Published in final edited form as:

Mol Ecol Resour. 2019 January ; 19(1): 296–305. doi:10.1111/1755-0998.12958.

Metapop: an individual-based model for simulating the evolution of tree populations in spatially and temporally heterogeneous landscapes

Jean-Paul Soularue^{1,*}, Armel Thöni¹, Léo Arnoux¹, Valérie Le Corre², and Antoine Kremer¹

¹BIOGECO, INRA, Univ. Bordeaux, 33610 Cestas, France

²INRA - UMR1347 Agroécologie, Dijon, France

Abstract

Metapop is a stochastic individual-based simulation program. It uses quantitative genetics theory to produce an explicit description of the typical life cycle of monoecious and hermaphroditic plant species. Genome structure, the relationship between genotype and phenotype, and the effects of landscape heterogeneity on each individual can be finely parameterized by the user. Unlike most existing simulation packages, Metapop can simulate phenotypic plasticity, which may have a genetic component, and assortative mating, two important features of tree species. Each simulation is parameterized through text files, and raw data are generated recurrently, describing the allelic state of each quantitative trait locus involved in phenotypic variability. The data can be generated in Genepop or Fstat format, and may thus be analyzed with other existing packages. Metapop also automatically computes a range of populations statistics, enabling the user to monitor evolutionary dynamics directly, from gene to metapopulation level.

Keywords

stochastic individual-based modeling; quantitative genetics; simulation; landscape heterogeneity; phenotypic plasticity; assortative mating; forest trees

Introduction

Forest trees are particularly exposed to environmental changes, because of their long generation time. The increasing availability of phenotypic and genetic data, now produced at the population scale, is providing opportunities to improve our understanding of the evolutionary dynamics of tree populations. However, despite the strong mathematical foundations of evolutionary theory, the direct application of existing theoretical models to concrete situations remains difficult (Chevin et al. 2010, Segelbacher et al. 2010). As a result, mechanistic explanations of experimentally characterized patterns of genetic variability are still lacking. These gaps in our knowledge make it difficult to infer the

*correspondence: jean-paul.soularue@inra.fr.

Data accessibility

All results can be recreated using Metapop version 2.2.5, Nemo version 2.3.44 and Quantinemo version 2.0.0, in combination with the analysis configuration files provided in the supplemental material.

evolutionary responses of tree populations to global changes (but see Aguilée et al. 2016) and limit the extent to which evolutionary theory can be refined on the basis of empirical data collected in forest tree species (Shou et al. 2015). The rapid increase in computing power over the last three decades has led to the development of genetically explicit individual-based simulation programs (see Hoban et al. 2012 and Peng et al. 2013 for reviews), a bottom-up modeling approach based on algorithmic descriptions of each of the individuals of the populations of interest. In these models, each individual carries a set of genes and undergoes a sequence of evolutionary processes during its lifetime, more commonly referred to as its life cycle. The resulting abstractions explicitly model the most relevant biological specificities of the species and the selection landscape for the hypothesis to be tested, while avoiding oversimplifying assumptions. Simulation models can also be extensively customized, as the software in which they operate can handle fine parameterizations of the model and calculate numerous outputs for monitoring evolution at the level of genes, traits or species demographics. Individual-based evolutionary modeling thus enables users to make use of the available experimental data collected at individual level, to test a wide range of hypotheses. This approach can provide new insight at the various levels of the systems studied, from gene to metapopulation level.

Metapop, a genetically and spatially explicit individual-based simulation program, was designed to model the evolution of tree populations in spatially and temporally heterogeneous environments. Metapop is a forward-in-time simulator: it explicitly simulates the life cycle of each individual over successive non-overlapping generations (Hoban et al. 2012, Schiffrers and Travis 2014). By contrast, backward-in-time simulators consider whole lineages and adopt the coalescent approach: starting from individual genes, they search for the most recent common ancestor and reconstruct gene trees (Kingman 1982, Hoban et al. 2012). Metapop, which is based on quantitative genetic theory, explicitly describes (i) the relationship between the genotypes and phenotypes of the individuals, and (ii) the effects of landscape heterogeneity on the traits studied. Several similar spatially explicit simulators based on quantitative genetics have been developed (*e.g.* Nemo- Guillaume and Rougemont 2006, Quantinemo-Neuenschwander et al. 2008, Forsim- Lambert et al. 2008, Pedagog-Coomb et al. 2010, Aladyn- Schiffrers and Travis 2014), and a generic programming environment called Simupop has also been established (Peng and Kimmel 2005). All were designed with flexibility and genericity in mind, but each has a unique set of functionalities and is based on its own set of assumptions, particularly as regards the biological life cycle simulated. These differences and assumptions are of the utmost importance, because they can lead to discrepancies when similar questions are addressed (see for instance table 1 in Balkenhol and Landguth 2011 and Soularue et al. 2017). Therefore, when relevant, simulation models must describe the biology of the species of interest and the specific features of the environment. Metapop is unique in that it simulates the evolution of phenotypic plasticity (see for instance Pigliucci 2001 for a definition) and assortative mating, two important characteristics of forest trees (Franjic et al. 2011, Soularue and Kremer 2014). Metapop is, therefore, highly suitable for use in studies of the evolutionary dynamics of monoecious or hermaphroditic sessile species in which male gametes and seeds may be dispersed over long distances. Table 1 provides an overview of the principal

differences in general features between Metapop and several of the most frequently cited spatially explicit forward-in-time simulators based on quantitative genetics.

