ARTICLE

The role of dispersal, selection, and timing of sampling on the false discovery rate of loci under selection during geographic range expansion

Paul Mayrand, Élise Filotas, Julian Wittische, and Patrick M.A. James

Abstract: Identifying adaptive loci is important to understand the evolutionary potential of species undergoing range expansion. However, in expanding populations, spatial demographic processes such as allele surfing can create spatial patterns of neutral genetic variation that appear similar to those generated through adaptive processes. As a result, the false discovery rate of adaptive loci may be inflated in landscape genomic analyses. Here, we take a simulation modelling approach to investigate how range expansion affects our ability to correctly distinguish between neutral and adaptive genetic variation, using the mountain pine beetle outbreak system as a motivating example. We simulated the demographic and population genetic dynamics of populations undergoing range expansion using an individual-based genetic model CDMetaPOP. We investigated how the false discovery rate of adaptive loci is affected by (*i*) dispersal capacity, (*ii*) timing of sampling, and (*iii*) the strength of selection on an adaptive reference locus. We found that a combination of weak dispersal, weak selection, and early sampling presents the greatest risk of misidentifying loci under selection. Expanding populations present unique challenges to the reliable identification of adaptive loci. We demonstrate that there is a need for further methodological development to account for directional demographic processes in landscape genomics.

Key words: range expansion, allele surfing, landscape genomics, CDMetaPOP, individual-based model, simulation.

Résumé : Il est important d'identifier les locus adaptatifs pour comprendre le potentiel évolutif des espèces qui connaissent une expansion de leur aire de distribution. Cependant, au sein de populations en expansion, les processus démographiques spatiaux comme le surf allélique peuvent générer des distributions de la variation allélique neutre qui sont semblables à celles issues de processus adaptatifs. En conséquence, le taux de faux positifs chez les locus adaptatifs est possiblement gonflé lors d'analyses en génomique du paysage. Dans ce travail, les auteurs utilisent la simulation pour étudier comment l'expansion de l'aire affecte la capacité de distinguer correctement la variation génétique neutre vs adaptative, en utilisant comme modèle l'épidémie du dendroctone du pin. Les auteurs ont simulé les dynamiques de la démographie et de la génétique populationnelle chez les populations connaissant une expansion de leur aire au moyen d'un modèle CDMetaPOP fondé sur les individus. Les auteurs ont étudié comment le taux de faux positifs était affecté par (*i*) la capacité de dispersion, (*ii*) le moment de l'échantillonnage et (*iii*) l'intensité de la sélection sur un locus adaptatif de référence. Les auteurs ont trouvé qu'une combinaison d'un faible taux de dispersion, une faible sélection et un échantillonnage précoce entraînait le plus grand risque d'identification erronée des locus sous sélection. Les populations en expansion présentent des défis uniques pour l'identification de locus adaptatifs. Les auteurs démontrent qu'il faut développer des méthodes pour tenir compte des processus démographiques directionnels en génomique du paysage. [Traduit par la Rédaction]

Mots-clés : expansion de l'aire de distribution, génomique du paysage, CDMetaPOP, modèle basé sur les individus, simulation.

Introduction

Range expansion is a complex spatial phenomenon that can be the result of directional species adaptation to environmental changes (Hassall et al. 2009), species in-

vasions (Brown et al. 2015; Lombaert et al. 2014), or population outbreaks (Cullingham et al. 2011; Raffa et al. 2013). The consequences of shifting species distributions have fundamental consequences for ecosystem function-

P. Mayrand, J. Wittische, and P.M.A. James. Université de Montréal, Département de sciences biologiques, CP 6128 Succursale Centre-Ville Montréal, QC H3C 3J7, Canada.
É. Filotas. TÉLUQ (Université du Québec), Département Science et Technologie, 5800 rue Saint-Denis, Montréal, QC H2S 3L5, Canada.
Corresponding author: Patrick M.A. James (email: patrick.ma.james@gmail.com).

ing and human well-being (Pecl et al. 2017). Understanding the eco-evolutionary mechanisms underlying range expansion is essential to develop new forecasting models and mitigation strategies for expanding species.

Range expansions pose significant challenges to the analysis of population genetic and genomic data because of the complex spatial patterns in allele frequencies they produce (Bialozyt et al. 2006; Excoffier et al. 2009). For example, successive founder effects and consequent allele surfing can produce clines in the frequencies of both neutral and adaptive alleles along an expansion axis (Klopfstein et al. 2006; Peter and Slatkin 2015; Slatkin and Excoffier 2012). Clinal variation in neutral alleles due to allele surfing can mimic spatial structure due to adaptive processes (De Villemereuil et al. 2014; Excoffier and Ray 2008) and increase the probability of incorrectly identifying neutral loci as adaptive (i.e., inflated False Discovery Rate (FDR); Storey and Tibshirani 2003). An inflated FDR (i.e., Type I error) of adaptive loci can impede our understanding of the evolutionary drivers and consequences of range expansion. As a result, efforts to better manage, predict, or prevent the spread of irruptive expanding species may be negatively affected.

Landscape genomics seeks to improve our understanding of these eco-evolutionary dynamics through the identification of potentially adaptive loci that are linked to environmental conditions (Holderegger et al. 2008). Here, we explore some of the consequences of range expansion on the application of landscape genomics methods. Many landscape genomic tools currently exist to identify genotype–environment correlations, including BAYENV (Coop et al. 2010), LFMM (Frichot et al. 2013), and SAM (Joost et al. 2007). Methods have also been developed to account for the effects of population genetic structure (Frichot et al. 2013), sampling design (Günther and Coop 2013), and linkage disequilibrium (François et al. 2016). However, to our knowledge, the potentially confounding effects of range expansion and allele surfing on identifying loci under selection have not been examined.

In this paper, we used the irruptive mountain pine beetle (*Dendroctonus ponderosae*) as a motivating example of how range expansion can affect our ability to identify loci under selection. We used a spatially explicit, individual-based model to simulate beetle range expansion and to investigate the conditions under which allele frequencies of neutral loci are more likely to exhibit spatial patterns (i.e., clines) similar to those expected of loci under selection. We specifically quantified how the probability of incorrectly identifying a neutral locus as adaptive varies as a function of (*i*) dispersal capacity, (*ii*) timing of sampling during range expansion, and (*iii*) strength of selection for a reference adaptive locus.

Mountain pine beetle range expansion

The mountain pine beetle is a native irruptive forest pest in western North America that attacks and kills pine species (Safranyik and Wilson 2006). Mountain pine bee-

tle outbreak dynamics are shaped by interactions among landscape topography (de la Giroday et al. 2011), dispersal (Safranyik et al. 1992), forest connectivity (Bone et al. 2013), and climate (Raffa et al. 2008). Outbreaks of the mountain pine beetle have significant consequences for the forest industry and forest dependent communities (Pedersen 2003), carbon budgets (Kurz et al. 2008), as well as the dynamics of future disturbances, such as fire (Harvey et al. 2014; Simard et al. 2011).

Spatial legacies in forest landscape structure due to forest management, in combination with climate warming, have resulted in an unprecedented outbreak and range expansion in Canada and the USA (Bentz et al. 2010; Raffa et al. 2008). In western Canada and the western United States, the outbreak has affected over 18 million ha (Nealis and Cooke 2014) and 8.9 million ha (USDA Forest Service 2012) of forest, respectively. Continued northward expansion of the mountain pine beetle into the Yukon territory, eastward expansion through Alberta towards Saskatchewan (Safranyik et al. 2010), and successful attack of the novel host jack pine (*Pinus banksiana*) in Alberta (Cullingham et al. 2011) pose a national-scale threat to the boreal forest.

It has been suggested that the recent mountain pine beetle range expansion may have been facilitated by rapid adaptation to new environmental conditions (Janes et al. 2014). Characterization of the geographic distribution of adaptive loci can improve our understanding of the evolutionary mechanisms driving mountain pine beetle range expansion which is essential to develop new tools to predict future outbreak risk. While landscape genomics methods are compelling tools to reach these goals, uncertainty remains regarding how the dynamics of spatially expanding populations might affect the false positive rate associated with the identification of putative adaptive loci.

