Long distance dispersal and the fate of a gene from the colonisation front

Julien Fayard, Etienne K. Klein, Francois Lefèvre

► To cite this version:

Julien Fayard, Etienne K. Klein, Francois Lefèvre. Long distance dispersal and the fate of a gene from the colonisation front. Journal of Evolutionary Biology, 2009, 22 (11), pp.2171-2182. 10.1111/j.1420-9101.2009.01832.x . hal-02668328

HAL Id: hal-02668328
https://hal.inrae.fr/hal-02668328v1

Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

Long distance dispersal and the fate of a gene from the colonisation front

J. Fayard [1]*, E.K. Klein [1,2], F. Lefèvre [1]

[1] INRA, UR 629 Ecologie des Forêts Méditerranéennes, Domaine Saint-Paul, Site Agroparc, F-84914 Avignon Cedex 9, France.

[2] INRA, UR 546 Biostatistique et Processus Spatiaux, Domaine Saint-Paul, Site Agroparc, F-84914 Avignon Cedex 9, France.

* Author for correspondence: julien.fayard@avignon.inra.fr

INRA, UR 629 Ecologie des Forêts Méditerranéennes, Domaine Saint-Paul, Site Agroparc, F-84914 Avignon Cedex 9, France. Tel : +33 4 32 72 21 45 Fax : +33 4 32 72 21 82

Short running title: LDD and the gene surfing phenomenon

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

Abstract

There is an increasing recognition that Long Distance Dispersal (LDD) plays a key role in establishing spatial genetic structure during colonisation. Recent works, focused on short distance dispersal, demonstrated that a neutral mutation arising at the colonisation front can either "surf" with the wave front and reach high frequencies or stay near its place of origin at low frequencies. Here, we examine how LDD, and more generally the shape of the dispersal kernel, modifies this phenomenon and how colonisation domain size affects the fate of the mutation. We demonstrate that when LDD events are more frequent, the "surfing phenomenon" is less frequent and the loss of diversity is attenuated. We also demonstrate that the width of the colonisation domain influences the fate of the mutation, wide spaces decreasing the probability of invasion. Overall, the genetic structure of diversity resulted not only from LDD but particularly from the shape of the dispersal kernel.

Keywords:

Long distance dispersal, range expansion, founder effect, spatial genetic structure, mixing of genes, surfing phenomenon

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

1. INTRODUCTION

Both empirical and theoretical studies of colonisation and the biological processes operating during colonisation (e.g. dispersal) have become crucial in analysing population biology. Many threats to biodiversity are directly related to the colonisation process (Hewitt, 2000). Colonisation is of major interest for predicting a species' response to global warming (McLachlan et al., 2005), designing conservation practices (Higgins et al., 1996, Trakhtenbrot et al., 2005) and managing invasive species (Higgins et al., 1996, Shigesada & Kawasaki, 1997). The study of colonisation has also helped understanding how trees recolonised continents so rapidly after the last glaciation (Clark et al., 1998) and it can potentially give insights about the current structure of forest diversity (Austerlitz & Garnier-Gere, 2003, Petit et al., 2004).

Long distance dispersal (LDD) is now accepted as a key factor in the colonisation process, which influences both population expansion dynamics and spatial structure of genetic diversity (Bohrer et al., 2005, Cain et al., 2000, Nathan & Muller-Landau, 2000). LDD events occur at low frequency with thin-tailed dispersal kernels (i.e. dispersal kernels with exponentially bounded tails, (Kot et al., 1996)) but are more frequent with fat-tailed dispersal kernels (i.e. dispersal kernels with non-exponentially bounded tails). Rather than characterizing LDD solely by the proportion of genes dispersed further than a fixed dispersal distance, the shape of the dispersal kernel, and particularly its tail, is now considered to be the main determinant of population expansion dynamics and genetic diversity (Clark et al., 2001, Klein et al., 2006, Kot et al., 1996). Dynamic models have shown that the recolonisation of the northern hemisphere by trees would have been impossible without the occurrence of LDD events (Clark et al., 1998, Davies et al., 2004, LeCorre et al., 1997), which significantly increased colonisation speed. Currently, expansion dynamics with LDD can be simulated

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

using models with fat-tailed dispersal kernels (Kot et al., 1996) or a mixture of thin-tailed kernels at several spatial scales, such as Gaussian mixture kernels (LeCorre et al., 1997).

From a population genetics point of view, LDD can have two opposite effects: it can either increase founder effects (Lambrinos, 2004) or promote gene mixing (or propagules, haplotypes, genotypes) at long distances from the sources (Klein et al., 2006). Because LDD events increase founder effects, they tend to reduce diversity in rectangular domains (i.e. corridors). This was illustrated with Gaussian mixture kernels (i.e. thin tailed dispersal kernels but with a significant amount of LDD) for which the founder effects can possibly lead to an almost total loss of diversity through an "embolism effect" (Petit et al., 2004): only one gene finally occupies the whole front in a corridor and prevents other genes from reaching empty spaces. Regardless of the total area of the simulation space, it is expected that it would be more difficult for a particular gene to block the progression of the other genes when simulation space width (i.e. the smallest dimension of the rectangular area) increases, because the time required to establish a sufficiently large population increases with simulation space width. Surprisingly, no effect of suitable domain width has been shown in the literature, even though some authors have directly tested for it (Bialozyt et al., 2006) and despite the fact that it seems critical to create an "embolism effect". As opposed to founder effects caused by LDD events, fat-tailed dispersal kernels are expected to improve gene mixing at long distances from the front (Klein et al., 2006). This result suggests that LDD can lead to the conservation of genetic diversity at the population level. Thus, the effect of LDD on the genetic structure of a population during a range expansion is not completely understood, potentially because of the complexity introduced by two opposite effects, founder effects and gene mixing, which occur at both local and global scales. Further attention should be paid to the choice of the dispersal kernel used in colonisation models and its possible effect on genetic structure (Ibrahim, 2004, Ibrahim et al., 1996).

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

Edmonds et al. (2004), Hallatschek et al. (2007) and Wei & Krone (2005) have shown that, without LDD, a neutral mutation arising at a colonisation front has only two possible fates: it can either stay near its place of origin at low frequency or travel with the front and colonize a large area thus reaching high frequencies in the newly colonized sections of the landscape. Using a set of simulated colonisations, Edmonds et al. (2004) obtained bimodal distributions for the mutant frequency at the end of colonisation and the distance travelled by the mutant centroïd (i.e. the mean position of mutant individuals) from the occurrence of the mutant to the end of the colonisation. The mechanism involved in this type of colonisation has been called the "surfing phenomenon" because mutants seem to travel with the colonisation front (Vlad et al., 2004a, Vlad et al., 2004b, Vlad et al., 2005, Wei & Krone, 2005). Edmonds et al. (2004) support that LDD is not required for diversity loss and that invasion by one gene can occur relatively frequently without LDD, whereas Petit et al. (2004) suggest that population invasion by one gene is due to LDD events, linking *de facto* LDD to a loss of diversity. This discrepancy might be related to the particular dispersal kernel used for modelling LDD by Petit et al. (2004), i.e. a Gaussian mixture kernel (thin-tailed kernel) rather than a fat-tailed kernel.