Main features

The core of Metapop (*i.e.* the model part), was developed in C++. Metapop also includes a range of R, Python and Lua scripts, which analyze user input files for parameterization of the compiled model, and process the data generated by the simulations. The complete source code is available under the GNU general public license from <https://quercusportal.pierroton.inra.fr/index.php?p=METAPOP>. The model consists of four major interacting components: (i) the genome of the individuals, (ii) the relationship between the phenotypes of the individuals and the genome, (iii) the life cycle, defining the sequence of evolutionary processes in each generation, (iv) the influence of the landscape on both the expression of individual traits and the fitness values of the individuals. We provide a brief description of each of these components here. For more information, see the software documentation also available on the website.

Genome

In Metapop, a genome is modeled as a linear succession of diploid loci, subdivided into chromosomes. Each chromosome is indicated by the position of its first locus within the linear succession. Rates of recombination can be set between each pair of successive loci along each chromosome. By default, all loci are considered neutral. Once users have set the genome length, the number of chromosomes and the rates of recombination between loci, they can define the quantitative trait loci (QTLs) contributing to the additive values of the targeted traits. Pleiotropic effects can be simulated by assigning a QTL to more than one trait. In this case the effects of the alleles at a pleiotropic QTL are independent.

Phenotype

The phenotypes of the individuals consist of one or several traits. Each trait Z has a specific genetic architecture, in which the contribution of the environment to the expression of the phenotype can be described in three different ways. In the simplest case, Z is not plastic and simply corresponds to the sum of the additive effects of the alleles at the underlying QTLs and a random microenvironmental contribution ϵ following the centered-reduced distribution $\mathcal{N}(0, 1)$:

$$Z = \sum_{l=1}^n (\alpha_i + \alpha_j) + \epsilon \quad (1)$$

where l is the locus subscript, n is the number of QTLs contributing to the trait, and α_i and α_j are the additive effects of the two alleles at each locus. In the second case, Z is plastic, and also includes a macroenvironmental component E associated with the site at which the population is found, making a uniform contribution to the phenotypic values of all individuals in the population:

$$Z = \sum_{l=1}^n (\alpha_i + \alpha_j) + E + \epsilon \quad (2)$$

In the third case, plasticity is genetically variable and can therefore evolve too, assuming a linear reaction norm between the trait modeled and the E values (see Lande 2009 for an example). In such cases, the expression of the plastic trait depends on two separate sets of loci, a and b , determining the intercept and the slope of the reaction norm, respectively:

$$Z = \sum_{l=1}^{n_a} (\alpha_i + \alpha_j) + \sum_{l=1}^{n_b} (\beta_i + \beta_j)E + \epsilon \quad (3)$$

where n_a and n_b denote the numbers of QTLs associated with the intercept and the slope of the individual reaction norm, respectively, and α and β are the allelic effects at loci a and b , respectively.

Landscape

The landscape is a collection of cells distributed over two dimensions. Each cell is defined by three components. The first component corresponds to the conditions of natural selection to which the individuals are subjected. Natural selection is defined throughout the landscape by assigning a phenotypic optimum Z_{opt} and an intensity of selection ω^2 to each cell and for each trait simulated. These values are used to simulate stabilizing selection in each cell. Divergent selection can be simulated at the landscape (between-cell) level by assigning different phenotypic optima to the cells (figure 1). The second component characterizing a cell is related to phenotypic plasticity, as the macroenvironmental effect E contributes to trait expression (equations 2 and 3). E corresponds to the biotic or abiotic factors affecting the expression of plastic traits, such as mean temperature, which affects the timing of bud burst, an important adaptive character in broadleaved trees (Vitasse et al. 2009). Note that E values differ from the microenvironmental ϵ values, which vary according to the microenvironmental stochastic effects associated with the position of each individual in a cell (equations 1, 2 and 3, above). The third component is the initial number of individuals N and the carrying capacity K assigned to each cell. The input system of Metapop allows users to initialize linear patterns of variations of all of these values rapidly, throughout the landscape, or to assign a value to each cell independently. It is also possible to modify landscape characteristics; the E , Z_{opt} and ω^2 values assigned to each trait at the beginning of the simulation can also be modified during the course of the simulation.