Methods

Modelling framework

We simulated the development of spatial genetic structure of a population of mountain pine beetle during range expansion using CDMetaPOP (Landguth et al. 2016). Each simulated beetle was represented by a genotype of 100, freely recombining, neutral bi-allelic (i.e., SNP) loci and one bi-allelic adaptive locus.

The simulation landscape was modelled as a homogeneous rectangular grid of 20 by 80 cells with absorbing boundary conditions. Mirroring the geographic extent of mountain pine beetle range expansion from northwestern British Columbia to northern Alberta within 10 years of the current outbreak (Janes et al. 2014), the simulated landscape can be interpreted as an area of 160 km by 640 km with cells of 8 km by 8 km (6400 ha). The carrying capacity of each cell was set to 30 individuals, which translates into a maximum possible population of 48 000 individuals on the simulated landscape. These

Table 1. Summary of all scenarios and associated experimental factor levels.

Scenario	Maximum dispersal capacity	Strength of selection	Maximum s-value (%)
1	Weak (1 cell; 8 km)	Weak	1
2	Weak	Intermediate	5
3	Weak	Strong	50
4	Intermediate (2 cells; 16 km)	Weak	1
5	Intermediate	Intermediate	5
6	Intermediate	Strong	50
7	Strong (8 cells; 64 km)	Weak	1
8	Strong	Intermediate	5
9	Strong	Strong	50

Note: We simulated three dispersal capacities and three strengths of selection in a fully crossed design for a total of nine scenarios, with 50 replicates each. Maximum dispersal capacity refers to the maximum distance that could possibly be selected from our dispersal kernel (eq. 1). Strength of selection indicates the maximum mortality rate for genotype *aa* at the far end of the geographic expansion grid. Maximum *s*-value expresses these mortality rates as fitness costs. This study reports on a fully crossed simulation experiment using these two three-level factors. See text for further details.

parameter values were selected as a compromise between realism and time-related computational limitations. While the size of the simulated population is much smaller than that of the actual mountain pine beetle outbreak, it was large enough to generate the complex eco-evolutionary dynamics (i.e., allele surfing and neutral clines) that are the subject of this study.

Three main processes determined the spatio-temporal population and genomic dynamics of this simulated population: (1) dispersal; (2) selection; and (3) reproduction. Below, we summarize each process and describe the associated parameters. Additional details regarding parameters and CDMetaPOP architecture are included in the Supplementary Materials[1].

Experimental design

We examined model sensitivity to three levels of dispersal (weak, intermediate, and strong), three strengths of selection (weak, intermediate, and strong) in a crossed design for a total of nine experimental scenarios (Table 1). We also examined the effect of the timing of sampling (i.e., time since the beginning of the expansion) within each scenario. Simulations were run for a time horizon of 150 non-overlapping, univoltine generations (i.e., 1 generation per year). The choice of 150 generations was motivated by a desire to allow the simulation grid to be filled or near filled in all scenarios while avoiding unnecessary simulation of additional time steps in scenarios that filled the grid quickly. Simulations were sampled at six different points in time: generations 25, 50, 75, 100, 125, and 150. Within each selected generation, we sampled all individuals from each of 200 regularly spaced sample locations. These sample locations were consistent for all scenarios and generations. Each scenario was replicated 50 times to capture a range of sto-

chastic variation in the simulated spatial processes (Table 1).

Dispersal

Dispersal was modelled using a thresholded and normalized dispersal kernel (Landguth et al. 2016) in which the probability of dispersal declined exponentially as a function of the Euclidean distance between cells

$$(1) \qquad P_{ij} = \frac{10^{-\alpha \cdot d_{ij}}}{\sum_{j=1}^{d_{max}} 10^{-\alpha \cdot d_{ij}}}$$

Here, d_{ij} is the distance between the centroids of cells i and j, expressed in cell units. We set $\alpha = 0.3$ to include the potential for within stand movement (Evenden et al. 2014; Robertson et al. 2009) as well as potential long-distance dispersal (Chen and Walton 2011). For any simulated individual dispersal event, the distance travelled was randomly selected from this kernel. The cell to which an individual travelled was selected randomly from the set of cells at the pre-selected distance. We investigated the effect of different maximal dispersal capacities (d_{max}) on spatial genetic structure by truncating this dispersal kernel at three different distances that represented weak, intermediate, and strong dispersal. Weak dispersal corresponds to a maximal dispersal distance (d_{max}) of one cell (i.e., 8 km), intermediate dispersal corresponds to a maximal dispersal distance of two cells (i.e., 16 km), and strong dispersal corresponds to a maximal dispersal distance of eight cells (i.e., 64 km). Although these truncated dispersal kernels do not directly reflect MPB movement dynamics, they contribute to an experimental framework that allows us to examine the consequences of different dispersal capacities and to

make generalizations on the role of movement on the development of neutral genetic clines.

Selection

We sought to identify simulated neutral loci that exhibited spatial patterns similar to an adaptive locus during different stages of a range expansion. To do so we compared the slopes of the relationship between neutral allele frequencies and latitude to that of a known adaptive locus with varying strengths of selection. This reference codominant adaptive locus had three genotypes: *AA*, *Aa*, and *aa*. The strength of selection (*s*) for this locus was controlled using different geographic mortality gradients associated with each of these three genotypes. Gradients varied in terms of the rate of change in the viability fitness (i.e., probability of survival) with increasing latitude due to the presence of the minor allele *a*. That is, individuals carrying the *a* allele had a greater probability of mortality as they moved from the bottom of the simulation area (i.e., south) to the top (i.e., north). In this way we simulated northward expansion of mountain pine beetle from a suitable to less suitable habitat. The viability fitness for the homozygous genotype *aa* at a given latitude was equal to $1 - s$, where *s* is the geographic selection coefficient (i.e., latitude). The *A* allele was advantageous such that the survival of the homozygous genotype *AA* was always equal to 1. The viability fitness of the heterozygote genotype *Aa* was half that of the homozygote *aa* such that survival was equal to $1 - 0.5s$. This framework resulted in a selection model in which there was no genetic trade-off between the two adaptive alleles (Agosta and Klemens 2009; Gompert et al. 2015), which is not uncommon for studies of range expansion (Burton and Travis 2008; Creech et al. 2017; Slatkin 1976).

The maximum fitness costs (i.e., mortality probabilities) at the extreme north end of the expansion grid for the homozygous genotype *aa* were set to 1%, 5%, and 50%. These values correspond to the maximum *s*-value for a given scenario and represent the weak, intermediate, and strong strengths of selection. Geographic gradients in fitness costs were generated by linearly interpolating from the most southern sites where the fitness cost was equal to 0, to the northernmost sites and their respective maximum cost (*s*). For this study, selection gradients were strictly linear along the south-north axis of the simulation grid.

The values for the weak and intermediate levels of selection were selected as they represent standard, plausible fitness costs across a spatial gradient (Landguth et al. 2016). The value for the strong selection scenario (50%) was chosen based on the known range of overwinter mortality rates across a gradient from the historical MPB range (i.e., southern British Columbia) into areas where the MPB is expanding (i.e., Alberta; Cooke 2009). The corresponding maximum fitness costs for heterozygotes (*Aa*) were 0.5%, 2.5%, and 25%. In this model we

assume that the dominance parameter $h = 0.5$ (i.e., no recessivity).

Reproduction

Reproduction was sexual between simulated diploid individuals and mates were selected from within the same cell after dispersal. The number of eggs produced by each female per mating event was selected from a Poisson distribution with mean of 60 (Safranyik and Wilson 2006). Sex-ratio was set to one male for two females (James et al. 2016). Males could reproduce with more than one female per generation, whereas mature females could only be mated by a single male per generation. Mutation rate was set to 0.0005, the default value for CDMetaPop (Landguth et al. 2016). Populations within each cell grew exponentially up to the carrying capacity of a cell (i.e., $K = 30$). Once the number of individuals exceeded the carrying capacity of the cell, individuals were randomly removed until the population was reduced to *K*. As individuals are removed randomly there is no effect on the relative frequencies of the adaptive alleles.

