Detecting temporal changes in genetic diversity: a new tool for molecular ecology studies with repeated surveys

Julian Wittische1, Pierre Legendre1, Patrick M. A. James1,2

1 Département de Sciences Biologiques, Campus MIL, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, QC, Canada, H3C 3J7

2 Institute of Forestry and Conservation, University of Toronto, 33 Willcocks St., Toronto, ON, Canada, M5S 2J5

\* Corresponding author, E-mail: [jwittische@gmail.com](mailto:jwittische@gmail.com)

### Abstract

Understanding spatiotemporal changes in biodiversity, including genetic diversity, is essential to track the effects of global change and to inform effective conservation plans. Although temporal questions are common in community ecology, they are less often investigated in landscape genetics. Indeed, detecting changes in local genetic diversity due to demographic events that exceed what one would expect under neutral conditions (e.g., drift) is challenging. Our capacity to detect such changes is also information dependent. Existing methods to detect meaningful genetic changes through time typically require large datasets containing information beyond simple allele counts, such as mutational models and fine-scale population structure. However, when such extensive information is not available, methods are still needed to detect and understand temporal changes in genetic diversity induced by demographic events. In this paper, we describe Temporal Genetic Indices (TGI), a new method for identifying significant changes in genetic diversity through time. This method uses permutations of genotypic matrices to test the significance of genetic temporal change at sites, given neutral genetic change at other sampling sites in the study landscape. TGI overcomes existing challenges to detecting temporal change in genetic data with minimal information. We demonstrate the utility of TGI for identifying the genetic legacies of important historical demographic events using demo-genetic simulation. We further demonstrate the ability of our TGI approach to identify such legacies under different levels of dispersal, spatial extent of the demographic events, and the timing of sampling relative to the events. Finally, we successfully apply TGI to an empirical dataset, with our application providing a straightforward test for genetic change and supporting previous conclusions about the data. An R function to implement the method is now available, as well as utility functions for those wishing to further simulate and analyse their simulations.

**Keywords:** Time; Significance testing; Population genetics; Monitoring; R software

### 1. Introduction

Global biodiversity at the gene, species, population, and ecosystem scales is being altered at an increasing rate, with significant consequences for ecosystem functioning and the long-term viability of the biosphere (Bellard et al. 2012; Dirzo et al. 2014; Leigh et al. 2019). Given these global changes and the inherent temporal variability of biological systems, conservation biologists are increasingly recognizing that it is no longer sufficient to study spatial patterns in biodiversity at a single point in time. Instead, trends in biodiversity must be observed across both space and time (Allendorf et al. 2010; Bradburd & Ralph 2019; Fenderson et al. 2019). Important events in a population’s history can be detected using genetic data (Bradburd & Ralph 2019), however, novel techniques are needed to quantify and track such events through both space and time.

Spatiotemporal variation in genetic diversity can provide important insights into the connectivity and demographic history of populations (Draheim et al. 2018; Moraes et al. 2017). Indeed, population genetics has proven essential for translating observed genetic variation into meaningful inferences that can inform conservation efforts (Allendorf et al. 2010; Harrisson et al. 2014; Segelbacher et al. 2010), and the causes and consequences of temporal variation in genetic diversity are at the crux of many conservation and public health issues (Díez-del-Molino et al. 2018; Lauterjung et al. 2019; Moraes et al. 2017). Researchers commonly explore patterns in spatiotemporal population genetic data (Banks et al. 2013) to quantify isolation-by-distance (Rousset 1997; Wright 1943), time since population bottlenecks (Gattepaille et al. 2013; Maruyama & Fuerstt 1985), rates of migration between isolated populations (Bezemer et al. 2019; Buschbom et al. 2011), and the timing and extent of outbreak expansions (James et al. 2015; Larroque et al. 2019; Wittische et al. 2019).

However, new approaches are needed to detect meaningful temporal variation in genetic diversity. The significance of genetic changes may be derived from the comparison to an expected reference distribution, or a null model based on typical background processes (e.g., drift, local gene, flow, mutation). Detecting such meaningful temporal changes represents a first step in elucidating the processes that govern demographically dynamic systems such as those found during population outbreaks (Fisher & Garner 2020; Maynard et al. 2017), major weather events (Poff et al. 2018), species invasions (Mack et al. 2000), or other disturbances such as a wildfire (Suárez et al. 2012). Temporal genetic analyses could similarly identify which populations, among a set of previously sampled populations, received migrants from long-distance dispersal events (Apodaca et al. 2013). Because temporal genetic variation reflects the evolutionary potential of a population and the probability of its persistence (Aeschbacher et al. 2017; Bolnick & Nosil 2007; Kremer et al. 2012), relating temporal genetic variation to landscape change can provide important insights about the eco-evolutionary dynamics of a species, and be used to inform conservation strategies (e.g., Landguth et al. 2017b).

There are currently two general approaches for investigating temporal genetic variation. The first suite of approaches uses statistical models to infer demographic history from genetic data obtained at a single time point (Excoffier et al. 2013; Gutenkunst et al. 2009; Kamm et al. 2020). This approach is often computationally intensive, requires high-quality microsatellite or extensive SNP datasets. This approach also requires extensive knowledge of the biological system, including information on recombination processes (Gattepaille et al. 2013) and ascertainment bias (Albrechtsen et al. 2010; Clark et al. 2005; Marth et al. 2004). The second suite of approaches compares genetic diversity between samples taken from the same sites over time using any genetic markers through either qualitative comparison or statistical models, ideally using a null reference distribution. These repeated-survey approaches are more readily usable in systems where less information is available, such as non-model species. Repeated-survey approaches can also be used in systems that were sampled in the past, with a goal of comparing contemporary to historical patterns (Moraes et al. 2017).

Despite our ability to compare genetic diversity at two points in time, several technical and conceptual challenges remain. One such challenge is determining which metric to use to meaningfully quantify and detect temporal changes. Some studies have used genetic differentiation metrics such as Jost’s *D* or *F*ST or its analogues (Knight et al. 2018; Larroque et al. 2019; Segura-García et al. 2019) to evaluate temporal changes between genetic datasets. However, translating our spatial understanding of these genetic differentiation indices to the temporal dimension is not straightforward (Bhatia et al. 2013). An additional challenge for temporal genetic analyses is disentangling spatial from temporal effects, because the additivity of genetic drift means that genetic differentiation can be associated with both space and time (Murray et al. 2016; Skoglund et al. 2014). Finally, repeated-survey analyses remain challenging because we lack sufficiently developed tools to distinguish natural temporal variation in genetic structure due to demographically-induced drift, local gene flow, and mutation, from the changes caused by external forces. In other words, we need to develop an adequate distribution of reference patterns of genetic variation based on common background processes prior to deciding whether a change is meaningful.

