created striking microgeographic

---

* Corresponding author. Tel.: +44 1330 826345; fax: +44 1330 823303.
E-mail addresses: jmj@ceh.ac.uk (J.M.J. Travis), thomas.ezard@imperial.ac.uk (T.H.G. Ezard).
1 Tel.: +44 44 75942447.
differentiation within the population accompanied by an overall increase in homozygosity. Since then a number of authors have used similar models to further explore the spatial distribution of genotypes (Doligez et al., 1998; Epperson, 1995; Epperson and Li, 1997; Goldstein and Holsinger, 1992) but like Turner et al. (1982) their simulations lasted for at most a few hundred generations. Here, we take advantage of the rapid processor speeds now available to investigate fixation times in contiguous, but viscous populations.

Most spatially explicit models tend to operate on square lattices (e.g.: Doligez et al., 1998; Epperson and Li, 1997). This is true regardless of whether the model is built to represent a single population, where each cell can support a single individual, or whether it is constructed to represent a metapopulation in which case each cell supports a population of individuals with occasional dispersal between cells. In reality, many populations (or metapopulations) are not well represented by a square lattice. The spatial structure of natural populations has a significant effect on a population’s genetic structure (Epperson, 1995). Wright (1943) predicted greater isolation by distance in essentially linear habitats, such as rivers, shorelines and woodland edges. Furthermore, Kawata (2001) showed that the shape of habitat affects the accumulation of deleterious mutations. Knutsen et al. (2001) emphasised genetic differentiation in linear habitats through investigation of genetic patterns of sea trout in eight streams on the southeast coast of Norway.

There has been surprisingly little work considering the potential consequences of habitat shape for processes such as genetic drift. The importance of genetic drift in small island

![Diagram of C++ computer programme to simulate the spread of a unique, selectively neutral mutation through a population on a homogeneous landscape of size n*n or l*w.]

Fig. 1 – Schematic illustration of C++ computer programme to simulate the spread of a unique, selectively neutral mutation through a population on a homogeneous landscape of size $n \times n$ or $l \times w$. 
populations has been shown for, amongst others, Australian bush rats (*Rattus fuscipes greyii*) (Hinten et al., 2003), sea stars (*Coscinasteria muricata*) (Perrin et al., 2004) and perennial herbs (Yan et al., 2003). Here we investigate how the shape of viscous populations influences time to fixation. We show that habitat geometry has a significant effect on the time to fixation of neutral mutant alleles.

2. Methods

2.1. Model structure

Individual-based models were constructed to examine the effects of population size and habitat shape on the rate of genetic drift in a population. The model is designed to simulate the spatio-temporal dynamics of three diploid genotypes (AA, Aa, and aa) at an arbitrary diallelic locus. In this model time is discrete, a hermaphrodite with annual lifecycle is assumed, and every patch on the landscape is occupied every generation (i.e. there are no population dynamics). The maternal and paternal parents of the offspring located at position \((x, y)\) are chosen at random from the nearest \(N_f\) and \(N_m\) neighbours respectively. The offspring’s genotype is determined by drawing one allele at random from each of the parent’s genotypes.

The landscape consists of a series of patches arranged on a lattice. Initially, square lattices divided into \(n \times n\) habitat patches were examined. In later simulations, length and width varied such that a rectangular habitat was divided into \(l \times w\) patches. It was assumed that every patch is occupied for all generations, so the population size was constant throughout (either \(n \times n\) or \(l \times w\)). Habitat across each grid was assumed to be suitable and of identical quality throughout.

Three methods of mating were examined: nearest neighbour, partially viscous and freely mixing. When a freely mixing population was assumed, the alleles possessed by an individual at location \((x, y)\) at time \(t+1\) are inherited with equal likelihood from any two individuals on the landscape at time \(t\) (i.e. \(N_f = N_m = n \times n\)). Simulations for partially viscous system assumed that one allele possessed by a focal individual at location \((x, y)\) at time \(t+1\) is selected at random from the focal individual and the other from another randomly selected individual on the landscape at time \(t\) (i.e. \(N_m = n \times n\)). Greater viscosity is introduced by amending the model as follows: the genotype of an individual in patch \((x, y)\) at time \(t+1\) is determined by selecting one allele at random from those possessed by the focal individual at time \(t\), and by selecting one allele at random from a randomly selected local neighbour of the focal individual (i.e. \(N_m = 8\)). In all cases, individuals are assumed to be incapable of self-fertilization.