Recent simulation studies have also shown an erosion of diversity with short distance dispersal during colonisation (Hallatschek et al., 2007, Hallatschek & Nelson, 2008). Hallatschek & Nelson (2008) formalized mathematically how the colonisation process gives rise to a gradual loss of diversity, due to repetitive samplings of lineages within the colonisation front. With LDD, we expect that this erosion would not take place at the same rate because individuals in the population core can also colonize empty spaces and thus maintain a higher level of diversity. As the final mutant frequency in one simulation run of the Edmonds' model measures the reproductive success of a single individual (or gene) sampled at the colonisation front, the distribution of mutant frequencies over many replicated

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x
Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

simulations represents the distribution of reproductive successes among individuals located at the colonisation front. Therefore, from a population genetics point of view, the bimodality of mutant frequencies obtained after many simulations in Edmonds et al. (2004) can be interpreted as a high variance in reproductive success of individuals located at the colonisation front, i.e. a reduced effective population size leading to diversity loss. As expected from population genetics theory, in the Edmonds et al. (2004) model, the variance of reproductive success increased with population growth rate and decreased with population carrying capacity (Klopfstein et al., 2006). The Edmonds' model provides an interesting framework for testing the effect of LDD events on genetic diversity during colonisation, with clear conclusions obtained using only short distance dispersal. It also provides an efficient method to evaluate variance in reproductive successes of individuals in the population.

In this study, we constructed an original model based on Edmonds et al. (2004) that can account for LDD, using a variety of dispersal kernels to simulate the colonisation of a rectangular grid of demes where progeny of a single individual can be traced. Our goal was to examine (i) how the fate of a neutral mutant arising at the colonisation front is affected by LDD, specifically by the weight of the tail of the dispersal kernel, and (ii) how the width of the colonisation domain where simulation takes place influences the probability of mutation success.

2. MATERIALS AND METHODS

(a) *Population Dynamics Model*

The model simulates haploid individuals (or maternally inherited genes) that reproduce, disperse and die, with non-overlapping generations. Individuals were distributed within a grid

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

of demes, with width (i.e. the number of rows) varying from 5 to 50 and a length (i.e. the number of columns) of 200 demes (for a total of 10 grid sizes). Individuals were not explicitly positioned in their demes, but only spatially characterized by the deme to which they belong. We assumed that (i) the quantity of offspring dispersed at any position inside a deme centered on (*x'*,*y'*) depended only on the position (*x'*,*y'*) of the recipient deme relative to the parental deme (*x*,*y*), based on the 2D dispersal kernel $\gamma(x'-x, y'-y)$, and (ii) the demographic processes inside a deme were identical for all demes. At each simulation step, all the offspring were dispersed from the centre of their grid cell. Individuals dispersed outside the grid were discarded. We then summed the number of individuals arriving in each cell to calculate the dispersal stage. We could thus write a reproduction-dispersal model:

$$(1) \qquad n_{res}(k,t+1) = \sum_{l \neq k} f(N_{res}(l,t), N_{tot}(l,t))\gamma(x_k - x_l, y_k - y_l)\Delta_{cell} + (1-m) f(N_{res}(l,t), N_{tot}(l,t))$$

$$(2) \qquad n_{mut}(k,t+1) = \sum_{l \neq k} f(N_{mut}(l,t), N_{tot}(l,t))\gamma(x_k - x_l, y_k - y_l)\Delta_{cell} + (1-m) f(N_{mut}(l,t), N_{tot}(l,t))$$

where $n_{res}(k, t)$ and $n_{mut}(k, t)$ are, respectively, the expected number of non-mutant and mutant individuals in deme *k* at time *t*; $N_{res}(k,t)$ and $N_{mut}(k,t)$ are the actual numbers of non-mutant and mutant individuals in deme *k* at time *t*; $N_{tot}(k,t)$ is the actual total number of individuals in deme *k* at time *t* ; *f* is a function describing the demography within a deme (see below) and γ is the 2D dispersal kernel ; Δ_{cell} is the area of a cell (1 in our system) and *m* is the emigration rate (i.e. the integral of the kernel outside of the parental cell, see below). In this model, individuals were either mutants or not and this trait was inherited by the progeny. Per simulation run, only one copy of the mutant was introduced, in a pre-determined cell of the grid (see below). Other copies originate only from reproduction, mutation is neglected.

The local demography within a deme is described by:

$$(3) \qquad f(N_{res}(k,t), N_{tot}(k,t)) = N_{res}(k,t)\frac{rK}{rN_{tot}(k,t) - N_{tot}(k,t) + K}$$

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

$$f(N_{mut}(k,t), N_{tot}(k,t)) = N_{mut}(k,t)\frac{rK}{rN_{tot}(k,t) - N_{tot}(k,t) + K} \quad (4)$$

where r is the intrinsic growth rate, K is the carrying capacity of each deme. Non-mutant and mutant individuals were thus indistinct for the resource competition. The values of the parameters were fixed at $r = 5$ and $K = 20$ in all simulations described below. Klopfstein et al. (2006) showed that an increase in growth rate increases the probability of surfing contrary to an increase in carrying capacity. Actually, the rate of mutant surfing was proportional to $\frac{r}{Km}$ (where m is the migration rate). With LDD, these results are expected to be unchanged since (i) high r values should still allow the new mutant to rapidly reach high local densities, and thus counteract non-mutants immigration and (ii) high values of Km decrease the intensity of genetic drift in the saturated parts of the corridor, and thus promote mutant survival at a low frequency. In this model, when the total number of individuals in one deme was higher than the carrying capacity, the net growth rate of the population was lower than 1 (i.e. the population size decreased) but larger than 0 (i.e. the population did not go extinct instantaneously). Because computational time did not allow us to investigate the effect of deme extinctions, we made this choice to reduce the number of deme extinctions due to negative net growth rate. Furthermore, we focus on the colonisation phenomenon and not the metapopulation behaviour once the space is filled. Demographic stochasticity was taken into account by assuming:

$$N_{res}(k,t+1) = P(n_{res}(k,t+1)) \quad (5)$$

$$N_{mut}(k,t+1) = P(n_{mut}(k,t+1)) \quad (6)$$

$$N_{tot} = N_{res} + N_{mut} \quad (7)$$

where $P(\lambda)$ is a Poisson distribution with mean λ.

(b) *Dispersal kernels*

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

Simulation of dispersal was performed using a 2D dispersal kernel, i.e. a function representing the probability for a gene to reach a 2D position relative to the emitting position (0,0). We only used isotropic kernels, meaning that γ(x,y) only depended on the distance *d* between (0,0) and (*x*,*y*). We used a wide variety of dispersal kernels that include LDD, i.e. which varied in terms of tail shapes or kurtosis coefficient. We used nine different dispersal kernels γ (Fig. 1 and Table 1): a Gaussian kernel to investigate only short distance dispersal (it is known to be a continuous model that behaves like the stepping-stone process (Mollison, 1977) used in Edmonds et al. 2004 and Klopfstein et al. 2006), two 2Dt kernels (Clark et al., 1998), two exponential power (EP) kernels (Clark et al., 1998, Klein et al., 2006) and four Gaussian mixture kernels (Austerlitz & Garnier-Gere, 2003, Bialozyt et al., 2006, LeCorre et al., 1997). Among the four Gaussian mixture kernels, two had a Gaussian kernel with large variance (i.e. equal to 50, scale parameter *b* = 10) and the other two had a smaller variance (i.e. equal to 12.5, scale parameter *b* = 5). Within these two groups, one Gaussian mixture kernel had a high (0.1) proportion of events following the Gaussian with high variance; the other had a low proportion (0.01). To make comparisons meaningful, we chose the parameters of each dispersal kernel so as to provide similar migration rate (*m*) and mean distance travelled (δ):

(8) $$(1-m) = \int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\gamma(x,y)dxdy = 0.8$$

(9) $$d = \int_{od}\gamma(x,y)\sqrt{x^2+y^2}dxdy = 3$$

where *m* is the emigration rate (the expected proportion of individuals emigrating from one deme to others), δ is the mean distance dispersal (the expected number of demes travelled), the first integral is an integration on the square area of the deme centred in *0* and the second integral is an integration on *od*, the remaining area, with $od = R^2 - [-1/2,1/2]$. Finally, by verifying $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)dxdy = 1$, all dispersal kernels are 2D density probability functions.