Although its objectives differ from those of population genetics, the field of community ecology has a history of explicitly examining change in community composition through time. Temporal beta-diversity indices (TBI; Legendre 2019) are used to quantify and assess temporal changes in ecological community composition using a dissimilarity index calculated between samples taken at different times and at several sites. The significance of these dissimilarities is then tested using a permutational procedure. The TBI approach has effectively demonstrated temporal variation in simulated community composition (Legendre 2019) and in dozens of empirical datasets, but the potential of a TBI-inspired tool to detect meaningful temporal changes in genetic diversity has not yet been examined. Given the conceptual similarity in data structure between species diversity in multi-species community composition data, and genetic diversity in multilocus genotype data, we sought to determine how TBI could be modified to identify significant variation in spatiotemporal genotypic data.

In this paper, we propose and evaluate a method for extending the TBI framework to spatiotemporal population genetic data. Our new framework, which we call temporal genetic diversity indices (TGI), is designed to identify significant temporal variation in spatial genetic diversity using relatively information-poor genetic data while accounting for confounding forces such as drift, local gene flow, and mutation by creating a reference distribution. We demonstrate the effectiveness and applicability of the TGI approach using simulated genetic data, where each simulation combined multiple scenarios in which portions of a landscape were affected by a non-selective demographic change.

Specifically, we assess TGI’s capacity to detect significant temporal variation in genetic diversity in three demographic factors: 1) species dispersal ability; 2) the number of populations affected by a demographic event (i.e., spatial extent of the event); and 3) the time between consecutive surveys. We predict that TGI’s ability to detect temporal genetic changes will be lower in populations with higher dispersal capacity because of the homogenizing effect of higher gene flow. We also predict that our ability to detect changes will decrease as the time between successive sampling events increases. Beyond the method to simulate the null distribution, the selection of the cut-off value may also influence the performance of our permutational significance testing. Indeed, selecting a proper significance threshold for the *p*-value calculated from the TGI test permutations, and therefore defining which changes in genetic diversity are significant or not, is important for balancing selectivity and power. Therefore, we describe how this potential trade-off occurs along a range of thresholds and discuss its implications. Finally, we illustrate how TGI provides a functional testing framework by applying it to a real genetic dataset representing a large landscape with many populations of the Northern tidewater goby, a threatened species of fish found in western North America. A better understanding of the factors that influence changes in genetic diversity through time and improved techniques to monitor these changes is essential to describe and understand global biodiversity losses in the context of the current sixth mass extinction. That is why we built a tool to make meaningful assessments of the significance of temporal genetic diversity changes relative to an informed null model based on common background processes.

### 2. Methods

#### 2.1. Adapting TBI for genetic data

Temporal beta-diversity indices (TBI) are calculated by computing dissimilarities in species composition between data surveyed at two different times for all sampling sites. TBI give local measures of the change in community composition at each site; significance of these indices is then tested through simultaneous permutations of the two site-by-species input matrices. To extend TBI to TGI, we substituted community dissimilarities with genetic distances calculated from site-level allele frequencies in order to compare two different temporal surveys (see **Table 3‑1**.-A) for a simple example showing how we transformed two temporal surveys into a genetic distance, for a two-site landscape). The null hypothesis in this case is that genetic composition between the two time points does not differ more than would be expected due to background processes typical to the landscape.

Neutral spatial processes such as genetic drift can produce temporal variation in spatial genetic structure. In designing TGI, our goal was to determine how to identify temporal changes that are significantly different from what would be expected under a scenario with drift, local gene flow, and mutation. Because there are no reference criteria for the changes in genetic diversity that would constitute a significant temporal genetic change, we used a permutation-based approach to generate a distribution of genetic distances to which an observed genetic distance can be compared.

**Table 3‑1** **:** Examples of A) the computation of the original TGI values for a biallelic marker (the method works for microsatellites too!) and B) the way we permutated input genotypic matrices to create a distribution to test TGI significance. The numbers in the table represent the number of copies of an allele in a sampled population.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Locus Allele** | **1**  **A** | **1**  **B** | **2**  **A** | **2**  **B** | **3**  **A** | **3**  **B** | **4**  **A** | **4**  **B** |
| **FIRST**  **SURVEY T1** | **Pop. 1** | 4 | 0 | 4 | 0 | 4 | 0 | 3 | 1 |
| **Pop. 2** | 0 | 4 | 1 | 3 | 2 | 2 | 3 | 1 |
| **Pop. 3** | 1 | 3 | 1 | 3 | 0 | 4 | 0 | 4 |
| **Pop. 4** | 0 | 4 | 0 | 4 | 1 | 3 | 1 | 3 |
|  |  |  |  |  |  |  |  |  |  |
| **SECOND**  **SURVEY T2** | **Pop. 1** | 2 | 2 | 3 | 1 | 3 | 1 | 2 | 2 |
| **Pop. 2** | 3 | 1 | 1 | 3 | 1 | 3 | 2 | 2 |
| **Pop. 3** | 1 | 3 | 2 | 2 | 1 | 3 | 0 | 4 |
| **Pop. 4** | 0 | 4 | 0 | 4 | 2 | 2 | 1 | 3 |

**A)**

TGI value for Population 1 between T1 and T2 (see Supp. Mat. 3-1):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Locus Allele** | **1**  **A** | **1**  **B** | **2**  **A** | **2**  **B** | **3**  **A** | **3**  **B** | **4**  **A** | **4**  **B** |
| **Pop. 1** | 4 | 0 | 4 | 0 | 4 | 0 | 3 | 1 |
| **Pop. 2** | 0 | 4 | 1 | 3 | 2 | 2 | 3 | 1 |
| **Pop. 3** | 1 | 3 | 1 | 3 | 0 | 4 | 0 | 4 |
| **Pop. 4** | 0 | 4 | 0 | 4 | 1 | 3 | 1 | 3 |

**B)**

Permutation example: this is done for both surveys and repeated 999 times to create the test for TGI. Permutation index 2, 4, 1, 3.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Locus Allele** | **1**  **A** | **1**  **B** | **2**  **A** | **2**  **B** | **3**  **A** | **3**  **B** | **4**  **A** | **4**  **B** |
| **Pop. 1** | 0 | 4 | 1 | 3 | 2 | 2 | 3 | 1 |
| **Pop. 2** | 0 | 4 | 0 | 4 | 1 | 3 | 1 | 3 |
| **Pop. 3** | 4 | 0 | 4 | 0 | 4 | 0 | 3 | 1 |
| **Pop. 4** | 1 | 3 | 1 | 3 | 0 | 4 | 0 | 4 |

**Table 3‑2 :** Two-factor simulation experiment with scenario abbreviations used throughout the manuscript. The third column indicates the number of affected populations with spatiotemporal population genetic legacies; scenarios with 0 affected populations represent the control scenarios. We ran 180 unique simulations (replicates) for each combination of factor levels, which amounts to 2160 simulations in total.