Initial conditions

All simulations began with a population of 120 individuals occupying four adjacent cells in the center of the southernmost part of the territory. Alleles were initially allocated among individuals randomly, resulting in maximum diversity (Landguth et al. 2016). A single burn-in generation, under panmictic conditions (i.e., random mating, no dispersal limitation), was run to remove artefacts of this initial genotype assignment and to establish Hardy-Weinberg Equilibrium (Hardy 1908; Nagylaki 1976). Each simulation began with its own random assignment of alleles and associated burn-in generation. Other simulation parameters are described in Table S1[1].

Identifying outlier loci

Loci were identified as potentially adaptive on the basis of the coefficient estimates (slope) of linear regressions of simulated allele frequencies on latitude. We compared the slopes of neutral loci to the simulated adaptive locus. When the absolute value of the slope of a known neutral locus exceeded that of the known adaptive locus it was considered a false discovery. We examined how the false discovery rate (FDR) of potentially adaptive loci varied in response to our experimental treatments involving dispersal, strength of selection, and the timing of sampling.

We calculated the FDR of adaptive loci at six points in time for each simulation scenario (Table 1). To do so, we calculated the mean of the slopes of the known adaptive locus over the 50 replicates at each time point within a scenario (Table 1). Neutral slopes that were to exceed this mean value were classified as "adaptive". We summed the number of neutral loci that met this condition and divided this value by the total number of loci (i.e., 101) to obtain our FDR. Finally, we used ANOVA to quantify the

Fig. 1. Changes in (*a*) isolation by distance (Mantel's *r*) and (*b*) spatial genetic structure (F_{st}) over time for three different simulated dispersal strengths. Points indicate mean values and error bars represent one standard deviation calculated using 50 replicates. Results show values obtained when the strength of selection was intermediate (5%). Patterns did not differ with different levels of selection.
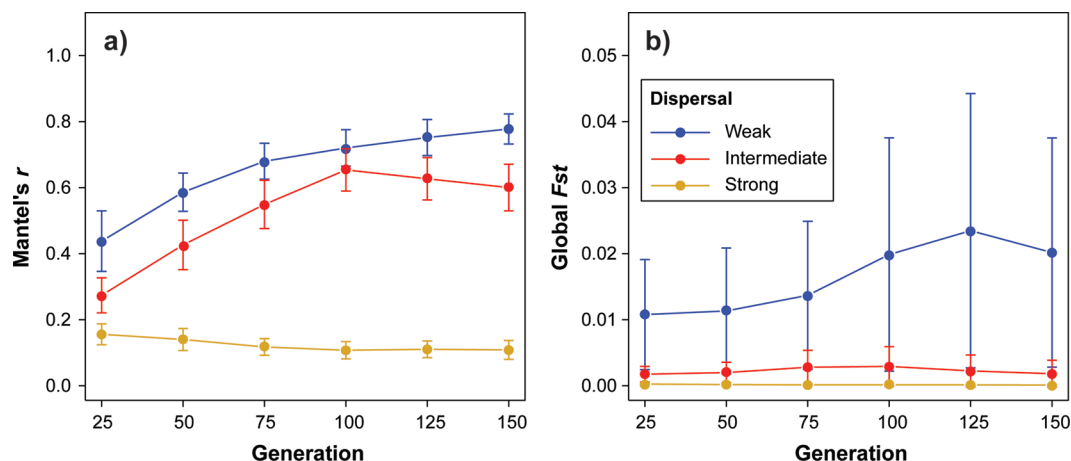
relative importance of the dispersal, selection, timing, as well as all two-way and three-way interactions on the FDR of adaptive loci.

Model validation

Prior to launching simulations, we verified that the model was producing the spatial patterns expected of neutral loci during range expansion. We specifically sought to verify that (*i*) spatial genetic structure and isolation by distance (IBD) changed as the expansion progressed, (*ii*) greater dispersal reduced population genetic structure, and (*iii*) allele surfing produced neutral clines along the main axis of expansion (Excoffier and Ray 2008; Vasemägi 2006).

The development of global spatial genetic structure (F_{st}) was assessed using the sample-size corrected global F_{st} (Weir and Cockerham 1984) calculated using the *wc* function in the *HIERFSTAT* package in R (Goudet and Jombart 2015). The temporal development of IBD was assessed using Mantel tests of the correlation between pair-wise genetic and geographic distance matrices. Pair-wise genetic distances (F_{st}) were calculated from the regular grid of sample of locations (*n* = 200) using log-transformed neutral allele frequencies. Mantel tests were calculated using the *mantel.rtest* function from the package ADE4 (Dray and Dufour 2007). Global F_{st} and IBD were assessed using data sampled at generations 25, 50, 75, 100, 125, and 150. The creation of neutral clines was verified by visually assessing the relationship between allele frequency and geographic location (latitude) for multiple randomly sampled neutral loci.

Results

Simulations

The time to fully populate the simulation grid varied with dispersal and selection strength. It took on average 160 (Standard Error; SE = 8) generations to fill the simulated landscape in weak dispersal scenarios, 109 (SE = 5)

generations in intermediate dispersal scenarios, and 58 (SE = 4) for strong dispersal scenarios. Slightly less time was required to fill the simulation grid under weak and intermediate strengths of selection because mortality rates were lower in these scenarios.

Model validation

Our model behaved as expected and produced genetic patterns characteristic of a range expansion (Figs. 1, 2). Isolation by distance (IBD; Fig. 1a) increased steadily through time for both weak and intermediate levels of dispersal but did not change meaningfully when dispersal capacity was strong. The strength of spatial genetic structure (F_{st}; Fig. 1b) remained low for intermediate and strong dispersal capacities. However, when dispersal capacity was weak, genetic structure increased through time and then decreased near the end of the simulation horizon. This final decrease reflects the onset of genetic homogenization of the simulation lattice. We also observed the emergence of spatial clines in neutral loci (Fig. 2) indicating that allele surfing was occurring. We also found that these clines were dynamic and changed as the expansion progressed (Fig. 2). As expected, strength of selection did not influence patterns in IBD, spatial genetic structure, or the emergence of neutral clines.

Identifying outlier loci

The slopes of the adaptive locus did not remain fixed through time. Instead, these slopes steadily decreased over the course of the simulation for all scenarios due to fixation (Fig. 3). These changes in the slope of the adaptive reference locus resulted in conservative classification of neutral loci as adaptive early in the simulation, and liberal classification later in the simulation (Fig. 3).

Among loci exhibiting neutral clines, many exhibited slopes that were greater than the adaptive threshold and could therefore be interpreted as under selection. Incor-

