**Detecting temporal changes in genetic diversity using limited information: a new tool for molecular ecology studies with repeated samples.**

**INTRODUCTION**

Global biodiversity in terms of genes, species, populations, and ecosystems is being lost at an increasing rate with significant consequences for ecosystem functioning and long term viability of the biosphere [1–3]. Novel monitoring techniques are needed to track these losses and to inform conservation efforts. Further, it is recognized that it is no longer sufficient to study spatial patterns in biodiversity loss at a single point in time. Instead, one must examine trends and patterns in biodiversity through both space and time [4,5].

Spatial and temporal variation in genetic diversity can tell us a great deal about population demography and connectivity [5,6]. Indeed, population genetics has proven essential to translate observed genetic variation into meaningful inferences that are essential for conservation efforts [7–9]. More specifically, temporal variation in genetic diversity, and its drivers, are at the crux of many conservation and public health issues [10–12]. Researchers commonly use patterns in spatio-temporal population genetic data [13], to quantify isolation-by-distance [14,15], time since population bottlenecks [16,17], rates of migration between isolated populations [18,19], and timing and extent of outbreak expansions [20,21]. Further development of approaches able to detect changes in genetic variation through time is needed to further elucidate the processes that govern demographically dynamic systems such as insect outbreaks, invasions, disease spread, and species declines [4,5,8]. Such tools are especially needed to identify populations that have experienced important historical demographic events which may include disturbances such as a wildfire, major weather events, invasions of predators or competitors, or outbreaks of disease [22–26]. Similarly, such analyses could identify which, among a set of previously sampled populations, received migrants from a long-distance dispersal event [27]. Because temporal genetic variation reflects the evolutionary potential of a population and the probability of its persistence [28–30], relating it to landscape change can give us important insights about the eco-evolutionary dynamics of a species, and be used to inform conservation strategies (e.g. Landguth, Holden, Mahalovich, & Cushman, 2017).

There are two types of approaches available to investigate temporal genetic variation. The first kind of approaches generally consists of a single time sampling scheme, combined with statistical approaches purpose-designed to directly infer demographic history from the single time sampling [32–34]. This approach often involves extensive knowledge about the biological system, very intensive computation, and requires high-quality microsatellite datasets, or a minimum of tens of thousands of loci, and further information such as recombination processes [17] and ascertainment bias [35–37]. The second kind of approaches consists in comparing genetic diversity between repeated (same sites) temporal samples. Such approaches are more readily usable in systems where less information is available such as non-model species or systems that were sampled historically, with and an interest in comparing contemporary patterns with these older data [10]. However, there are several technical and conceptual challenges associated with the latter kind of approaches.

Despite our ability to compare genetic diversities evaluated at two points in time, how to meaningfully quantify and detect temporal changes remains unclear. Some studies have used genetic differentiation metrics such as Jost’s D, FST or its analogs [21,38,39] to evaluate temporal changes between genetic datasets. However, translating our spatial understanding of such differentiation indices the temporal dimension is not straightforward [40]. Additionally, disentangling spatial from temporal effects is a challenge because the additivity of genetic drift means than genetic differentiation can be associated with both space and time [41,42]. Distinguishing between natural variation in temporal genetic structure due to the processes of recombination, mutation, and demographically-induced genetic drift from the changes brought by external landscape variation remains challenging.

Temporal Beta-diversity Indices (TBI; Legendre 2019) are used to quantify and assess the significance of changes in ecological community composition through time using a dissimilarity index. The significance of these dissimilarities is then tested using permutation. The TBI approach has been extensively tested on simulated community composition data [43], 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 between the question of how diversity in multi-species communities and multi-locus genetic diversity change through time, we sought to determine if TBI could be successfully applied to spatio-temporal genotypic data. We also set out to determine how different demographic contexts influence the ability of this method to identify atypical temporal genetic changes in diversity.

