Received: 19 December 2020 | Revised: 2 June 2022 | Accepted: 16 June 2022

DOI: 10.1111/1755-0998.13676

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Using *gridCoal* to assess whether standard population genetic theory holds in the presence of spatio-temporal heterogeneity in population size

Enikő Szép[1] | Barbora Trubenová[1,2] | Katalin Csilléry[3]

[1]IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Austria

[2]Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland

[3]Biodiversity and Conservation Biology, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

Correspondence
Barbora Trubenová, IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Austria.
Email: barbora.trubenova@env.ethz.ch

Funding information
H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 704172; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: CRSK-3_190288; Swiss National Science Foundation; Horizon 2020; European Union; IST Austria

Handling Editor: Alana Alexander

Abstract

Spatially explicit population genetic models have long been developed, yet have rarely been used to test hypotheses about the spatial distribution of genetic diversity or the genetic divergence between populations. Here, we use spatially explicit coalescence simulations to explore the properties of the island and the two-dimensional stepping stone models under a wide range of scenarios with spatio-temporal variation in deme size. We avoid the simulation of genetic data, using the fact that under the studied models, summary statistics of genetic diversity and divergence can be approximated from coalescence times. We perform the simulations using *gridCoal*, a flexible spatial wrapper for the software *msprime* (Kelleher et al., 2016, *Theoretical Population Biology*, *95*, 13) developed herein. In *gridCoal*, deme sizes can change arbitrarily across space and time, as well as migration rates between individual demes. We identify different factors that can cause a deviation from theoretical expectations, such as the simulation time in comparison to the effective deme size and the spatio-temporal autocorrelation across the grid. Our results highlight that F_{ST}, a measure of the strength of population structure, principally depends on recent demography, which makes it robust to temporal variation in deme size. In contrast, the amount of genetic diversity is dependent on the distant past when N_e is large, therefore longer run times are needed to estimate N_e than F_{ST}. Finally, we illustrate the use of *gridCoal* on a real-world example, the range expansion of silver fir (*Abies alba* Mill.) since the last glacial maximum, using different degrees of spatio-temporal variation in deme size.

KEYWORDS
coalescence, demography, spatially explicit simulations, island model, 2Dstepping stone model, Fst

Enikő Szép and Barbora Trubenová contributed equally.

1 | INTRODUCTION

The distribution and dynamics of genetic diversity within species are shaped by a myriad of evolutionary and ecological processes acting across different spatial and temporal scales (Ellegren & Galtier, 2016). Although the role of space and, in particular, spatial autocorrelation in allele frequencies have been recognized since the dawn of population genetics (Felsenstein, 1976; Malécot, 1948; Sokal & Wartenberg, 1983; Wright, 1943), disproportionately more theoretical and methodological developments have been focused on understanding the effect of temporal changes in population size and gene flow among populations without spatial structure (e.g. Hey & Nielsen, 2007). Further, most statistical tools have been developed to detect past population size changes, either by testing different hypotheses such as exponential growth and bottlenecks (e.g. Excoffier et al., 2013), or by using Bayesian methods to detect arbitrary population size changes from whole genome sequences (e.g. Drummond et al., 2005). Researchers in landscape genetics have aimed to overcome the limitation imposed by population genetics methods that rely on the assumption of non-spatial and discrete populations (Manel et al., 2003). However, the field has been mostly influenced by meta-population models (Hanski & Gilpin, 1991) and by spatial statistics and geo-statistics (e.g. Forester et al., 2016; Guillot et al., 2005; Smouse et al., 2008), rather than by population genetic theory.

There is increasing evidence that ignoring space can lead to biases and erroneous inferences (Bradburd & Ralph, 2019). Indeed, simulation studies have shown that ignoring isolation by distance can lead to false positives in efforts to detect hierarchical population structure and loci under selection (Meirmans, 2012). Similarly, ignoring space can severely bias common population genetics summary statistics, especially when the local effective population size (i.e. neighbourhood size) is small (Battey et al., 2020). However, spatially explicit models are often mathematically intractable and theoretical predictions are valid only under limited conditions (Barton et al., 2002; Bradburd & Ralph, 2019; Kelleher et al., 2014; Slatkin, 1985). This is particularly true for spatially continuous models. For example, the coalescence process under the continuous space isolation-by-distance (IBD) model (Malécot, 1948; Wright, 1943) can be approximated using the Lambda-Fleming-Viot algorithm (Barton, Etheridge, & Véber, 2010; Barton, Kelleher, & Etheridge, 2010). However, results are inconsistent with large-scale patterns and often predict lower diversity than expected from census numbers (Barton, Etheridge, & Véber, 2010), although some of these issues have been solved by the subsequently introduced model of extinction and recolonization (Kelleher et al., 2014). Discrete spatial models are worse at capturing reality but are mathematically more tractable (Cox & Durrett, 2002), and several equivalences have been shown across island models, two-dimensional (2D) stepping stone models (Kimura, 1953) and IBD models assuming infinite or finite populations and the absence or presence of mutations (Felsenstein, 1976; Malécot, 1975; Slatkin, 1985). In particular, a 2D stepping stone model can approximate the decrease in genetic correlation with increasing distance of continuous space (Kimura & Weiss, 1964; Malécot, 1955), and when a sufficiently large lattice is used, it can produce summary statistics similar to those from a continuous space model (Battey et al., 2020).

Efficient spatially explicit simulators have recently been developed, both those using a forward in time approach, such as *SLiM* (Haller & Messer, 2019), and those using a mixture of forward and coalescent approaches, such as *SPLATCHE 3* (Currat et al., 2019). These developments have increasingly enabled the inclusion of space in population genetics applications (e.g. Battey et al., 2020; González-Serna et al., 2019; Ortego & Knowles, 2020; Quilodrán et al., 2019). However, these spatial simulators can be challenging to parametrize. This is particularly true for forward simulations, as they require background knowledge on the demography, mating system and dispersal patterns. Furthermore, forward simulations need to track large numbers of individuals, which is memory intensive and thus not practical, or even feasible, for long timescales. Backward, coalescent simulations have the advantage of allowing likelihood calculations while only tracing back the genealogy of sampled individuals (Felsenstein, 1992). Nevertheless, they still also require that past population size changes are known or follow a predictable pattern, such as constant size, expansion, decline or bottleneck. Ecological models, such as species distribution models coupled with recently developed paleo-climatic databases (e.g. Cook et al., 2015; Karger et al., 2021; Lima-Ribeiro et al., 2015), may be used to predict past species distributions in a spatially and temporally explicit manner (e.g. Lima-Rezende et al., 2019; Tallavaara et al., 2015; Wang et al., 2017). Such time series of species distribution maps can provide potential input parameters for spatially explicit coalescent simulations (He et al., 2013).

The aim of this work is to explore the properties of island models and 2D stepping stone models under a wide range of scenarios with spatio-temporal variation in population size. To this end, we first develop a spatially explicit coalescent wrapper, *gridCoal*, for the most efficient coalescent simulator currently available, *msprime* (Kelleher et al., 2016). In *gridCoal*, we implement the 2D stepping stone model with population sizes that may vary in space and time, and with migration rates that may differ between demes. *gridCoal* is different in several ways from *SPLATCHE 3*, the spatially explicit coalescent simulator that is currently used most frequently. Most importantly, in *gridCoal* (i) there is no forward simulation phase; and (ii) demes do not follow a logistic growth model (as in *SPLATCHE 3*), instead instantly increasing or decreasing to user-defined deme sizes. Further, unlike *SPLATCHE 3*, *gridCoal* does not simulate genetic marker data. Instead, we exploit the fact that under the 2D stepping stone model, summary statistics of genetic diversity and divergence between populations can be approximated from the coalescence times (Ralph et al., 2020; Slatkin, 1991). After developing the coalescent wrapper, we use *gridCoal* to simulate various scenarios of spatial and temporal changes in population size and compare their outcome with theoretical expectations of the island models and 2D stepping stone models. In particular, we compare simulations with expectations for the mean coalescence time, which is proportional

to the effective population size N_e and the amount of genetic diversity, for a measure of the strength of population structure F_{ST}, and for isolation-by-distance patterns. Our simulated scenarios include simplified and biologically realistic cases of population movement and expansion, where the spatial and temporal autocorrelation are decoupled. Finally, we illustrate the use of *gridCoal* on a real-world example, the range expansion of silver fir (*Abies alba* Mill.) since the last glacial maximum, using different degrees of spatio-temporal variation in deme size.

2 | MATERIALS AND METHODS

2.1 | The simulation tool: *gridCoal*

We developed a 2D stepping stone coalescent simulation tool, *gridCoal* (Appendix A), based on *msprime* (Kelleher et al., 2016). Space is represented by a rectangular grid (size $L \times L$ in most of our simulations). Each grid cell contains a single panmictic population, hereafter referred to as a deme, whose size (N) can change in time at equally spaced time points comprising a given number of generations. A forward migration matrix defines the fraction of individuals that migrate from one deme to its four neighbouring demes. Forward migration rates (m) are independent from deme sizes, and they can be symmetric or asymmetric between demes, and homo- or heterogeneous across space. The backward migration matrix, required for the coalescent process, contains elements that specify the fraction of individuals in a given deme that have a parent in another deme. Backward migration rates are calculated for each time point based on the deme sizes and the forward migration matrix.