Life cycle

There is no overlap between the simulated generations. Thus, in each generation, the parents are totally replaced by the offspring generated, and all individuals simultaneously undergo the successive steps of the life cycle defined as follows:

(1) *Demography*. Determines the number of offspring N_t to be generated in each cell at generation t . N_t is calculated from the logistic function:

$$N_t = N_{t-1} + g \times N_s \times \left(1 - \frac{N_{t-1}}{K}\right) + \gamma \quad (4)$$

where N_{t-1} is the census population number at generation $t-1$, N_s is the number of seeds reaching the cell, and g and K , which are set by the user, are the growth rate of the species and the carrying capacity of the cell, respectively. Finally, γ is a random variable with a uniform distribution $U(0,1)$.

(2) *Fecundity selection*. Fitness values weight the probability of each individual mating and generating offspring. Stabilizing selection is simulated within each cell. For a phenotype consisting of a single trait Z , the fitness value $W(Z)$ of an individual is calculated according to the deviation of its phenotypic value from the phenotypic optimum Z_{opt} and an intensity of selection ω^2 (Turelli 1984):

$$W(Z) = \exp\left[\frac{-(Z - Z_{opt})^2}{2\omega^2}\right] \quad (5)$$

Both Z_{opt} and ω^2 values are assigned to the cells. Divergent selection between cells can, therefore, be defined by assigning different Z_{opt} values to the cells according to a spatial pattern. If the phenotype of the individual is defined by multiple traits, the user must assign Z_{opt} and ω^2 values to each cell, for each trait, and the fitness W of each individual is calculated as follows (Reeve 2000):

$$W = \exp\left(-0.5 \times (P - \theta)^T \times \Omega^{-1} \times (P - \theta)\right) \quad (6)$$

where P is a vector of the trait values making up the phenotype, and θ is a vector of the optimal trait values assigned to the cell. Ω is a symmetric matrix with a diagonal consisting of intensities of selection for each trait, and off-diagonal values scale the intensity of correlational selection.

(3) *Gene flow*. Seeds and pollen are dispersed within and between cells, according to the Wright's Island migration model (Wright 1931), the stepping-stone migration model (Kimura et al. 1964) or any spatial pattern defined by the user in matrix notation. The values entered by the user correspond to forward migration rates.

(4) *Reproduction*. Two modes of reproduction are simulated in Metapop: random mating and assortative mating according to a single trait. In the case of random mating within a population p , a female parent is randomly drawn from the individuals of p , and each male parent is randomly drawn from the individuals contributing to the pollen cloud reaching p . In both cases, the draws are weighted by the fitness values of the individuals (*i.e.* fecundity).

selection), making up the pools of possible female or male parents. In the case of assortative mating, a correlation ρ between Z_F and Z_M , the phenotypic values of the female and male parents, respectively, is required. For each mating, $Z_M \in [Z_F - \delta, Z_F + \delta]$, with δ following a Gaussian distribution $\mathcal{N}(0, \sigma_\delta)$ where σ_δ is obtained from ρ and the total standard deviation of the trait value in the starting metapopulation $\sigma_{Z_0}^2$:

$$\sigma_\delta = \sqrt{\sigma_{Z_0}^2 \times \left(\frac{1}{\rho^2} - 1 \right)} \quad (7)$$

For each assortative mating in population p , a female parent is first randomly drawn from the individuals of p . The male parent is then drawn from the set of individuals (i) contributing to the pollen cloud reaching population p and (ii) with a phenotypic value Z_M sufficiently close to the phenotypic value of the female parent Z_F such that $Z_M \in [Z_F - \delta, Z_F + \delta]$. As under random mating, the fitness values $W(Z)$ also systematically weight the random draws of female and male parents under assortative mating. Metapop can also simulate a user-specified rate of selfing combined with assortative or random mating.

(5) *Recombination*. Crossing-over events occur during gametogenesis at each locus, at a probability defined by the user.

(6) *Mutation*. For each offspring produced, the alleles at each locus mutate with a probability initially set according to the k -allele mutation model (Peng et al. 2012). Under this model, a mutation changes an allele to any of the other possible $k-1$ alleles. k , the maximum number of allelic states per locus is set by the user. The variance of the allelic effects initially drawn is scaled by the heritability level and the genetic architecture specified for each trait.

Input/Output

The model and the simulations are parameterized through user-authored text files. Each simulation generates raw data corresponding to the genotypes of all individuals at all loci (QTL or neutral), recurrently saved at the different steps in the simulations, and phenotypic and genetic values for the traits. By default, the genotypes are saved in Metapop format, but additional conversion scripts are available within Metapop, enabling the user to make use of the Fstat (Goudet 1995) and Genepop (Rousset 2008) packages to analyze the raw data further. Interestingly, files consisting of genotypic arrays (resulting from Metapop, Fstat or Genepop) can also be used as input source in Metapop, in order to load the starting metapopulation for any other new simulation. In this latter case, the settings of the new simulation can be entirely redefined. By default, each simulation also generates aggregated data calculated from the raw data for all of the replicates of a given scenario. The summary files provide information about demographics, quantitative genetics (genes and traits), or population genetic statistics that can be read directly by the user. Each summary file can easily be plotted with any relevant tool, such as R (Team RC 2000) or Matplotlib (Hunter 2007).