To address these questions, we extend the TBI framework to spatio-temporal population genetic data. The objective of our new framework, Temporal Genetic diversity Indices (TGI), is to identify significant temporal variation in spatial genetic diversity using relatively information-poor genetic data, despite confounding forces such as drift. We demonstrate the effectiveness and applicability of the approach using simulated genetic data generated using a spatially-explicit demo-genetic simulator (*CDMetaPOP*; Landguth, Bearlin, Day, & Dunham, 2017). Our general approach was to simulate multiple scenarios in which portions of a landscape are affected by a non-selective demographic change. We then used TGI to measure changes in the genetic diversity of our populations under these different demographic contexts. Specifically, we explored how dispersal ability, the number of populations affected by a demographic event (i.e., spatial extent), and time between two sampling efforts affected our capacity to detect significant temporal variation in genetic diversity. Performance was quantified using standard false positive/negative rates binary classification [45]. We predict that our ability to detect temporal genetic changes would be lower with increasing dispersal ability because of the homogenizing effect of higher gene flow. We also predict that the longer the time between successive sampling, regardless of when an event occurred between them, the harder it will be to identify where and when atypical genetic change occurred. 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 a threatened vertebrate.

**METHODS**

*Adapting TBI for genetic data*

Calculating TBI involves computing dissimilarities in species composition between the data sampled at two different times at all sampling sites, and testing the significance of these indices through simultaneous permutations of the two site-by-species input matrices. In extending TBI to TGI, we used genetic distances (see Fig. 1 A for a simple example showing how we transformed two temporal samples into a genetic distance, for a two-site landscape) calculated from site-level allele frequencies instead of community dissimilarities to compare the two different temporal samples. In this case, the null hypothesis is that genetic composition does not differ between the two points in time that were sampled, more than is expected by processes typical to the landscape.

Indeed, Background genetic processes such as drift can also result in temporal differences in genetic structure. The challenge is therefore to identify temporal changes that are significantly different from what one would expect under a drift only scenario. Because there are no reference distributions for what constitutes a significant temporal genetic change, we use a permutation-based approach to generate a distribution of values to which an observed value can be compared.

Testing the significance of TGI involved permuting the input genotypic matrices. Here, we permuted the genotypes at each locus in the same way in both time samples (see Fig. 1 B for a simple example showing how we conducted the permutations, for a two-site landscape) using the *poppr* R package (see *Software*). This permutation was chosen because it maintains allelic structure and heterozygosity [46]. We used 999 permutations in all analyses.

*Genetic distance*

Genetic distances between points in time for a given location were calculated using the Rogers’ genetic distance [47], which is similar to the Euclidean genetic distance (see Annex A). It makes no assumptions about base-pair substitutions or time since separation and is suitable to study short-term dynamics. It has recently been used to investigate spatial genetic structure in a pond turtle [48] and a fungus [49]. We computed the distance using the *dist.genpop* function from the *adegenet* R package (see *Software*).

*Simulation framework*

To simulate the dynamics of population genetic changes through time and test the performance of TGI, we used the spatially-explicit gene flow simulator *CDMetaPOP* [44]. *CDMetaPOP* simulates dispersal and mating of individuals across a landscape and allows the user to define initial genetic structure, spatial distribution of individuals, dispersal characteristics, and life-history traits of the population. The physical landscape we simulated was a homogeneous and interconnected square grid containing 5 × 5 cells, each cell representing a population. Each population had a maximum carrying capacity of 50 individuals. Distance between populations was simply set as the Euclidean geographic distance. The populated landscape, therefore, contained a maximum of 1250 individuals. The genotypic information of each individual was recorded and consisted of 100 neutral, unlinked, bi-allelic SNP loci. Sampling was done before and after the event unless otherwise specified. The mutation rate was set at 10-8 [50].

We simulated 180 replicates for each scenario (see below). For each replicate, we initialized the simulation with random and unique allocations of alleles among individuals, therefore approaching maximum diversity [51]. Those parameters were chosen as a compromise between realism and computational time limitations, and they seemed appropriate to produce the complex evolutionary dynamics necessary to produce reasonably realistic and useful simulated genetic data. Each simulation was run for 100 generations before a demographic event (see below) was imposed on up to three populations in the landscape. 10 more generations were simulated after this event.