Edge effects have potentially significant consequences for non-global mating systems. To be able to ascertain how much of any differences in results obtained for freely mixing and viscous populations are due to edge effects, we run the simulations utilising a nearest-neighbour method for both “wrapped” and “unwrapped” lattices. A schematic illustration of the model structure is shown in Fig. 1.

Fig. 2 – The effect of the different mating neighbourhoods on time to fixation for square lattices. Points plotted are mean values for all occasions when the unique mutant gene increased in frequency to fixation. The line is Kimura and Ohta’s (1969) theoretical expectation. Triangles are results obtained under a freely mixing mating system, empty circles for a partially viscous system and solid circles for the nearest-neighbour system.

Fig. 3 – Histograms of fixation times for four square lattices. There is no significant change in the shape of the distribution despite large changes in population size and time to fixation. The dimensions of each lattice are indicated in bold type.
2.2. The simulations

The simulation programme was written in C++ and run on both desktop PCs using the Windows operating system and on a Beowulf cluster using MPI parallel programming and the Linux operating system. Random numbers were generated using the computer clock as a seed. Thus, a different seed was used for each simulation.

All individuals were assumed to be homozygous for the same allele, A. A selectively neutral mutation, a, was then introduced in one randomly selected individual in the population. The model was run until the new, unique mutation either increased in frequency to fixation or was lost. Time to fixation was recorded, excluding the occasions when the mutation became extinct. The unlikely scenario of a further mutation reversing the effects of the initial one (Kawata, 2001) was not considered. This process was repeated 1,000,000 times for each lattice size and shape in order to obtain a distribution of fixation times. The mean time to fixation and number of mutant allele fixations were then calculated and recorded. The model was run for a range of square and rectangular lattice (i.e. population) sizes. In the case of simulations using nearest-neighbour mating systems both wrapped and unwrapped grids were investigated.

2.3. Statistical analysis

Statistical analyses were carried out in R. A gamma generalized linear model (GLM) with identity link was fitted to the data for time to fixation on square lattices. The gamma

Fig. 4 – Fitting statistical models to the data. (a) Grey points represent fixation times for a simulation run when the unique mutant gene increased in frequency to fixation. The line is the fitted line under a gamma GLM identity link function. (b) Standard residuals under the simple linear model and (c) Pearson residuals for the gamma GLM with identity link function. Note different scales on the y-axis. The simple model shows increased variance with later index values, i.e. larger lattice size. This suggests violation of the key assumption of constant variance. There appears no such pattern for the Pearson residuals of the gamma GLM.

Fig. 5 – Significant changes in mean fixation time result from reducing the width of the lattice. Differences are greatest for very narrow lattices. The relationship between population size and fixation time is not linear for non-square lattices. The relationship is more curved for very narrow lattices (relative to lattice length). Solid circles are square grids, empty circles are rectangles with fixed width = 20, solid triangles are rectangles with fixed width = 10, empty triangles are rectangles with fixed width = 5 and solid squares are rectangles with fixed width = 2.
distribution was chosen because the data is bounded below by zero and continuous.

The extent of genetic patchiness throughout the time to fixation, or “clumping”, was investigated by join-count statistics using Rook’s Case Software version 0.9.6 (http://www.ai-geostats.org/software/Geostats_software/ROOKS_CASE.htm). Join-count statistics test for spatial autocorrelation in a population and have been frequently used to examine genetic patchiness through time (e.g. Epperson, 1995; Epperson and Li, 1997; Doligez et al., 1998). The Z_{BW} statistic tests the boundary between the different genotypes. A two-sided z-test is appropriate, as no direction of departure from spatial randomness is specified. Null hypotheses are tested at the 5% level. The critical region is |Z| > 1.645. The Queen’s Case method, which tests joins between the focal individual and its eight neighbours, is used as it mirrors the fertilization neighbourhood employed.