The coalescent process consists of two phases: a scattering phase in the recent past with the fully defined demographic history of individual demes, and a collecting phase in the more distant past assuming panmictic population(s). While spatial structure is important in the scattering phase, its effect becomes smaller and even negligible in the collecting phase, which can be thus approximated by the standard coalescent process (Wakeley, 1998, 1999). This implies that it is unnecessary to run the spatially explicit simulations until all lineages coalesce; before that point, the lineages can instead be combined to a single or a few spatially non-explicit panmictic populations. It is, nevertheless, possible to specify multiple ancestral populations with low migration among them, and thus account for the spatial aspect of the collecting phase.

Time is managed in *gridCoal* using three parameters: (1) the number of time points T when the deme sizes are defined, (2) the time period that elapses between two time steps dt (in years, or other suitable time units), and (3) the generation time gt (in years or other time units, compatible with dt; see also Table 1). Thus, $T \times dt$ determines the length of the spatially explicit phase. To achieve the highest efficiency, this time should be equal to the scattering phase. After this phase, all lineages are combined into one or more panmictic, spatially non-explicit, so-called ancestral population(s). This non-spatial phase ensures that all lineages coalesce even when the product of the effective population size and migration rate (Nm) is small, and it facilitates the simulation of different refugial populations that may colonize different parts of the grid. Note that *gridCoal* does not explicitly model mutations. The coalescent approximations of genetic diversity and divergence provided by *gridCoal* assume that mutation is a weak force (Slatkin, 1991).

TABLE 1 Symbols and terms and their definitions

Symbol	Term	Definition
d	Deme	Panmictic population in a single grid cell
	Map	Grid with a defined distribution of deme sizes
L	Grid size	Number of rows (columns) in a square grid
N	Deme size	Number of individuals in a deme
N_b	Neighbourhood size	Size of a focal deme and its four neighbours
T	Number of time points	Number of time points with the defined demographic history
gt	Generation time	The interval between the birth of an individual and the birth of its offspring
dt	Time step	Time between two defined time points, in years
m	Migration rate	Fraction of population moving from the ancestral cell to a neighbouring cell
T_W	Within-deme coalescence time	Coalescence time between two lineages drawn from the same deme
T_B	Between-deme coalescence time	Coalescence time between two lineages drawn from different demes
T_T	Average coalescence time	Average coalescence time of any two lineages drawn from the grid

2.2 | Simulated scenarios

Here, we provide a brief summary of the simulated scenarios, while more details can be found in Appendix B. Across all scenarios, we used a forward migration rate that is constant in time and homogeneous across the grid. Simulations were run with an average deme size of $N \in (10, 50, 100, 250, 500)$, with migration rate $m \in (10^{-5}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0})$ between neighbouring cells (see Table 1 for explanations of terms and symbols). To analyse the effect of spatial heterogeneity, we simulated various maps differing in the amount of spatial variation and autocorrelation in deme size (Figure 1). Our simulated scenarios ranged from a homogeneous map, where all demes have the same size, to a map with large variance in deme size, with deme sizes drawn from a uniform distribution. To investigate the effect of temporal changes, we simulated scenarios with various demographic histories: static scenarios with fixed deme sizes in time; simple demographics, where all demes changed in the same manner on average, such as undergoing an expansion, decline or bottleneck; and more biologically realistic scenarios of colonization from one side or from "seeds" (such as refugia), or range expansion and shift (Figure 1). For each scenario and combination of N and m in a factorial design, we ran 1000 replicates.

We sampled lineages across the grid in two different ways. In order to estimate the within-deme coalescence time, we sampled two lineages from each deme. In contrast, for studying between-deme coalescence times, we sampled lineages along a row of L demes in the middle of the grid (see Figure 1, Static scenarios). Since the calculation of between-deme coalescence times would have been computationally expensive, this solution assured that all distance classes are represented in each replicate. Further, our changing demographic scenarios were designed such that using demes of the middle row provided a representative sample of demes with different demographic histories (see e.g. side colonization or range expansion and shift, Figure 1).

2.3 | Summary statistics

Hudson (1990) and Slatkin (1991) noted the close relationship between the probabilities of identity by descent and coalescence times, which makes it possible to bypass the simulation of genetic data, instead estimating diversity and divergence statistics directly from coalescence times (Ralph et al., 2020). Additionally, for such calculations, it is sufficient to simulate the genealogies of two lineages per deme.

2.3.1 | Coalescence times

For low and high Nm, the individual demes and the grid as a whole, respectively, are nearly panmictic, the distribution of coalescence times is close to exponential, and most lineages coalesce within the scattering phase. Under these conditions, the maximum likelihood estimate of the mean coalescence time is the sample mean. In contrast, for intermediate Nm, the probability that lineages migrate away from their present-time demes before coalescing is high, but the probability that they meet again and coalesce within the spatially explicit phase of the simulation is low. As a result, the distribution of coalescent times is no longer exponential and the sample mean is an incorrect estimate of the coalescent time. In order to still consider these intermediate Nm values, we use the sample median, which is expected to be less sensitive to the missing tail values (Figure S1).

We calculated the expected within-deme coalescence time (T_W) as the coalescence time of two lineages from the same deme. Assuming a mutation model, measures of within-population genetic diversity can be calculated from T_W. Here, we simply used T_W as a proxy for within-deme diversity and plotted it as a heatmap across the grid. T_B is the coalescence time between any two lineages from two different demes, and T_T is the coalescence time of any two lineages across the grid.

2.3.2 | Population structure (F_{ST})

F_{ST} was introduced by Wright (1951) and provides a measure of population structure. Under the infinite island model, the extent of population structure can be described as in Whitlock and McCauley (1999):

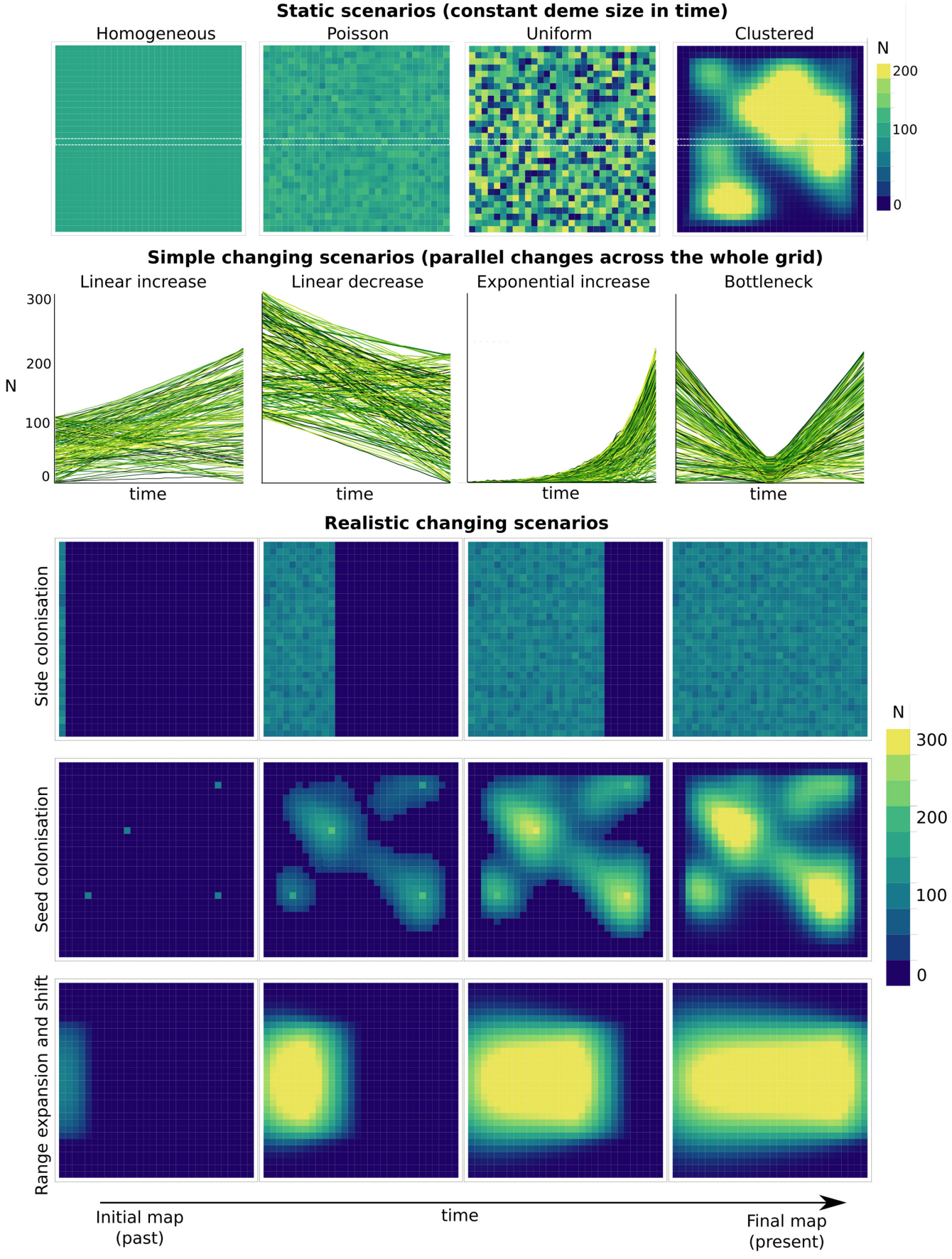
$$F_{ST} = \frac{1}{4Nm + 1}. \quad (1)$$

Under Kimura's 2D stepping stone model, given a homogeneous migration rate and equal sized demes, F_{ST} can be defined as follows (Cox & Durrett, 2002; Maruyama, 1977):

$$F_{ST} = \frac{\frac{L^2 \log L}{2\pi\nu\sigma^2}}{\frac{L^2 \log L}{2\pi\nu\sigma^2} + 2NL^2}, \quad (2)$$

where $\sigma = 1/2$ is the average axial parent–offspring distance, $\nu = 4\,m$ and L is the grid size. The value $2\pi\nu\sigma^2$ is the neighbourhood size, which is the local panmictic unit that determines the amount of variation between populations at the migration–drift equilibrium; thus, it is

FIGURE 1 The three different groups of scenarios simulated. Static scenarios: Demes had a constant size across the spatially explicit phase of the simulations. Simple changing scenarios: The size of all demes changed in a correlated manner. In the present time step, all scenarios were identical to the deme sizes drawn from a uniform distribution. Realistic changing scenarios: Deme sizes changed in space and time to model a colonization event. The grid size was 30 across all scenarios. To estimate T_W two lineages were sampled in each deme, and to estimate T_B two lineages were sampled from 30 demes in a row in the middle of the grid



equivalent to Nm in the island model. Note that when $\frac{logL}{2\pi\nu\sigma^2} < < 2N$, Equation (2) simplifies to $F_{ST} = \frac{logL}{4N\pi\nu\sigma^2}$ (Cox & Durrett, 2002).