Multiple demographic processes affect the spatial apportionment of genetic variation. For this study, we chose to investigate the influence of immigration from a previously isolated population. When modelling immigration, we simulated immigration from a population that was separate from our 5×5 grid (i.e., population #26). Our goal was to apply the TGI approach to detect historical population changes in genetic data due to immigration. This independent source population otherwise shared the same attributes as other populations in our simulated landscape. Only during simulated demographic events were individuals from the 26th isolated population allowed to disperse into the simulation grid.

Using this model, we examined the influence of dispersal and the spatial extent of demographic events (number of affected populations) on the persistence of genetic spatial legacies. We examined three levels of dispersal (see below), and three different numbers of populations affected (1, 2, 3) for a total of 9 unique scenarios, each of which was replicated 180 times, for a total of 1620 (9 × 180) unique simulations for this experiment, aside from the control simulations described below (Table 1). In the next sections, we detail how we modelled the two experimental factors.

*Dispersal*

To model dispersal, we weighted the geographic distances between populations using a power law function, , where *B* represents how difficult it is to disperse and high values of *B* correspond to low dispersal capacity (elaborated below). We then rescaled the values, using the maximum and the minimum (0) distances possible in this virtual landscape, as described in the *CDMetaPOP* [44] user manual (p.63). This produced values in the [0,1] range, which are considered to represent probabilities that an individual disperses to a cell located at that distance (Fig.1). We chose this way of modelling dispersal to allow both within-population movements and long-distance dispersal.

The population to which an individual disperses was selected randomly from the set of populations available at the given distance. Individuals always stay within our simulated landscape, and any individual may disperse to one of our 25 populations at each generation. To investigate the effect of different levels of dispersal, we used three different values of *B*: low (*B* = 2), moderate (*B* = 1.301) and high (*B* = 0.6015) dispersal levels (Fig.1; Table 1).

*Spatial extent*

We also wanted to evaluate how the spatial extent of the demographic event affected the performance of our method. To achieve this, we triggered demographic events at 1 to 3 populations randomly selected among the 25. When only 1 population was affected, we partitioned the 180 replicates of that scenario equally among six populations in the landscape. We decided to vary the position of where the demographic event would happen in the landscape because deme topology may influence the outcomes of population genetics analyses [52]. The positions of these six populations were randomly selected once and identical across runs. Indeed, because our landscape is square and homogeneously resistant to movement, it is therefore symmetric and only six positions need to be assessed. When several (2 or 3) populations underwent a demographic event, we randomly sampled 1 position among the six previously described and randomly picked 1 or 2 additional populations directly adjacent (when possible) to it. We did this six times (30 replicates for each different set of populations). We chose to pick populations this way to respect the spatial autocorrelation often exhibited in demographic events.

*Time*

To assess how the time since the simulated demographic event affects our ability to detect genetic changes, we used TGI on simulated data collected up to 9 years before, and after, the event, and compared them with data collected respectively right after, and right before the event year. We chose 9 years as the maximum time between samplings as this time lag would represent most of the “before/after” population genetic studies we encountered, and because most long-term ecological research programs monitor at a shorter time interval. For our analyses concerning the timing of sampling, we chose the 0.05 *p*-value threshold as it was a good compromise between different performance metrics in our earlier results.