Fig. 2. The slopes of neutral clines change as an expansion progresses due to allele surfing. Points and fitted lines illustrate the development of a neutral cline for a single selected locus (locus 36) along a center transect at three different points in time (*a*, *b*, and *c*). Data were taken from a single replicate of the weak selection and weak dispersal scenario (scenario 1; Table 1). The slopes of each individual alleles relationship with latitude were compared to the slope of a single reference adaptive locus. Neutral slopes that exceeded the adaptive reference slope (not shown here) could be considered (incorrectly) to be under selection.
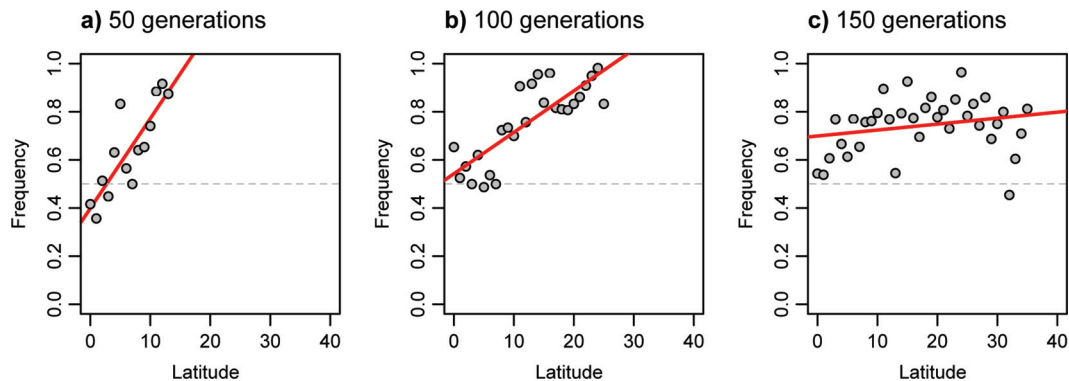
Fig. 3. Changes in the average slope of the reference adaptive locus through time as a function of different simulated dispersal and selection strengths. Averages were calculated over 50 replicates. The X-axis has been log-scaled to better illustrate the important change in the tendency of the mean slope after 10 generations. As a result of this simulation artefact, the classification of neutral loci relative to an adaptive reference is likely to be overly conservative early in the simulation. The slope of the simulated adaptive locus tends to zero in all scenarios due to fixation but does so at different rates.
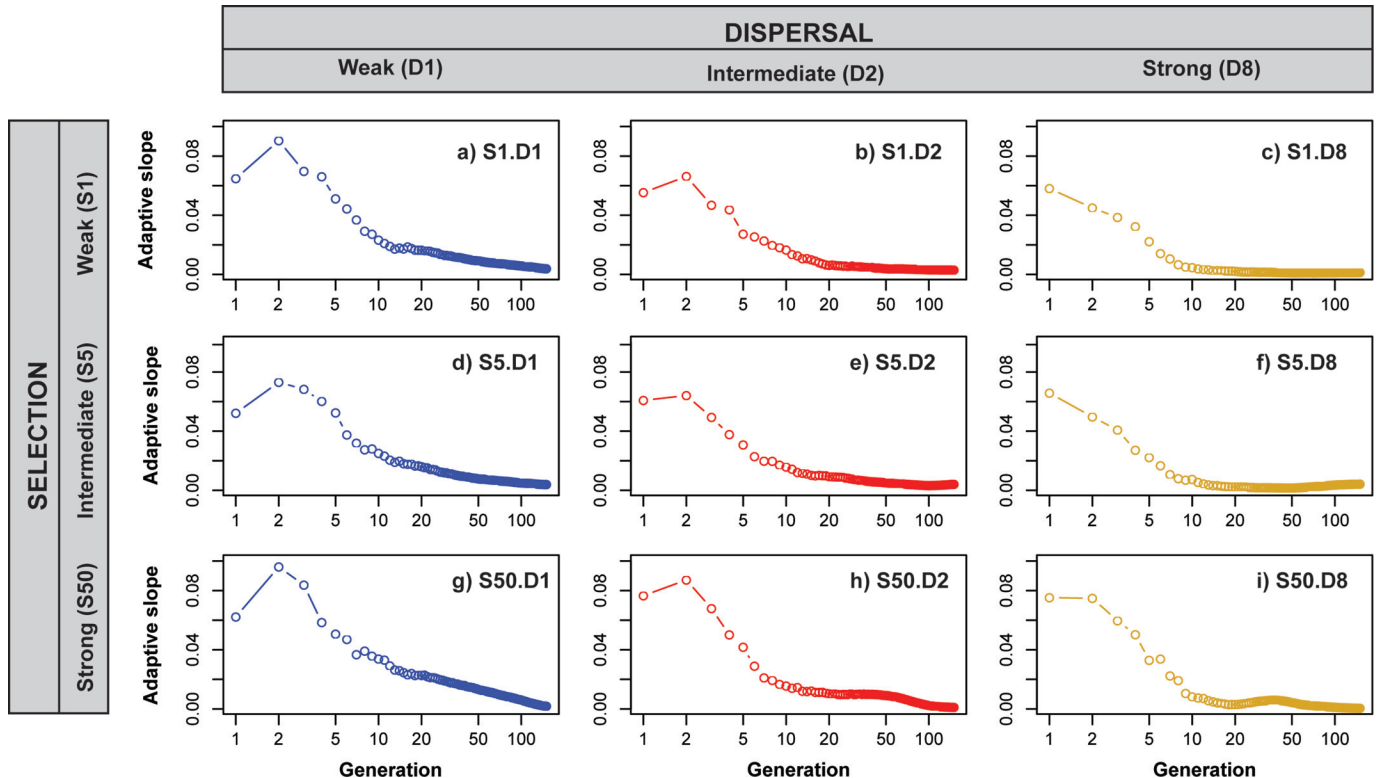
rectly identified outlier neutral loci were found in all scenarios (Fig. 4). However, the frequency of such incorrect classification tended to be higher in scenarios with weak dispersal (Figs. 4*a*–4*c*), with early sampling relative to the start of the expansion (Figs. 4*f*, 4*i*), and when the adaptive locus used as a reference for adaptation was subject to weak selection (Figs. 4*a*–4*c*). Because of the effects of fixation in the adaptive reference locus, the

rate of which is affected by selection strength, it is most informative to compare scenarios with the same level of selection (i.e., the rows in Fig. 4).

Interpretation of individual effects was not straightforward as we also identified a significant three-way interaction effect among dispersal, selection, and timing (Table 2) on the FDR of neutral loci as adaptive. Because statistical significance is sensitive to sample size, and

Fig. 4. Percentage of neutral loci classified as potentially adaptive through time as a function of simulated dispersal and selection strengths. Box and whisker plots summarize variation over 50 simulation replicates. The risk of confounding neutral variation for adaptive is consistently higher with weak dispersal capacity and weak selection in the reference locus. Under a high strength of selection, the increase in the number of outlier loci later in the simulation is the consequence of fixation of the adaptive reference locus (see text for details).
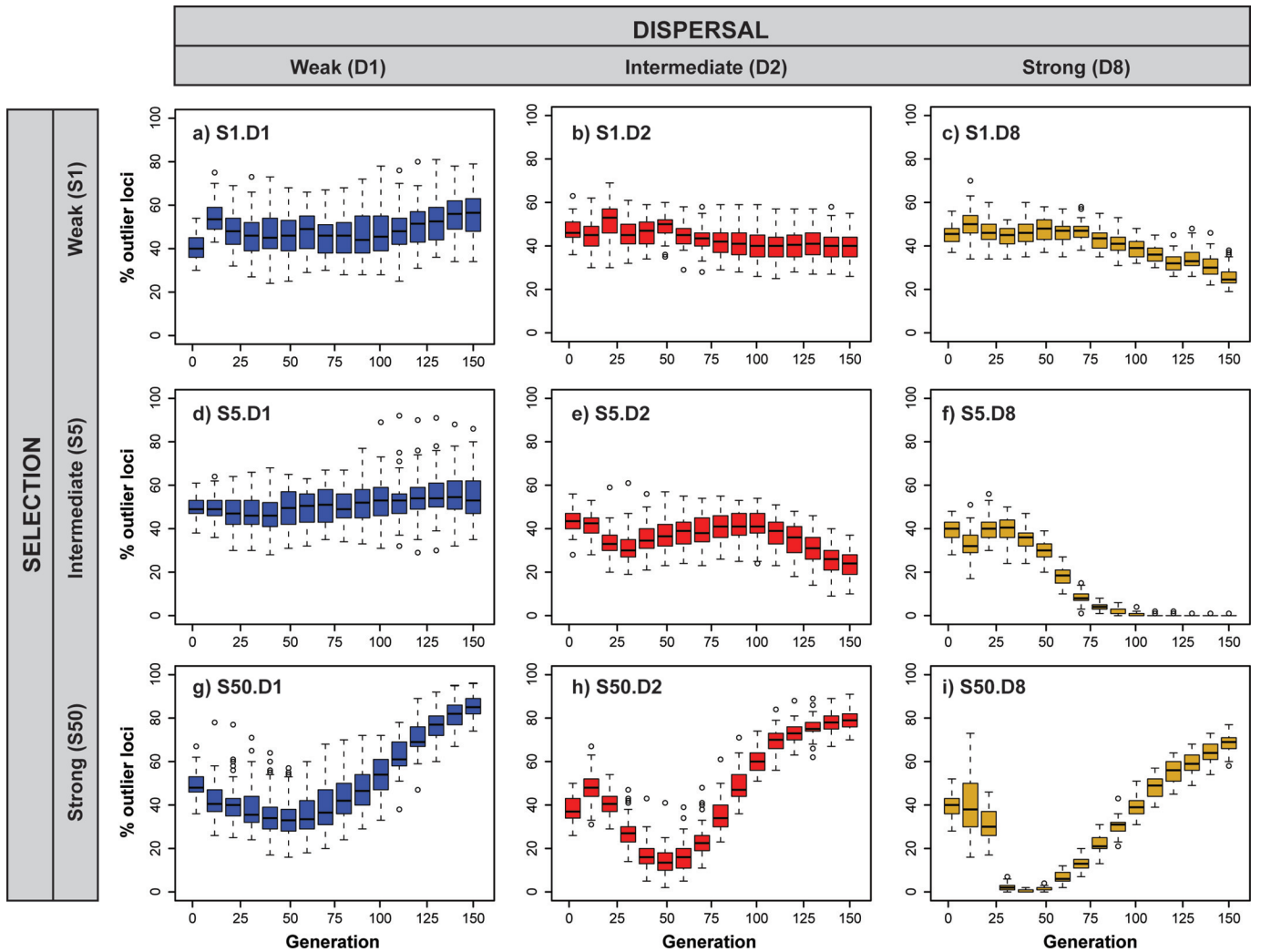
Table 2. ANOVA table summarizing the influence of strength of selection, dispersal, timing of sampling, and their interactions on the proportion of neutral loci misidentified as adaptive (i.e., FDR).