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

The choice of $\delta = 3$ was made to ensure a sufficient resolution of the dispersal kernel when it is discretized on a grid (for instance, rapid variations of dispersal probabilities typical of peaked kernels are poorly resolved with $\delta = 1$). Since the value of δ does not vary in this study, only the values of the domain width matter for the fate of the mutant. The individuals dispersing out of the grid were discarded, which results in a loss of less than 2% of individuals when the thinner dispersal kernel (Gaussian kernel) is discretised with a domain width of 5. This loss is less than 4% when the fatter dispersal kernel (exponential power (b=0.25)) is discretised with the same domain width.

(c) *Simulation Design*

Each simulation started by placing individuals at carrying capacity (*K*) in all demes located in the first left column of the grid. The deme in which the mutation will occur is decided (*longitude* = 5th or 30th column; *latitude* = 25 or 50% of the maximum latitude size). The run was then divided into two temporal parts:

(i) in a first part, we let non-mutant individuals reproduce and disperse, with a given *initial* dispersal kernel, until the deme where mutation should occur is reached, in order to establish an initial population;

(ii) then, we placed one mutant individual in the initial population and chose another dispersal kernel for both mutant and non-mutant individuals, called *colonisation* kernel.

This scheme was chosen to isolate the effect of the spatial pattern of the initial population where the first mutant individual appeared (generated through the initial kernel) from the effect of LDD during colonisation after the appearance of the mutant. We did not directly manipulate the initial population's spatial pattern because we wanted to use a realistic pattern,

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript

i.e. generated by individuals reproducing and dispersing. We simulated all 9x9 combinations of each initial kernel and each colonisation kernel.

For each simulation, we chose the deme where we would place the first mutant individual from four positions that varied in their longitudinal and latitudinal coordinates (*longitudes* = 5th or 30th column; *latitudes* = 25 or 50% of the maximum latitude size), in order to control the initial population depth (longitude) and edges effect (latitude). The longitudinal axis was parallel to colonisation direction and latitudinal axis perpendicular to it. Then we used one given initial kernel to disperse individuals until at least one individual had reached the chosen deme. For each initial population thus obtained, we replaced one non-mutant individual by one mutant in this deme. Then we used one colonisation kernel (which could be different from the initial kernel) to disperse individuals until they reached at least half the demes located in the right-hand column of the grid. For each combination of initial and colonisation dispersal kernels, we counted the number of runs where no mutant individual persisted, which made it possible for us to compute survival probabilities ($surv = \frac{success}{Total}$). Then, at the end of each successful repetition, i.e. when at least one mutant individual persisted until the end of the colonisation process, we measured individual mutant frequencies over all demes, the number of demes colonized by the mutant (i.e. demes with at least one mutant at the end of the colonisation) and the distance travelled by the mutant centroïd (the mean position of the mutant individuals). Within successful simulations, we also counted the percentage of simulations where mutant frequencies were superior to 50% and called this the probability of surfing. We generated ten initial populations for each of the nine initial kernels and ten replicates of colonisation for each colonisation kernel and each initial population to take into account the variability of colonisation histories. This design resulted in 10 (grid sizes) x 4 (mutant positions) x 9 (initial kernels) x 10 (initial populations) x 9 (colonisation kernels) x 10 (repetitions), i.e. a total of 324000 simulations.

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript

Since our focus was the colonisation process itself, we defined the end of the colonisation process as the time when at least half of the demes situated at the right edge of the grid were occupied. To follow the structure of genetic diversity over that time and in the generations after the end of colonisation, we ran additional simulations where we computed the fixation index (Fst) at each generation as follows:

$$Fst = \frac{\bar{p}(1-\bar{p}) - \overline{p(1-p)}}{\bar{p}(1-\bar{p})} \quad (10)$$

where p is the mutant frequency per deme. $\bar{p}(1-\bar{p})$ is the average number of pairs of different individuals (mutant – non-mutant) inside one deme and $\overline{p(1-p)}$ is the number of pairs of individuals that are different throughout the whole population. These simulations were conducted as described above except that we used only one grid width of 25 x 200 demes, two dispersal kernels, the thinnest and the fattest kernels (Gaussian and exponential power) and the same kernels both for creating initial populations and colonizing. We let each simulation run for 1000 generations (the mean number of generations needed to colonize the grid was approximately 200 for both kernels). This design resulted in 2 (kernels) x 100 (repetitions) = 200 simulations. Continuing the simulations long after the end of the colonisation aimed at evaluating if the spatial genetic structure designed by the colonisation process lasts long.

As a neutral model, we compared our results with that of a sparse population growing to fill an almost empty rectangular domain. We ran similar simulations using a single grid size (25 x 200 demes) and two dispersal kernels (Gaussian and exponential power (b=0.25) used as both initial and colonisation kernels) with all demes of the initial population shuffled over the whole grid from the moment when we introduced the mutant individual (however, we did not change the deme where the first mutant individual was introduced). We let each

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

simulation run for 200 generations and we measured the frequency of mutant individuals over all demes. 10000 simulations were performed for each dispersal kernel.

To characterize the effect of the initial kernel on the spatial structure of the population around the deme for the first mutant individual, we also simulated 9000 initial populations (1000 with each dispersal kernel). Here, we used a single grid size of 25 x 200 demes. We measured the number of occupied demes and the total number of individuals within a neighbourhood of 3 demes around the first mutant individual (a total of 24 demes). We tested for the effect of the dispersal kernel on these variables using a Kruskal-Wallis test.

Finally, we ran simulations with various positions for the first mutant individual, not always in the colonisation front. In these simulations, we used only one grid size of 25 x 200 demes and two dispersal kernels (Gaussian and exponential power kernel (b=0.25)), using the same kernel to set up the initial population and to colonize. The first mutant was always introduced when at least one individual had reached the 30th column. A first mutant was only introduced if the deme was not empty or if a non-empty deme in the same row existed. We chose four classes of longitudes for the occurrence of the first mutant individual (1st-10th columns, 11th-20th columns, 31st-40th columns, 41st-50th columns). Inside these classes of demes, we randomly chose the longitude and the latitude of the first mutant individual. We chose these classes because the two first classes of longitudes are in the part of the population at carrying capacity and the two last classes of longitudes are in the part of the population where individual density decreases from K to 0 (lower individual density than above). We used ten different initial populations and, for each initial population, we made 5000 repetitions for each position of the first mutant individual. This last design resulted in 2 (kernels) x 10 (initial populations) x 5000 (repetitions) = 100000 simulations.