Slatkin (1991) introduced an estimate of F_{ST} based on coalescence times:

$$F_{ST} = \frac{T_T - T_W}{T_T}, \quad (3)$$

where T_T is the average total coalescence time and T_W is the within-deme coalescence time averaged across demes. The advantage of this approximation is that it depends on purely demographic processes, such as genetic drift and migration, when mutation is a weak force. In this work, we refer to Equation (3) as the *global* (population-wide) F_{ST}, which measures the strength of the population structure and can be compared across different simulated scenarios, and we use this definition throughout the manuscript. Note that approximating summary statistics of genetic diversity and F_{ST} from coalescence times holds only when migration is only possible among neighbouring demes (Slatkin, 1985).

2.3.3 | Genetic distance (F^*)

While F_{ST} can be computed for pairs of populations, we note that Wright actually rejected F_{ST} as a genetic distance measure because it fails to satisfy the triangle inequality (Wright, 1978; p. 89). Nevertheless, it is widely accepted that F_{ST} leads to a useful measure of genetic similarity if the goal is to infer patterns of gene flow (Slatkin, 1991). Here, we used a measure of genetic distance based on coalescence times in order to investigate the genetic differentiation between pairs of demes and its relationship with physical distance following Slatkin (1993). If only two demes are considered, Equation 3 transforms into:

$$F^* = \frac{T_B - T_W}{T_B + T_W}, \quad (4)$$

where T_B is the mean coalescence time for two lineages sampled from different demes, and T_W is the mean within-deme coalescence time. Slatkin (1993) pointed out that this equation is not appropriate to asses the strength of population structure in general, but it is a useful measure of the genetic distances between demes. We used F^* between all pairs of sampled demes against the physical (Euclidean) distance between demes to assess isolation-by-distance patterns across the grid. Note that if $T_W \approx T_B$ (which is the case for large Nm), $F^* = \frac{F_{ST}}{2 - F_{ST}}$ and F^* thus does not provide any more information about the population structure than F_{ST}.

2.3.4 | Effective population size (N_e)

Under the island model, N_e is

$$N_e = Ns\left(1 + \frac{(s-1)^2}{4Nvs^2}\right), \quad (5)$$

where $\nu = 4\,m$ (the total migration rate for each grid cell), and s is the number of demes in the island model. While, N_e under two dimensional stepping stone model can be calculated as Cox and Durrett (2002):

$$N_e = \frac{L^2\log(L)}{4\pi\sigma^2\nu}. \quad (6)$$

Effective population sizes predicted from simulations were obtained by halving the coalescence time of lineages from the same deme.

3 | RESULTS

3.1 | Coalescence times

We used our simulations to explore the effect of spatial heterogeneity in deme size on theoretical expectations, so we first recall the predictions of the island and 2D stepping stone models. The expected coalescence time of two samples drawn from the same deme is $T_W = 2N$, where N is the total number of diploid individuals in the deme. This result is independent of the migration matrix if all demes are connected by migration. Under the island model with d demes each containing N individuals, the expected coalescence time for two samples from the same deme is $2Nd$ (Strobeck, 1987), which is higher than $2N$ as a result of lineages escaping before coalescence occurs. Under the 2D stepping stone model, the expected coalescence time is $2NL^2$ (Cox & Durrett, 2002). However, in a 2D stepping stone model, where only the four neighbours are connected, strong local differentiation across demes occurs when Nm <1 (Kimura & Maruyama, 1971). Nevertheless, when Nm >1, local differentiation is less pronounced, and when Nm ≥4, the whole grid behaves like a single panmictic population (Kimura & Maruyama, 1971).

Our simulations showed that spatial heterogeneity can be ignored when $m = 0$ and thus each individual deme behaves like a panmictic population. In these cases, the expected coalescence time for two samples taken from the same deme is independent of the spatial heterogeneity of the grid, and thus our simulations confirmed the expected value of 2 N (Figure 2a and d). At the other extreme, when m was one, the whole grid behaved like a single panmictic population. Here, T_W was decoupled from the local deme size and was, on average, equal to $2NL^2$ (Figure 2c). Additionally, when there was large spatial variance in deme size across the grid, as in the uniform map, the coalescence time was systematically underestimated (Figure 2f). This was because the spatial heterogeneity decreased the total effective population size across the grid.

For low migration rates, the expected coalescence time of two samples from the same deme is expected to be $T_W = 2NL^2$, independent of m (Maruyama, 1971). When m is very small, lineages coalesce mostly within demes, on average, in time 2 N. However, very rarely, they escape and coalesce only in time 2 NL^2/m, which results in a mean coalescence time of $T_W = 2NL^2$. The spatially explicit phase of our simulations was not long enough for these samples to coalesce after they escaped from the deme. Since we forced these lineages to coalesce sooner by pulling them to an ancestral population, we

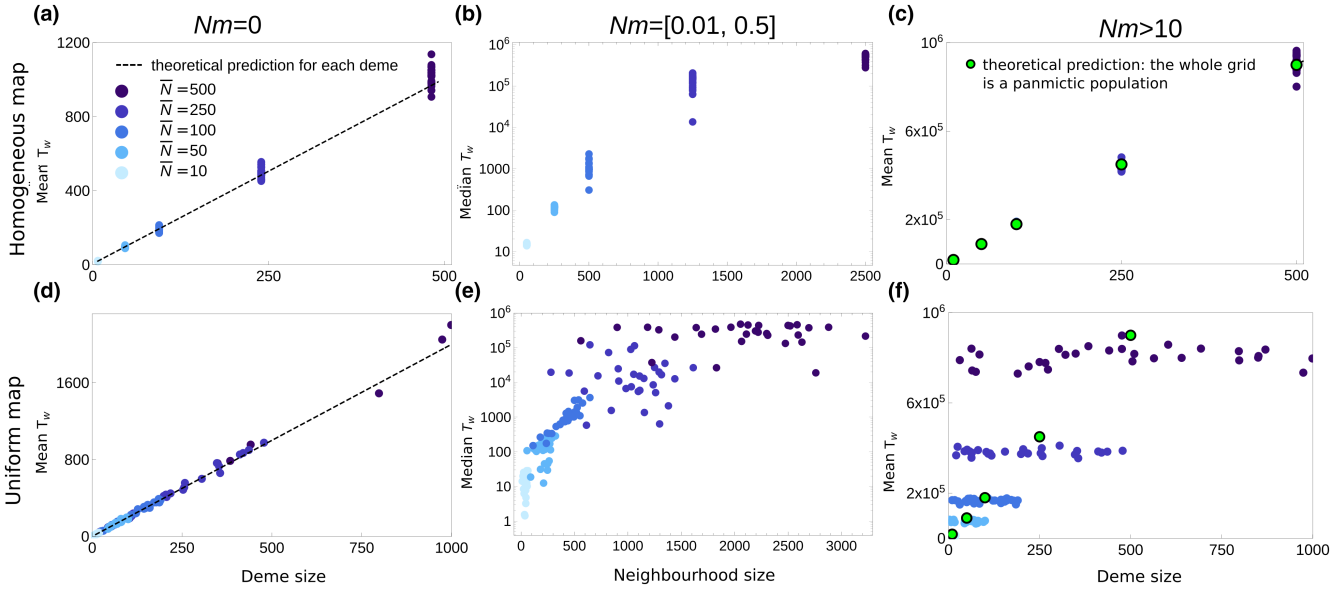


FIGURE 2 Within-deme coalescence times (T_W) for different ranges of Nm from zero (a, d) to $m = 1$ (panmictic grid) (c, f) across two maps: Homogeneous (a, b, c) and uniform (d, e, f). The theoretical predictions are for a Wright-Fisher model with N_i, where i is an index for demes (a, d) and for $2L^2\overline{N}$, where L is the grid size and $\overline{N}$ is the average deme size across the grid. There is no theoretical prediction for intermediate Nm(b, e) because for those parameter ranges the median T_W is shown. Each parameter combination (N and m) is represented by 30 dots showing T_W for individual demes