Factor	df	SS	Mean SS	F
Selection (S)	2	53275	26637	516.875
Dispersal (D)	2	58783	29392	570.318
Timing (T)	1	26184	26184	508.067
S:D	4	31002	7751	150.392
S:T	2	166637	83318	1616.718
D:T	2	7362	3681	71.425
S:D:T	4	12761	3190	61.903
Residuals	880	45351	52	

Note: Results are based on 50 simulation replicates. All F-statistics were significant ($p < 0.001$).

one can manipulate sample size in simulation experiments, significance alone is not sufficient to interpret results. To further aid interpretation, we chose to examine the F-statistics associated with each factor (Nakagawa and Cuthill 2007). F-ratios give a better impression of which predictors and interactions are of greater relative importance; larger values indicate greater inconsistency with the null hypothesis of no relationship between or among factors. We found that dispersal was the most influential factor on FDR, followed closely by strength of selection, and timing of sampling (Table 2). The interaction between selection and sample timing was the strongest ($F > 1500$; Table 2), whereas the other interactions seemed to be less important ($F < 150$; Table 2).

Fixation of adaptive locus

Scenarios that included strong selection for the adaptive reference locus (i.e., scenarios 7, 8, and 9; Table 1)

showed a notable increase in FDR near the end of the simulation horizon due to fixation of the adaptive reference locus (Figs. 4g–4i). We determined that as the simulated adaptive allele became fixed, its slope approached zero (Fig. 3). At this point, most, if not all, neutral loci are classified as potentially under selection using our simple detection method as the absolute values of their slopes are greater than zero. This fixation-related misclassification accounts for the upward trajectories in FDR later in the simulations (Figs. 4g–4i). The effect of fixation on the FDR was most pronounced in the strong dispersal scenarios (Fig. 4), in which fewer generations were needed for fixation to occur relative to the weak dispersal scenarios (Fig. 4).

Discussion

We investigated how range expansion affects our ability to identify loci under selection using the mountain pine beetle as a case study. We employed a spatially explicit individual-based simulation modelling approach (CDmetaPOP: Landguth et al. 2016) to model the population expansion of thousands of individuals represented as spatially-referenced genotypes that included 100 neutral loci and 1 adaptive locus. We sought to identify how often the slope of spatial clines of known neutral loci exceeded that of our reference adaptive locus and would be therefore potentially classified as under selection (i.e., FDR) under different dispersal and selection scenarios at different sample timings. We found that dispersal, strength of selection, and timing of sampling can all influence the probability of misclassifying neutral loci as adaptive. As with all simulation studies, these results should be seen as the consequence of interactions among assumptions, hypotheses, and initial conditions rather than an attempt to predict future population genetic dynamics. Therefore, it is the differences among scenarios that are most informative regarding the challenges to identifying loci under selection during range expansions.

Dispersal

We found that spatial patterns in neutral loci were less likely to be misidentified as adaptive when dispersal is strong. Greater dispersal has been shown to result in greater genetic connectivity which in turn limits the creation of strong neutral clines (Bialozyt et al. 2006; Polechová and Barton 2015), and it reduces the likelihood of allele surfing (Excoffier and Ray 2008). In contrast, reduced dispersal capacity and consequent limited gene flow facilitates allele surfing (Excoffier et al. 2009) and creates stronger neutral clines that can be more easily confounded with adaptive patterns. Our results confirm these previous studies.

Differentiating between neutral spatial genetic patterns and those that arise due to adaptive processes becomes more challenging when a species has a range of dispersal capacities, as is the case of highly mobile organ-

isms such as the mountain pine beetle (Safranyik et al. 1992). Most beetles disperse locally within a stand after emergence (Chen and Walton 2011), however some individuals succeed in dispersing passively hundreds of kilometers (Safranyik et al. 1992). Such long-distance dispersal events, although rare, can have an important effect on range expansion and consequent spatial genetic structure (Neubert and Caswell 2000). Here, we simulated variation in dispersal as a maximal distance possible for a given static dispersal kernel. Greater understanding of the risk of misidentifying neutral loci as adaptive in expanding systems would benefit from knowledge of the mean and maximal dispersal capacities of the species in question (Kot et al. 1996). Given our findings, one would expect that the risk of misidentifying neutral loci as adaptive would be increased in species with weak, or otherwise limited, dispersal capacity.

Timing of sampling

Timing of sampling had a significant effect on the patterns observed (Figs. 3, 4). This effect of timing was in part due to how the cline in our reference adaptive locus changed as the expansion progressed as a consequence of fixation (see below). As a result, transient dynamics in the probability misidentifying neutral loci were observed, partly due to dynamic nature of the reference locus (Fig. 3). In this context we can parse our results into two themes: first, the selection model we chose resulted in a dynamic adaptive reference slope. Such temporally varying adaptive slopes during range expansion may have consequences for the identification of true adaptive loci (i.e., false negatives) and should be investigated further. More germane to the objectives of this paper, we found that our ability to make meaningful inference regarding the influence of dispersal and timing on the development of strong neutral clines (i.e., the false positives) is limited to a short temporal window before fixation has occurred.

Sampling during the early stages of range expansion (i.e., <50 generations) resulted in a higher FDR, particularly when dispersal and selection are weak (Fig. 4). We simulated an expansion in which populations were initially comprised of few individuals. With fewer individuals, populations are more sensitive to genetic drift and there is a greater chance that allele surfing will result in neutral clines that can be misidentified as adaptive (Slatkin and Excoffier 2012). As the expansion progresses, and population size and gene flow increases, the initial genetic signal of range expansion, including strong spatial genetic structure and neutral clines, decays (Lotterhos and Whitlock 2015). This means that newly identified rapidly expanding populations, such as the mountain pine beetle, or recently introduced invasive species, are also those that present the highest risk of misidentifying loci as adaptive. That said, it can be difficult to know for certain the age of an outbreak or expansion, although observational data of first observation or

spatial variation in expected heterozygosity can be used (Swaegers et al. 2015). These findings present new research challenges and imperatives as outbreaking and newly introduced invasive species, about which we tend to know the least, also tend to be those of greatest eco-evolutionary and management interest.

The effect of timing on the FDR of adaptive loci varied with both selection and dispersal (Table 2; Fig. 4). These effects were most pronounced when dispersal was strong (Figs. 4c, 4f, 4i). Under strong dispersal and an intermediate level of selection, the FDR of adaptive loci decreased significantly after approximately 50 generations (Fig. 4f). When the strength of selection was strong, this decrease happened after approximately 25 generations (Fig. 4i). Consequent increases in the FDR after these generations reflects fixation of the adaptive locus, rather than an increase in the strength of neutral clines (Figs. 3, 4i).