3. RESULTS

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d’auteur / Author manuscript
Manuscrit d’auteur / Author manuscript
Manuscrit d’auteur / Author manuscript

Grid width was the main determinant of survival probability and surfing phenomenon. As grid width increased (and thus population size), the probability of a mutation travelling with the colonisation front decreased (fig.2). When the width of the domain was not too large (< 15-20 rows), the distribution of mutant frequencies at the end of colonisation was bimodal when aggregating the results from all nine dispersal kernels, confirming results from Edmonds et al. (2004). For widths larger than 20 rows, bimodality was lost in all nine dispersal kernels (results not shown). These observations reveal the effect of grid size on the “surfing phenomenon” (mutant individuals cannot preclude the progression of non mutant individuals in large grids) and on the survival probability of the mutation (in large grids, individuals have a higher survival probability and mutant individuals are more often conserved at low frequencies, results not shown).

The effect of the dispersal kernel was examined at the level of the initial kernel and the colonisation kernel. We found no effect for the initial kernel, either on the “surfing phenomenon” (i.e. the probability of a mutation colonizing a large area and reaching high frequencies) or on the survival probability of the mutation. In fact, the various initial kernels did not yield any significantly different spatial aggregation of non-empty demes near the first mutant individual nor any different numbers of individuals near the first mutant individual (Kruskal-Wallis test on the number of occupied demes, $p = 0.535$, d.f. = 8, $\chi^2 = 7.01$, Kruskal-Wallis test on the number of individuals, $p = 0.525$, d.f. = 8, $\chi^2 = 7.11$).

For the fattest-tailed colonisation kernels (2Dt with $b = 2$ or EP with $b = 0.25$), the bimodality of the distribution of mutant frequencies at the end of the colonisation process was not clearly visible (fig.3), even if some mutants still succeed in reaching high frequencies. The same pattern held true for the number of demes colonized or the distance travelled by the centroid of mutant individuals. These fattest-tailed kernels also induced the lowest probability

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

of surfing (i.e. the ratio between the number of simulations where the mutant frequencies at the end of the colonisation was higher than 50% and the total number of simulations where at least one mutant survived), followed by fat-tailed kernels *sensu stricto* (i.e. exponential power and 2Dt) and finally by Gaussian mixture kernels and the Gaussian kernel (fig.4). Fatter-tailed kernels also significantly increased mutation survival (fig.5).

Although fatter-tailed kernels always led to smaller probabilities of success for the mutant, the effect of the tail of the dispersal kernel was stronger for wider grids (Fig. 4). For grids wider than 25 rows, the probabilities of success for the 2 fattest tails are 10 times smaller than that for the Gaussian. At the opposite, mixture of Gaussian kernels always provided probabilities of success at most 2 times smaller than that of the Gaussian.

In 75% of our simulations where at least one mutant survived at the end of the colonisation, at least one mutant individual survived until twice the mean colonisation time for both thin- and fat-tailed dispersal kernels (Gaussian and exponential power, results not shown). On average, the fixation index F_{st} decreased by 10^{-2} each 100 generations while values of 0.38 ± 0.1 were observed at the end of successful colonisation. After the end of the colonisation, the regression slope of the fixation index over time (in number of generations) was -1.10^{-4} for both kernels (results not shown).

In the absence of a directional colonisation dynamic (i.e. neutral model), the frequency distribution of mutant individuals at the end of the colonisation was not bimodal. The maximum frequency of mutant individuals obtained after 10000 repetitions was 0.0056 for the Gaussian and 0.0051 for the exponential power kernels.

Using a Gaussian kernel, in 5000 simulations, no situation was found where the frequency of mutant individuals at the end of colonisation was higher than 0.015 when the first mutant individual occurred between the 1st and the 10th columns. When the first mutant individual occurred between the 11th and the 20th columns, only one simulation led to a

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript

frequency of mutant individuals at the end of colonisation which was higher than 0.015 (frequency was 0.375). When the first mutant individual occurred between the 31st and the 51st columns, the distributions of mutant individual frequencies at the end of colonisation presented the same pattern as that obtained when it occurred within the colonisation front (30th column), with a clear bimodality.

Using an exponential power kernel, we found high frequencies of mutant individuals at the end of colonisation whatever the position of the first mutant individual. The maximum frequency of mutant individuals reached 0.827 when the first mutant individual occurred between the 1st and the 10th columns, 0.585 when it occurred between the 11th and the 20th columns, 0.989 between the 31st and the 40th columns and 0.960 between the 41st and the 50th columns.

The position of the deme where the mutation occurred also influenced the frequencies of mutant individuals, with mutations arising near an edge (latitude effect) and later during colonisation (longitude effect), leading to lower mutant frequencies (Table 2).

Discussion

Our original explicit colonisation model demonstrated the effect of LDD on the genetic structure established during a colonisation. Here, we show that the whole shape of the dispersal kernel influences the rate of surfing. Thus, the way LDD is taken into account in colonisation models is not a trivial choice, as previously demonstrated for demographic aspects of colonisation (Kot et al., 1996, Shaw, 1995, Wingen et al., 2007). In particular, Gaussian mixture kernels generate results that do not seem to be applicable to all kernels with LDD. The results we obtained concerning the probability of a "surfing" event using this

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

kernel were more similar to those obtained using a Gaussian kernel rather than a fat-tailed kernel. LDD is often modelled by a Gaussian mixture kernel, but this dispersal kernel is not fat-tailed *sensu stricto*. And our results show that an important difference between Gaussian and fat-tailed kernels is not only presence vs. absence of LDD events, but the relative frequencies of the different distances travelled by seeds (i.e. the whole shape of the tail) that determine if the accentuation of founder effect is compensated by gene mixing (Klein et al., 2006).

A major result of this study is that the bimodality of mutant frequency distributions only appeared with thin-tailed dispersal kernels, i.e. the Gaussian and the Gaussian mixture kernel, and not with fat-tailed dispersal kernels. LDD is the underlying mechanism that allows migrants to jump above a surfing gene and establish a new focus (i.e. a new population far ahead of the colonisation front). However, if LDD is sufficiently frequent it also prevents the "surfing phenomenon" from occurring, i.e. prevents an individual at the colonisation front from colonizing a large area and reaching high frequencies. This confirms the analytical results obtained by Klein et al. (2006) and the predictions made by Ibrahim (2004). When dispersal only occurs at short distances, the individuals located at the colonisation front are the only ones to contribute to the next generation of individuals located at the colonisation front, whereas all individuals in the population participate in the creation of the next generation of individuals located at the colonisation front in fat-tailed kernels. This result was also confirmed when the first mutant individual was placed far inside the initial population. The fat-tailed dispersal kernel provided examples of genes very far from the front that finally reached high frequency (but not as often as individuals at the front with short distance dispersal). This was probably because a long distance event established an offspring in the uncolonized area, whereas no such case was observed among simulations with the Gaussian kernel.

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

The distribution of mutant frequency at the end of colonisation represents the distribution of the reproductive successes of individuals at the colonisation front (each simulation run representing the reproductive success of a particular individual). Therefore, the distribution of mutant frequencies at the end of colonisation is directly linked to effective population size (Ne) and to intensity of genetic drift that drives the genetic diversity observed at the end of colonisation. Distributions with high variance indicate a potential loss of diversity (low Ne) due to the invasion of a single lineage, whereas distributions with low variance indicate the potential conservation of diversity (several lineages can survive until the end of the colonisation).