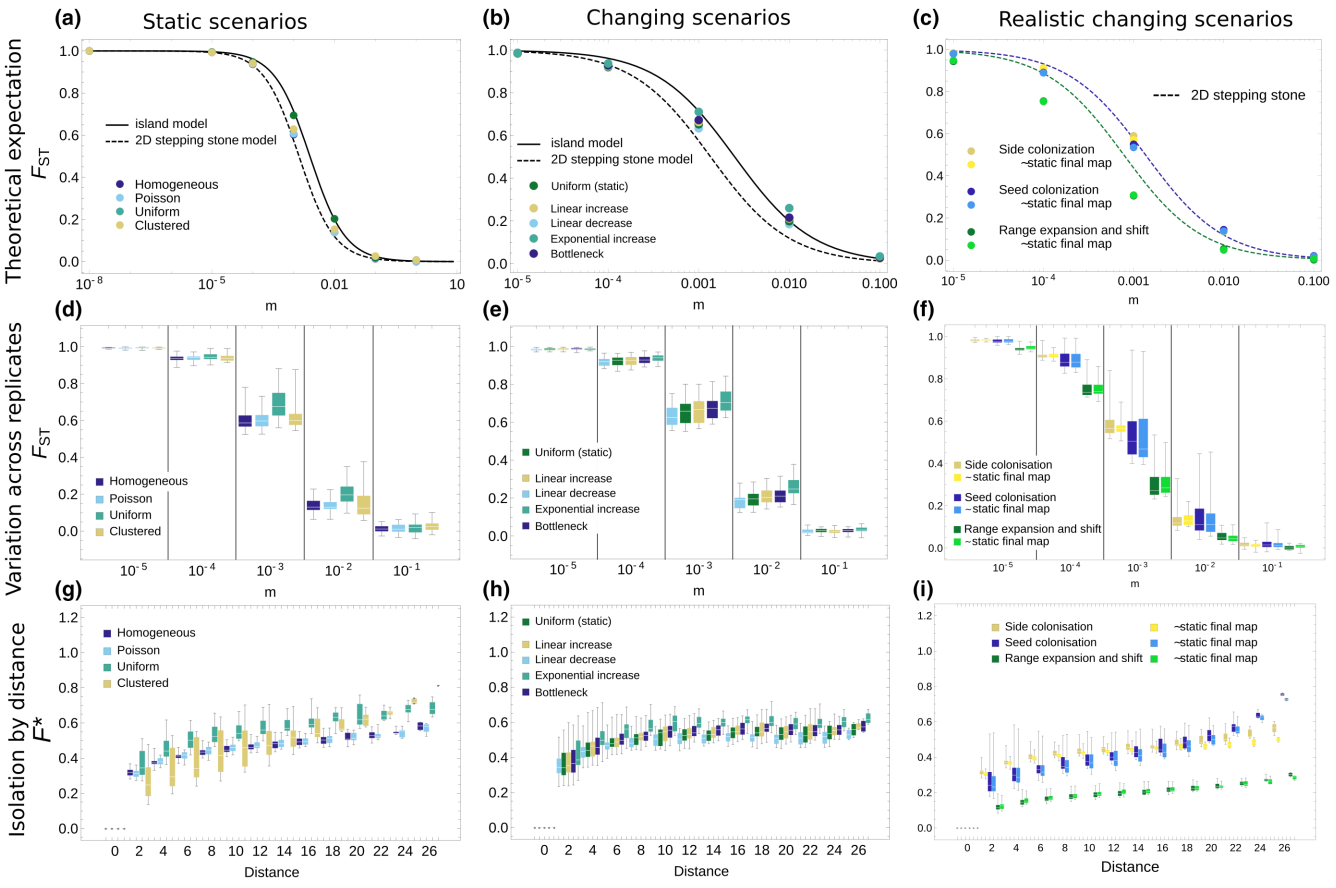


FIGURE 3 Comparison of F_{ST} from different scenarios with theoretical predictions of the island model and the 2D stepping stone model (a, b, c). Variation in F_{ST} across 1000 replicate simulations of the same map (d, e, f). Isolation-by-distance patterns characterized as F^* against distance for different scenarios (g, h, i). Simulation parameters: $L = 30$, $T = 30$, $dt = 50{,}000$, $gt = 25$

observed that the average coalescence time was underestimated (Figure 2b,e). The problem of escaping lineages matters the most in the transition phase from low to high *Nm*. Recall that in this part of the parameter space we could not estimate the theoretical mean of the coalescence time of the samples; therefore, we show the median T_W instead (Figure 2). Note that these results should be treated with caution and cannot be compared with any theoretical expectations. We found that for $Nm \leq 0.05$, T_W was best predicted by $2\bar{N}$ (Figure 2b; $m = 0.001$ and $N = 50$). Then, for $Nm = 0.1$, our simulations showed that the median coalescence time was best predicted by twice the neighbourhood size, that is, the size of the deme plus that of its four neighbours (Figure 2e; $m = 0.001$ and $N = 100$). However, we found that already at $Nm = 0.5$ the coalescence time was best predicted by $2NL^2$ (Figure 2; see also Figure S1b,e), suggesting that the transition phase is fast, which is in agreement with previous observations by Kimura and Maruyama (1971).

3.2 | Global F_{ST}

The island (Equation 1) and 2D stepping stone models (Equation 2) provide expectations for the strength of population structure (F_{ST}) in subdivided populations. Here, we explored the robustness of these predictions with respect to the spatio-temporal heterogeneity in deme size. We found that all simulated scenarios deviated the most from theoretical predictions for intermediate migration rates (or *Nm*), where the predictions of the two models also differed the most (Figure 3a, b and c). Not surprisingly, the island model provided, on average, a better approximation than the spatially explicit 2D stepping stone model when the deme sizes were drawn from a uniform distribution across the grid, thus when there was no spatial autocorrelation in deme size (Figure 3a). In contrast, when deme sizes were homogeneous across space, and thus the spatial autocorrelation was maximized, F_{ST} was closer to the 2D stepping stone model predictions (Figure 3a). F_{ST} of the clustered and low variance maps were in between the two model predictions. F_{ST} also varied considerably across replicates, with the largest variation occurring for the uniform map among all static scenarios considered (Figure 3d).

Under realistic scenarios of change, we observed a consistent bias: scenarios where the mean deme size decreased over time (declining population) gave a lower F_{ST}, while scenarios where the mean deme size increased (expanding population) gave a higher F_{ST} in comparison to the static equivalent scenarios (Figure 3b,c). Similar to the uniform static map, F_{ST} was relatively close to the predictions of the island model under realistic scenarios of change that ended in a uniform map (Figure 3b). More unexpectedly, under realistic scenarios of change, where we decoupled the spatial and temporal autocorrelation, on average, F_{ST} did not deviate more from the island model prediction than the simple changing scenarios for the studied parameter combinations (Figure 3c). These realistic changing scenarios also provided a relatively close match to their static equivalents (Figure 3c). An exception is the range expansion and shift scenario. This is because here the theoretical expectation is shown for the average *N* across the grid, which is lower than the row of sampled demes situated in the middle of the grid (Figure 1). Finally, the variation in F_{ST} across replicates was important, and F_{ST} for different realizations of the same map overlapped between values of *Nm* that were one order of magnitude different, especially for low and intermediate values (Figure 3d, e and f). The sampling variance in F_{ST} also increased with spatial variance in deme size across the grid, with the highest values corresponding to the two clustered maps (Figure 3d and f).

3.3 | Genetic distance (F^*)

Varying *N* and *m* across a homogeneous map showed that increasing the deme size and/or the migration rate led, as expected from Equation (1), to weaker differentiation between demes (Figure S2). The degree of spatial variance in deme size affected both the average genetic distance between demes and the shape of the isolation by distance curves (Figure 3g). Maps with homogeneous deme size had the lowest and flattest isolation by distance curves. Note that these can be treated as a baseline expectation under the 2D stepping stone model (Slatkin, 1993). The uniform map gave higher F^* values across all the distance classes, that is, the isolation by distance curve was shifted upwards, because the compared pairs of demes had, on average, a different size. The clustered map resulted in a lower mean F^* for small distance classes and a higher F^* for larger distance classes, meaning that the isolation by distance curve was steeper. This was because pairs of demes located close to each other tended to have similar sizes, and those for large distance classes often had different sizes. Varying deme classes also caused a large variance in the genetic distance across replicates (Figure 3g–i).

Demographic, that is temporal, changes introduced a bias in the same direction as in F_{ST}: scenarios where the mean deme size decreased over time had a lower F^*, while scenarios with increasing average deme size had a larger F^* value in comparison to a static uniform map. However, the shape of the isolation by distance curve did not change (Figure 3h). The realistic scenarios of change had all increasing deme sizes, so we observed the same upward bias as before (Figure 3i). The fact that the population sizes were changing had the strongest influence on F^* when the spatial and temporal autocorrelation was the most decoupled, that is for the side colonization scenario, and for large distance classes (Figure 3i).