In the case of the mountain pine beetle in western Canada, we know that the population outbreak and associated expansion began around 2005. Given the known long distance dispersal capacity of this species and the age of the expansion, our study suggests that the loci previously identified as potentially adaptive using landscape genomics (Batista et al. 2016; Janes et al. 2014) could benefit from further validation with additional analyses that take the effects of allele surfing into account.

Selection

We simulated three plausible levels of selection on a single bi-allelic codominant locus using a latitudinal mortality gradient. The slope of the relationship between allele frequency and location of this reference locus was used as a threshold to classify neutral loci as adaptive or not. We used this simple method of classifying loci as potentially under selection as it is logically consistent with the theory of allele surfing and neutral clines (Excoffier and Ray 2008), is conceptually clear, and was computationally straightforward. Several more sophisticated landscape genomic tools exist for the identification of adaptive loci (e.g., Coop et al. 2010; Forester et al. 2016; Frichot et al. 2013). A comparison of the relative performance of these different tools under the different scenarios we examined here is beyond the scope of this paper. However, future work should certainly examine the sensitivity of these methods to unique demographic situations such as range expansion.

We found that weaker selection in our reference adaptive locus resulted in smaller adaptive slopes, confirming the results of other studies (Yeaman and Whitlock 2011). As a result, the FDR of potentially adaptive loci was higher when selection was weak (Fig. 4), in particular around generation 50. This generation seems to mark the end of an initial period of transient dynamics, while at the same time is before the bias introduced by adaptive locus fixation has come into effect. As a consequence, we consider comparisons among scenarios at

generation 50 to be the most informative. Beyond this time, the slope of the reference adaptive allele tends to zero and most neutral alleles, due to either chance or allele surfing, will have a slope that exceeds this reference value.

When selection was stronger, fewer neutral loci were misidentified as adaptive because the reference adaptive slope was sufficiently high. Thus, our ability to correctly identify meaningful "outlier" loci is strongly contingent on the spatial distributions of true neutral and true adaptive slopes, which are unfortunately generally unknown. This finding also suggests that when the strength of selection is weak, arbitrary threshold tests using landscape genomics approaches may not be able to identify true adaptive loci.

Stronger selection might be expected in expansion scenarios involving large differences in climatic conditions of the home range relative to the new range (Colautti and Barrett 2013). Such is the case for the MPB where there is a large range of overwinter mortality rates (>50% increase) across a gradient from the historical to the recently expanded range (Cooke 2009).

This means that the steepness of the gradient in habitat suitability between historic and expanded ranges will affect our ability to identify adaptive loci potentially involved in successful expansion. Of course, the challenge here remains to identify the most relevant gradient between historic and novel ranges, which is ostensibly the goal of searching for adaptive loci in the first place. Practically speaking, we hypothesize that the most challenging situations to distinguish truly adaptive loci from those exhibiting allele-surfing induced neutral clines to be those in which the differences in habitat suitability between the two ranges are subtle.

Models of selection

We modelled selection as a linear latitudinal gradient, in the style of a temperature gradient (Keller et al. 2013). Range expansion was modelled as individuals leaving suitable to less suitable areas along this gradient. This model was chosen because it captures the hypothesized spatial dynamics of mountain pine beetle range expansion. The current outbreak has expanded north- and eastward, where mean temperatures are generally lower, which is expected to reduce larvae winter survival rates (Bentz et al. 2010; Cullingham et al. 2011; Safranyik and Linton 1991). Contrary to these expectations, the mountain pine beetle has continued to spread into novel areas, and it has been suggested that rapid adaptation may have played an important role (Janes et al. 2014).

Use of this selection model within our simulation framework resulted in an overly conservative threshold early in the simulation and an overly liberal threshold later in the simulation. This variation is due to the fact that the threshold adaptive reference slopes also vary through time (Fig. 3). The dynamic pattern of FDR of outlier loci is a direct consequence of the interacting

spatio-temporal dynamics of both the neutral loci and the adaptive locus that is used to classify them. To our knowledge, this is the first exposition of the role of transient dynamics in identifying loci potentially under selection during range expansion.

Our results show that early in simulations (i.e., 10–20 generations; Fig. 3) individuals do not experience much of the selection gradient because selection depends directly on latitude. That is, when expansion begins, very little geographic space (i.e., latitude) is occupied and a limited range of the selection gradient is accessible which produces adaptive slopes that are higher relative to those observed later in simulations. These transient dynamics may result in an underestimation of the number of true adaptive loci early during expansion. In contrast, our selection model results in an overestimation of the number of true adaptive loci later in the simulation due to fixation of the adaptive reference locus (Figs. 3, 4). In this simulation experiment, we were limited in our use of a single adaptive locus. In empirical analyses, known adaptive loci that become fixed would not be used to distinguish between neutral and putatively adaptive loci.

This model of selection is common in studies of range expansion or adaptive allele spread. Slatkin (1976) provided one of the first analyses of how an adaptive allele spreads without assuming a genetic trade-off and used one advantageous and one neutral allele. Burton and Travis (2008) also used a non-trade-off scenario to demonstrate that a maladapted allele could still spread during range expansion. Recently, Creech et al. (2017) used a simulation approach to model the spread of an advantageous allele in bighorn sheep employing a selection system with a locus with one advantageous allele and one neutral allele. Similar to our research, these studies focused on shorter, ecological time scales for which a selection model without trade-offs was appropriate.

Although there can be opposite selective advantages for two alleles along an environmental gradient, a strong genetic trade-off is not always the case. Indeed, a newly mutated allele may be advantageous everywhere in a landscape. For example, Gompert et al. (2015) found that most alleles affecting the survival of butterflies on one host had little to no effect on survival on the other host. Agosta and Klemens (2009) showed that loci under selection for resource specialization did not present a situation where alleles were advantageous for one resource and disadvantageous for another in a moth species with multiple host plants.

While we acknowledge that simulating a selection scenario where an adaptive allele is selected against in parts of the range would produce long-term stable clines and would avoid fixation, we sought to model rapid range expansion, as seen in outbreaks, rather than expansion on longer time scales (e.g., post-glacial expansion). Therefore, a scenario without trade-offs and with fewer assumptions was considered the most suitable for our research question.

Future work should involve exploration of alternative models of selection and multiple adaptive reference loci. One model option would be antagonistic pleiotropy, which promotes balancing selection and would prevent allele fixation (Jones et al. 2013). Such a model could also apply to the mountain pine beetle system given the hypothesis that cold tolerant alleles might be less advantageous in historic well-suited habitat, perhaps due to other life-history trade-offs. However, such a model also would require additional parameters. Our choice of a simple selection model revealed that the rate of misidentification of putatively adaptive alleles can vary as the adaptive cline varies and suggests that the rate of incorrectly selecting true adaptive loci will also be affected (Fig. 3). The consequences of varying effective population sizes, different assumptions regarding dispersal, as well as non-continuous selection surfaces typical of heterogeneous landscapes should also be explored in the future (Gauffre et al. 2015; Ortego et al. 2012). These questions would also benefit from a simulation modelling approach as it provides an opportunity to explore and sample this complex parameter space in greater detail than would be practical or possible using field-based approaches.

Conclusion

Our goal was to examine the conditions in which neutral loci can mimic the spatial patterns of adaptive loci during range expansion. Although we used the mountain pine beetle as a motivating example, these findings can be applied to other expanding populations, including invasive species. We found that weak dispersal capacity, early timing of sampling relative to the onset of the expansion, and a weak strength of selection in a reference adaptive locus all have the potential to increase the false discovery rate of putatively adaptive loci. A priori understanding of the ecological and demographic context in which range expansion is occurring should be used to improve inferences regarding the evolutionary dynamics of range expansions. Nonetheless, our results illustrate that it may not be possible to distinguish between neutral and adaptive patterns at times. In particular, our study indicates that the probability of incorrectly identifying adaptive loci changes over the course of an expansion and may also depend on the dispersal ability of the species in question.