Within this framework, the results presented here contradict the intuitive idea that LDD leads to a loss of diversity during colonisation due to founder effects (Haag et al., 2006). Here, we found higher effective population size associated with fatter-tailed kernels, showing that LDD can lead to a better conservation of diversity. First, even if some mutations could eventually reach high frequencies at the end of the colonisation, this was much less frequent than for thin-tailed kernels as denoted by the absence of clear bimodality in the mutant frequency distribution (Fig. 3). Second there was no position in the population that gave individuals in this position a zero chance of "surfing", when the dispersal kernel was fat-tailed. This was exemplified by the fact that individuals at the population core could also reach high frequencies with fat-tailed kernels, as did individuals at the front. Third, mutants had a higher probability of survival with fat-tailed kernels. Altogether, these results seem to indicate that fat-tailed kernels lead to a better conservation of diversity. Furthermore, our results emphasize the role that long distance gene mixing plays in structuring genetic diversity during colonisation by tempering the impact of founder effects in the set-up of a spatial genetic structure. This result contrasts with the conclusion of Bialozyt et al. (2006) who only used a Gaussian mixture kernel to model LDD and who did not control the mean dispersal

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

distance. That conclusion is demonstrated here in the case of a colonisation were the non-equilibrium of demography plays a crucial role in the definition of the variance of fitness of the individual of the population. Our simulations of a growing population not colonizing a corridor show that genetic drift is much weaker in that situation and that the differences between kernels is negligible.

From a biological point of view, some species are known to disperse without LDD such as humans (Edmonds et al., 2004), bacteria (Hallatschek et al., 2007), land snails and bushcricket (Excoffier & Ray, 2008). Some other dispersal behaviours can sometimes be modelled with a mixture of Gaussian for taking into account two different processes of dispersal. Some patterns of mutant invasions were actually observed experimentally (e.g. for bacteria Hallatschek et al. 2007) or largely supported by population genetics data (e.g. for humans, see Edmonds et al. 2004). At the opposite, numerous plant species (trees in particular) and fungus species are known to disperse with highly leptokurtic kernels, and a relative conservation of diversity is observed over wide areas (Petit et al., 2004). Even if the diversity observed is not only due to the processes investigated here (for other explanations see Petit & Hampe, 2006), we showed that the effect of LDD on the maintenance of diversity could be even more important than demonstrated before only with mixture of Gaussian kernels (Austerlitz & Garnier-Gere, 2003, LeCorre et al., 1997, Petit et al., 2004). This results means that the dispersion syndrome might partly determine the risk of confusion between a selective shift and a neutral variant that benefited from the surfing phenomenon (Edmonds et al., 2004, Excoffier & Ray, 2008, Foll & Gaggiotti, 2008). Our results tend to support that species with frequent LDD events, or spatially unstructured dispersal events are less subject to false positive results for selection patterns detection. This case could include most invasive species that results from multiple introductions and long-distance transports by human activities (Rosenthal et al., 2008). Also, the question should be investigated of whether the

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

differences of spatial genetic structure we found among dispersal kernels are sufficiently strong to allow inferences about the dispersal kernel from multi-locus localised genotypes. However, it seems difficult to infer the dispersal kernel from spatial genetic structures since the spatial genetic structure is the result of many stochastic processes (i.e. not easily repeatable) and because the effect of habitat heterogeneities on a spatial genetic structure remains to be investigated. A first approach could be to investigate this question using simple experimental systems (e.g. bacteria in Hallatschek et al., 2007) where the intensity of LDD and the heterogeneity of the environment could be controlled easily.

We further showed that the simulation space in which the colonisation takes place plays a major role in the evolution of genetic diversity. A large simulation space (relative to dispersal capacities) promotes the conservation of genetic diversity. We argue that this result is not only due to a "dilution effect" (i.e. the first mutant individual represents a smaller proportion of the population in large grids than in narrow ones). Indeed, another active phenomenon is that surfing alleles can be stopped by more frequent LDD events in a larger simulation space than in a narrow one, because the probability of a LDD event falling into a large uncolonized area is higher. This explanation also supports the result that differences among dispersal kernels are stronger for wider grids. Actually, the effective dispersal function (i.e. the one after removal of individuals falling outside the grid) is more different of the simulated dispersal functions for more narrow grids because more LDD events are subtracted in this case. This demonstrates that the geometry of the colonized area is of importance for predicting the spatial genetic structure and that patterns obtained in a corridor might be different than patterns obtained in an angular area, or in a real 2D open area. This result emphasizes that habitat fragmentation, caused by human activities in particular, could also contribute to diversity loss: local reductions in the width of colonisation domain create short and narrow corridors, thus facilitating the local fixation of genes (Rees et al., 2009). Further

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

studies are needed to account for habitat heterogeneity and assess its impact on genetic diversity.

We used a specific simulation design, which made it possible to change the dispersal kernel after the mutation had occurred to be sure that the effect of the dispersal kernel on the probability of surfing was not due to the initial population structure (the initial kernel) or to dispersal capacity of the mutant (from the colonisation kernel). Not only did we find no effect of the initial dispersal characteristics but we also found no significant differences in deme occupancy measured near the first mutant individual among initial patterns generated by the nine dispersal kernels. This could be due to our definition of the colonisation front as the part of the population where the mean density of individuals decreased from K to 0 and not as the part of the population formed by the furthest forward individuals. Using the latter definition, we would expect to find higher differences among initial kernels for the spatial structure around the furthest forward individuals, with more demes occupied using thin tailed rather than fat tailed kernels. With LDD, the colonisation front cannot be defined properly (i.e. the population density do not decrease monotonically with distance) and thus, we cannot use the same approach as Hallatschek et al. (2007) with PDE models.

As shown with the fixation index, the colonisation dynamics generates a spatial genetic structure that is likely to persist for a long time in the population (see also Austerlitz & Garnier-Gere, 2003). Indeed, the decrease in Fst after the end of the colonisation is very low compared to Fst observed at the end of the colonisation, and in most simulations (75%), mutant individuals were still present after twice the time needed for colonizing the whole domain. Furthermore, the spatial genetic structure established during a colonisation is very particular compared to spatial genetic structures obtained without colonisation dynamics. We investigated the specificity of this situation by shuffling demes (i.e. many initial foci in an almost empty space) and showed that the expansion of the focus with the mutant is rapidly

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

blocked by its neighbouring foci, leading to a drastically reduced variance of the mutant fitness. We argue that the stochastic events (such as LDD events) occurring during the first generations of a colonisation determine the genetic structure at the end of the colonisation, which is likely to persist for a long period of time. These events also depend on the environmental structure, such as heterogeneities in habitat, and further attention should be paid to the modalities of these stochastic events.

To conclude, our model provides a better understanding of the drivers of genetic diversity and structure during a colonisation process. However, other factors deserve attention using a similar framework. First, our results were obtained using a unique migration rate ($m = 0.2$), due to computational time constraints. Since the migration rate is a synthetic parameter controlling the mutant and the non-mutant dispersal, a change in migration rate could affect the critical values (domain width, ...) at which LDD effect operates. Second, growth rate and carrying capacity, as shown by Klopfstein et al. (2006), interact with the surfing phenomenon and lead to different outcomes. Third, Bialozyt et al. (2006) have also shown that the effect of LDD on the genetic structure depends critically on the amount of LDD events and Petit & Hampe (2006) have also reviewed the genetic consequences of the particular life cycle of trees (e.g. conservation of high population diversity due to the length of the juvenile phase). The results obtained with our generic model, not representing a realistic and thus specific situation observed in nature, reveal a simple and clear understanding of processes occurring during colonisations. They remain to be validated by natural observations (e.g. genetic studies of diversity) and should be used to help collect experimental data for such purposes.

