**Identifying genetically exceptional populations through space and time**

Julian Wittische, Pierre Legendre, Patrick M. A. James

Département de Sciences Biologiques, CP 6128 Succursale Centre-Ville, Université de Montréal, Pavillon Marie-Victorin, Montréal, QC, Canada, H3C 3J7

Correspondence: Julian Wittische; E-mail: [jwittische@gmail.com](mailto:jwittische@gmail.com)

Running title: Spatio-temporal genetic uniqueness

INTRODUCTION

*Finding out when and where disturbance happens*

One of the main challenges for biologists and managers is to detect and predict where and when extraneous disturbance events influence the ecological dynamics and the evolution of species. Changes in diversity, e.g. genetic diversity, can be the result of natural or anthropogenic disturbance at any spatial or temporal scale, from a local and abrupt change like a wildfire to a global and gradual change like climate warming. However, it is rarely possible to observe the effects of these events instantaneously and researchers are often left with their legacies which may be cryptic. Populations genetics can be used to uncover the magnitude of a change in genetic diversity.

*Why population genetics can help?*

Population genetics is the scientific study of genetic variation within populations and seeks to model changes in the frequencies of alleles over space and time (REF). When a disturbance does not constitute a selective pressure, alleles are randomly transferred from a generation to the next and genetic drift happens leading to a loss of diversity. Common examples of situations where genetic drift occurs include geographic isolation and population bottleneck (REF). Massive migrations from previously isolated populations may also affect the gene pool (REF). The result of such events in a local population tend to increase the genetic distance of this population with surrounding populations (REF). This can constitute a clue for population geneticists that a disturbance event happened.

*A lack of specific tools*

While the development of tools to identify aberrant loci (REFs LFMM, PCAdapt...) or classify population samples in genetically coherent clusters (REFs DAPC, STRUCTURE) is thriving, there are few options offering to test whether a sample is truly different from others. Ordination (e.g. PCA) offers clues as to which samples are different and is very valuable as an exploratory technique. Ordination has been used in classification tools such as DAPC (REF) which as it seeks to group samples may indicate which samples are different although the focus is not on finding singularly outlying samples. Furthermore, the relevance and performance of DAPC on temporal datasets, where the objective is to find which population has indeed changed more significantly than others in the landscape, has not evaluated.

*A glimmer of hope brought by community ecology*

Two statistical inference methods have recently been proposed and tested for the analysis of spatial-temporal changes in community composition. Namely they are Local Contributions to Beta Diversity (LCBD; Legendre, P., & De Cáceres 2013) for the spatial question of whether there are sites where the difference in community composition between samples seems exceptionally large and Temporal Beta diversity Indices (TBI; Legendre 2019) for the temporal question of whether there are sites where the difference in community composition between survey times seems exceptionally large. Legendre (2019) suggested that comparing genetic data at two different dates separated by a known event may indicate the locations where the event had strong effects, or that in the absence of an *a priori* event these sites should be investigated to learn about what caused the exceptional genetic change. Using genetic data to shed light on the causes and/or consequences of a local genetic change can also be done for LCBD.

*What this paper is about*

Our paper seeks to describe how to find out find out which parts of a landscape have undergone atypical and substantial genetic change after a disturbance event. We simulated two scenarios where part of the landscape is affected by non-selective demographic changes mimicking the effects of common disturbance events. We then used TBI to measure changes in the gene pool of our subpopulations and used a permutation-based approach to distinguish exceptionally different sites.

METHODS

*Simulation experiments*

a) Simulation parameters

We modeled the effects of disturbance on the genetic diversity using the spatially-explicit gene flow simulation software CDMetaPOP (Landguth et al. 2017). CDMetaPOP simulates dispersal and mating of individuals across a landscape, and allows to define the initial genetic structure, spatial distribution of individuals, dispersal characteristics, and life history traits of the population. For each scenario we simulated 1000 replicates, with 25 interconnected populations, 100 bi-allelic loci and maximum carrying capacity of 100 individuals per population. X% (5 vs 50 atm) of individuals within a population may migrate at each generation. 220 generations. Dispersal kernel equation.

b) Simulation scenarios

The first scenario involves modelling a massive extraneous migration from a previously isolated 26th population. This population was simulated during the same number of generations and the cost distance between the 13th (central) and the 26th (isolated) populations is set to 0 between the 200th and 201st generations, mimicking a mass migration event between the two. The 26th population is then isolated again by resetting the cost distance to an unreachable number.

The second scenario involves modelling a demographic bottleneck through massive mortality. To do that, the carrying capacity of the 13th population (central), was set to 10% of its original value between the 200th and 201st generations.

Massive extraneous migration/Bottleneck

*Genetic dissimilarity*

Chord distance has been commonly used in both community ecology (Orlóci 1967; Legendre & Borcard 2018) and population genetics (Cavalli-Sforza & Edwards 1967; Balkenhol et al. 2016). We chose chord distance because it has already been tested for use with TBI with non-genetic data (Legendre 2019) and because it may be more appropriate than other indices of genetic dissimilarity when most of the variation among populations is due to recent changes (Takezaki & Nei 1996; Kalinowski 2002) as it does not assume populations are in drift-mutation equilibrium.

*Estimating type I and type II errors*

a) Permutation approach

b) Two-step criterion

RESULTS

DISCUSSION

A few systems are consistently monitored through time (e.g. REFs LTER...) and/or exhaustively sampled in space (e.g. REFs Fushan...).

REFERENCES