*Statistical performance*

We used the False Positive Rate (FPR) and False Negative Rate (FNR) to assess the statistical performance of the TGI testing procedure. A false positive is a population that we know *a priori* did not undergo the demographic change we imposed but has been found to have done so by the TGI test. A false negative is a population that did experience a demographic event but was not found to have done so. FPR represents the number of false positives over the total number of negative tests (i.e., true negatives and false positives). FNR represents the number of false negatives over the total number of positive tests (i.e., true positives and false negatives). A high FPR means that we often select the wrong population(s) as having changed substantially, and researchers generally want to keep it as low as possible when there are, for example, limited resources available for conservation efforts. The higher the FPR, the lower the selectivity of our testing procedure. In contrast, a high FNR means that we often fail to identify the population(s) that were actually affected. The higher the FNR, the lower the power of our testing procedure. Researchers may want to minimize the FNR in situations where finding the right population is the most important aspect, for example, if there is limited time to take conservation action. Selecting a proper threshold to compare to the p-value calculated through the TGI test permutations, and therefore defining positives and negatives, is often important to identify a compromise between power (1 – FNR) and selectivity (1 – FPR). To characterise this compromise, we evaluated the statistical performance of TGI using a range of thresholds for FPR and FNR: 0.0001, 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

*Controls*

Control simulations were run in which no populations were affected by demographic events and were therefore only subject to the processes of gene flow, drift, and mutation. Dispersal was the only parameter that varied among the control simulations, resulting in three control scenarios (Table 1). We evaluated the FPR of these control scenarios; there was no need to compute FNR because there were no true positives/false negatives so it was always equal to 0. When describing the performance of other scenarios with similar dispersal parameters, we always used these control values as reference.

*Software  
CDMetaPOP* runs on *Python 2.7* [44]. We used the *R* software [53] in the *RStudio* IDE [54] for all analyses and illustration. We used the *adegenet* [55,56], *pegas* [57], *poppr* [58,59] and *adespatial* [60] *R* packages for calculations. Our *TGI* function coded in *R* is available in the appendix.

*Application example*

To showcase TGI’s general ability to provide sensible information about temporal change in a real system with conservation implications, we chose to apply it to a threatened vertebrate: the Northern tidewater goby. We chose this example because it uses a different genetic data type than we the one we simulated, to make it clear TGI is usable on a diversity of markers, and because the authors suggested that one population had undergone change among other more stable local populations, which allows to test a real hypothesis, and go beyond a simple illustration of our method [61]. The dataset was available through DRYAD (doi:10.5061/ dryad.871db). We used 9999 permutations for this application.

**RESULTS**

We were able to translate the TBI framework to TGI by accommodating it to the specific structure of genetic data. TGI provided good and stable performance according to the analyses of a wide array of simulations. Although results support its efficacy and warrant the use of TGI on empirical datasets, there are a few specific results to consider, especially related to dispersal, spatial extent of the demographic event producing the genetic change, and time lag between sampling and event.

*Dispersal*

As hypothesized, dispersal capacity influenced our ability to detect temporal changes in genetic diversity. FNR generally increased with dispersal intensity (Fig. 3). Results were generally satisfying and encouraging for the use of TGI as only one scenario (H3; Table 1) exhibited values above a very conservative FNR limit of 1%, regardless of the p-value threshold used (Fig. 3). Among the four scenarios never reaching an average FNR of 0 (L3, M3, H2, H3), two involved high dispersal. Taking the average from scenarios sharing the same dispersal parameters for the ubiquitous 0.05 threshold (e.g.one value for L1, L2, and L3 grouped together), we had mean FNRs of 0.0037 (0.0007 - 0.0066; 95% 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. FNR values overall decreased with threshold, with a sharp decrease (most notable for H3) before 0.025 followed by a slower decrease until 0.1.

FPR did not substantially change as a function of dispersal capacity (Fig. 4). Low dispersal did not consistently result in higher FPR than moderate dispersal, or high dispersal (Fig. 4). Indeed, given the same number of affected populations we can see that L1 has slightly higher values than M1 and H1, L2 has slightly lower values than M2 and H2, and L3 has intermediate values, between that of M3 and H3. FPR averages from scenarios sharing the same dispersal parameters, for the 0.05 threshold, 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. FPR values overall increased with threshold, with a sharp increase at low thresholds followed by a continued but saturating increase until threshold 0.1.