|  |  |  |  |
| --- | --- | --- | --- |
| **Number** | **Dispersal** | **Affected populations** | **Code** |
| 1 | Low | 1 | L1 |
| 2 | Moderate | 1 | M1 |
| 3 | High | 1 | H1 |
| 4 | Low | 2 | L2 |
| 5 | Moderate | 2 | M2 |
| 6 | High | 2 | H2 |
| 7 | Low | 3 | L3 |
| 8 | Moderate | 3 | M3 |
| 9 | High | 3 | H3 |
| 10 | Low | 0 | CL |
| 11 | Moderate | 0 | CM |
| 12 | High | 0 | CH |
|  |  |  |  |

This randomly generated distribution of genetic distances allows us to approximate a large number of landscapes subject only to drift, local gene flow, and mutation. For each of the two input genotypic matrices representing the two temporally distinct surveys containing data about the same alleles, loci and sites, we permuted the allele counts at each locus (**Table 3‑1**.-B). Permutations were performed using the *poppr* R package (see *Software*) to maintain allelic structure and heterozygosity (Agapow & Burt 2001). We used 999 random permutations in all analyses. To be clear, we created an adequate null model for testing purposes by combining realistic simulations of drift, local gene flow, and mutation with a permutation which allowed us to shuffle allele counts between populations, to find out truly remarkable genetic distances. That way, the permutation procedure unlinks the genetic histories of different populations, which is akin to standardizing the recent effects of gene flow and removing the spatial autocorrelation in genetic drift.

#### 2.2. Genetic distance

Genetic distances between time points for a given location were calculated using Rogers’ genetic distance (Rogers 1972), which is similar to the Euclidean genetic distance (**Supp. Mat. 3‑1**). Rogers’ distance makes no assumptions about base-pair substitutions or time since separation and is therefore appropriate to study short-term dynamics influenced by environmental forces. One advantage is that Rogers’ distance is simplifiable for biallelic markers and is related to the well-known Euclidean distance (**Supp. Mat. 3‑1**, **Supp. Mat. 3‑2**). Rogers’ distance has been used previously to investigate variation in genetic diversity in several taxa (Bennett & Stone 2019; Pereira et al. 2018). We computed Rogers’ distance using the *dist.genpop* function from the *adegenet* R package (see *Software*). While we chose to illustrate TGI using Rogers’ distance, our approach and functions may also be used with other true distance metrics (e.g., angular distance/Edward’s distance) and measures of differentiation (e.g., *F*ST), although testing should ensure their relevance (see *Discussion*).

#### 2.3. Simulation framework

To simulate population genetic changes through time and test the performance of TGI, we used the spatially-explicit gene flow simulator *CDMetaPOP* (Landguth et al. 2017a). *CDMetaPOP* simulates dispersal and mating of individuals across a landscape and allows the user to define the initial genetic structure, spatial distribution of individuals, dispersal characteristics, and life history traits of the populations. The physical landscape we simulated was a homogeneous, interconnected 5 × 5 square grid with each of the 25 cells representing a population. Each population had a maximum carrying capacity of 50 individuals; the populated landscape therefore contained a maximum of 1,250 (25 × 50) individuals. Distance between populations was set as the Euclidean geographic distance. The genotypic information of each individual consisted of 100 neutral, unlinked, bi-allelic single nucleotide polymorphism (SNP) loci. We included mutation to increase the realism of our simulations, because mutation is a common source of temporal genetic variation. The nuclear mutation rate (per base pair per generation) was set at 10-8 (Allio et al. 2017). Although this empirical estimate seems rather low, the total number of simulated loci over the simulation study far exceeds it and we, therefore, expect mutation events in some of our simulations.

We investigated the influence of a single demographic event on the spatial-temporal apportionment of genetic variation. Specifically, we tested the capacity of our TGI approach to detect demographic and genetic changes due to immigration from an isolated population. This isolated population was separate from our 5 × 5 grid (i.e., population #26). This independent source population shared the same attributes as other populations in our simulated landscape. During simulated demographic events, individuals from population #26 were only allowed to disperse into the 5 × 5 simulation grid.

Using *CDMetaPOP*, we examined the influence of dispersal (movement among our 25 populations) and the spatial extent of a demographic event (number of populations that received immigrants from population #26) on the persistence of spatial genetic legacies. Here, persistence of genetic legacies was defined as our ability to identify them using TGI (see *Section 2.6*). We examined three levels of dispersal capacity (described below) and three different numbers of affected populations (1, 2, or 3) for a total of nine unique scenarios (**Table 3‑2**). Each scenario was then simulated 180 times, for a total of 1,620 (9 × 180) unique replicates for this experiment, excluding the control simulations (**Table 3‑2**).

For each replicate, we initialized the simulation by randomly distributing alleles among individuals, therefore approaching the maximum (*p*=*q*=0.5 maximizes *p×q*) genetic diversity (Landguth et al. 2017a). Those parameters were chosen as a compromise between realistic allele distributions and computational limitations and were appropriate for producing simulated genetic data that could reasonably recreate the complex evolutionary dynamics in real populations. Each simulation was run for 100 generations before the demographic event was imposed on up to three populations in the landscape. Ten additional generations were simulated after the event. Sampling was performed up to 9 generations before and after the event.

#### 2.4. Dispersal

Dispersal was modelled using the weighted geographic distance between populations using a power law function, , where *B* represents the difficulty of dispersal. High values of *B* correspond to low dispersal capacity. Within a simulation run, *B* was constant while we randomly picked distances at which individuals dispersed based on the power law function. To do so, we rescaled the values of all distances in the landscape, using the maximum and the minimum (0) distances possible in this virtual landscape, as described in the *CDMetaPOP* (Landguth et al. 2017a) user’s manual (p. 63). This produced values in the [0,1] range. Rescaled values were considered to represent probabilities that an individual disperses to a cell located at that distance (**Table 3‑1**). We chose this way of modelling dispersal to allow for both within-population movement and landscape-wide dispersal (Mayrand et al. 2019).

Chart

Description automatically generated

**Figure 3‑1 :** Probability of dispersal of an individual as a function of geographic distance, in three different dispersal scenarios.

The population to which an individual dispersed was selected randomly from the set of populations available at the distance which was randomly picked following the probability distribution described above. We set the landscape boundary conditions to impermeable and individuals, therefore, always stayed within our simulated landscape. Any individual could disperse to any one of the 25 populations at each generation with probability of dispersal decreasing with distance. To investigate the effect of different levels of dispersal, we ran separate simulations using three different values of *B*: low (*B* = 2), moderate (*B* = 1.301) and high (*B* = 0.6015) dispersal capacity (**Figure 3‑1**; **Table 3‑2**).