3.4 | Effect of spatial and temporal resolution

Our simulations were carried out on a finite square grid of $L \times L$ (not a torus), which implies a finite number of demes and that demes on the edge of the grid had only two or three neighbours. Not surprisingly, we found that F_{ST} estimated from a larger grid provided a better fit to the predictions of the island model and the 2D stepping stone model (Figure 4a). Further, we found that there was an edge effect, which led to the overestimation of F^* for demes that were *L* or nearly

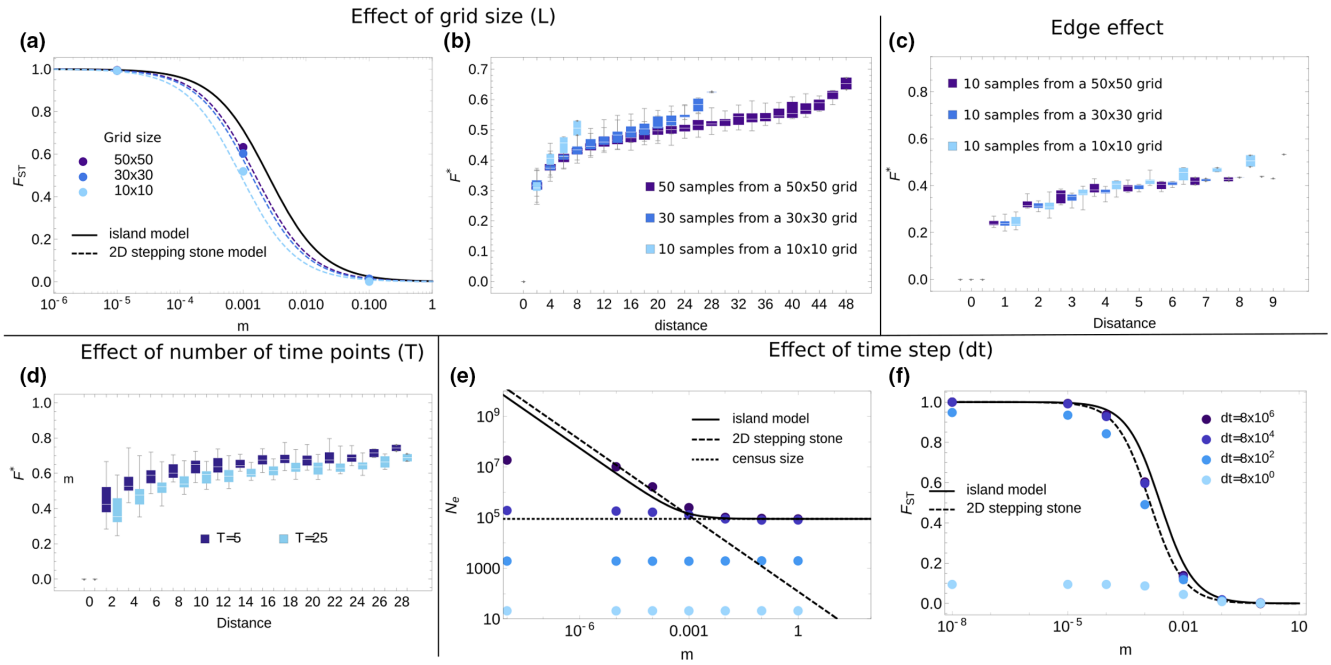


FIGURE 4 Sensitivity of F_{ST} and F^* to L(a, b), T(d), and N_e (estimated as $T_T/2$, where T_T is the average coalescence time) and F_{ST} to dt(e, f) using a homogeneous map, $\overline{N} = 100$, and $m = 10^3$. Unless otherwise specified, $dt = 2\times10^8$, $gt = 25$, and $T = 5$. The edge effect (c) was explored using inner demes sampled along a line in the middle of the grid

L steps away from each other (Figure 4b). Analyzing F^* against distance from samples in the middle the grid allowed us to disentangle the effect of grid size from the edge effect. We found that close to the edges the genetic distance is overestimated between the demes, mainly due to edge effects, while grid size principally influences the precision of the estimates, that is, larger grids provide more precise estimates of F^* for a given distance class (Figure 4c).

We investigated the effect of temporal resolution in the case of a simple changing scenario where all demes increased linearly in size (Figure 1). The coarser time resolution ($T = 5$) did not have a noticeable effect on the estimation of the mean coalescence time within demes, but the between-deme coalescence times were systematically overestimated (not shown). As a result, the genetic distances between demes were also overestimated (Figure 4e). This is because when $T = 5$, the population size at any time is larger than in the finer time-resolution scenario ($T = 25$). Time resolution is also important in more complex setting such as range expansion and shift (Figure 1).

The time necessary for lineages to coalesce during the spatially explicit phase of the simulations may become a limitation in practical applications. When the spatially explicit phase is too short compared with the deme sizes, the coalescence time between lineages is determined by the non-spatial coalescence process of the panmictic ancestral population. Extremely long simulations may be required to reliably estimate the coalescence time when the deme sizes are large. Figure 4(e) shows estimates of N_e calculated as half of the mean total coalescence time. In contrast, Figure 4(f) demonstrates that it is possible to obtain relatively precise estimates of F_{ST} with much shorter simulation times. This is because F_{ST} is defined as a ratio of coalescence times and the biases cancel out. Indeed, both the estimation of within-deme and total coalescence times are biased because of the same process, that is the limited length of the spatially explicit phase, which means that their distributions are missing the same amount from the tails on the right side. This result also highlights that F_{ST} is dependent only on recent demographic events and is independent of the deeper ancestry, which makes it a useful measure.

3.5 | Application example: *Abies alba* post-glacial colonization history

Silver fir (*Abies alba* Mill.) is a coniferous tree species that has progressively colonized the mountainous regions of Europe from different refugia since the last glacial maximum (LGM, 21 kyrs BP). While the exact location of the refugia are debated, it is generally agreed that the Central and/or Northern Apennines hosted the largest populations in pre-LGM times, with other important populations occurring on the Balkan Peninsula (Tinner et al., 2013). Mitochondrial DNA variation clearly suggests the presence of two haplotypes corresponding to the Italian and Balkan Peninsulas (Liepelt et al., 2009; Ziegenhagen et al., 2005; Figure S3a).

The demographic history of silver fir over the past 22 kyrs BP was obtained from the LPX-Bern dynamic global vegetation model with a resolution of 1° by 1° Lat/Lon (Ruosch et al., 2016; Sitch et al., 2003). The model was forced with climate anomalies and included competition between common tree species and plant functional types. The output of LPX-Bern is the Foliar Projective Cover (FPC), which is the fraction of a grid cell that is covered by silver

fir. We estimated the number of trees (N) in each deme from FPC, assuming that a mature tree occupies 40 m^2, and that $N/N_e = 0.001$ (an arbitrary but realistic value [Waples et al., 2011]). The full input data consisted 221 time points spaced at 100 year (i.e. four generations) intervals on a 53 × 24 grid. In the following we shall refer to one grid cell of LPX-Bern as one deme. While the population size of the whole species (i.e. all demes) showed an overall increasing trend with time (post-LGM colonization), the size fluctuations of individual demes were highly variable (see Figure 5a). We used the expected coalescence time for two samples taken from the same deme as an approximation for the genetic diversity in a deme, thus assumed that mutations can be neglected. Finally, we note that LPX-Bern has several shortcomings and does not predict the current distribution of silver fir accurately. However, the objective of this example was not to make predictions for the expected levels of genetic diversity in silver fir, but to study the effect of spatio-temporal heterogeneity in population size in a biologically realistic scenario.

We performed four simulated scenarios. First, we used a homogeneous deme size (i.e. N_e) in space and time, which represented our null model. We fixed the deme size to its average size based on the last step of the LPX-Bern data. Second, we included the spatial variation in deme size, represented by the last step of the LPX-Bern data, but kept deme sizes constant in time. Third, we used the full LPX-Bern input data, thus considering realistic deme sizes changing both in space and in time. Fourth, we explored the effect of having two ancestral populations, that is using pre-LGM historical information. For this, we used simulations identical to the third scenario, but at the oldest time point (i.e. 22 kyrs BP), we combined the demes into one of the two most plausible ancestral populations based on the spatial distribution of mtDNA haplotypes in contemporary samples (Figure S3a). We achieved this by simply assigning each deme with mtDNA data to the dominant haplotype (i.e. more than 50% Balkan or Italian type) or to the origin of the nearest deme of known origin, in case of missing data (Figure S3b).