Finally, another very encouraging result is that experimental FPR values consistently stayed much below any control FPR values regardless of dispersal parameters (Fig. 4). This means that in the presence of an atypical demographic event, we were always more likely to correctly identify a population as having been affected. The higher the dispersal, the lower the control FPR (Fig. 4). Control FPR values were generally at least twice as high as the maximum experimental FPR values (L1, M1), regardless of the threshold used. This means that even for the worst scenarios, TGI was much more effective at avoiding false positives, in the presence of an event, than in its absence.

*Number of populations affected*

Although TGI displayed good performance overall, the number of populations affected by a demographic event also influenced our ability to detect meaningful temporal change. Scenarios with fewer populations affected exhibited reduced FNRs, and an increased FPR (Figs. 3, 4).

Scenarios in which a single population was affected (i.e., L1, M1, H1) showed perfect FNR (0; Fig. 3), while L2 and M2 reached this perfect FNR only for liberal thresholds (i.e., >0.03; Fig 3). Taking the average from scenarios sharing the same number of populations affected, for the ubiquitous 0.05 threshold (e.g.one value for L1, M1, and H1 grouped together), we had mean FNRs of 0 for scenarios with one affected population as previously stated, 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. FNR values overall decreased with increasing threshold, with a sharp decrease in FNR (most notable for H3) before 0.025 followed by a slower decrease until 0.1.

As shown by the overlap of FPR values across scenarios with similar numbers of populations affected, rather than across scenarios with similar dispersal, the number of populations affected generally influenced the FPR performance more than dispersal, for the levels we used in the simulations (Figs. 3, 4). FPR averages from scenario sharing the same number of populations affected, for the 0.05 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.

*Time*

We found that the genetic signal of the demographic event decays over time, and that TGI may be useful for identifying significant changes in genetic diversity at a time scale of 1-10 years. The longer the interval between a pre- and a post-event sampling, the lower the power available to detect the demographic event, as evidenced by the increase in false positives and false negatives for several scenarios (Fig. 5, 6). However, this effect is strongly affected by the dispersal capacity and the extent of the event.

The timing of sampling prior to a simulated event was, as expected, generally less important than that of the posterior sampling. The decrease in signal, which would be found with any comparative method – not just TGI – was purposely strong in our simulations. For example, for second samplings done 9 years after a first sampling done before the event, we observe rather high FNR values, especially with high and moderate dispersal scenarios (Fig. 5). FNR also increased with time lag for the posterior sampling for low dispersal scenarios, for which the increase was more linear, and values never reached 30% in the scope of our analyses, even after 9 years, (Fig. 5). One very interesting point is that the number of populations affected is the main factor increasing FNR values in older first samplings (3>2>1; left side of Fig. 5), while dispersal capacity is the main factor increasing FNR values with the time lag of second samplings (H>M>L; right side of Fig. 5). For a same number of populations affected, moderate dispersal scenarios showed the worst performance for first sampling time lags, while high dispersal scenarios generally showed the worst performance for second sampling time lags (Fig .5). FNR changed the least for the L1 scenario, and the most for the H3 scenario (Fig. 5).

While the relative differences in FPR performance given different time lags were not as high as for FNR, FPR nonetheless increased with sampling time lag, for example, up to more than twice the value of the threshold for first samplings for L1, and even higher for the largest time lags possible for second samplings (Fig. 6). There are no clear strong patterns for whether dispersal or the number of populations affected most influenced FPR change for pre-event samplings (Fig .6). However, for time gaps attributed to the second sampling, dispersal became the main factor in driving FPR (Fig. 6). The previously described strong relationship between FPR and the number of populations affected by the demographic event, therefore, changed as dispersal became more influential. FPR did not change much for the L1 scenario, and changed the most dramatically for the H3 scenario (Fig. 6), again denoting differences in the influence of time between our two most extreme scenarios, which could be useful for potential TGI users.

When considering the scenario most likely to preserve the signal according to earlier results on FNR and FPR (L1), the TGI approach was still able to keep false negatives below 15% and false positives below 10%, even for a second sampling done 9 years after the event (Fig. 5, 6), regardless of whether the first or second sampling is responsible for the time lag.