#### 2.5. Spatial extent

We also evaluated how the spatial extent of the simulated immigration event affected the performance of our TGI method. To do this, we allowed individuals from population #26 to immigrate into one, two, or three populations that were randomly selected from the original 25. We varied the position of where the demographic event occurred in the landscape among simulation replicates because deme topology may influence the outcomes of population genetic analyses (Robledo-Arnuncio & Rousset 2010). For scenarios in which only one population was affected, we partitioned the 180 replicate simulations equally among six populations in the landscape. Because our landscape is square and homogeneously resistant to movement, it is symmetric and there are only six unique positions. Systematically choosing those six positions as a pool of potential targets for our demographic event therefore covered all possible spatial patterns in our simulations. One or more populations among these six populations were randomly selected once and were identical across runs. When multiple (two or three) populations underwent a demographic event, we randomly chose one of these six geographically unique populations and randomly picked one or two additional populations directly adjacent to it. We chose to pick adjacent populations to respect the spatial autocorrelation often exhibited in demographic events. For each of the two- and three-population simulations, we repeated this population selection procedure six times and ran 30 replicate simulations for each set of populations.

#### 2.6. Statistical performance

We assessed the statistical performance of our TGI testing procedure using the false positive rate (FPR) and false negative rate (FNR). In our study, a false positive was a population that we knew did not undergo the demographic change we imposed, but was found to have done so using the TGI test, whereas a false negative was a population that experienced the demographic event but exhibited no significant change according to the TGI test. The FPR is expressed as the ratio of false positives to the total number of negative tests (i.e., true negatives and false positives), and the FNR is expressed as the ratio of false negatives to the total number of positive tests (i.e., true positives and false negatives).

A high FPR would indicate that our TGI measure often selected the wrong population(s) as having changed significantly and that our testing procedure was less selective. Researchers generally want to minimize the FPR when there are, for example, limited resources available for conservation efforts. In contrast, a high FNR would mean that we often failed to identify the population(s) that were actually affected, and that our testing procedure had low discriminatory power. Researchers may want to minimize the FNR in situations where finding all affected populations is the most important aspect, for example, if there is limited time to take conservation action.

TGI’s performance may vary with cut-off values, and a trade-off between power and selectivity is expected to occur. To characterize this compromise, we evaluated the statistical performance of TGI using a range of significance thresholds for calculating FPR and FNR: 0.001, 0.005, 0.01, 0.015, 0.020, 0.025, 0.030, 0.035, 0.040, 0.045, 0.050, 0.055, 0.060, 0.065, 0.070, 0.075, 0.080, 0.085, 0.090, 0.095, and 0.1. We did not conduct corrections for multiple tests (e.g., Bonferroni) because in our simulation framework, use of such a test would simply translate into another range of thresholds. However, choosing to correct for multiple tests is possible in our function. Note that 0.001 is the lowest possible value when using 999 permutations. A compromising threshold value for the trade-off, as evaluated through this question, was used in the analyses focussed on time lag.

#### 2.7. Time

We sought to assess how the time since the simulated demographic event influences the performance of the TGI method in each of our nine dispersal/spatial extent scenarios (**Table 3‑2**). To do so, we calculated the TGI for simulated data collected up to nine generations before and after the event and compared it to the TGI calculated from data collected immediately before or after the event generation. We chose nine generations as the maximum time between samplings (nine generations before the event, or nine generations after the event) because this timeline is longer than most “before/after” population genetic studies in the literature (e.g., Bezault et al. 2011; Kinziger et al. 2015; Moraes et al. 2017) and most long-term ecological research programs monitor at a shorter time interval (e.g., Hobbie et al. 2003; Knapp et al. 2012; Kuemmerlen et al. 2016). Comparisons between TGI results were based on the FPR and FNR calculated at a significance threshold of *p* ≤ 0.05, as this threshold was a good compromise between different performance metrics as indicated by the results to the trade-off analysis described above.

#### 2.8. Controls

Control simulations were run in which populations were only affected by local gene flow, drift, and mutation. No demographic events were included. Dispersal was the only parameter that varied among the control simulations, resulting in three control scenarios (**Table 3‑2**). We only evaluated the FPR of these control scenarios; because there were no true positives or false negatives for populations affected by the demographic event, the FNR was always equal to 0. The performance of experimental scenarios was always compared to the control scenario with the same dispersal capacity.

#### 2.9. Software

*CDMetaPOP* runs on *Python 2.7* (Landguth et al. 2017a). We used the *R* software (R Core Team 2019) in the *RStudio* IDE (RStudio Team 2018) for all analyses and illustrations. We used the *adegenet* (Jombart 2008; Jombart & Ahmed 2011), *pegas* (Paradis 2010), *poppr* (Kamvar et al. 2014, 2015) and *adespatial* (Dray et al. 2019) *R* packages for calculations. Our *TGI* function is available in the supplementary material (**Supp. Mat. 3‑2**) as an *R* script. The function enables the user to choose among five different genetic distance metrics for the calculations of TGI indices via the “*method”* argument.

#### 2.10. Applied example: an endangered fish

To demonstrate that our TGI measure provides valuable information about temporal change in a real system with conservation implications, we applied it to real genetic data from a study of a threatened vertebrate, the Northern tidewater goby (Kinziger et al. 2015). We chose this example because it uses a different type of genetic data than we used for our simulations, thus demonstrating that TGI is applicable to a variety of genetic markers. In addition, the study authors suggested that one goby population had undergone more genetic change than the other, more stable local populations, allowing us to test a real hypothesis and go beyond a simple illustration of our method (Kinziger et al. 2015). The dataset was downloaded from DRYAD (doi: 10.5061/dryad.871db). Significance testing of temporal change in these data used 9,999 permutations.

### 3. Results

We were able to translate the TBI framework to TGI by adapting it to the specific structure of genetic data. Although our results in the present section support the general efficacy of TGI and warrant its use on empirical datasets, the performance of the TGI approach was sensitive to dispersal capacity, spatial extent of the demographic event producing the genetic change, and the time difference between surveys and the demographic event.

Experimental FPR values were consistently lower than control FPR values, regardless of dispersal parameters (**Figure 3‑2**). This suggests that when using TGI, we are less likely to misidentify a significant genetic change in presence of an actual demographic, than we are in its absence. Among the control simulations, runs with higher dispersal capacity had a lower FPR (**Figure 3‑2Error! Reference source not found.**). Control FPR values were generally at least twice as high as the maximum experimental FPR values encountered (L1, M1), regardless of the significance threshold used. This means that, even for the lowest-performing scenarios in our simulations, TGI was much more effective at avoiding false positives in the presence of an event than in the absence of one.

#### 3.1. Dispersal

Dispersal capacity influenced our ability to detect temporal changes in genetic diversity, as the FNR generally increased with dispersal capacity (**Figure 3‑3**). However, only one scenario (H3; **Table 3‑1**) exhibited FNR values above a very conservative limit of 1%, regardless of the *p*-value threshold used (**Figure 3‑3**). Of the four scenarios that did not achieve an average FNR of 0 (L3, M3, H2, and H3), two involved high dispersal. When we averaged the FNR values calculated at the traditional p ≤ 0.05 threshold across scenarios sharing the same dispersal parameters (e.g.,averaging FNR value for L1, L2, and L3 grouped together), the mean FNRs were 0.0037 (0.0007 – 0.0066; 95% confidence interval [CI]) for low dispersal, 0.0049 (0.0015 – 0.0083; 95% CI) for moderate dispersal, and 0.0108 (0.0055 – 0.0161; 95% CI) for high dispersal.