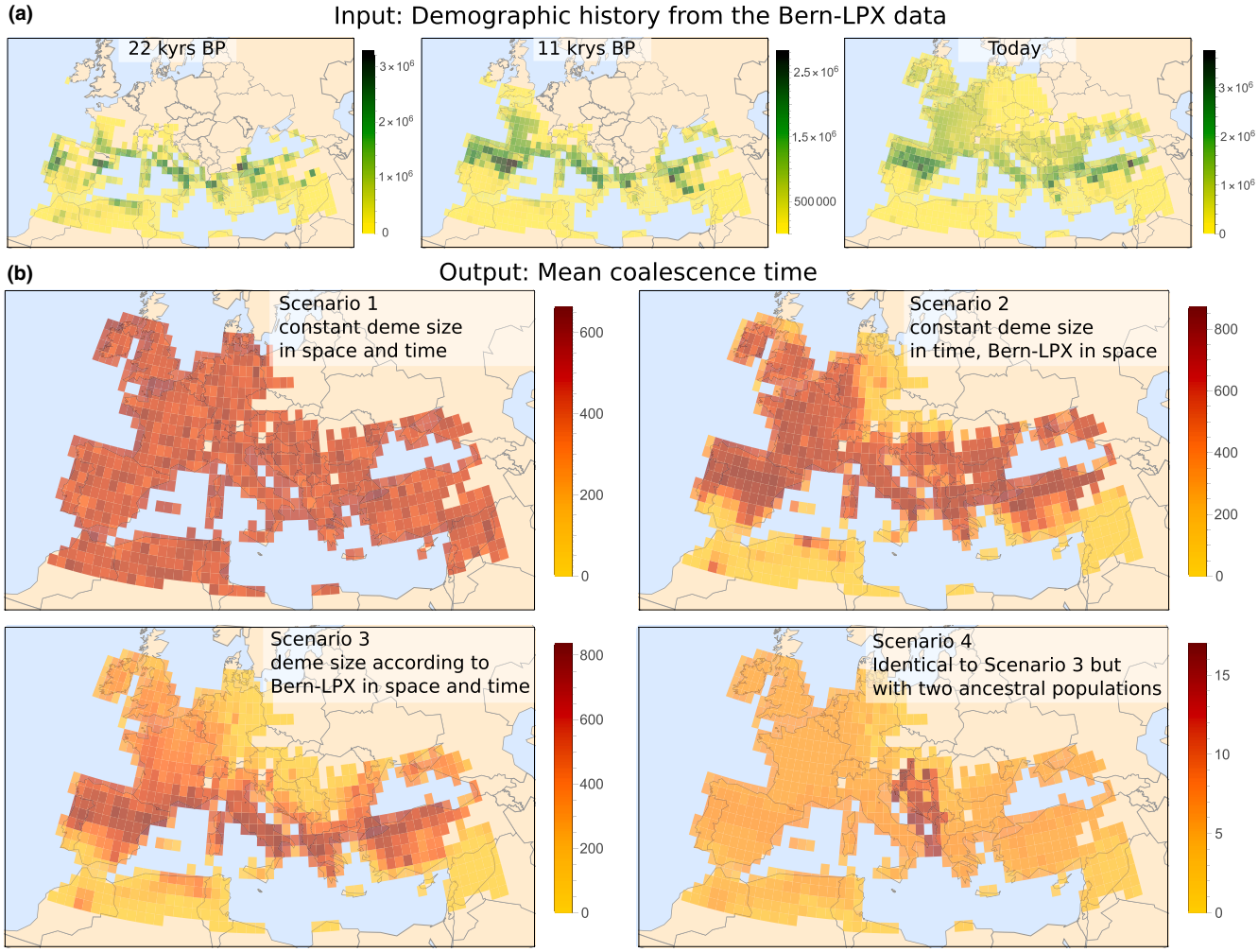


FIGURE 5 Real-world example: Range expansion of silver fir (*Abies alba* mill.) since the last glacial maximum (LGM, 22 kyrs BP). (a) Raw input data for *gridCoal*: The demographic history from the global dynamic vegetation model LPX-Bern. Three time points are shown out of the 220: The LGM, the beginning of the Holocene, and today. (b) Mean coalescence time from the simulated scenarios with increasing complexity in terms of spatio-temporal variation in deme size from the top left to the bottom right panel. Note that the colour scale differs between maps

We found that both space and time had an effect on the coalescence times, and thus on the distribution of genetic diversity in space (Figure 5b). As expected, when the deme size was constant in space and time, the distribution of genetic diversity only reflected stochastic effects of the coalescence process (Figure 5b, Scenario 1). Spatial variation in deme size introduced variation in the expected levels of genetic diversity, which was also proportional to the deme size (Figure 5a and b, Scenario 2). When deme size varied both in space and time, the spatial variation in the mean coalescence time became even stronger. In particular, the recently colonized areas of Northern Europe had a lower expected level of genetic diversity (Figure 5b, Scenario 3). Finally, when we assumed two ancestral populations, their contact zone had much higher levels of expected genetic diversity (Figure 5b, Scenario 4). This is because there was a much longer waiting time for the two ancestral populations to coalesce, which is determined by the size of these populations and also by the migration rate between them. For a real data application, calibration of these two parameters would be necessary to match the observed genetic diversity data. Alternatively, the match between simulated and observed data could be used to estimate the divergence time between the two mtDNA haplotypes (e.g. Hickerson et al., 2007).

4 | DISCUSSION

4.1 | The role of spatial and temporal autocorrelation

Using a wide range of simplified and biologically realistic simulations, we have identified several factors that may cause a deviation from theoretical expectations of the island model and the 2D stepping stone model. We found that non-spatial null models, such as the island model, are inappropriate in the presence of spatial autocorrelation in deme size (Figure 3). Most real-life situations involve some degree of spatial autocorrelation. Previous studies have already demonstrated the limitations of non-spatial null models, for example in the presence of isolation by distance (Meirmans, 2012; Wang & Whitlock, 2003), due to population structure or biased sampling schemes (Chikhi et al., 2010), or to local variation in deme size or barriers to gene flow across the landscape (Duforet-Frebourg & Blum, 2014). Here, we show that the 2D stepping stone model can account for spatial autocorrelation, at least when it is homogeneous across the landscape, and to some extent when there is local variation in deme size (clustered scenario) (Figure 3). Thus far, the 2D stepping stone model has rarely been used as a null model (but see Duforet-Frebourg & Blum, 2014 and Battey et al., 2020), partly due to the lack of a simulation tool. *gridCoal* could facilitate more widespread use of the 2D stepping stone model to generate the null distributions of neutral statistics, such as genetic diversity (assuming a non-zero mutation rate) or F_{ST}, in the presence of spatial autocorrelation in population size.

Demography, or temporal change in population size, is well known to contribute to deviations from theoretical expectations of the island model, and can limit the validity of statistical procedures that are based on this model, such as F_{ST}-outlier tests used to detect loci under selection (e.g. Bierne et al., 2013; Chikhi et al., 2010; De Mita et al., 2013; Lotterhos & Whitlock, 2014). Here, we simulated realistic scenarios with the presence of both spatial and temporal heterogeneity in deme size, and observed that deviations from the theoretical expectations are strongest when the spatial and temporal autocorrelation in deme size are decoupled (Figure 3c, Range expansion and shift). Our simulations also demonstrate that neutral F_{ST} is well below the theoretical expectations for such a range expansion and shift (Figure 3c). This result is in agreement with previous findings. Wegmann et al. (2006) also studied a range expansion and shift scenario and found that spatial heterogeneity in the carrying capacity of demes leads to an increased population differentiation. Further, Lotterhos and Whitlock (2014) showed that spatial autocorrelation in deme size or recent range expansion resulted in the largest number of false positives for most methods in efforts to detect spatially divergent selection. Spatio-temporal trends in population size are expected to be common in nature, especially in the Northern hemisphere, where the demographic history is often dominated by expansion from glacial refugia and a shift towards the north (e.g. Excoffier et al., 2009). Our example application also illustrates such a case (Figure 5).

F_{ST} and F^* are based on the same information, but F_{ST} is a more integrative and therefore more robust measure, while F^* is more sensitive to local differentiation patterns (Figure 3g–i). Note that our F^* is closely related to $\widehat{M}$ of Slatkin (1993), which has the advantage of being independent of the mutation rates when they are small across loci. Based on a wide range of scenarios, we found that spatial and temporal variation in deme size can influence the steepness of the isolation-by-distance curve. In agreement with Duforet-Frebourg and Blum (2014), we found that local variation in population size, as in our clustered map, caused large variance in local F^* (Figure 3g). The most complex range expansion and shift scenario led to a relatively flat isolation-by-distance curve (Figure 3i). Indeed, Slatkin (1993) already proposed that the lack of an isolation-by-distance pattern in a natural population can indicate non-equilibrium populations or recent colonization, a pattern that has been confirmed through empirical studies (e.g. De Kort et al., 2014; Leblois et al., 2000).

4.2 | *gridCoal*: Guidelines for future users

gridCoal is a wrapper for the most efficient algorithm to simulate genealogies, *msprime* (Kelleher et al., 2016). It complements the existing arsenal of spatially explicit simulators (Becheler et al., 2019; Currat et al., 2019; Dellicour et al., 2014; Guillaume & Rougemont, 2006; Haller & Messer, 2019; Landguth & Cushman, 2010). The main merit of gridCoal is that it facilitates the use of spatial null models for empirical biologists and/or for users, who are less familiar with *msprime*. *gridCoal* generates the fairly complex input data needed for *msprime* simulations (multiple backward matrices and demographic events) from simple inputs with clear biological meaning, such as generation

time, population size, or migration rates. Therefore, we hope that it will make spatially and temporally explicit coalescence simulations available for a wide range of audience. Nevertheless, the choice of parameters and model calibration are essential for using *gridCoal*; thus, here we provide some guidelines for future users.

The spatially explicit phase (given by the number of steps *T* and the time step *dt*) should be long enough so that lineages coalesce during this phase, but also short enough to avoid wasting computational time. The choice of *dt* should be driven by the particular biological question. For example, throughout this paper we used a combination of parameter values (number of steps *T*, time step *dt* and generation time *gt*) such that most lineages coalesced in the spatially explicit phase across all combinations of *N* and *m* (Figure 4d–f). Note that the largest *dt* was necessary for intermediate values of *Nm*, where lineages can escape and take a long time to coalesce. We suggest that users perform test simulations with the required values of *N* and *m* to choose an appropriate *dt*. This is particularly important if it is necessary that all lineages coalesce during the spatially explicit phase, e.g. for estimating genetic diversity maps such as those shown in the example of post-glacial colonization of silver fir (Figure 5). In contrast, if the question concerns a particular organism with a given generation time and across a particular time period, the parameters can be chosen accordingly. For example, setting $dt = 100$ and using 210 time points takes the ancestral population back to the last glacial maximum (LGM, 21 kya), which could be a suitable parameter combination for several species that expanded after the LGM.

gridCoal avoids the simulation of genetic data and instead simulates summary statistics that can be derived from coalescence times, that is, gene diversity, the strength of population structure (F_{ST}), and the genetic distance between pairs of demes (F^*). We emphasize that approximating summary statistics of genetic diversity and F_{ST} from coalescence times holds only when the mutation rate is low and when migration is possible to neighbouring demes only (Slatkin, 1985). Further, for comparing *gridCoal* simulations to real data, a calibration of N_e and mutation rate is necessary because these parameters are non-identifiable. Such a calibration can be achieved by using additional information about the mutation rate of particular genetic markers used and by estimating N/N_e (Waples et al., 2011). Finally, simulations from *gridCoal* are closer to that of a continuous space model, and thus to biological reality, for large grid sizes. Nevertheless, at least for small neighbourhood sizes, a grid of 50×50 already appears to be sufficient to accurately approximate a continuous space process for many commonly used summary statistics (see details in Battey et al., 2020). In empirical applications, the grid size will be determined by the resolution of the input data. For example, in the example of post-glacial colonization of silver fir, we used a grid size of 53×24 (Figure 5), which was the resolution of the LPX Bern data.