The use of Metapop in the evolutionary genetics of trees

From its earliest development in the late 1990s, the embryonic version of Metapop was used to address key issues in the population and evolutionary quantitative genetics of forest trees. We provide a brief overview here of some of the issues considered, the ways in which they were addressed by simulation, and the major outcomes. A key question in tree microevolution during the Holocene concerns the shaping of the current distribution and structure of genetic diversity by postglacial colonization dynamics. Metapop simulations highlighted the importance of rare long-distance dispersal events during postglacial migration (Le Corre *et al.* 1997). They helped to explain the Reid paradox in the case of European oaks, highlighting the discrepancy between the observed postglacial migration velocity of trees and the migration velocity inferred from traditional dispersal mechanisms and vectors (Clark *et al.* 1998). More specifically, the simulations showed how new populations created by rare long-distance migration events, subsequently grew and coalesced, thereby greatly accelerating the overall progress of the migration front. Austerlitz *et al.* (1997), with a derived version of Metapop, subsequently showed how population differentiation due to these founding events was progressively erased in trees, resulting in the weak population structure observed today. Indeed, simulations made it possible to determine how the long juvenile phase of trees facilitated the progressive reduction of founder effects, through the addition of new migrants to the founder population. Metapop also made a major contribution by resolving the decoupling between the genetic differentiation of traits and their underlying genes in a metapopulation undergoing divergent selection. Even before genetic diversity surveys of candidate genes in natural populations became available, there was some concern that the genes underlying adaptive traits would be much less differentiated than the traits to which they contribute (Lewontin, 1984). Latta (1998) subsequently suggested that the genetic covariances between additive values at the loci controlling the trait might account for more of the between-population variance of the trait than the genetic variance at individual loci. These expectations were confirmed by simulations conducted with Metapop in a broad spectrum of evolutionary scenarios (Le Corre & Kremer 2003, Le Corre & Kremer, 2012). These authors also provided an analytical relationship between differentiation of the trait on the one hand, and differentiation at the loci and their linkage disequilibria on the other hand, under a simplifying hypothesis. Simulations have shown that this relationship holds under most realistic evolutionary scenarios. Kremer and Le Corre (2012) subsequently explored the mechanisms underlying the decoupling related to the genetic architecture of the trait and the transient dynamics of decoupling under a non-equilibrium scenario. Using Metapop, they showed that, for outcrossing species displaying extensive gene flow, such as trees, divergent selection initially captures beneficial allelic associations for the various traits controlling the adaptive trait, and subsequently targets changes in allelic frequencies. A final major contribution relates to analyses of the ways in which genetic clines of phenological traits in trees are shaped by environmental gradients and assortative mating. Empirical results from common garden experiments in trees have shown that phenological traits display strong genetic clines along temperature gradients (Alberto *et al.* 2013). Such patterns were traditionally interpreted as a consequence of divergent selection along the gradient. However, Metapop simulations have indicated that the reproduction system also has a strong influence over

genetic differentiation for the timing of bud burst (TBB) between populations. Indeed, Soularue and Kremer (2012, 2014) showed that in conditions of fixed non-heritable plasticity (equation 1), assortative mating according to TBB reduces pollen flow between distant populations, leading to a spatial sorting of alleles along environmental gradients. This filtering effect results in genetic differentiation between populations along environmental gradients even in the absence of selection, the slope of the resulting genetic clines being related to (i) the degree of correlation between the TBB of mating partners and (ii) the slope of the environmental gradient (Soularue and Kremer 2012, see also Stam 1983). Moreover, Soularue and Kremer (2014) showed that assortative mating favors the migration of beneficial alleles into populations under selection when plasticity is fixed (equation 1) and adaptive (co-gradient variation; Conover 1995), thereby amplifying the genetic differentiation between populations for TBB over that observed for random mating. Conversely, when plasticity is non-adaptive (counter-gradient variation; Conover 1995) assortative mating impairs the population response to selection, slightly decreasing the slope of the genetic cline relative to random mating (Soularue and Kremer 2014). Along the same lines, research efforts are currently directed towards integrating the evolution of phenological trait plasticity into models, to anticipate future population responses to ongoing climate change.

Comparative analysis

We simulated three evolutionary scenarios with Metapop (version 2.2.5) and compared the results obtained with theory. When possible, we also compared the results of the simulations with those produced by two other simulation programs: Nemo (version 2.3.44) and Quantinemo (version 2.0.0). The other simulators cited in Table 1 were excluded from these tests for two different reasons: either they did not allow us to accurately initialize all the parameters required, or they did not provide the necessary output. For each scenario, we simulated 30 independent replicates. The three scenarios can be recreated from the input files provided in the supplementary materials.