3. **Results**

3.1. **Probability of fixation**

As expected, the probability of fixation is unaffected by changes in spatial structure (mating system) or habitat shape (results not presented). The probabilities of fixation found throughout this study conform to existing theory (see Kimura, 1983) that the probability of fixation for a unique, neutral mutation is \( u(p) = \frac{1}{2N_e} \), i.e. \( \frac{1}{2n^* n} \) or \( \frac{1}{2l^* w} \).

3.2. **Impact of population viscosity**

Times to fixation for square grids under random mating, partially viscous and nearest-neighbour systems are given in Fig. 2. Time to fixation increases linearly with increases in population size. Simulations for populations under random mating concur with the theory presented by Kimura and Ohta (1969). They demonstrated that, in the case when a new selectively neutral mutation does “fix” (assuming the negligible probability of its effect being reversed by some subsequent mutation) around \( 4N_e \) generations are required for an initially unique neutral mutation to reach fixation in a population with a finite effective population size \( N_e \). Our results show that when population viscosity is introduced the time to fixation increases markedly. This is true even for partial viscosity (Fig. 2).

3.3. **Statistical analysis**

It is important to consider the distribution of fixation times as well as mean time to fixation. The marginal distribution of time to fixation has a gamma distribution. Time to fixation is

**Fig. 6** – Distribution of AA (yellow), Aa (orange), aa (red) through time after 5000 (top left), 6000 (top right), 7000 (bottom left) and 8000 (bottom right) generations.
non-negative and continuous. The distribution of fixation times is the same approximate shape for all population sizes (Fig. 3). These distributions are characterized by substantial positive skew.

A gamma GLM (Fig. 4a) was chosen in favour of a simple linear model because of cogent graphical evidence and formal tests against the key assumption of constant variance. Quantile-Quantile plots (not shown) and a formal (Wilk-Shapiro) test \( p < 2.2\times10^{-16} \) provided evidence that the errors are not normally distributed. Comparison of normal and Pearson residual plots reveals trends in magnitude for a simple linear model but not for the gamma GLM (Fig. 4b and c respectively). The absolute values of the standard residuals increase in magnitude as the index, i.e. lattice size, increases. This suggests that larger lattices have larger errors. A formal (Breusch-Pagan) test provided compelling evidence \( p=0 \) against the null hypothesis of constant variance.

An identity link function was chosen for the gamma GLM because extreme data was less influential for the model fit than for a reciprocal link gamma model. Although an identity link function produces more influential points for smaller lattices (not shown) these do not affect the fitted line as much as influential points when a reciprocal link function is used. An example of this is the outlying point for a 100×100 grid (top right hand corner in Fig. 4a and b). This point, whilst influential for both link functions, has increased significance in the GLM fitted with a reciprocal link function.

3.4. Distributions of fixation time

Each set of simulations features examples where a significantly larger number of generations are required for a new, unique mutation to reach fixation in a population (Figs. 3 and 4a provide examples). As population size increases, the variance in fixation times increases. This is due to genetic patchiness by greater isolation by distance (Wright, 1943).

3.5. Non-square lattices

Mean time to fixation is not the same for populations of identical size occupying different shaped habitat (Fig. 5). For a given population size, the time to fixation increases markedly as the habitat becomes increasingly linear. This effect is especially pronounced for larger population sizes. For example, a population of size 1600 has a mean fixation time of 13,022 on a square lattice, but this increases to 30,011 on a 10×160 lattice and 459,900 on a 2×800. There is no significant difference in distribution shape for those rectangles investigated for populations of the same size, despite very large changes in mean time to fixation (results not shown).

When we run a simulation experiment extending the length of linear habitats, the relationship between population size and mean time to fixation is not linear for rectangular lattices (Fig. 5). Curvature of the relationship between mean time to fixation and population size depends upon the width of the habitat. In a separate experiment that fixed the length:width ratio while varying the population size, we found that the linear relationship is retained (results not shown). This indicates that it is the changing length:width ratio that introduces the non-linearity observed in Fig. 5.