4.3 | *gridCoal* for eco-evolutionary data fusion

gridCoal might be useful for empirical applications of eco-evolutionary data-fusion approaches, such as integrative Distributional Demographic Coalescent (iDDC) approach (Brown & Knowles, 2012; He et al., 2013). In this context, one key feature of *gridCoal* is that it is not only spatially but also temporally explicit. Temporal explicitness means that the exact population size of each deme has to be set by the user at regularly placed time intervals. In this way, *gridCoal* is fully deterministic in terms of the forward-time demography, and stochastic in terms of the backward coalescence events. Although this feature may appear as a limitation in some situations, it is necessary for applications that make use of species distribution data issued from ecological models and paleological data (Gavin et al., 2014; Svenning et al., 2011). This feature also represents an important contrast to *SPLATCHE 3* (Currat et al., 2019), where each deme follows a logistic growth model. As a result, in *SPLATCHE 3*, user-provided population sizes are only approximately achieved, no population declines, and only local extinctions are possible. Indeed, to set up explicit temporal changes in population size, Ortego and Knowles (2020) updated the population sizes only three times from 21 ky BP to the present, which is a rough approximation of actual population size changes and may bias the isolation-by-distance patterns (Figure 4d).

There is a wide range of possible input data sets that can be used for eco-evolutionary data-fusion approaches. First, the availability of global paleo-climatic databases, such as in CHELSA (Karger et al., 2021), opened the possibility for predicting species distributions in the past using species distribution models (SDMs; Elith & Leathwick, 2009; Sexton et al., 2009). SDMs can be used to generate a time series of species abundances (Calabrese et al., 2014), which may be interpreted as relative deme sizes. Second, process-based dynamic vegetation models (DVMs; Pereira et al., 2010) hold a great potential for use in data-fusion approaches. This is because unlike SDMs that are based on simple empirical correlations of presence and climate, DVMs also integrate biological processes such as competition or dispersal. Even though DVMs also suffer from limitations related to their complex parametrization, they are continually improving as the quality and richness of climatic, remote sensing, and other biological data increases (e.g. Hartig et al., 2012). Third, fossil data is increasingly being organized in databases (Davis et al., 2013; Peters et al., 2019). The most abundant type of fossil data is pollen, particularly from forest trees, which has been used to reconstruct past population size fluctuations (e.g. Kaufman et al., 2020; Ruosch et al., 2016). Indeed, our example of post-glacial colonization history in silver fir (see Figure 5) could also benefit from more realistic estimates of relative deme sizes by using the spatio-temporal interpolation of pollen records (Ruosch et al., 2016).

AUTHOR CONTRIBUTIONS

Katalin Csilléry designed the research. Enikő Szép and Barbora Trubenová developed *gridCoal*, ran the simulations and analysed the data. Enikő Szép, Barbora Trubenová and Katalin Csilléry wrote the manuscript.

ACKNOWLEDGEMENTS

ES was supported by an IST studentship provided by IST Austria. BT was funded by the European Union's Horizon 2020 research

and innovation programme under the Marie Sklodowska-Curie Independent Fellowship (704172, RACE). This project received further funding awarded to KC from the Swiss National Science Foundation (SNSF CRSK-3_190288) and the Swiss Federal Research Institute WSL. We thank Nick Barton for many invaluable discussions and his comments on the thesis chapter and this manuscript. We thank Peter Ralph and Jerome Kelleher for useful discussions and Bisschop Gertjan for comments on this manuscript. We thank Fortunat Joos for providing us with the raw data from the LPX-Bern model for silver fir, and Willy Tinner for helpful insights about the demographic history of silver fir. We also thank the editor Alana Alexander for useful comments and advice on the manuscript. Open access funding provided by Eidgenossische Technische Hochschule Zurich.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Detailed information about the software, a user manual, and example simulations are available for download at https://github.com/Trubenova/gridCoal.

ORCID

Barbora Trubenová https://orcid.org/0000-0002-6873-2967
Katalin Csilléry https://orcid.org/0000-0003-0039-9296

REFERENCES

Barton, N., Etheridge, A., & Véber, A. (2010). A new model for evolution in a spatial continuum. *Electronic Journal of Probability*, *15*, 162–216.

Barton, N. H., Depaulis, F., & Etheridge, A. M. (2002). Neutral evolution in spatially continuous populations. *Theoretical Population Biology*, *61*, 31–48.

Barton, N. H., Kelleher, J., & Etheridge, A. M. (2010). A new model for extinction and recolonization in two dimensions: Quantifying phylogeography. *Evolution*, *64*, 2701–2715.

Battey, C., Ralph, P. L., & Kern, A. D. (2020). Space is the place: Effects of continuous spatial structure on analysis of population genetic data. *Genetics*, *215*, 193–214.

Becheler, A., Coron, C., & Dupas, S. (2019). The quetzal coalescence template library: A C++ programmers resource for integrating distributional, demographic and coalescent models. *Molecular Ecology Resources*, *19*, 788–793.

Bierne, N., Roze, D., & Welch, J. J. (2013). Pervasive selection or is it ...? Why are FST outliers sometimes so frequent? *Molecular Ecology*, *33*, 2061–2064.

Bradburd, G. S., & Ralph, P. L. (2019). Spatial population genetics: It's about time. *Annual Review of Ecology, Evolution, and Systematics*, *50*, 427–449.

Brown, J. L., & Knowles, L. L. (2012). Spatially explicit models of dynamic histories: Examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American Pika. *Molecular Ecology*, *21*, 3757–3775.

Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, *23*, 99–112.

Chikhi, L., Sousa, V. C., Luisi, P., Goossens, B., & Beaumont, M. A. (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*, *186*, 983–995.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., van der Schrier, G., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., ... Zang, C. (2015). Old world megadroughts and pluvials during the common era. *Science Advances*, *1*, e1500561.

Cox, J. T., & Durrett, R. (2002). The stepping stone model: New formulas expose old myths. *Annals of Applied Probability*, *12*, 1348–1377.

Currat, M., Arenas, M., Quilodran, C., Excoffier, L., & Ray, N. (2019). SPLATCHE3: Simulation of serial genetic data under spatially explicit evolutionary scenarios including long-distance dispersal. *Bioinformatics*, *35*, 4480–4483.

Davis, B. A., Zanon, M., Collins, P., Mauri, A., Bakker, J., Barboni, D., Barthelmes, A., Beaudouin, C., Bjune, A. E., Bozilova, E., & Bradshaw, R. H. (2013). The European modern pollen database (EMPD) project. *Vegetation History and Archaeobotany*, *22*, 521–530.

De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O., & Mergeay, J. (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology*, *23*, 4709–4721.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*, 1383–1399.

Dellicour, S., Kastally, C., Hardy, O. J., & Mardulyn, P. (2014). Comparing phylogeographic hypotheses by simulating dna sequences under a spatially explicit model of coalescence. *Molecular Biology and Evolution*, *31*, 3359–3372.

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, *22*, 1185–1192.

Duforet-Frebourg, N., & Blum, M. G. (2014). Nonstationary patterns of isolation-by-distance: Inferring measures of local genetic differentiation with Bayesian kriging. *Evolution*, *68*, 1110–1123.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677–697.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*, 422–433.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*, e1003905.

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 481–501.

Felsenstein, J. (1976). The theoretical population genetics of variable selection and migration. *Annual Review of Genetics*, *10*, 253–280.

Felsenstein, J. (1992). Estimating effective population size from samples of sequences: Inefficiency of pairwise and segregation sites as compared to phylogenetic estimates. *Genetical Research*, *56*, 139–147.

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, *25*, 104–120.

Gavin, D. G., Fitzpatrick, M. C., Gugger, P. F., Heath, K. D., Rodríguez-Sánchez, F., Dobrowski, S. Z., Hampe, A., Hu, F. S., Ashcroft, M. B.,

Bartlein, P. J., Blois, J. L., Carstens, B. C., Davis, E. B., Lafontaine, G., Edwards, M. E., Fernandez, M., Henne, P. D., Herring, E. M., Holden, Z. A., … Williams, J. W. (2014). Climate refugia: Joint inference from fossil records, species distribution models and phylogeography. *New Phytologist*, *204*, 37–54.

González-Serna, M. J., Cordero, P. J., & Ortego, J. (2019). Spatiotemporally explicit demographic modelling supports a joint effect of historical barriers to dispersal and contemporary landscape composition on structuring genomic variation in a red-listed grasshopper. *Molecular Ecology*, *28*, 2155–2172.