Scenario 1: mutational variance

We considered a quantitative character determined additively by 20 independent QTLs and a random micro-environmental effect (equation 1). In this first scenario we were interested in the amount of genetic variability accumulated by mutation in a single population free of selection during 5 000 generations. As reference, we used the model (Lande 1975):

$$V_m = 2n\mu\sigma_a^2 \quad (8)$$

where V_m is the amount of genetic variance introduced each generation by mutation in the absence of drift and selection, n the number of loci, μ the mutation rate and σ_a^2 is the variance of each mutational effect. In our simulations, we set the size of the population to 5 000 in order to limit the effects of drift. Mutation occurred at a rate of 10^{-4} per generation, according to two distinct mutation models: the k -allele model and the Random mutation model (see table 1 for a definition). The k -allele model is proposed by default by Metapop,

while the Random mutation model is proposed by default by Nemo. Quantinemo provides both mutation models. The k -allele model was set here with $k = 256$. In all cases, the variance of the distributions used to draw the allelic and/or the mutational effects was set to 0.12. The starting population was monomorphic and homozygous at each QTL.

As predicted by Lande's model, the genetic variance increased in the simulations. The increases in genetic variance simulated by Nemo and Quantinemo (Random mutation model only) fitted particularly well the predictions of Lande's model during the first 1 200 generations, but then, the genetic variance was accumulated at a lower rate than predicted (figure 1). The accumulation of genetic variance simulated by Metapop was first slightly higher than the prediction of Lande's model, it became then very close to the theoretical predictions between generations 2 000 and 3 000, and also subsequently decreased (figure 1). The rapid decrease of the rate of accumulation of mutational variance observed after thousands of generations in all the simulations resulted from the effects of drift which is not accounted for by Lande's model. Indeed, in finite populations, the variance is predicted to decrease by a rate of $1/Ne$ (Ne being the effective population size). Besides, we also simulated the same scenario with the k -allele mutation model of Quantinemo. The pattern obtained was similar, however the increase generated remained very limited compared to the theoretical prediction and the results of the other simulations (figure 1). This is because, in Quantinemo, the mutational effects generated by the k -allele mutation model are uniformly distributed over an interval defined from the allelic variance specified. For instance, in these simulations, the 256 allelic effects per locus were evenly spaced over the interval $[-0.12, +0.12]$. By contrast, the mutational effects are normally distributed in Nemo and Metapop, which is congruent with the underlying hypothesis of Lande's model.

Scenario 2: mutation-selection balance

In a second step, we monitored the evolution of the genetic variance under mutation and stabilizing selection, starting from the populations obtained at generation 5 000 in scenario 1. The strength of the stabilizing selection simulated was strong ($\omega^2=5$). The settings related to the definition of the trait and mutation were the same than those defined in scenario 1. We simulated 3 000 generations and compared the asymptotic value of the genetic variance with the mutation-selection equilibrium predictions yielded by the "House of Card model" (Turelli 1984, Johnson and Barton 2005):

$$V_g = 4n\mu V_s \quad (9)$$

V_g being the mean genetic variance of the population, n , the number of loci, μ , the mutation rate, and V_s corresponding to the intensity of selection.

As expected, the genetic variance decreased towards asymptotic values close to the theoretical prediction in all cases. Interestingly, the uniform distribution of the mutational effects simulated by the k -allele mutation model in Quantinemo did not alter the mutation-selection balance predicted by the House of Card model (figure 2).

Scenario 3: evolution of phenotypic plasticity

The last scenario is related to the effect of gene flow between populations on the evolution of plasticity across an environmental gradient. When plasticity is heritable (equation 3) and not subject to constraints, the evolution of optimal phenotypic plasticity allows genotypes to express the optimal phenotype in any habitat (Bradshaw 1965, Scheiner 2012). In this case, we expect (i) the convergence of individual reaction norms towards the optimal reaction norm, and (ii) very limited genetic variation for individual reaction norms, at both the within- and between- population levels (Pigliucci 2001). Assuming a linear reaction norm, this situation corresponds to (i) mean intercept and slope of the individual reaction norms close to the optimal values (a_{opt} and b_{opt} respectively), and (ii) very low Q_{ST} -values (see Soularue and Kremer 2014 results section for a definition) at the quantitative trait loci determining the slope and the intercept. We used Metapop to simulate the evolution of an adaptive trait with a linear reaction norm. The slope and the intercept of individual reaction norms each depended on 10 QTLs. We simulated 55 populations homogeneously distributed over 11 latitudes (see Soularue and Kremer 2012 for an illustration) and interconnected by gene flow according to the Wright's Island migration model. The landscape was characterized by phenotypic optima (Z_{opt} values) varying linearly with latitude, with a slope of 0.7, and macro-environmental effects (E values) influencing the expression of the trait (equations 2 and 3) varying linearly with a slope of 0.5 over the entire landscape. As a result the Z_{opt} and E values ranged from -3.5 to 3.5 and from -2.5 to 2.5 respectively, the central value (latitude 5) being 0 in both cases. Hence, in this scenario, the optimal reaction norms was defined by the intercept $a_{opt}=0$ and the slope $b_{opt}=1.4$. We simulated 30 replicates per scenario and table 3 summarizes all the settings used to run the simulations. The simulations performed with Metapop confirmed that the optimal evolution of plasticity across a spatially heterogeneous landscape requires high levels of gene flow between populations (figures 3 and 4) (Tufto 2000; Scheiner 2012).