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

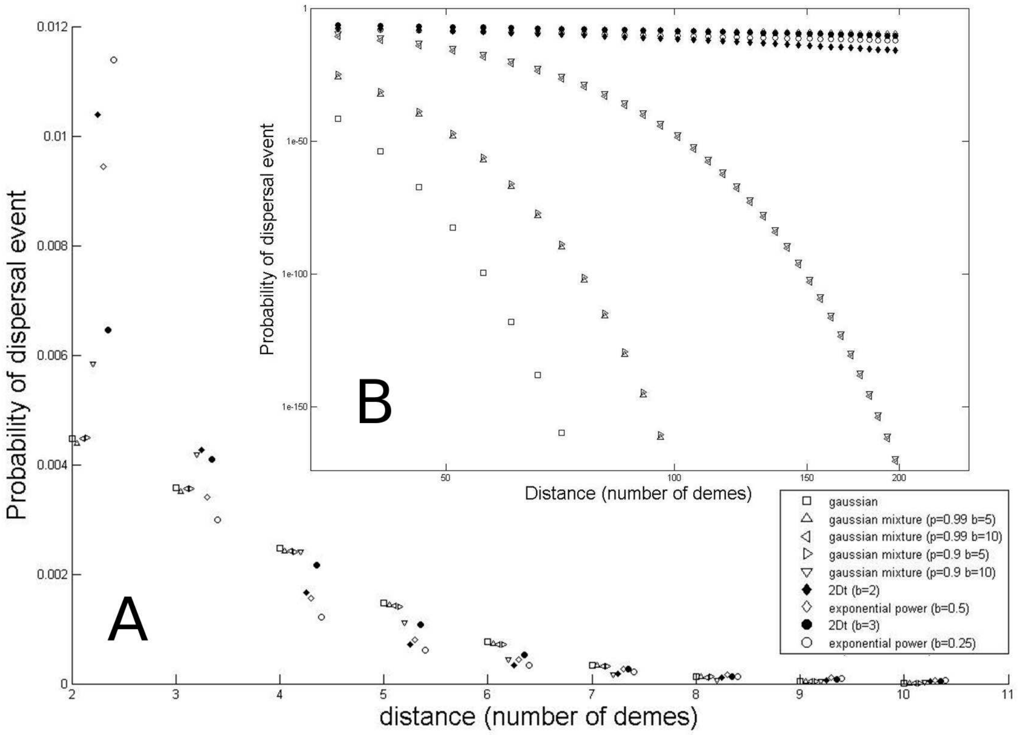

Fig. 1 – Differences among dispersal kernels. For the 9 colonisation kernels used, sorted from thinner-tailed to fatter-tailed, the probability of a dispersal event is represented against the distance. A- We plotted the probability of a dispersal event against the distances up to 10 demes. B- We plotted the logarithm of the probability of a dispersal event against the logarithm of distances up to 200 demes. Notice that fat-tailed kernels (2Dt and exponential power kernels) have higher probabilities of dispersal until the 4th deme and after the 40th deme and lower probabilities of dispersal between the 4th deme and the 40th deme.

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

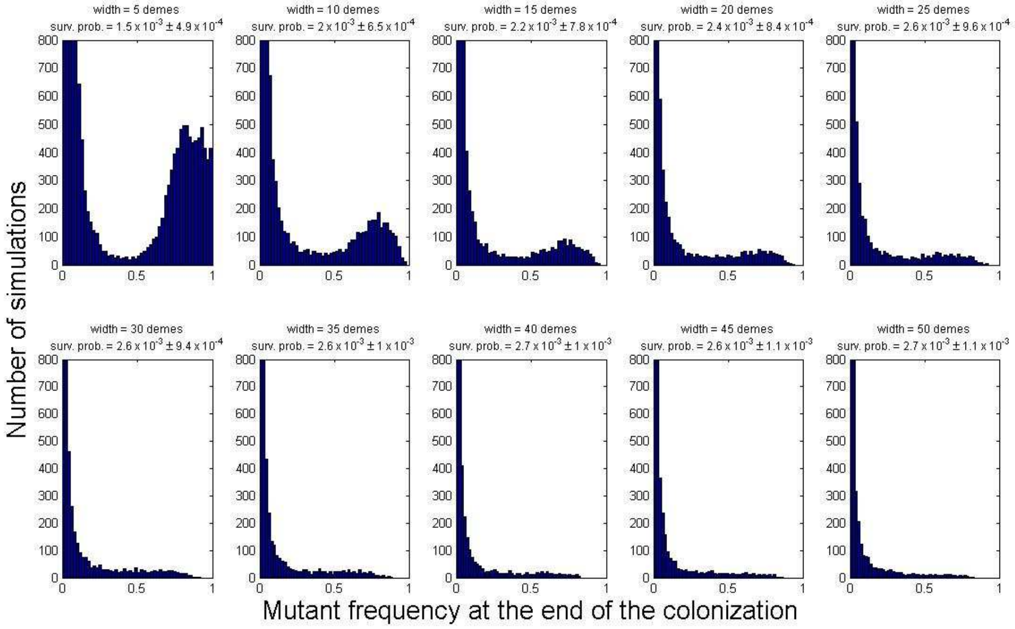

Fig. 2 – Distributions of mutant frequency at the end of colonisation for all grid widths. We plotted the number of successful simulations (*y*-axis) that ended with a given proportion of mutant individuals over the whole grid (*x*-axis). The scale of the *y*-axis was adjusted to stress the bimodality of the distribution. The survival probability of mutants for each width is indicated at the top of each figure. All colonisation kernels and initial kernels were pooled together, resulting in 32400 repetitions per histogram.

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

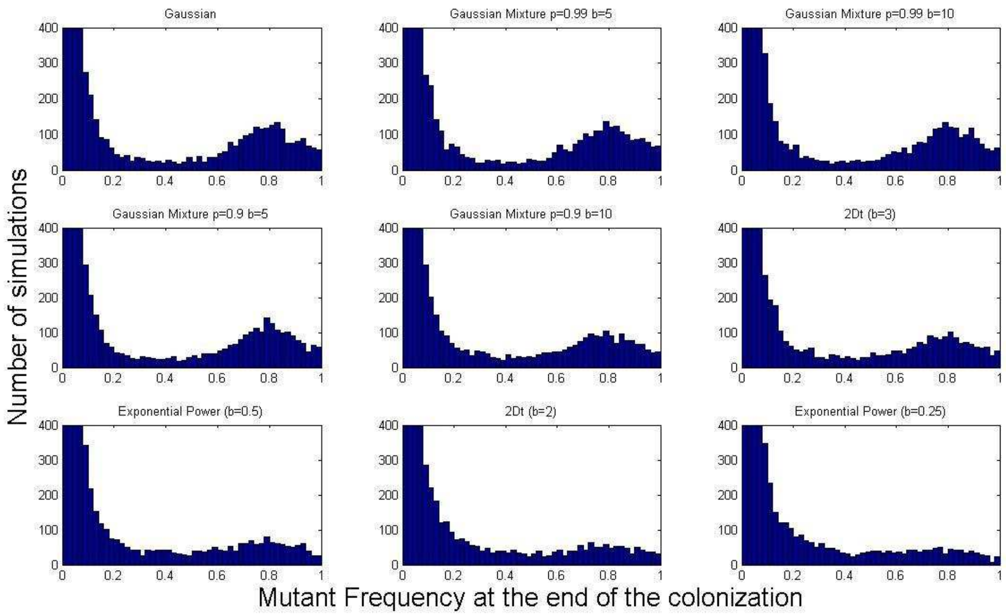

Fig. 3 – Distributions of mutant frequency at the end of the colonisation for 9 dispersal kernels (Fig 1 and Table 1). We plotted the number of successful simulations (*y*-axis) that ended with a given proportion of mutant individuals over the whole grid (*x*-axis). The scale of the *y*-axis was adjusted to stress the bimodality of the distribution. All grid sizes were pooled together, resulting in 36000 repetitions per histogram.

