The effect of local population dynamics on patterns of isolation by distance

M. Björklund a,⁎, S. Bergek a,b, E. Ranta c,1, V. Kaitala c

a Department of Animal Ecology, Evolutionary Biology Centre (EBC), Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden
b Institute of Coastal Research, Swedish Board of Fisheries, Institute of Coastal, Skölgtägen 6 SE-742 42 Öregrund, Sweden
c Integrative Ecology Unit, Department of Biological and Environmental Sciences, P.O. Box 65 (Vilinkääri 1), FIN-00014 University of Helsinki, Finland

A R T I C L E   I N F O

Article history:
Received 8 October 2009
Received in revised form 21 December 2009
Accepted 22 December 2009

Keywords:
Isolation-by-distance
Wright’s Island model
Effective population size
Stepping-stone model
FST
Computer simulations

A B S T R A C T

Isolation-by-distance (IBD) is a widely used model explaining population structure and how gene flow decreases with increasing distance. It is biologically intuitive that populations which rarely exchange individuals should drift apart genetically. However, the model is based on the assumptions that populations are large, equal in size and stable over time – conditions that are unlikely to occur in natural conditions. The model has been challenged in the past, for example, in the light of metapopulations or variance in reproductive success. However, an appraisal of the assumption of a large and stable population size per se is lacking. We investigate the robustness of the results concerning IBD patterns when smaller and fluctuating population sizes, or differences in population size are allowed. Through computer simulations we show that allowing for different population sizes and random fluctuations lead to unpredictable patterns regarding the results concerning gene flow and IBD. A pattern of IBD could be the result of high gene flow or no gene flow at all, depending on how populations differ in size and how they fluctuate. Adding environmental noise (white, red and blue noise corresponding to random, positive and negative autocorrelation respectively) gives even more unpredictable results concerning patterns of IBD. Our results have important implications for genetic and conservation research. Interpreting an IBD pattern, or lack thereof, is not as easy as earlier thought and needs to be more thoroughly explored.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

One of the most influential results in population genetics is that populations affected by drift and gene flow diverge over time, resulting in physically distant populations being less related and proximate populations more closely related. This pattern of increasing population differences with increasing time or physical distance is incorporated in the concept referred to as isolation-by-distance (IBD; Wright, 1943). The model is intuitively appealing as there are good biological reasons as to why populations rarely receiving migrants from each other should drift apart genetically with time, while populations exchanging migrants at a more regular basis should be fairly similar genetically.

There is a substantial empirical support for IBD in a large set of taxa using a wide variety of molecular markers (Sharbel et al., 2000; Pogson et al., 2001; Rudh et al., 2007; Pope et al., 2006; McDonald et al., 1999). Interestingly, the pattern of increasing genetic divergence with increasing distance is not always found (Castric et al., 2001; Clark-Tapia and Molina-Freaner, 2003; Erich and Stenseth, 2001; Johnson and Black, 1995; Bergek and Björklund, 2007, 2009). The standard interpretation of a departure from the expected pattern is usually that a high gene flow between some, or all, of the populations obscures the relationship.

A feature of the theoretical models that predict an IBD relationship is the assumption of large population sizes that are equal between populations and constant over time (Wright, 1943; Rousson, 2004). This is a biologically stringent assumption which is likely to be wrong in most cases (Whitlock and McCauley, 1999). The problem with the unrealistic assumptions of the IBD model has been raised before, for example, in the case of metapopulations or source-sink systems (see Whitlock and McCauley, 1999 for examples) and non-equilibrium (between gene flow and drift) cases (Hutchinson and Templeton, 1999).

Natural populations vary in size, and are in many cases much smaller than assumed in the models. Although it is accepted that fluctuations in population size have importance for the operation of drift and differentiation within a single population (Wright, 1931, 1938), the effect on large-scale patterns such as IBD of independent fluctuations in a set of populations has not been explored. Such considerations raise the important question of how robust the results concerning IBD patterns are if smaller and fluctuating population sizes or differences in size between populations are allowed. Alternatively stated, if we allow each population to have their own dynamics and gene flow with other populations, can we still expect to find IBD? If not, what kind of patterns will we find?