*Application example*

The Northern tidewater goby (*Eucyclogobius newberryi*) is a small endangered fish which lives in brackish estuaries and lagoons along the coast of California. Beyond the need to further our understanding of its ecology to protect it, this species represents an interesting model because dispersal between suitable habitat patches occurs during rare, discrete events. Previous work investigated extinction–colonization dynamics through an evaluation of genetic diversity across the landscape at several points in time [61]. The Elk River population was a population where unexpected temporal genetic change was suggested to have taken place [61]. TGI was a relevant approach to analyse this dataset as it provided a rigorous test of whether significant temporal genetic change occurred in any population of this landscape.

Using TGI, we found that the Elk River population of Northern tidewater goby [61] has indeed changed significantly relative to the other populations sampled in the study area (permutation p-value = 0.0005), even after using strict adjustments (Holm-Bonferroni adjusted permutation p-value = 0.004). The results describe a loss of genetic diversity in that population between 2006 and 2011, supporting the suggestion made in a previous study, through a proper testing framework.

**DISCUSSION**

*A new and useful framework*

TGI fills a gap: it provides a way to compare temporally-repeated samples and test whether a change in genetic diversity that is unusual relative to the ubiquitous landscape-wide change associated with genetic drift, has occurred, even when the number of markers available is limited. We succeeded in our general goal of adapting TBI to TGI, which required a general translation of a species-by-site approach to a genotype-by-site approach, and which also notably required changing the permutation algorithm to accommodate the specific structure of various genetic data formats. One main contrast between our approach and previous investigations of the performance of TBI, which used community composition data, is that we focused on the timing of sampling. Although community composition data generally varies at a larger time scale than genetic data, we encourage future investigations of the influence of timing on TBI performance. Beyond the description of our new framework, we also evaluated its power and specificity.

We investigated how dispersal, the spatial extent of a demographic event, and the timing of sampling affects our ability to identify populations having undergone significant changes in genetic diversity. Performance of our permutation-based TGI procedure was evaluated using a spatially-explicit gene flow simulator [44]. We were able to identify the scope of successful application of our TGI procedure by evaluating its capacity to correctly identify populations affected by disturbance in various demographic scenarios and with an array of different time lags between samples.

*Dispersal and spatial extent*

The ability of our method to detect temporal genetic changes is sensitive to the dispersal capacity of the organism of interest. Indeed, false negatives increased with dispersal, whereas false positives did not show a trend (Fig. 3, 4). The influence of dispersal on the false negative rate increases with increasing time lags between an event and the subsequent sampling. The effect of dispersal on the FNR is already present with only one generation separating two temporal samples (right before and right after the event), and increases as the time between samplings increases. Landscape connectivity and dispersal are functionally linked as connectivity represents the degree to which a landscape constrains dispersal [62]. Considering that we used dispersal ability as a proxy for landscape functional connectivity, this result suggests that well-connected landscapes might require more frequent sampling to overcome the negative effect of connectivity on our ability to correctly identify affected populations.

The spatial extent of an event increases our ability to correctly reject populations that have not truly changed (lower FPR), but it decreases our ability to correctly detect populations that have truly changed (higher FNR). This trade-off is apparent for all scenarios and its strength appears to depend on dispersal. Although the spatial extent of an event may help researchers detect it, as the chance of sampling an affected area increases, it may also increase the risk of not identifying the genetic legacy of the event at all, especially in high dispersal landscapes. Given the way it was designed, TGI is better suited to analyze systems in which disturbances are discrete and affect a limited part of the landscape, *i.e.* not landscape-wide and gradual disturbances. When several populations were affected, we chose adjacent populations; whether lowering the degree of spatial autocorrelation in the spatial genetic legacy, that is targeting populations not necessarily adjacent to each other, influences detection, has not been investigated in our paper. Spatial autocorrelation may greatly affect many genetic analyses, and solutions are being developed to integrate it within them [63]. We believe that explicitly taking spatial autocorrelation into account in temporal analyses of genetic diversity [5] represents a promising and challenging avenue of research.