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

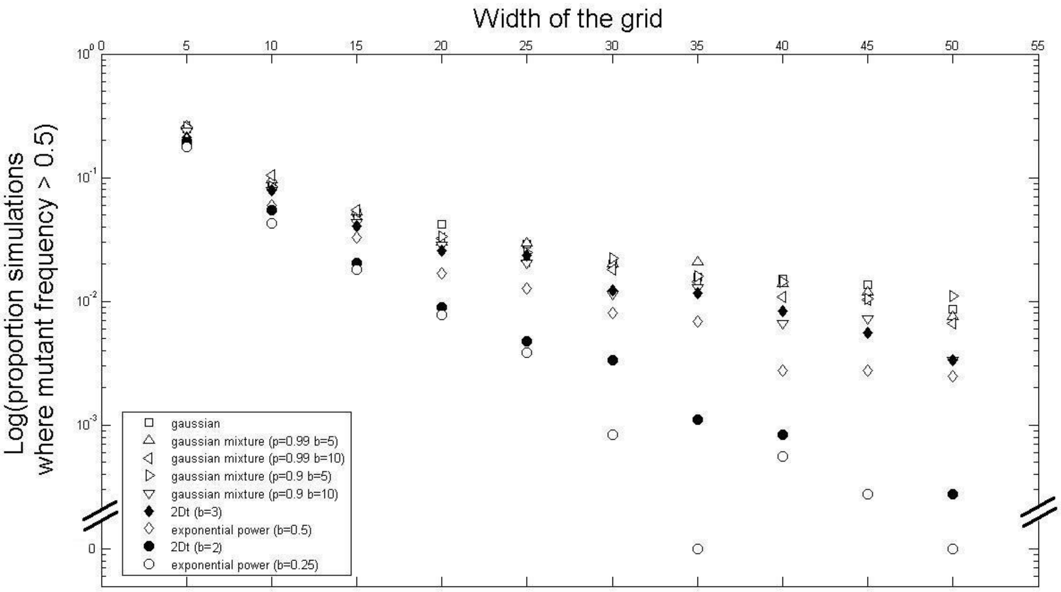

Fig. 4 – Proportion of successful simulations where mutant frequencies at the end of the colonization were higher than 0.5 for 9 colonisation dispersal kernels (Fig.1 and Table 1) and 10 grid widths. We plotted the logarithm of the proportion of successful simulations where mutant frequencies at the end of the colonisation were higher than 0.5 (*y*-axis) against grid sizes (*x*-axis). One different symbol was used for each different dispersal kernel. We used a logarithmic representation of the values on the *y*-axis to stress the differences between dispersal kernels. The *y*-axis was cut to show all values, even proportions of simulations equal to zero (exponential power b=0.25).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

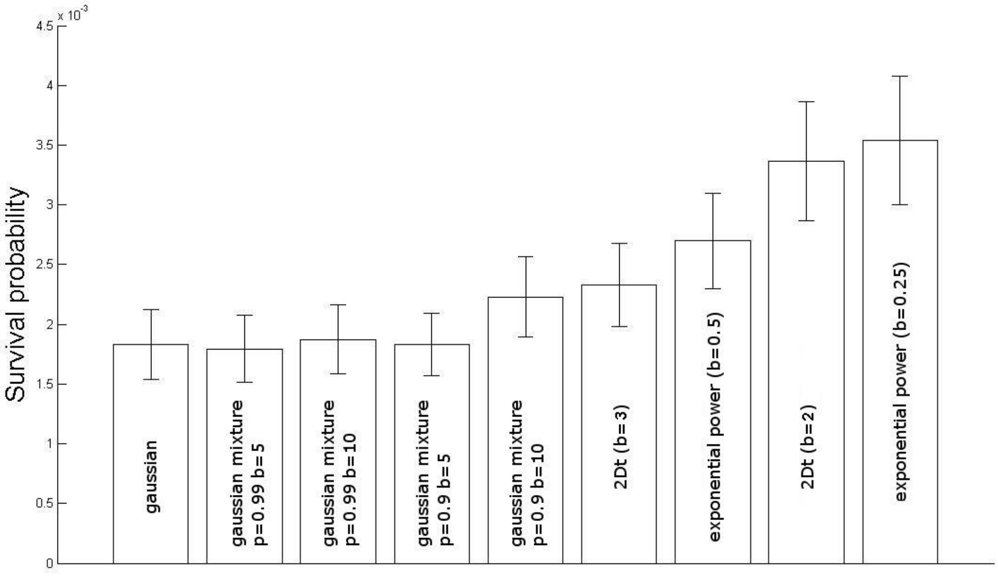

Fig. 5 – Effect of the dispersal kernel on the probability of mutation survival. For the 9 colonisation kernels used, sorted from thinner-tailed to fatter-tailed, the proportion of simulations where the mutation was still present at the end of the colonisation is represented. The error bars represent ± 2 Standard Deviation.

Version définitive du manuscrit publié dans / Final version of the manuscript
published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

Dispersal kernel	**Form**	**Parameters**	**Fatness**
Gaussian	$\frac{k}{\pi a^2}\exp(-\frac{r^2}{a^2})$	a = 3,68 k = 0,20	---
Gaussian mixture	$p(\frac{k}{\pi a^2}\exp(-\frac{r^2}{a^2}))+(1-p)(\frac{k}{\pi b^2}\exp(-\frac{r^2}{b^2}))$	a = 2,98 b = 10 p = 0,90 k = 0,20 a = 3,53 b = 5 p = 0,90 k = 0,20 a = 3,61 b = 10 p = 0,99 k = 0,20 a = 3,66 b = 5 p = 0,99 k = 0,20	+
2Dt	$\frac{k(b-1)}{\pi a^2}(1+\frac{r^2}{a^2})^{-b}$	a = 4,16 b = 3 k = 0,21 (fatter) a = 2,09 b = 2 k = 0,22	++/+++
Exponential power	$\frac{kb}{2\pi a^2\Gamma(2/b)}\exp(-\left(\frac{r}{a}\right)^b)$	a = 4.10^{-4} b = 0,25 k = 0,27 (fatter) a = 0,16 b = 0,50 k = 0,23	++/+++