3.6. Impact of “wrapped” and “unwrapped” lattices

All results presented thus far have been for unwrapped lattices. The time to fixation was also investigated for wrapped lattices (i.e., the top and bottom as well as the left and right sides of the lattice were joined together). Unsurprisingly, fixation times for wrapped lattices were smaller than for unwrapped lattices (results not shown). In both cases on a square lattice, the relationship between population size and fixation time was linear.

3.7. Genetic patchiness

Different genotypes dominate different areas on larger lattices (Figs. 6 and 8). This is consistent with previous studies investigating nearest-neighbour mating systems (Epperson, 1995; Turner et al., 1982), but is not the case for freely mixing, random mating populations that are not subject to isolation by distance. This “quasi-stationary phase” persists for many generations (Fig. 6), although not thousands as suggested by Epperson and Li (1997). Patches of homozygous genotypes do indeed persist for thousands of generations, but not in the same (approximate) locations (Fig. 6). Homozygote patches are more likely to persist for longer periods of time in patches near the edge of the lattice.

Fig. 7 – An illustration of how numbers of genotypes and Z_{BW} statistic change through time. High genotype numbers do not necessarily imply that fixation is imminent. Z_{BW} statistics provide strong evidence that genotypes cluster together, as found previously (Epperson, 1995; Epperson and Li, 1997). Continuous and dotted lines indicate the homozygotes AA and aa respectively, and the dashed line represents the heterozygotes.
Habitat geometry affects the “quasi-stationary” phase. Although the system remains dynamic on a narrow lattice, the increased number of border cells provide greater stability, and a genotype can dominate an area of the grid for extended periods of time, certainly in the thousands of generations as suggested by Epperson and Li (1997) (Fig. 8).

Z_{BW} test statistics are significant at almost any level for all types of patch (see Fig. 7). The heterozygote value is much greater than the homozygote test statistic. This is because heterozygotes cluster along the borders of large patches of homozygotes (Figs. 6 and 8; Epperson, 1995; Turner et al., 1982). “Patches” of heterozygotes are essentially linear, and under the employed join-count analysis have fewer heterozygote neighbours.

4. Discussion

This paper investigates rates of fixation of unique, new and selectively neutral mutations due to genetic drift. The first main result of this paper is that the time to fixation depends critically on how well-mixed the population is. In viscous populations with limited fertilization and dispersal neighbourhoods, the time to fixation will be much longer than in a similar sized freely mixing population. While a few studies have looked at the spatial genetic structuring of viscous populations using similar models to that presented here (e.g. Turner et al., 1982; Epperson, 1995), they did not go on to look at fixation times.

Simulation studies that have considered genetic drift in viscous populations have focused on populations living on square lattices. The second main result of this paper is that changing the shape of the lattice has a significant effect on time required for a selectively neutral mutation to increase in frequency to fixation in a population. Our results indicate the shape of habitat occupied by a population can be as important as population size in determining how long fixation takes. This may have important implications for conservation biology, where maintenance of genetic diversity is frequently a goal (Reed and Frankham, 2003). Higher genetic diversity is likely to be present in populations that occupy narrower habitats (due to the increased fixation times), than in similar sized populations occupying squarer patches. This might indicate that for some species, given a free choice regarding the shape of a reserve, there may be genetic benefits in protecting (or creating) long thin patches of habitat. An obvious extension to the work presented in this paper would be to allow repeated mutations, resulting in a potentially infinite number of different neutral alleles and then investigate how the level of genetic diversity varies with population viscosity and habitat geometry.

The marginal distribution of time to fixation (Fig. 4a) has a gamma distribution characterized by substantial positive skew (Fig. 3). A priori, we speculated that for very large populations this distribution would become normal, and that we would lose the right skew. This was not observed for any of the populations observed here. It remains an open question whether the distributions of fixation time are right skewed because they are bounded by 0, or for another reason. If processor speeds continue to increase as rapidly over the next few years as they have in the past, it will be possible to repeat a similar simulation experiment for populations with much longer fixation times and this may resolve the issue. Alternatively, the development of spatial moment closure techniques such as pair-wise approximation (e.g. Harada et al., 1995; Eames and Keeling, 2003) that incorporate genetics may offer a method for exploring the dynamics of very large populations with lengthy fixation times. Moment closure techniques are facilitating the proper exploration of parameter space for a wide range of ecological and epidemiological models, and in combination with simulation experiments are yielding many valuable insights (Bolker and Pacala, 1999; Keeling, 2000; Law et al., 2003 Murrell et al., 2004). To the best of our knowledge, moment closure techniques similar to those being developed in ecology have not yet been extended to incorporate genetic processes, and this definitely represents an unexplored area deserving of attention.