Guillaume, F., & Rougemont, J. (2006). Nemo: An evolutionary and population genetics programming framework. *Bioinformatics*, *22*, 2556–2557.

Guillot, G., Estoup, A., Mortier, F., & Cosson, J. F. (2005). A spatial statistical model for landscape genetics. *Genetics*, *170*, 1261–1280.

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright–fisher model. *Molecular Biology and Evolution*, *36*, 632–637.

Hanski, I., & Gilpin, M. (1991). Metapopulation dynamics: Brief history and conceptual domain. *Biological Journal of the Linnean Society*, *42*, 3–16.

Hartig, F., Dyke, J., Hickler, T., Higgins, S. I., O'Hara, R. B., Scheiter, S., & Huth, A. (2012). Connecting dynamic vegetation models to data–an inverse perspective. *Journal of Biogeography*, *39*, 2240–2252.

He, Q., Edwards, D. L., & Knowles, L. L. (2013). Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution*, *67*, 3386–3402.

Hey, J., & Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences*, *104*, 2785–2790.

Hickerson, M. J., Stahl, E., & Takebayashi, N. (2007). msBayes: Pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics*, *8*, 268. https://doi.org/10.1186/1471-2105-8-268

Hudson, R. R. (1990). Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology*, *7*, 44.

Karger, D. N., Nobis, M. P., Normand, S., Graham, C. H., Zimmermann, N. E. (2021) Chelsa-trace21k v1. 0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the past discussions*, pp. 1–27.

Kaufman, D., McKay, N., Routson, C., Erb, M., Davis, B., Heiri, O., Jaccard, S., Tierney, J., Dätwyler, C., Axford, Y., Brussel, T., Cartapanis, O., Chase, B., Dawson, A., de Vernal, A., Engels, S., Jonkers, L., Marsicek, J., Moffa-Sánchez, P., … Zhilich, S. (2020). A global database of Holocene paleotemperature records. *Scientific data*, *7*, 1–34.

Kelleher, J., Etheridge, A., & Barton, N. H. (2014). Coalescent simulation in continuous space: Algorithms for large neighbourhood size. *Theoretical Population Biology*, *95*, 13–23.

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, *12*, e1004842.

Kimura, M. (1953). "stepping-stone" model of population. *Annual Reports National Institute of Genetics*, *3*, 62–63.

Kimura, M., & Maruyama, T. (1971). Pattern of neutral polymorphism in a geographically structured population. *Genetics Research*, *18*, 125–131.

Kimura, M., & Weiss, G. H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, *49*, 561–576.

Landguth, E. L., & Cushman, S. (2010). CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources*, *10*, 156–161.

Leblois, R., Rousset, F., Tikel, D., Moritz, C., & Estoup, A. (2000). Absence of evidence for isolation by distance in an expanding cane toad (bufo marinus) population: An individual-based analysis of microsatellite genotypes. *Molecular Ecology*, *9*, 1905–1909.

Liepelt, S., Cheddadi, R., de Beaulieu, J. L., Fady, B., Gömöry, D., Hussendörfer, E., Konnert, M., Litt, T., Longauer, R., Terhürne-Berson, R., & Ziegenhagen, B. (2009). Postglacial range expansion and its genetic imprints in *Abies alba* (mill.) - a synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, *153*, 139–149.

Lima-Rezende, C. A., Rocha, A. V., Júnior, A. F. C., Martins, É. S., Vasconcelos, V., & Caparroz, R. (2019). Late Pleistocene climatic changes promoted demographic expansion and population reconnection of a neotropical savanna-adapted bird, *Neothraupis fasciata* (Aves: *Thraupidae*). *PLoS One*, *14*, e0212876.

Lima-Ribeiro, M. S., Varela, S., González-Hernández, J., de Oliveira, G., Diniz-Filho, J. A. F., & Terribile, L. C. (2015). EcoClimate: A database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodiversity Informatics*, *10*. https://doi.org/10.17161/bi.v10i0.4955

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, *23*, 2178–2192.

Malécot, G. (1948). *Mathématiques de l'hérédité*. Masson et Cie.

Malécot, G. (1955). Remarks on decrease of relationship with distance, following paper by M. Kimura. *Cold Spring Harbor Symposium Quantitative Biololgy*, *20*, 52–53.

Malécot, G. (1975). Heterozygosity and relationship in regularly subdivided populations. *Theoretical Population Biology*, *8*, 212–241.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*, 189–197.

Maruyama, T. (1971). An invariant property of a structured population. *Genetical Research*, *18*, 81–84.

Maruyama T. (1977) Maruyama lecture notes in biomathematics. 17. Stochastic problems in population genetics.

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, *21*, 2839–2846.

Ortego, J., & Knowles, L. L. (2020). Incorporating interspecific interactions into phylogeographic models: A case study with Californian oaks. *Molecular Ecology*, *29*, 4510–4524.

Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P. W., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. L., Chini, L., Cooper, H. D., Gilman, E. L., Guénette, S., Hurtt, G. C., Huntington, H. P., Mace, G. M., Oberdorff, T., Revenga, C., … Walpole, M. (2010). Scenarios for global biodiversity in the 21st century. *Science*, *330*, 1496–1501.

Peters, K. J., Saltré, F., Friedrich, T., Jacobs, Z., Wood, R., McDowell, M., Ulm, S., & Bradshaw, C. J. A. (2019). FosSahul 2.0, an updated database for the late quaternary fossil records of Sahul. *Scientific Data*, *6*, 1–7.

Quilodrán, C. S., Nussberger, B., Montoya-Burgos, J. I., & Currat, M. (2019). Hybridization and introgression during density-dependent range expansion: European wildcats as a case study. *Evolution*, *73*, 750–761.

Ralph, P., Thornton, K., & Kelleher, J. (2020). Efficiently summarizing relationships in large samples: A general duality between statistics of genealogies and genomes. *Genetics*, *215*, 779–797.

Ruosch, M., Spahni, R., Joos, F., Henne, P. D., Van der Knaap, W. O., & Tinner, W. (2016). Past and future evolution of *Abies alba* forests in Europe–comparison of a dynamic vegetation model with palaeo data and observations. *Global Change Biology*, *22*, 727–740.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 415–436.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant

geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*, 161–185.
Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, *16*, 393–430.
Slatkin, M. (1991). Inbreeding coefficients and coalescence times. *Genetics Research*, *58*, 167–175.
Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*, 264–279.
Smouse, P. E., Peakall, R., & Gonzales, E. (2008). A heterogeneity test for fine-scale genetic structure. *Molecular Ecology*, *17*, 3389–3400.
Sokal, R. R., & Wartenberg, D. E. (1983). A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics*, *105*, 219–237.
Strobeck, C. (1987). Average number of nucleotide differences in a sample from a single subpopulation: A test for population subdivision. *Genetics*, *117*, 149–153.
Svenning, J. C., Fløjgaard, C., Marske, K. A., Nógues-Bravo, D., & Normand, S. (2011). Applications of species distribution modeling to paleobiology. *Quaternary Science Reviews*, *30*, 2930–2947.
Tallavaara, M., Luoto, M., Korhonen, N., Järvinen, H., & Seppä, H. (2015). Human population dynamics in Europe over the last glacial maximum. *Proceedings of the National Academy of Sciences*, *112*, 8232–8237.
Tinner, W., Colombaroli, D., Heiri, O., Henne, P. D., Steinacher, M., Untenecker, J., Vescovi, E., Allen, J. R. M., Carraro, G., Conedera, M., Joos, F., Lotter, A. F., Luterbacher, J., Samartin, S., & Valsecchi, V. (2013). The past ecology of *Abies alba* provides new perspectives on future responses of silver fir forests to global warming. *Ecological Monographs*, *83*, 419–439.
Wakeley, J. (1998). Segregating sites in Wright's Island model. *Theoretical Population Biology*, *53*, 166–174.
Wakeley, J. (1999). Nonequilibrium migration in human history. *Genetics*, *153*, 1863–1871.
Wang, J., & Whitlock, M. C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, *163*, 429–446.
Wang, S., Xu, X., Shrestha, N., Zimmermann, N. E., Tang, Z., & Wang, Z. (2017). Response of spatial vegetation distribution in China to climate changes since the last glacial maximum (LGM). *PLoS One*, *12*, e0175742.
Waples, R. S., Do, C., & Chopelet, J. (2011). Calculating n_e and n_e/n in age-structured populations: A hybrid Felsenstein-Hill approach. *Ecology*, *92*, 1513–1522.
Wegmann, D., Currat, M., & Excoffier, L. (2006). Molecular diversity after a range expansion in heterogeneous environments. *Genetics*, *174*, 2009–2020.
Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, *82*, 117–125.
Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.
Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, *15*, 323–354.
Wright, S. (1978). *Evolution and the genetics of populations* (Vol. *4*). University of Chicago Press.
Ziegenhagen, B., Fady, B., Kuhlenkamp, V., & Liepelt, S. (2005). Differentiating groups of abies species with a simple molecular marker. *Silvae Genetica*, *54*, 123–126.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Szép, E., Trubenová, B., & Csilléry, K. (2022). Using *gridCoal* to assess whether standard population genetic theory holds in the presence of spatio-temporal heterogeneity in population size. *Molecular Ecology Resources*, *22*, 2941–2955. https://doi.org/10.1111/1755-0998.13676