Although greater knowledge of expansion history and dispersal will aid in the interpretation of potentially naive landscape genomic analyses, further work is still required to (i) fully test the performance of existing statistical tools in different demographic and range expansion contexts, and (ii) develop new tools to explicitly integrate these concepts. Until these questions are resolved, and new methods become available, great care

must be taken when searching for adaptive loci using landscape genomics in systems undergoing range expansion.

Acknowledgements

References

Agosta, S., and Klemens, J. 2009. Resource specialization in a phytophagous insect: no evidence for genetically based performance trade-offs across hosts in the field or laboratory. J. Evol. Biol. **22**(4): 907–912. doi:10.1111/j.1420-9101.2009.01694.x. PMID:19220649.

Batista, P.D., Janes, J.K., Boone, C.K., Murray, B.W., and Sperling, F.A. 2016. Adaptive and neutral markers both show continent-wide population structure of mountain pine beetle (*Dendroctonus ponderosae*). Ecol. Evol. **6**(17): 6292–6300. doi:10.1002/ece3.2367. PMID:27648243.

Bentz, B.J., Régnière, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A., et al. 2010. Climate change and bark beetles of the western United States and Canada: direct and indirect effects. Bioscience, **60**(8): 602–613. doi:10.1525/bio.2010.60.8.6.

Bialozyt, R., Ziegenhagen, B., and Petit, R.J. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. J. Evol. Biol. **19**: 12–20. doi:10.1111/j.1420-9101.2005.00995.x. PMID:16405572.

Bone, C., White, J.C., Wulder, M.A., Robertson, C., and Nelson, T.A. 2013. Impact of forest fragmentation on patterns of mountain pine beetle-caused tree mortality. Forests, **4**: 279–295. doi:10.3390/f4020279.

Brown, G.P., Phillips, B.L., and Shine, R. 2015. Directional dispersal has not evolved during the cane toad invasion. Funct. Ecol. **29**: 830–838. doi:10.1111/1365-2435.12397.

Burton, O.J., and Travis, J.M. 2008. The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing. Genetics, **179**(2): 941–950. doi:10.1534/genetics.108.087890. PMID:18505864.

Chen, H., and Walton, A. 2011. Mountain pine beetle dispersal: spatiotemporal patterns and role in the spread and expansion of the present outbreak. Ecosphere, **2**(6): art66. doi:10.1890/ES10-00172.1.

Colautti, R.I., and Barrett, S.C. 2013. Rapid adaptation to climate facilitates range expansion of an invasive plant. Science, **342**(6156): 364–366. doi:10.1126/science.1242121. PMID:24136968.

Cooke, B.J. 2009. Forecasting mountain pine beetle-overwintering mortality in a variable environment. Mountain Pine Beetle Working Paper 2009-03. Natural Resources Canada. Canadian Forest Service, Pacific Forestry Centre.

Coop, G., Witonsky, D., Di Rienzo, A., and Pritchard, J.K. 2010. Using environmental correlations to identify loci underlying local adaptation. Genetics, **185**: 1411–1423. doi:10.1534/genetics.110.114819. PMID:20516501.

Creech, T.G., Epps, C.W., Landguth, E.L., Wehausen, J.D., Crowhurst, R.S., Holton, B., and Monello, R.J. 2017. Simulating the spread of selection-driven genotypes using landscape resistance models for desert bighorn sheep. PLoS ONE, **12**(5): e0176960. doi:10.1371/journal.pone.0176960. PMID:28464013.

Cullingham, C.I., Cooke, J.E.K., Dang, S., Davis, C.S., Cooke, B.J., and Coltman, D.W. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. Mol. Ecol. **20**: 2157–2171. doi:10.1111/j.1365-294X.2011.05086.x. PMID:21457381.

de la Giroday, H.-M.C., Carroll, A.L., Lindgren, B.S., and Aukema, B.H. 2011. Incoming! Association of landscape features with dispersing mountain pine beetle populations during a range expansion event in western Canada. Landsc. Ecol. **26**: 1097–1110. doi:10.1007/s10980-011-9628-9.

De Villemereuil, P., Frichot, E., Bazin, E., François, O., and Gaggiotti, O. 2014. Genome scan methods against more complex models: when and how much should we trust them? Mol. Ecol. **23**: 2006–2019. doi:10.1111/mec.12705. PMID:24611968.

Dray, S., and Dufour, A.-B. 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. **22**(4): 1–20. doi:10.18637/jss.v022.i04.

Evenden, M., Whitehouse, C., and Sykes, J. 2014. Factors influencing flight capacity of the mountain pine beetle (Coleoptera: Curculionidae: Scolytinae). Environ. Entomol. **43**(1): 187–196. doi:10.1603/EN13244. PMID:24367930.

Excoffier, L., and Ray, N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. Trends Ecol. Evol. **23**(7): 347–351. doi:10.1016/j.tree.2008.04.004. PMID:18502536.

Excoffier, L., Foll, M., and Petit, R.J. 2009. Genetic consequences of range expansions. Annu. Rev. Ecol. Evol. Syst. **40**: 481–501. doi:10.1146/annurev.ecolsys.39.110707.173414.

Forester, B.R., Jones, M.R., Joost, S., Landguth, E.L., and Lasky, J.R. 2016. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. Mol. Ecol. **25**: 104–120. doi:10.1111/mec.13476. PMID:26576498.

François, O., Martins, H., Caye, K., and Schoville, S.D. 2016. Controlling false discoveries in genome scans for selection. Mol. Ecol. **25**: 454–469. doi:10.1111/mec.13513. PMID:26671840.

Frichot, E., Schoville, S.D., Bouchard, G., and François, O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol. Biol. Evol. **30**(7): 1687–1699. doi:10.1093/molbev/mst063. PMID:23543094.

Gauffre, B., Mallez, S., Chapuis, M.-P., Leblois, R., Litrico, I., Delaunay, S., and Badenhausser, I. 2015. Spatial heterogeneity in landscape structure influences dispersal and genetic structure: empirical evidence from a grasshopper in an agricultural landscape. Mol. Ecol. **24**: 1713–1728. doi:10.1111/mec.13152. PMID:25773398.

Gompert, Z., Jahner, J.P., Scholl, C.F., Wilson, J.S., Lucas, L.K., Soria-Carrasco, V., et al. 2015. The evolution of novel host use is unlikely to be constrained by trade-offs or a lack of genetic variation. Mol. Ecol. **24**(11): 2777–2793. doi:10.1111/mec.13199. PMID:25877787.

Goudet, J., and Jombart, T. 2015. hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. Available from https://CRAN.R-project.org/package=hierfstat.

Günther, T., and Coop, G. 2013. Robust identification of local adaptation from allele frequencies. Genetics, **195**: 205–220. doi:10.1534/genetics.113.152462. PMID:23821598.

Hardy, G.H. 1908. Mendelian proportions in a mixed population. Science, **28**(706): 49–50. doi:10.1126/science.28.706.49. PMID:17779291.

Harvey, B.J., Donato, D.C., and Turner, M.G. 2014. Recent mountain pine beetle outbreaks, wildfire severity, and postfire tree

regeneration in the US Northern Rockies. Proc. Natl. Acad. Sci. **111**(42): 15120–15125. doi:10.1073/pnas.1411346111. PMID: 25267633.

Hassall, C., Thompson, D.J., and Harvey, I.F. 2009. Variation in morphology between core and marginal populations of three British damselflies. Aquat. Insects, **31**(3): 187–197. doi:10.1080/01650420902776708.

Holderegger, R., Herrmann, D., Poncet, B.N., Gugerli, F., Thuiller, W., Taberlet, P., et al. 2008. Land ahead: Using genome scans to identify molecular markers of adaptive relevance. Plant Ecol. Divers. **1**(2): 273–283. doi:10.1080/17550870802338420.

James, P.M.A., Janes, J.K., Roe, A.D., and Cooke, B.J. 2016. Modeling landscape-level spatial variation in sex ratio skew in the mountain pine beetle (Coleoptera: Curculionidae). Environ. Entomol. **45**(4): 790–801. doi:10.1093/ee/nvw048. PMID: 27209334.