Whilst the work presented here provides further insight into how genetic drift causes mutations to fix in populations, many other factors that affect gene frequency have not been discussed. The probability of a reversible mutation (Kawata, 2001) has not been considered. An important assumption made in this study is that of selectively neutral mutations. Evidence shows that the number of harmful mutations that occur in each generation is surprisingly high (Crow, 1999; Eyre-Walker and Keightley, 1999). Investigation of slightly deleterious mutations, which occur more frequently in real life than neutral mutations (Ohta, 1992; Ridley, 2004) and/or significantly less common advantageous or Darwinian mutations...
(Kimura and Ohta, 1974) is of interest to further elucidate the evolutionary process. Another important consideration for future work is the impact of heterozygote advantage: how would it modify fixation times, and to what extent would it alter the spatial patterns observed?

The model developed can be readily extended to look at the potential genetic consequences of habitat loss and climate change. Many authors have used similar patch-occupancy approaches to model the ecological dynamics of populations living on fragmented landscapes (e.g. Bascompte and Sole, 1996; Dytham, 1995; Travis and Dytham, 1999; With and King, 1999). Results indicate the existence of threshold responses to habitat loss: up to a critical threshold of loss a population maintains high occupancy of available patches, but beyond that threshold the population rapidly declines towards extinction. The position of the threshold is sensitive to the pattern of habitat loss, and in general populations appear more able to tolerate habitat loss, if that loss is spatially autocorrelated (With and King, 1999). Combining the methods presented in this paper with those used in the ecological literature to simulate habitat loss would enable integration of the key concepts of landscape ecology with population genetics. Recently, Travis (2003) extended the patch occupancy model to simulate the range shifting dynamics of hypothetical species in response to a changing climate. Again, ecological threshold responses were observed: populations are able to keep pace with climate change maintaining high patch occupancy until a threshold rate of change is reached. The position of this threshold is dependent upon the extent of habitat loss. Incorporating genetic processes within the framework described by Travis (2003) would be an interesting avenue for future work, and might reveal how we should expect genetic diversity to change as populations’ spatial distributions become increasingly dynamic.

In this paper we have demonstrated that when individuals are not freely mixing the rate of genetic processes can be considerably slowed. Analogously, if subpopulations (within a metapopulation) are not equally linked, it is likely that the rate of metapopulation genetic dynamics will be reduced. It is highly likely that dispersal of individuals is more likely to occur between subpopulations that are relatively close to one another, and many metapopulations are likely to be extremely viscous. Since Pannell and Charlesworth’s (2000) review of the genetic consequences of metapopulation processes, considerable progress has been made in theoretical metapopulation genetics (e.g. Cherry, 2003; Roze and Rousset, 2003; Whitlock, 2002). Roze and Rousset (2003) derived fixation probabilities of new mutations as, amongst others, a function of their dominance coefficient. Cherry (2003) showed that extinction and recolonization increases fixation probability for deleterious alleles and decreases fixation probability for advantageous alleles. The recurrent fixation of deleterious mutations, sufficient to expect a decline in fitness, by genetic drift is facilitated by small population size (Whitlock, 2002). These metapopulation models are all spatially implicit, assuming that each subpopulation is equally connected to every other subpopulation. There is an urgent need for these approaches to be extended to incorporate the influence of localised dispersal. We suspect that the results may in many cases be very sensitive to the exact form of dispersal incorporated.

Acknowledgments

We thank Monique MacKenzie for assistance with statistical analysis as well as the Natural Environment Research Council for financial assistance.

REFERENCES