Chart, scatter chart

Description automatically generated

**Figure 3‑2** **:**FPR across all threshold and scenarios. Control experiments are shown with dashed lines. Those values are for samplings done at the 100 and 101 generations 100 and 101, i.e., right before and after the migration event. 95% confidence intervals of the FPR estimates are displayed by vertical bars.

Chart

Description automatically generated

**Figure 3‑3****:** FNR across all threshold and scenarios. There are no control experiment results displayed for FNR because there are no possible true positives in control experiments, hence no false negatives either. Those values are for samplings done at generations 100 and 101, i.e., right before and after the migration event. 95% confidence intervals of the FNR estimates are displayed by vertical bars. Very low thresholds (i.e., < 0.01) are so conservative that they sometimes lead to no population being selected. Symbols overlap for some scenarios (those with only one affected population for example), which reach a null FNR.

In contrast, dispersal capacity did not substantially affect the FPR (**Figure 3‑2**). There were no consistent trends in FPR when comparing scenarios with different dispersal capacities but the same number of affected populations: L1 had slightly higher values than M1 and H1; L2 had slightly lower values than M2 and H2; and L3 had intermediate values between those of M3 and H3. Average FPR values for scenarios sharing the same dispersal parameters, calculated using FPRs at the p < 0.05 threshold as before, were 0.0599 (0.0558 – 0.0641; 95% CI) for low dispersal, 0.0621 (0.0580 – 0.0662; 95% CI) for moderate dispersal, and 0.0600 (0.0562 – 0.0638; 95% CI) for high dispersal (**Figure 3‑2**).

#### 3.2. Spatial extent

The number of populations affected by a demographic event also influenced our ability to detect meaningful temporal change. Scenarios in which fewer populations were affected exhibited a reduced FNR and an increased FPR (**Figure 3‑2**, **Figure 3‑3**). Scenarios in which a single population was affected (i.e., L1, M1, H1) had a perfect FNR (0;**Figure 3‑3**), while scenarios L2 and M2 only reached this perfect FNR at more liberal significance thresholds (i.e., above p ≤ 0.03; **Figure 3‑3**). The mean FNRs at p ≤ 0.05, averaged across scenarios sharing the same number of affected populations (e.g.,one averaged value for L1, M1, and H1 grouped together), were zero for scenarios with one affected population, 0.0028 (0 – 0.0059; 95% CI) for scenarios with two affected populations, and 0.0167 (0.0105 – 0.0228; 95% CI) for scenarios with three affected populations.

The number of affected populations influenced the FPR more than dispersal in our simulations. FPR values were consistent across scenarios with different dispersal levels but the same number of affected populations, rather than across scenarios with similar dispersal levels but different numbers of affected populations (**Figure 3‑2**, **Figure 3‑3**). The average FPRs from scenarios with the same number of affected populations, determined at the *p* ≤ 0.05 significance threshold, were 0.0820 (0.0778 – 0.0863; 95% CI) for scenarios with one affected population, 0.0553 (0.0516 – 0.0591; 95% CI) for scenarios with two affected populations, and 0.0447 (0.0413 – 0.0481; 95% CI) for scenarios with three affected populations.

#### 3.3. Time

We found that the genetic signal of the demographic event decayed over time, but that the TGI test was still able to identify significant changes in genetic diversity at a time scale of 1–9 generations. However, as the time interval between pre- and post-event surveys increased, the ability of TGI to detect the demographic event decreased, evidenced by the increase in false positives and false negatives for several demographic scenarios (**Figure 3‑4**, **Figure 3‑5**). The effect of time between surveys on the sensitivity of TGI was strongly affected by dispersal capacity and the extent of the event.

The timing of sampling prior to a simulated event was, as expected, generally less important than the timing of the post-event survey. The decrease in genetic signal over time — which would be found with any comparative method, not just TGI — was considerably strong in our simulations. For example, if the second (post-event) survey was taken nine generations after the first (pre-event) survey, we observed high FNR values that approached 75–90% in high- and moderate-dispersal scenarios (**Figure 3‑4**). The FNR also increased with the time lag in low-dispersal scenarios, but the increase was more linear, and values never reached 30%, even after nine generations (**Figure 3‑4Error! Reference source not found.**). One interesting result was that the number of affected populations was the main factor driving increasing FNR values with the age of the pre-event survey (3>2>1; left side of **Figure 3‑4Error! Reference source not found.**), while dispersal capacity was the main factor driving increasing FNR values when the time between the event post-event survey increased (H>M>L; right side of **Figure 3‑5**). For scenarios with the same number of affected populations, moderate-dispersal scenarios showed the worst performance with pre-event survey time lags, whereas high-dispersal scenarios generally showed the worst performance with post-event survey time lags (**Error! Reference source not found.**). Over our nine-generation sampling window, the FNR changed the least for the L1 scenario and the most for the H3 scenario (**Figure 3‑5Error! Reference source not found.**).

Chart, diagram

Description automatically generated with medium confidence

**Figure 3‑4 :** FNR from TGI tests performed between surveys carried out up to 9 generations before or after the migration event (arrow) when compared with surveys done the generation after the event for prior sampling, or the generation before the event for posterior sampling. 95% confidence intervals are displayed by bars.



**Figure 3‑5 :** FPR from TGI tests performed between surveys executed up to 9 generations before or after the event (arrow) when compared with surveys done the generation after the event for prior sampling, or the generation before the event for posterior sampling. 95% confidence intervals are displayed by bars.

While the relative differences in FPR performance given different time lags were not as high as for FNR, FPR nonetheless increased with the survey time lag. There were no clear patterns for whether dispersal or the number of affected populations most influenced the change in FPR associated with pre-event sampling time (**Figure 3‑5**); however, dispersal was the main factor driving FPR for time gaps associated with post-event sampling (**Figure 3‑5Error! Reference source not found.**). The strong relationship that we observed between FPR and the number of populations affected by the demographic event therefore became less pronounced as dispersal became more influential. As with the FNR, the FPR did not change much for the L1 scenario and changed the most dramatically for the H3 scenario (**Figure 3‑5**). These differences in how time affects our two most extreme scenarios is a useful consideration for potential TGI users.

The simulation that was most likely to preserve the signal of the demographic event was the low-dispersal scenario with a single affected population (L1). In this scenario, the TGI approach was still able to keep false negatives below 15% and false positives below 10%, even when the second survey was done nine generations after the event (**Figure 3‑4**, **Figure 3‑5**) and regardless of whether the first or second survey was responsible for the time lag with the event.

#### 3.4. Thresholds

A trade-off based on significance threshold values between FNR and FPR was present across scenarios. FNR values decreased with the chosen significance threshold, with a sharp decrease (most notable for H3) before 0.025 followed by a slower decrease until 0.1. FPR values increased with the chosen significance threshold, with a sharp increase at low thresholds followed by a continued but saturating increase until p ≤ 0.10.