⁎ Corresponding author.
E-mail address: mats.bjorklund@ebc.uu.se (M. Björklund).
1 Deceased during the preparation of the paper.
Biological populations are also strongly influenced by random variations in their environment. Random variations are modelled by different environmental noise (white, red and blue, corresponding to random, positive and negative autocorrelation respectively). There is a possibility for autocorrelation in environmental forcing which has shown to be positive in natural systems (Vasseur and Yodzis, 2004). It has also been shown that noise results in higher genetic differentiation among populations (Ranta et al., 2008). Hence, the effect of environmental variability is important on the genotypic level and could thus also affect the model and patterns of IBD.

In this paper we analyse the effect of relaxing the assumption of large populations that are equal in size and constant over time and show by simple computer simulations that incorporating population size differences and local population dynamics (environmental noise) has a significant effect on the predictions of IBD.

2. Methods

In this part we describe the features that are similar to all, or most simulations, with additional details described when necessary. We created a set of ten populations in a linear fashion. The distance between the neighbouring populations were equal so that the distance between population 1 and 2 is X and between 1 and 3 2X, and so on. Population size N was set to 100 for each population. We assigned all ten populations ten loci each with two alleles, with initial frequencies randomly drawn from a uniform distribution between 0.1 and 0.9. We also conducted tests using 50 loci, but as the rare alleles were lost relatively quickly and the system effectively became a two-allele system (the most common, and the remaining few) and the results did not differ more than in minute numerical details, we only report the results using ten loci. Allele frequencies were subject to drift using binomial sampling. We did not allow mutations since the time period is quite short, and the population sizes are fairly small. This means that the probability for a mutation to go to fixation is rather small and consequently drift has a larger effect on the allele frequencies.

All runs lasted 500 generations, after which we calculated $F_{ST}$ and the regression of pairwise $F_{ST}$ and distance, where the slope is a measure of IBD (Rousset, 2004). We used raw distances rather than log-transformed distances as we deal with a linear habitat where we can expect $F_{ST}$ and distance to be related in a linear fashion (Rousset, 2004: 40). We tested the significance of the regression with 10000 randomisations (Mantel test). The P-value reported is the proportion of runs where we found a positive regression significant at the 5% level. We used the actual allele frequencies in each population, i.e. sample size was 100 in each case. In this way we have minimized the possible effects of sampling error on the test procedure. In a real situation sampling will add to the uncertainty.

Population fluctuations were simulated using the standard Ricker model where $N_{t+1} = N_t \exp[r(1 - N_t/K)]$, where $K$ is the carrying capacity and $r$ is the intrinsic rate of growth. We used three levels of $r$, 0.5, 1.8 and 2.4, where two first ones produce undercompensating and overcompensating stable dynamics, respectively, and 2.4 produces two-point cycles that are unstable (Ranta et al., 2006). Environmental stochasticity was incorporated via the carrying capacity, $K$, according to the model by Ripa and Lundberg (1996). $K$ was allowed to change between 50 and 150 with a mean of 100. Three different patterns of environmental noise were used as this has previously been shown to be important in affecting the dynamics of allele frequency changes (Ranta et al., 2008). The noise was set to be either negatively autocorrelated in time ($-0.8$, blue noise), positively autocorrelated in time ($0.8$, red noise), or not correlated at all ($0$, white noise). We did not use a demographically explicit model accounting for variance in reproductive success among individuals. All simulations were repeated 1000 times.

We then allowed for gene flow in both directions from each population, except in the two outermost populations where only one direction was possible. We used a simple stepping-stone model where individuals only disperse to the most adjacent deme, even though more complicated dispersal models would be biologically realistic for many species. In all cases the number of migrants in each generation was 0.01 * N, otherwise the simulations were as above. This value was based on Waples and Gaggiotti (2006), who show that when migration rate, $m$, is lower than $1/N_e$ ($N_e$ = effective population size) drift within populations is so important that populations can be considered independent, while if $m \gg 1/N_e$ gene flow is the by far dominating force and thus the result will be a more or less panmictic population. Since we are interested in the interplay between drift and gene flow that creates the pattern of IBD we use a value of migration between these boundaries. The suggestions by Waples and Gaggiotti (2006) was confirmed by preliminary runs using values of $m$ of 0.001 and 0.1, thus we decided to use $m = 0.01$.