Conclusion

Metapop is a stochastic individual-based program that has already a long history in the evolutionary genetics of forest trees. Metapop is complementary to other analog simulation programs. In its current state, Metapop will be particularly useful in theoretical studies involving heritable phenotypic plasticity or assortative mating in spatially and temporally heterogeneous landscapes. Interestingly, the comparative analysis conducted here indicated that Metapop produced results congruent with theoretical predictions and other simulators.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the European Research Council through the Advanced Grant Project TREEPEACE (#FP7-339728), and by the French National Research Agency (ANR) in the frame of the Investments for the future Programme, within the Cluster of Excellence COTE (ANR-10-LABX-45), and through the research project ANR MeCC (ANR-13-ADAP-006_02). We thank Didier Baradat and Nathalie Machon who initiated the development of Metapop. We are grateful to Frédéric Austerlitz, Pauline Garnier-Géré, Bernard Godelle, Pierre-Henri Gouyon, Stéphanie Mariette and Frédéric Raspail for their contributions that facilitated the development of the current

extended version. We thank Ophélie Ronce for useful discussions during model development. We thank François Ehrenman and Thierry Labbé for their help in building the web page hosting Metapop. Computer time for this study was provided by the computing facilities MCIA (Mésocentre de Calcul Intensif Aquitain) of the Université de Bordeaux and of the Université de Pau et des Pays de l'Adour. We thank the reviewers and the editor for their constructive comments and suggestions.

References

- Aguilée R, Raoul G, Rousset F, Ronce O. Pollen dispersal slows geographical range shift and accelerates ecological niche shift under climate change. *Proceedings of the National Academy of Sciences*. 2016; 113(39):E5741–E5748.
- Alberto FJ, Aitken SN, Alía R, González-Martínez SC, Hänninen H, Kremer A, et al. Savolainen O. Potential for evolutionary responses to climate change—evidence from tree populations. *Global Change Biology*. 2013; 19(6):1645–1661. [PubMed: 23505261]
- Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics*. 2000; 154(3):1309–1321. [PubMed: 10757772]
- Austerlitz F, Garnier-Géré PH. Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity*. 2003; 90(4):282. [PubMed: 12692580]
- Balkenhol N, Landguth EL. Simulation modelling in landscape genetics: on the need to go further. *Molecular Ecology*. 2011; 20(4):667–670. [PubMed: 21370553]
- Bradshaw AD. Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics*. 1965; 13:115–155.
- Chevin LM, Lande R, Mace GM. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology*. 2010; 8(4):e1000357. [PubMed: 20463950]
- Clark JS, Fastie C, Hurtt G, Jackson ST, Johnson C, King GA, et al. Schupp EW. Reid's paradox of rapid plant migration: dispersal theory and interpretation of paleoecological records. *BioScience*. 1998; 48(1):13–24.
- Conover DO, Schultz ET. Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*. 1995; 10(6):248–252. [PubMed: 21237029]
- Coombs JA, Letcher BH, Nislow KH. pedagog: software for simulating eco-evolutionary population dynamics. *Molecular Ecology Resources*. 2010; 10(3):558–563. [PubMed: 21565058]
- Franji J, Sever K, Bogdan S, Škvorc Ž, Krstonoši D, Aleškovi I. Fenološka neujedna enost kao ograničavajući imbenik uspješnoga oprašivanja u klonskim sjemenskim plantažama hrasta lužnjaka (*Quercus robur* L.). *Croatian Journal of Forest Engineering*. 2011; 32(1):141–154.
- Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*. 1995; 86(6):485–486.
- Guillaume F, Rougemont J. Nemo: an evolutionary and population genetics programming framework. *Bioinformatics*. 2006; 22(20):2556–2557. [PubMed: 16882649]
- Hunter JD. Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*. 2007; 9(3):90–95.
- Hoban S, Bertorelle G, Gaggiotti OE. Computer simulations: tools for population and evolutionary genetics. *Nature Reviews Genetics*. 2012; 13(2):110–122.
- Kimura M, Weiss GH. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*. 1964; 49(4):561. [PubMed: 17248204]
- Kingman JFC. The coalescent. *Stochastic Processes and their Applications*. 1982; 13(3):235–248.
- Kremer A, Le Corre V. Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*. 2012; 108:375–385. [PubMed: 21915150]
- Lambert BW, Terwilliger JD, Weiss KM. ForSim: a tool for exploring the genetic architecture of complex traits with controlled truth. *Bioinformatics*. 2008; 24(16):1821–1822. [PubMed: 18565989]
- Lande R. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical research*. 1975; 26(03):221–235. [PubMed: 1225762]