#### 3.5. Applied example

The Northern tidewater goby (*Eucyclogobius newberryi*) is a small, endangered fish that lives in brackish estuaries and lagoons along the coast of California. This species represents an interesting model for population genetic studies because dispersal between suitable habitat patches only occurs during rare, discrete events. A previous study investigated extinction–colonization dynamics in the tidewater goby by evaluating genetic diversity across the landscape at several points in time (Kinziger et al. 2015). These authors suggested that the Elk River goby population had experienced unexpected temporal genetic change between 2006 and 2011 (Kinziger et al. 2015). We used the TGI method to re-analyse these data and determine if significant temporal genetic changes had indeed occurred in any population in this landscape.

Using our TGI measure, we found that the genetic structure of the Elk River population (**Supp. Mat. 3‑3**) of Northern tidewater goby (Kinziger et al. 2015) had indeed changed significantly relative to the other populations surveyed in the study area (permutation *p*-value = 0.0004), even after using strict *p*-value adjustments (Holm-Bonferroni adjusted permutation *p*-value = 0.0032). The average expected heterozygosity decreased by 0.046 in the Elk River population from an original value of 0.2646, which represents a loss of around 17%. Using TGI, we were able to quantify the qualitative findings of the previous study that there was a loss of genetic diversity in the Elk River population.

### 4. Discussion

In this study, we investigated how dispersal, the spatial extent of a demographic event, and the timing of sampling affected our ability to identify populations that have experienced significant changes in genetic diversity using a novel statistical tool: TGI. The factors we chose to investigate are directly relevant to conservation studies. Dispersal is a key element in understanding population connectivity (Kool et al. 2013; McRae 2006). The spatial extent of significant demographic changes is relevant to examine as recent studies have advocated for a more comprehensive integration of space in evolutionary ecology research (Battey et al. 2020; Bradburd & Ralph 2019; Velázquez et al. 2016). The timing of sampling is key, first to make sure we capture the effects of a potential disturbance, and second, to better grasp the decay in its legacy signal.

#### 4.1. TGI: a new and useful framework

TGI provides a novel and robust framework for testing whether observed changes in genetic diversity through time are significant relative to variation associated with genetic drift, local gene flow, and mutation. Our successful application of TBI to genetic data involved translating a site-by-species approach to a site-by-genotype approach and changing the permutation algorithm to accommodate the specific structure of various genetic data formats such as SNPs in our simulations and microsatellites in our application. In addition to describing our new framework, we also evaluated its power and specificity and found that TGI is functional over a wide range of parameter values. One main contrast between our new TGI approach and previous investigations of the performance of TBI (Legendre 2019; Winegardner et al. 2017), which was developed for community composition data, is that we also examined how the timing of surveys, and its interaction with demographic parameters, may affect the downstream conclusions. Our results indicate that TGI consistently and accurately identifies populations that have experienced a demographic event.

#### 4.2. Dispersal and spatial extent

Detection of temporal genetic changes was sensitive to dispersal; false negatives increased with dispersal capacity, although false positives did not show a clear trend (**Figure 3‑2**, **Figure 3‑3**). The influence of dispersal on the FNR was also affected by the time lag between an event and the subsequent sampling effort; the effects of different dispersal capacities were evident even when surveys were separated by only one generation (i.e., samples were collected immediately before and after the event) and were magnified as the time between surveys increased. The effects of sampling time and dispersal capacity on the FNR suggest that species with high dispersal capacity in well-connected landscapes, such as many forest pests (e.g., Larroque et al. 2019; Wittische et al. 2019), might require more frequent sampling to overcome the negative effect of gene flow on our ability to correctly identify affected populations.

The spatial extent of a demographic event increased our ability to correctly identify populations that have not truly changed (lower FPR), but it also decreased our ability to correctly identify populations that did change (higher FNR). The magnitude of this trade-off varied with dispersal capacity. Although a broader spatial extent may help researchers detect an event, as the chance of sampling an affected population increases, it may also increase the risk of not identifying the genetic legacy of the event at all, especially in high-dispersal landscapes. It is less effective for analyzing gradual, landscape-wide disturbances. In addition, when multiple populations were affected in our simulations, we always chose to affect adjacent populations; we did not investigate whether lowering the degree of spatial autocorrelation (Dale & Fortin 2014; Legendre & Legendre 2012) in the spatial genetic legacy (e.g., two independent catastrophic events, a pollution and a flood for example, affecting the landscape) influenced our ability to detect the events.

#### 4.3. Time between surveys

As expected, spatial genetic legacies decayed over time, mostly due to gene flow and drift. Specifically, in this study we found that TGI was suitable for identifying changes over 1-9 generations (e.g., years) depending on landscape and demographic parameters. Two main points emerged from our analysis of how the timing of sampling affected the detection of significant genetic changes. First, when comparing an old survey to a survey realized soon after a demographic event, the spatial extent of the disturbance affected the power of TGI, with smaller spatial extents preserving high power even with large time gaps. Second, when comparing a survey realized immediately before a disturbance to one collected several years after, dispersal was the most important factor driving the performance of TGI, with low-dispersal scenarios better preserving the performance of TGI in the context of decay brought by background processes through time. High-dispersal systems could lead to as many as 10% of false positives even when sampling only a few years after an event. This result has serious implications: arbitrary and potentially inappropriate significance thresholds may result in misallocation of resources to monitoring or treating unaffected populations while missing some affected populations. In contrast, by considering the population dynamics and, if possible, planning relatively simple and short model-specific simulations, one can enhance the usefulness of TGI and proceed with a more appropriate sampling/monitoring strategy. Given the fact that FNR reach high values at the highest time gaps for some scenarios (**Figure 3‑4**), we believe that our choice of a maximum of 9 generations (e.g., years) between surveys was appropriate.

#### 4.5. Empirical application

We successfully applied TGI to an empirical dataset from an endangered fish, the Northern tidewater goby, for which temporal genetic change had been described but not quantitatively tested (Kinziger et al. 2015). The authors of the original publication hypothesized that one goby population had undergone meaningful genetic change relative to the rest of the landscape; our application of TGI supported this hypothesis. We therefore clearly showed that the straightforward TGI testing procedure can be used to strengthen the results from temporal genetic studies that use repeated surveys.

#### 4.6. Considerations about the use of TGI

Different empirical datasets and research objectives may require TGI users to customize our procedure, but the TGI function is transparent and flexible (Culina et al. 2020), and different permutation and genetic distance algorithms could easily be used by simply changing a few characters or lines of R code in the annotated TGI function provided in the supplementary material (**Supp. Mat. 3‑2**). Evaluating the performance of distance metrics with different statistical properties (Legendre & De Cáceres 2013) could be an avenue for future work. TGI can also readily be used on other types of genetic data, such as microsatellites. TGI provides a robust statistical framework more trustworthy than arbitrary comparison of pairwise genetic dissimilarities, or node-based genetic diversity values.