To investigate the effect of spatial differences in population sizes on the model of IBD a gradient in population size was introduced. The outermost population on one side was assigned a population size of 100, which then increased with an arbitrary factor 1.78 for each subsequent population, such that the outermost population at the other end of the line had a population size of 17 800.

3. Results

3.1. Stable and equally sized populations

The first simulation concern a set of populations with equal and stable population sizes over time, but with no gene flow. Thus, the only factor affecting the system in this part is drift in each population. This is the expected baseline level if there is no pattern of IBD at all. We found that the slope of the IBD regression was close to zero (Fig. 1), and the probability of finding a significant result was very close to 5%, as expected (Fig. 2). When allowing for gene flow, the slope of the regression increased as predicted above zero up to 0.016 (Fig. 1) and the probability of finding a significant regression was close to one (Fig. 2).

3.2. Stable populations with a population size gradient

In this part we investigated the effect of populations being of different sizes. Even if no gene flow was allowed, the IBD slope increased.

![Fig. 1.](image-url)
above zero (Fig. 1) and the probability of finding a significant regression was clearly higher than 5%, in fact as high as 26% (Fig. 2). The lower 2.5% interval from the simulation did not cover zero (0.0016), which means that the estimate of IBD is quite seriously biased upwards in this scenario. When we allowed for gene flow in the gradient, the IBD slope was about the same as the earlier simulation (Fig. 1) but the probability of finding a significant result was now slightly below one (Fig. 2).

3.3. Fluctuating population sizes — no gradient

In this part of the simulations, we started off with equal population sizes and then introduced a fluctuation (see Methods). When no gene flow was allowed, the IBD slope was again very close to the expected zero (Fig. 3) and the probability of finding a significant regression near 5% (Fig. 4). This was the same regardless of what environmental noise or r-values were used (Fig. 3, Fig. 4). When gene flow was allowed, the regression slope rose to maximum 0.048 (Fig. 3). Again different noise did not give different results, but higher r-values increased the slope to at most 0.032 (Fig. 3, Fig. 4).

3.4. Fluctuating population sizes — gradient

We further investigated the effect of fluctuating population sizes by incorporating a population size gradient. When no gene flow was allowed, the mean IBD slope rose to between 0.006 and 0.012 (Fig. 5). This is higher than expected since no gene flow was allowed and this pattern was approximately the same regardless of what environmental noise or r-values was used (Fig. 5, Fig. 6). In fact, in only one case (white noise, r = 2.4), did the 95% interval cover zero; in all other cases the lowest was 2.5% larger than zero. When gene flow was allowed, mean IBD slope rose to 0.022 but different intrinsic rate of growth affected the pattern differently (Fig 5). Allowing for more drastic fluctuations (higher r-values) lead to a large drop in finding IBD patterns (Figs. 5, Fig. 6).

4. Discussion

We have shown that violations of the assumption of equal and stable population sizes have great impact on the probability of finding a pattern of IBD. When the basic assumptions were implemented we found what is expected from theory; if there is no gene flow the slope (IBD) was very close to zero, and the probability of finding a significant result at the 5% level (α) was 5%. When we allowed for gene flow, the slope rose to be significant in all cases. However, dissimilar population sizes affected the pattern of IBD as did random fluctuations between years as well as the intrinsic rate of growth. This is biologically very relevant since almost any species will have variations in population sizes and different rates of growth. Inevitably, the environment is not stable either, which has been shown to affect the preservation of genotypes in populations (Ranta et al., 2008). We found in particular that if sampling has occurred along a population size gradient, the probability of finding a significant pattern of IBD is substantially elevated, even if there is no gene flow at all between the populations. Hence, a pattern of IBD can mean high levels of gene flow or no gene flow at all, depending on how population sizes differ and how they fluctuate. Adding environmental noise gives even more
unpredictable results where almost all simulated cases of noise gave a significant regression between geographical and genetically distances, even when no gene flow was allowed.