- Lande R. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology*. 2009; 22(7):1435–1446. [PubMed: 19467134]
- Latta RG. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *American Naturalist*. 1998; 151:283–292.
- Le Corre V, Kremer A. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*. 2003; 164:1205–1219. [PubMed: 12871925]
- Le Corre V, Kremer A. The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*. 2012; 21:1548–1566. [PubMed: 22332667]
- Lewontin RC. Detecting population differences in quantitative characters as opposed to gene frequencies. *American Naturalist*. 1984; 123:115–124.
- Neuenschwander S, Hospital F, Guillaume F, Goudet J. quantiNemo: an individual-based program to simulate quantitative traits with explicit genetic architecture in a dynamic metapopulation. *Bioinformatics*. 2008; 24(13):1552–1553. [PubMed: 18450810]
- Peng B, Kimmel M. simuPOP: a forward-time population genetics simulation environment. *Bioinformatics*. 2005; 21(18):3686–3687. [PubMed: 16020469]
- Peng, B, Kimmel, M, Amos, CI. Forward-time population genetics simulations: methods, implementation, and applications. John Wiley & Sons; 2012.
- Peng B, Chen HS, Mechanic LE, Racine B, Clarke J, Clarke L, et al. Feuer EJ. Genetic Simulation Resources: a website for the registration and discovery of genetic data simulators. *Bioinformatics*. 2013; 29(8):1101–1102. [PubMed: 23435068]
- Pigliucci, M. Phenotypic plasticity: beyond nature and nurture. JHU Press; 2001.
- Reeve JP. Predicting long-term response to selection. *Genetics Research*. 2000; 75(1):83–94.
- Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*. 2008; 8(1):103–106. [PubMed: 21585727]
- Scheiner SM, Barfield M, Holt RD. The genetics of phenotypic plasticity. XI. Joint evolution of plasticity and dispersal rate. *Ecology and Evolution*. 2012; 2(8):2027–2039. [PubMed: 22957203]
- Schiffers KH, Travis JM. ALADYN—a spatially explicit, allelic model for simulating adaptive dynamics. *Ecography*. 2014; 37(12):1288–1291. [PubMed: 25698848]
- Segelbacher G, Cushman SA, Epperson BK, Fortin MJ, Francois O, Hardy OJ, et al. Manel S. Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*. 2010; 11(2):375–385.
- Shou W, Bergstrom CT, Chakraborty AK, Skinner FK. Theory, models and biology. *Elife*. 2015; 4:e07158. [PubMed: 26173204]
- Soularue JP, Kremer A. Assortative mating and gene flow generate clinal phenological variation in trees. *BMC evolutionary biology*. 2012; 12(1):79. [PubMed: 22681626]
- Soularue JP, Kremer A. Evolutionary responses of tree phenology to the combined effects of assortative mating, gene flow and divergent selection. *Heredity*. 2014; 113(6):485. [PubMed: 24924591]
- Soularue JP, Robin C, Desprez-Loustau ML, Dutech C. Short rotations in forest plantations accelerate virulence evolution in root-rot pathogenic fungi. *Forests*. 2017; 8(6):205.
- Team RC. R language definition. Vienna, Austria: R Foundation for Statistical Computing; 2000.
- Tufto J. The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. *The American Naturalist*. 2000; 156(2):121–130.
- Turelli M. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology*. 1984; 25(2):138–193. [PubMed: 6729751]
- Vitasse Y, Delzon S, Bresson CC, Michalet R, Kremer A. Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*. 2009; 39(7):1259–1269.
- Wright S. Evolution in Mendelian populations. *Genetics*. 1931; 16(2):97–159. [PubMed: 17246615]

Figure 1.

Genetic variance accumulated by mutation in a population showing no initial genetic variation. The black line stands for the genetic variance simulated by Metapop (k -allele mutation model), red is used for Nemo (Random mutation model) and blue is used for Quantinemo (two lines: plain for Random mutation model, dashed for k -allele mutation model). The purple dashed-dotted-line is the theoretical prediction provided by Lande's model (Lande 1975).

Figure 2.

Genetic variance at mutation-selection equilibrium in a single population. The intensity of stabilizing selection was strong ($\omega^2=5$). In each case, the starting population was the population resulting from scenario 1. The black line corresponds to the genetic variance obtained with Metapop (k-allele mutation model), red is used for Nemo (Random mutation model) and blue is used for Quantinemo (two lines: plain for Random mutation model, dashed for k-allele mutation model). The purple dashed-dotted-line corresponds to the theoretical prediction provided by House of Card model (Turelli 1984).