Janes, J.K., Li, Y., Keeling, C.I., Yuen, M.M.S., Boone, C.K., Cooke, J.E.K., et al. 2014. How the mountain pine beetle (*Dendroctonus ponderosae*) breached the Canadian Rocky Mountains. Mol. Biol. Evol. **31**(7): 1803–1815. doi:10.1093/molbev/msu135. PMID:24803641.

Jones, M.R., Forester, B.R., Teufel, A.I., Adams, R.V., Anstett, D.N., Goodrich, B.A., et al. 2013. Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. Evolution, **67**(12): 3455–3468. doi:10.1111/evo.12237. PMID:24299400.

Joost, S., Bonin, A., Bruford, W., Després, L., Conord, C., Erhardt, G., and Taberlet, P. 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol. Ecol. **16**: 3955–3969. doi:10.1111/j.1365-294X.2007.03442.x. PMID:17850556.

Keller, I., Alexander, J.M., Holderegger, R., and Edwards, P.J. 2013. Widespread phenotypic and genetic divergence along altitudinal gradients in animals. J. Evol. Biol. **26**: 2527–2543. doi:10.1111/jeb.12255. PMID:24128377.

Klopfstein, S., Currat, M., and Excoffier, L. 2006. The fate of mutations surfing on the wave of a range expansion. Mol. Biol. Evol. **23**(3): 482–490. doi:10.1093/molbev/msj057. PMID: 16280540.

Kot, M., Lewis, M.A., and Van den Driessche, P. 1996. Dispersal data and the spread of invading organisms. Ecology, **77**(7): 2027–2042. doi:10.2307/2265698.

Kurz, W.A., Stinson, G., Rampley, G.J., Dymond, C.C., and Neilson, E.T. 2008. Risk of natural disturbances makes future contribution of Canada's forests to the global carbon cycle highly uncertain. Proc. Natl. Acad. Sci. U.S.A. **105**(5): 1551–1555. doi:10.1073/pnas.0708133105. PMID:18230736.

Landguth, E.L., Bearlin, A., Day, C., and Dunham, J. 2016. CDMetaPOP: an eco-evolutionary metapopulation simulation model for population viability analysis in landscape and riverscape genetics. v1.0.

Lombaert, E., Estoup, A., Facon, B., Joubard, B., Grégoire, J.-C., Jannin, A., et al. 2014. Rapid increase in dispersal during range expansion in the invasive ladybird *Harmonia axyridis*. J. Evol. Biol. **27**: 508–517. doi:10.1111/jeb.12316. PMID:24444045.

Lotterhos, K.E., and Whitlock, M.C. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol. Ecol. **24**: 1031–1046. doi:10.1111/mec.13100. PMID:25648189.

Nagylaki, T. 1976. The evolution of one-and two-locus systems. Genetics, **83**(3): 583–600. PMID:955405.

Nakagawa, S., and Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol. Rev. **82**: 591–605. doi:10.1111/j.1469-185X.2007.00027.x. PMID:17944619.

Nealis, V., and Cooke, B. 2014. Risk Assessment of the threat of mountain pine beetle to Canada's boreal and eastern pine forests. Canadian Council of Forest Ministers Forest Pest Working Group. Canadian Forest Service, Natural Resources Canada, Ottawa, Ont.

Neubert, M.G., and Caswell, H. 2000. Demography and dispersal: Calculation and sensitivity analysis of invasion speed for structured populations. Ecology, **81**(6): 1613–1628. doi:10.1890/0012-9658(2000)081[1613:DADCAS]2.0.CO;2.

Ortego, J., Riordan, E.C., Gugger, P.F., and Sork, V.L. 2012. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. Mol. Ecol. **21**: 3210–3223. doi:10.1111/j.1365-294X.2012.05591.x. PMID: 22548448.

Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.-C., et al. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science, **355**(6332): eaai9214. doi:10.1126/science.aai9214. PMID: 28360268.

Pedersen, L. 2003. How serious is the mountain pine beetle problem? From a timber supply perspective. *In* Proceedings of the Mountain Pine Beetle Symposium: Challenges and Solutions, pp. 10–20.

Peter, B.M., and Slatkin, M. 2015. The effective founder effect in a spatially expanding population. Evolution, **69**(3): 721–734. doi:10.1111/evo.12609. PMID:25656983.

Polechová, J., and Barton, N.H. 2015. Limits to adaptation along environmental gradients. Proc. Natl. Acad. Sci. U.S.A. **112**(20): 6401–6406. doi:10.1073/pnas.1421515112. PMID:25941385.

Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G., and Romme, W.H. 2008. Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. BioScience, **58**(6): 501–517. doi:10.1641/B580607.

Raffa, K.F., Powell, E.N., and Townsend, P.A. 2013. Temperature-driven range expansion of an irruptive insect heightened by weakly coevolved plant defenses. Proc. Natl. Acad. Sci. U.S.A. **110**(6): 2193–2198. doi:10.1073/pnas.1216666110. PMID:23277541.

Robertson, C., Nelson, T.A., Jelinski, D.E., Wulder, M.A., and Boots, B. 2009. Spatial-temporal analysis of species range expansion: the case of the mountain pine beetle, *Dendroctonus ponderosae*. J. Biogeogr. **36**: 1446–1458. doi:10.1111/j.1365-2699.2009.02100.x.

Safranyik, L., and Linton, D.A. 1991. Unseasonably low fall and winter temperatures affecting mountain pine beetle and pine engraver beetle populations and damage in the British Columbia Chilcotin Region. J. Entomol. Soc. B.C. **88**: 17–21.

Safranyik, L., and Wilson, B. 2006. The mountain pine beetle: a synthesis of biology, management, and impacts on lodgepole pine. Natural Resources Canada, 304.

Safranyik, L., Linton, D.A., Silversides, R., and McMullen, L.H. 1992. Dispersal of released mountain pine beetles under the canopy of a mature lodgepole pine stand. J. Appl. Entomol. **113**: 441–450. doi:10.1111/j.1439-0418.1992.tb00687.x.

Safranyik, L., Carroll, A.L., Régnière, J., Langor, D.W., Riel, W.G., Shore, T.L., et al. 2010. Potential for range expansion of mountain pine beetle into the boreal forest of North America. Can. Entomol. **142**(5): 415–442. doi:10.4039/n08-CPA01.

Simard, M., Romme, W.H., Griffin, J.M., and Turner, M.G. 2011. Do mountain pine beetle outbreaks change the probability of active crown fire in lodgepole pine forests? Ecol. Monogr. **81**(1): 3–24. doi:10.1890/10-1176.1.

Slatkin, M. 1976. The rate of spread of an advantageous allele in a subdivided population. *In* Population genetics and ecology. Elsevier. pp. 767–780.

Slatkin, M., and Excoffier, L. 2012. Serial founder effects during range expansion: a spatial analog of genetic drift. Genetics, **191**: 171–181. doi:10.1534/genetics.112.139022. PMID:22367031.

Storey, J.D., and Tibshirani, R. 2003. Statistical significance for

genomewide studies. Proc. Natl. Acad. Sci. U.S.A. **100**(16): 9440–9445. doi:10.1073/pnas.1530509100. PMID:12883005.

Swaegers, J., Mergeay, J., van Geystelen, A., Therry, L., Larmuseau, M.H.D., and Stoks, R. 2015. Neutral and adaptive genomic signatures of rapid poleward range expansion. Mol. Ecol. **24**: 6163–6176. doi:10.1111/mec.13462. PMID:26561985.

USDA Forest Service. 2012. Areas with tree mortality from bark beetles: Summary for 2000–2011, Western US. 3 pp.

Vasemägi, A. 2006. The adaptive hypothesis of clinal variation revisited: single-locus clines as a result of spatially restricted gene flow. Genetics, **173**: 2411–2414. doi:10.1534/genetics.106.059881. PMID:16849603.

Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution, **38**(6): 1358–1370. doi:10.2307/2408641. PMID:28563791.

Yeaman, S., and Whitlock, M.C. 2011. The genetic architecture of adaptation under migration-selection balance. Evolution, **65**(7): 1897–1911. doi:10.1111/j.1558-5646.2011.01269.x. PMID:21729046.