Despite these advantages, there are still several important considerations for the effective use of TGI tests. The implementation of TGI in new systems will ultimately be more successful if researchers have an a priori understanding of the population dynamics of their system and the nature and scale of possible disturbances in their study area. Indeed, this prior knowledge could guide researchers in the choice of survey intervals after important historical demographic events (Anderson et al. 2010; Fenderson et al. 2019). FNR and FPR values ultimately represent trade-offs in potential conservation costs (Moilanen et al. 2009; Welch et al. 2020), and it is therefore essential that researchers grasp their importance and choose these values deliberately. Stricter (lower) values for the TGI *p*-value threshold expectedly result in a lower FPR but may also result in a higher FNR (lower power). Identifying the most sensible threshold for a chosen objective would be valuable to better understand the trade-offs of different sampling schemes in specific empirical systems. Purpose-designed spatially explicit simulations can be used to address this challenge (Epperson et al. 2010; Haller & Messer 2019; Landguth et al. 2017a). Insome cases, it may be desirable to minimize false negatives relative to false positives – thus ensuring that we detect all the affected populations no matter the cost of detecting, and therefore monitoring and preserving, some populations that do not need preservation.

TGI was not developed as an alternative for inferring demographic history from large genetic datasets collected at a single time (e.g., Leblois et al. 2014). Instead, it was designed to help research teams collecting repeated surveys from non-model organisms with limited genotypic information (Draheim et al. 2018; Kinziger et al. 2015; Moraes et al. 2017), and teams wanting to compare new surveys to older ones. Nonetheless, more studies in different types of genetic systems, involving different historical demographic events, are needed to explore how the performance of TGI varies with factors that were not tested in our simulations, including 1) the chosen genetic distance algorithm; 2) spatiotemporal autocorrelation in genetic legacies; 3) effective population size; and 4) spatial heterogeneity in landscape resistance to movement.

Using a different null model approach could lead to different performance and insight, and permutation is but one of many possible ways to create reference distributions which can be used to build a test by comparison with the observed value. Although comparing different null model approaches was beyond the scope of this paper, preliminary work for this study showed that some permutation algorithms were generally less adequate for FPR and FNR, such as was the case for TBI (Legendre 2019). Future research may notably include evaluating more complex, multilevel, permutation approaches, and exploring null models which explicitly preserve some components of genetic structure such as isolation-by-environment (Wang & Bradburd 2014). Previous spatiotemporal research in ecology has shown that null models with increasing levels of complexity provided different insight (James et al. 2010), but that the usefulness of complex null models could be offset very long computation time (Leblois et al. 2014), and by the risk of overfitting (Bohl et al. 2019). The performance of more complex null models may also be more sensitive to threshold choice (Merckx et al. 2011), which is another reason why evaluating a range of threshold values, such as in this study, is pertinent.

TBI has inspired TGI but there are some reciprocal opportunities brought by insights generated through it’s porting to genetic data. We showed that dispersal, through gene flow, had profound effects on our ability to detect change. Dispersal is also a key element of community dynamics (Hubbell 2011), and we know that it influences community assembly and mediates community response to an environmental change (Catano et al. 2017; Condit et al. 2002; Conradi et al. 2017; Qian 2009). Dispersion limitation is not constant across different species and preserving the range of dispersal abilities through a change in the permutation algorithm may increase TBI performance, especially for fine grain/large extent landscapes, and for communities with strong distance decay in similarity. Another insight brought by our study, which also interacts with dispersal, is the window of the time during which the signal of community-changing event may be detectable using TBI. This is especially important to consider as species having different dispersal abilities may differently recolonize and the gains and losses part of a TGI analysis (Legendre 2019) may therefore be dependent on the time since the event. Integrating TGI and TBI in a multi-scale community genetics (James et al. 2011a) change approach, and testing the trajectory (De Cáceres et al. 2019; Sturbois et al. 2021) of the genetic diversity or community composition across more than two dates, represent two promising avenues of research.

#### 4.7. Conclusions

At the crux of many conservation biology questions, identifying changes in genetic diversity, beyond the expected changes due to background processes, can help researchers identify locations or populations that have experienced important past demographic events. These events could be detrimental (e.g., loss of diversity or maladaptation) or beneficial (e.g.,higher effective population size or genetic rescue), and are often relevant for conservation efforts. Such locations and populations could then be prioritized for increased monitoring and further investigation into the origin of these changes. As shown in our application of TGI to empirical data from the endangered Northern tidewater goby, our method provides a framework for detecting and pinpointing exceptional temporal genetic changes. Our approach to detect temporal genetic differentiation does not require extensive genomic information and can therefore be used to explore the temporal dynamics of genetic diversity changes using relatively small genetic datasets (e.g., hundreds of SNPs). We believe that the TGI approach is a promising tool for the spatiotemporal analysis of wild, non-model organisms for which extensive genomic resources are yet to be developed.

TGI and its future developments will, therefore, be of primary value to better manage the landscape in the context of global change and stronger, and more frequent, perturbations. Monitoring the genetic health of already at-risk populations (Draheim et al. 2018; Kinziger et al. 2015; Moraes et al. 2017) is but one of the uses TGI can offer. Climate change will displace populations which are differently adapted to warmer and drier environments (Masson-Delmotte et al. 2021) and being able to detect this genetic change is of primary value to better manage species affected by climate change. Given that the relative importance of tolerance, migration, and adaptation, may vary in different parts of a landscape under climate change (Sork et al. 2010), TGI could be used as a sentinel tool to assess whether populations have changed substantially more than the general response of this species to climate change in that landscape. TGI could also be used to detect a new contact between two previously distinct genetic clusters in the landscape, which could alert managers that a shift is happening (Jay et al. 2012; Pérez-Portela et al. 2019). TGI could also be used to study the dynamics and the synchrony of outbreaks (Larroque et al. 2019), by, for example, detecting a large migration event from a previously isolated genetic cluster, which may confirm the high connectivity synonymous with serious outbreaks. Biological invasions may be slowed, or fail, because of genetic factors such as successive reduction of genetic diversity following introduction, or the spatial analog of genetic drift at the leading edges of the expanding range (Austerlitz et al. 2000; Schrey et al. 2014; Slatkin & Excoffier 2012). A local gain or a loss of genetic diversity, as identified by TGI, could predict the evolutionary potential of a population which may affect the further spread of that invasive species (Lawson Handley et al. 2011), because of genetic load, inbreeding depression, and drift load (Charlesworth & Willis 2009; Szucs et al. 2014; Willi et al. 2013). Similarly, TGI could be used to detect illegal human-mediated translocations of non-native individuals (Dufresnes et al. 2016; Frantz et al. 2006, 2017) which can be an issue because of maladaptation, outbreeding depression, and pathogen introduction in naïve populations.