Figure 3.

Mean additive value at the loci determining the intercept (up) and the slope (bottom) of the individual reaction norms, across the landscape. The means were calculated from individual values for the five populations located at each latitude. Left and right panels (A and B) corresponds to the mean values obtained under limited ($Nm = 1$) and high ($Nm = 10$) gene flow respectively. The dash-dotted lines correspond to the optimal intercept (up) and slope (bottom).

Figure 4.

Changes in genetic differentiation at the intercept (plain lines) and slope (dotted lines) of the individual reaction norms under limited (A, $Nm = 1$) and extensive migration (B, $Nm = 10$). The high Q_{ST} values observed in A correspond to the maintenance of locally specialist genotypes at each site, with differences in intercept and slope between the latitudes. The low Q_{ST} values observed in B correspond to the optimal evolution of plasticity, *i.e.* the predominance of a single generalist genotype expressing the optimal reaction norm across the landscape.

Table 1

Overview of the features of the main spatially explicit simulation models based on quantitative genetics available. **Migration.** IMM: island migration model, SSM: stepping stone migration model. UDM: user-defined dispersal model from migration matrix, dispersal rates or adaptations of IMM/SSM. QTD: quantitative trait determining dispersal. SMR: sex-specific migration rates. AMR: age-specific migration rates. BI: influence of landscape boundaries on dispersal. **Mutation.** RMM: random mutation model, the mutational effect of the new allele is drawn randomly from a given distribution and added to the existing allelic value. IMU: increment mutation model, the mutational effect depends on the original state of the mutant allele. DAM: di-allelic mutation model, mutation effects can only increase or decrease a given value, resulting in symmetric effects. KAM: K-allele mutation model, which assumes that there are no more than k alleles at a given locus. Each mutation replaces the existing allele with another allele from a predefined set of alleles, ranging from 0 to k . SSM: given a predefined index of alleles, the single step mutation model replaces a mutating allele k with one of the “adjacent” alleles, $k+1$ or $k-1$, at random. LM: lethal mutations. SMM: stepwise mutator model for microsatellite markers. IAM: infinite allele model. UDP: user-defined probability model: each allele has a user-defined probability of mutating to yield another given allele. **Multi-traits:** n corresponds to a positive integer. **Mating system.** Assortative: assortative mating. Random: random mating. Sex-specific: possibility of defining the number of matings per individual on the basis of sex.

	Genetic map	xy coordinates	Migration	Selection	Mutation	Multi-traits	Mating system	Trait plasticity	Life cycle events
Metapop	Yes	Population on lattice	IMM SSM UDM	Stabilizing Divergent	KAM	n	Random Assortative Selfing	Heritable Non heritable	Fixed order
Nemo	Yes	Population on lattice	IMM SSM UDM QTD SMR	Stabilizing Divergent Fixed model	RMM DAM deleterious	n	Random Selfing Cloning Sex-specific	No	Flexible
Quantinemo	Yes	Population on lattice	IMM SSM UDM SMR BI	Stabilizing Divergent Directional Combination of models	KAM SSM RMM IMU UDP	n	Random Selfing Cloning Sex-specific	No	Fixed order but individual cycles can be left out
Aladyn	No	Individual in continuous landscape	lognormal dispersal kernel BI	Stabilizing Divergent	RMM	2	Random	No	Fixed order
Forsim	Yes	Population on lattice	UDM	Stabilizing Divergent Directional Truncation User-defined	RMM IMU lethal	n	Random Assortative	Heritable at family level only	Fixed order
Pedagog	No	Population on lattice	UDM SMR AMR	Stabilizing Divergent Directional	IAM SMM	n	Sexual selection Sex-specific	No	Fixed order

Table 2

Settings used in scenarios 1 and 2.

Parameter	Value
Number of populations	1
Number of QTLs (intercept)	20
Intensity of stabilizing selection ω^2	10^9 (no selection), then 5
Number of individuals per population	5 000
Mutation rate μ	10^{-4}
Selfing rate τ	0.02

Table 3

Settings used in scenario 3.

Parameter	Value
Number of latitudinal levels	11
Number of populations	55
Pollen and seed migration model	Wright's island migration model
Pollen (m_p) vs. seed (m_s) migration rates, scaled by α	$m_p = \alpha \cdot m_s$ with $\alpha=100$
Number of a QTLs (intercept)	10
Number of b QTLs (slope)	10
Slope of environmental gradient S_E	0.5
Slope of the gradient of the phenotypic optimum S_{Zopt}	0.7
Level of gene flow Nm	10
Intensity of stabilizing selection ω^2	50
Number of individuals per population	500
Mutation rate μ	10^{-4}
Selfing rate τ	0.02