Table 1 – Expression of the 2D dispersal kernels γ(*x*,*y*) outside of the central deme, where $r=\sqrt{x^2+y^2}$. For all dispersal kernels, except Gaussian mixture kernel, *a* is a scale parameter and *b* is a shape parameter. *k* is a constant used for keeping constant the migration rate (m = 20%) and the mean dispersal distance (δ = 3). $k=\frac{1-m}{1-m_0}$, where *m* is the expected migration rate (m = 20%) and m_o would be the migration rate for the unscaled kernel (i.e. without *k*).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

		Mean frequency of mutants	**Proportion of mutant frequencies > 0.5**	**Survival probability**
Latitude	½	$0.0516 \pm 1.78.10^{-2}$	0.0467	$0.0021 \pm 9.8.10^{-4}$
	¼	$0.0453 \pm 1.53.10^{-2}$	0.0399	$0.0020 \pm 9.7.10^{-4}$
Longitude	5	$0.0561 \pm 1.80.10^{-2}$	0.0492	$0.0021 \pm 9.7.10^{-4}$
	30	$0.0408 \pm 1.42.10^{-2}$	0.0374	$0.0020 \pm 9.7.10^{-4}$

Table 2 – Effect of the position of the first mutant on both the mean frequency of mutants at the end of the colonisation and the survival probability. The position on the y-axis of the first mutant (latitude) can be in the centre of the grid (½ of the grid width) or near the bottom edge (¼ of the grid width). The position on the x-axis of the first mutant (longitude) can be near an edge (5 demes after the edge) or farther on the grid (30 demes after the edge).

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript Manuscrit d'auteur / Author manuscript

Literature cited

Austerlitz, F. & Garnier-Gere, P. H. 2003. 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* **90**: 282-290.

Bialozyt, R., Ziegenhagen, B. & Petit, R. J. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**: 12-20.

Bohrer, G., Nathan, R. & Volis, S. 2005. Effects of long-distance dispersal for metapopulation survival and genetic structure at ecological time and spatial scales. *Journal of Ecology* **93**: 1029-1040.

Cain, M. L., Milligan, B. G. & Strand, A. E. 2000. Long-distance seed dispersal in plant populations. *American Journal of Botany* **87**: 1217-1227.

Clark, J. S., Fastie, C., Hurtt, G., Jackson, S. T., Johnson, C., King, G. A., Lewis, M., Lynch, J., Pacala, S., Prentice, C., Schupp, E. W., Webb, T. & Wyckoff, P. 1998. Reid's paradox of rapid plant migration - Dispersal theory and interpretation of paleoecological records. *Bioscience* **48**: 13-24.

Clark, J. S., Lewis, M. & Horvath, L. 2001. Invasion by extremes: Population spread with variation in dispersal and reproduction. *American Naturalist* **157**: 537-554.

Davies, S., White, A. & Lowe, A. 2004. An investigation into effects of long-distance seed dispersal on organelle population genetic structure and colonization rate: a model analysis. *Heredity* **93**: 566-576.

Edmonds, C. A., Lillie, A. S. & Cavalli-Sforza, L. L. 2004. Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 975-979.

Excoffier, L. & Ray, N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution* **23**: 347-351.

Foll, M. & Gaggiotti, O. 2008. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* **180**: 977-993.

Haag, C. R., Riek, M., Hottinger, J. W., Pajunen, V. I. & Ebert, D. 2006. Founder events as determinants of within-island and among-island genetic structure of Daphnia metapopulations. *Heredity* **96**: 150-158.

Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. 2007. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 19926-19930.

Hallatschek, O. & Nelson, D. R. 2008. Gene surfing in expanding populations. *Theoretical Population Biology* **73**: 158-170.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907-913.

Higgins, S. I., Richardson, D. M. & Cowling, R. M. 1996. Modeling invasive plant spread: The role of plant-environment interactions and model structure. *Ecology* **77**: 2043-2054.

Klein, E. K., Lavigne, C. & Gouyon, P.-H. 2006. Mixing of propagules from discrete sources at long distance: comparing a dispersal tail to an exponential. *BioMed Central Ecology* **6**.

Klopfstein, S., Currat, M. & Excoffier, L. 2006. The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* **23**: 482-490.

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Journal of Evolutionary Biology, 2009, DOI: 10.1111/j.1420-9101.2009.01832.x

Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript
Manuscrit d'auteur / Author manuscript

Kot, M., Lewis, M. A. & vandenDriessche, P. 1996. Dispersal data and the spread of invading organisms. *Ecology* **77**: 2027-2042.

Lambrinos, J. G. 2004. How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology* **85**: 2061-2070.

LeCorre, V., Machon, N., Petit, R. J. & Kremer, A. 1997. Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genetical Research* **69**: 117-125.

McLachlan, J. S., Clark, J. S. & Manos, P. S. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* **86**: 2088-2098.

Mollison, D. 1977. Spatial Contact Models for Ecological and Epidemic Spread. *Journal of the Royal Statistical Society Series B-Methodological* **39**: 283-326.

Nathan, R. & Muller-Landau, H. C. 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution* **15**: 278-285.

Petit, R. J., Bialozyt, R., Garnier-Gere, P. & Hampe, A. 2004. Ecology and genetics of tree invasions: from recent introductions to Quaternary migrations. *Forest Ecology and Management* **197**: 117-137.

Petit, R. J. & Hampe, A. 2006. Some evolutionary consequences of being a tree. *Annual Review of Ecology Evolution and Systematics* **37**: 187-214.

Rees, E. E., Pond, B. A., Cullingham, C. I., Tinline, R. R., Ball, D., Kyle, C. J. & White, B. N. 2009. Landscape modelling spatial bottlenecks: implications for raccoon rabies disease spread. *Biology Letters*.

Rosenthal, D. M., Ramakrishnan, A. P. & Cruzan, M. B. 2008. Evidence for multiple sources of invasion and intraspecific hybridization in Brachypodium sylvaticum (Hudson) Beauv. in North America. *Molecular Ecology* **17**: 4657-4669.

Shaw, M. W. 1995. Simulation of Population Expansion and Spatial Pattern When Individual Dispersal Distributions Do Not Decline Exponentially with Distance. *Proceedings of the Royal Society of London Series B-Biological Sciences* **259**: 243-248.

Shigesada, N. & Kawasaki, K. 1997. *Biological Invasions: Theory and Pratice*. Oxford University Press.

Trakhtenbrot, A., Nathan, R., Perry, G. & Richardson, D. M. 2005. The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions* **11**: 173-181.

Vlad, M. O., Arkin, A. & Ross, J. 2004a. Response experiments for nonlinear systems with application to reaction kinetics and genetics. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 7223-7228.

Vlad, M. O., Cavalli-Sforza, L. L. & Ross, J. 2004b. Enhanced (hydrodynamic) transport induced by population growth in reaction-diffusion systems with application to population genetics. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 10249-10253.

Vlad, M. O. I., Szedlacsek, S. E., Pourmand, N., Cavalli-Sforza, L. L., Oefner, P. & Ross, J. 2005. Fisher's theorems for multivariable, time- and space-dependent systems, with applications in population genetics and chemical kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 9848-9853.

Wei, W. & Krone, S. M. 2005. Spatial invasion by a mutant pathogen. *Journal of Theoretical Biology* **236**: 335-348.

Wingen, L. U., Brown, J. K. M. & Shaw, M. W. 2007. The population genetic structure of clonal organisms generated by exponentially bounded and fat-tailed dispersal. *Genetics* **177**: 435-448.