### Acknowledgements

This research was supported by a grant to PMAJ and the TRIA Network from the Natural Sciences and Engineering Research Council of Canada (grant no. NET GP 434810-12), with contributions from Alberta Agriculture and Forestry, fRI Research, Manitoba Conservation and Water Stewardship, Canadian Forest Service (Natural Resources Canada), Northwest Territories Environment and Natural Resources, Ontario Ministry of Natural Resources and Forestry, Saskatchewan Ministry of Environment, West Fraser, and Weyerhaeuser. This research was also supported by a Discovery Grant to PMAJ. JW was also supported by a scholarship from the Forest Complexity Modelling (FCM) NSERC CREATE. Computations were made on the supercomputer CEDAR managed by Compute Canada ([www.computecanada.ca](http://www.computecanada.ca)). We thank Jeremy Larroque, Hinatea Ariey and Charlotte Van Engeland for their comments on an earlier version of the manuscript.

DATA ACCESSIBILITY

All simulation data used for this paper will be deposited online upon acceptance. Functions used to analyze the simulations will be available on a public repository on *GitHub*.TGI, the function that would be most useful to potential users of our approach, will continue to be maintained and developed and may be contributed to a CRAN package in the near future.

AUTHOR CONTRIBUTIONS

J.W. designed the study, created the simulation inputs, ran the simulations, transformed the TBI function to TGI, and performed the analyses. P.L. and P.M.A.J. provided advice on the study design, analysis, and the visualization. J.W., P.L. and P.M.A.J. wrote the paper.

### Supplementary material

**Supp. Mat. 3‑1 :** Rogers’ genetic distance

Given the number of loci, the number ofalleles observed in locus *k*, and and the proportions of allele *j* at locus *k* in populations 1 and 2, respectively:

The sum of the allele proportions for any locus in any population is 1.

For biallelic markers such as SNP, the calculation becomes even simpler as seen in the example in **Table 3‑1**, A. Indeed, because the proportion of an allele is always 1 minus the proportion of the other allele, the sum per locus is always twice the term for the a single allele and the equation simplifies to:

**Supp. Mat. 3‑2 :** TGI function and output.

# mat1: the genotypic matrix associated with the first sampling; must be a genind object

# mat2: the genotypic matrix associated with the second sampling; must be a genind object

# nperm: the the number of permutations used in the evaluation of significance

# seed.: you may specify a seed by using this argument

# method : a number between 1 and 5. Five genetic distances are available in function dist.genpop # of the adegenet package.

# They are : (1) Nei’s D, (2) Edwards’ angular D, (3) Reynolds’ coancestry coefficient, (4)

# Rogers’ D, (5) Prevosti’s absolute genetic D. Methods 2, 3 and 4 produce Euclidean distances,

# whereas methods 1 and 5 produce non-Euclidean distances, which produce negative eigenvalues and # complex eigenvectors in principal coordinate analysis.

# correc: correction for multiple # inference; see ?p.adjust

# thresh\_for\_GL: indicate here the threshold you want to use

TGI <- function (mat1, mat2, nperm = 999, replace = FALSE, seed. = NULL, method = 4, correc = "holm", thresh\_for\_GL = 0.05) {

#### genind to genpop objects

mat1p <- genind2genpop(mat1)

mat1p <- mat1p[,order(colnames(mat1p@tab))]

mat2p <- genind2genpop(mat2)

mat2p <- mat2p[,order(colnames(mat2p@tab))]

##### Function to compute genetic distances

dissim <- function(mat1p, mat2p, method) {

dis <- vector(mode = "numeric", length = nrow(mat1p@tab))

for (i in 1:nrow(mat1p@tab)){

if (i == 1){

trick <- 2

} else {

trick <- 1

}

temp\_genpop <- mat1p

temp\_genpop@tab[trick,] <- mat2p@tab[i,]

dis[i] <- dist.genpop(temp\_genpop[c(trick, i),], method = method)

}

list(dis = dis)

}

##### Initialization of seed, tolerance

if (!is.null(seed.)){

set.seed(seed.)

}

epsilon <- sqrt(.Machine$double.eps)

##### Dimensions check

n <- nrow(mat1p@tab)

p <- ncol(mat1p@tab)

if ((nrow(mat2p@tab) != n) | (ncol(mat2p@tab) != p)){

stop("The matrices are not of the same size!")

}

##### Empirical genetic distances

tmp <- dissim(mat1p, mat2p, method)

dis.ref <- tmp$dis

##### Permutations

if (nperm > 0) {

my.vec <- sample(1:(10 \* nperm), size = nperm)

outlier.count = rep(1, n)

for (iperm in 1:nperm) {

set.seed(my.vec[iperm])

mat1.perm <- mat1p

mat1.perm <- shufflepop(mat1.perm, method=4)

set.seed(my.vec[iperm])

mat2.perm <- mat2p

mat2.perm <- shufflepop(mat2.perm, method=4)

tmp <- dissim(mat1.perm, mat2.perm, method)

dis.perm <- tmp$dis

ge <- which(dis.perm + epsilon >= dis.ref)

if (length(ge) > 0) {

outlier.count[ge] <- outlier.count[ge] + 1

}

}

p.dist <- outlier.count/(nperm + 1)

}

p.adj <- p.adjust(p.dist, method = correc)

##### Gain or loss?

n.pop1 <- seppop(mat1)

n.pop2 <- seppop(mat2)

mean.hexp1 <- do.call("c", lapply(n.pop1, function(x) mean(summary(x)$Hexp)))

mean.hexp2 <- do.call("c", lapply(n.pop2, function(x) mean(summary(x)$Hexp)))

mean.hexp1[is.nan(mean.hexp1)] <- NA

mean.hexp2[is.nan(mean.hexp2)] <- NA

simple\_diff <- mean.hexp2 - mean.hexp1

output <- list(TBI = dis.ref, p.TBI = p.dist, p.adj = p.adj, gainloss = simple\_diff[p.adj < thresh\_for\_GL])

class(output) <- "TGI"

return(output)

}

#############################################################################

**> goby\_test <- TGI(goby\_first, goby\_second, nperm = 9999, method = 4)**

**> goby\_test**

$TBI #index values

[1] 0.06432131 0.06089485 **0.15212258** 0.02258920 0.07247326 0.04463856 0.06672004 0.02238467

$p.TBI #unadjusted permutation p-values

[1] 0.4283 0.4891 **0.0004** 0.9943 0.3188 0.7756 0.3947 0.9949

$p.adj #adjusted p-values

[1] 1.0000 1.0000 **0.0032** 1.0000 1.0000 1.0000 1.0000 1.0000

$gainloss #difference in expected heterozygosity

**ELK**

**-0.04567755**

attr(,"class")

[1] "TGI"



20km

kkkkkkk km

N

N

**Supp. Mat. 3‑3 :** Satellite map of the Californian sampling stations used in the goby analysis. The red circle marks the Elk population which was the only one found to have significantly changed.