In all simulations where population size differed according to a gradient and where migration was not allowed, IBD patterns could nevertheless be found. Further, the intrinsic growth of rate decreased the slope at high levels. \( F_{ST} \) measures the balance between gene flow and drift (Barton, 2001), thus drift can easily overcome the effect of gene flow if population sizes are small enough, or vice versa. This can explain why the probability of finding a pattern of IBD was lowered when we imposed a high level of density-compensation. When populations drastically fluctuate, effective population size becomes small as it is dominated by the harmonic mean of population size, and thus drift becomes more important than gene flow. A spatial pattern in population size can easily be predicted by biological events such as range expansion, founder events and habitat fragmentation. Having a complete knowledge of the demography of populations is hard but as we have shown in this study, neglecting a pattern of IBD only because of a failure to find a significant result can be totally misleading.

Fluctuating population sizes per se did not seem to have any drastic effect on the model of IBD. In simulated cases where no gene flow was allowed and the populations fluctuated around the mean, the slope was around 0 and the chance of finding a significant pattern was as expected around 5%. When gene flow was allowed, a pattern of

**Fig. 4.** Probability of finding a significant pattern of IBD in the equal populations that fluctuates in size.

**Fig. 5.** Estimation of mean IBD slope in populations that have different population sizes according to a gradient which fluctuates. Different fluctuations were simulated by adding different environmental noise (blue, white, and red) and \( r \)-values (0.5, 1.8, and 2.4). Vertical bars denote 95% confidence intervals.
IBD was indeed found. The intrinsic growth of rate and/or noise seemed to influence to a certain degree though since the slope increased with increasing rate of growth, and was always larger than in the simulation with stable populations (Fig. 3).

Note that in all simulations we used the same migration rate ($m = 0.01$), but still the slope differed substantially depending on the setup (i.e. population size, environmental noise and intrinsic growth of rate). For example, when population sizes were allowed to fluctuate around the same mean ($N = 100$), and with blue noise affecting $K$, the slope was 0.016 when $r = 0.5$, but rose to 0.045 when $r = 2.4$. The slope is equal to $\frac{\text{var}(K)}{\text{var}(N)}$ (Rousset, 1997), where $N$ is population size, $\epsilon$ is the distance between subpopulations and $m$ is migration rate. Since $\epsilon$ and $m$ was constant the threefold increase in slope must be due to the differences in intrinsic growth rate and the resulting fluctuations in population size. Consequently, inferences on the level of gene flow from the slope will be affected by the pattern of local population dynamics, and inferences are valid only under the assumption of stable and equally sized populations. This has important implications for the interpretation of patterns (or lack thereof) of IBD. Suppose that we in a particular species find a significant IBD, while in another species we do not. The natural conclusion would be that there is a certain level of gene flow in the first species, but not in the other. However, the opposite might be true as the results may purely be a result of differences in population sizes.

The effect of population dynamics on effective population size has long been realized a theoretical possibility (Wright, 1931, 1938), while in reality this effect has largely been ignored. If local population dynamics is included, strong effects on the genetic differentiation between populations can be observed (Whitlock and Barton, 1997; Ranta et al., 2008), and it has also been shown that adding competing species can increase genetic differentiation (Ranta et al., 2009). Since the degree of differentiation is determined by several stochastic factors, prediction becomes very hard, as well as an understanding of why a certain level of differentiation has evolved. We want to stress that our findings are only a result of relaxing the unrealistic assumptions of very large and stable population sizes in the traditional IBD model. Apparently, this can lead to highly variable and unpredictable results far from the patterns predicted from the basic assumptions.

In many species in temperate and northern parts of the world 500 generations (as used in our simulations) can mean 2–3000 years, periods during which climatic conditions and distributional ranges can change dramatically, leading to a difference in population size. Thus, this is a reasonable, and possibly even overly generous, time frame of this kind of analysis. Another simplification we have used is concerning the spatial organization of populations. In real life, habitats and populations are not necessarily organized spatially in a linear fashion, although sampling efforts could be so. However, rather than going into details, which would result in a infinite number of different models (since the details of the loci in combination with the geographical patterns are infinite), we restricted ourselves to a simplified model in order to highlight the fact that the assumption of equal population sizes is rather crucial, and probably in most cases in nature not upheld.

Acknowledgements

We thank S. South and B. Rogell for comments on earlier drafts of the manuscript. MB was funded by NorFa/NordForsk as a visiting professor in the Nordic Centre of Excellence in Population Biology in Helsinki.

References


Fig. 6. Probability of finding a significant pattern of IBD in populations with different population sizes that fluctuates over time.