Overall, the implications of our results regarding the influence of dispersal and spatial extent are that different performances are to be expected for demographically-different systems, and that researchers will greatly benefit from taking that into account when interpreting TGI and choosing an appropriate threshold value given their specific objectives. For example,prioritizing avoiding false negatives rather than false positives – that is making sure we detect all the affected populations no matter the cost of detecting, and monitoring/preserving, some populations which do not need it.

*Time between sampling efforts*

As expected, spatial genetic legacies tend to decay over time and TGI appears generally suitable to identify changes at a temporal scale of 1-10 years, with our landscape and demographic parameters. However, two main points emerged from our analysis of the timing of sampling required to detect significant genetic change. First, when comparing an old sample to a new sample done after a potential disturbance, the spatial extent of the disturbance will affect the power of TGI, with smaller spatial extents preserving high power even with large time gaps, which should be reassuring to potential users planning to use TGI in similar systems. Second, when comparing a sample collected right before a disturbance to one collected several years after, dispersal becomes the most important factor in driving the performance of TGI, with low dispersal scenarios better preserving the performance against the decaying assaults of time and the associated genetic drift. This means that researchers may for example have to accept as many as 10% of false positives after sampling only a few years after an event, if their system presents high dispersal. This has serious implications: in case researchers are unaware of the dynamics of their system, and arbitrarily choose an inappropriate threshold, they might systematically spend a larger part of their resources than what they expect on monitoring or treating unaffected populations, while missing some affected populations.

*Empirical application of the method*

We successfully applied TGI on an empirical dataset featuring an endangered vertebrate for which temporal genetic change had been described, but not tested. One population had been hypothesized to have undergone atypical change, relative to the rest of the landscape, in the original publication and our application of TGI supported it. We therefore showed how to strengthen the results for future temporal genetic studies using repeated samples, through the straightforward TGI testing procedure. Different empirical datasets, and different research objectives may require TGI users to tweak our procedure. The TGI function is transparent and flexible, and using different permutation and genetic distance algorithms would be straightforward. This could be done by simply changing a few lines in the commented TGI function we provide as supplementary material.

*Considerations about the use of TGI*

Our TGI procedure has a now demonstrated place in a researcher’s arsenal to study genetic change, but there are certain considerations to keep in mind when using it. For example, stricter values (lower values) for the TGI *p*-value threshold expectedly bring a better FPR but may also bring a worse 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 systems. TGI can also readily be used on other types of genetic data, such as microsatellites.

TGI provides a way to go beyond arbitrarily comparing pairwise genetic differentiation, or node-based genetic diversity values. TGI was not designed as an alternative to sophisticated and well-established methods seeking to infer demography from large genetic datasets collected at a single time. Instead, it was designed to help research teams possessing repeated samples for non-model organisms with limited genotypic information, especially those wanting to compare new samples to older ones. Nonetheless, more work is needed to explore how the performance of TGI varies with other factors such as: 1) genetic distance algorithm; 2) spatial autocorrelation in genetic legacies; 3) effective population size; and 4) spatial heterogeneity in landscape resistance to movement.

Successful implementation of TGI in new systems can benefit from *a priori* understanding of their population dynamics and the nature and scale of possible disturbances in the study area. Indeed, this could help researchers better pick a range of useful threshold values to use. These values ultimately represent trade-offs in potential conservation costs and it important that researchers grasp their importance. Simulation is a powerful tool for investigating how demography and spatial context influence population genetic dynamics [64], and can be used to help identify those appropriate threshold values, as we have demonstrated here. Several programs such as *CDMetaPOP* [44], *Nemo* [65], *SPLATCHE* [66], or *SLIM* [67] provide very flexible and sophisticated ways to implement such simulations. We expect greater sensitivity to threshold selection in systems that exhibit dramatic demographic fluctuations, as is the case in outbreaking or invasive species. Those systems would therefore greatly benefit from of a simulation study designed to pick the most appropriate threshold for the analysis of empirical data, instead of using arbitrary threshold.

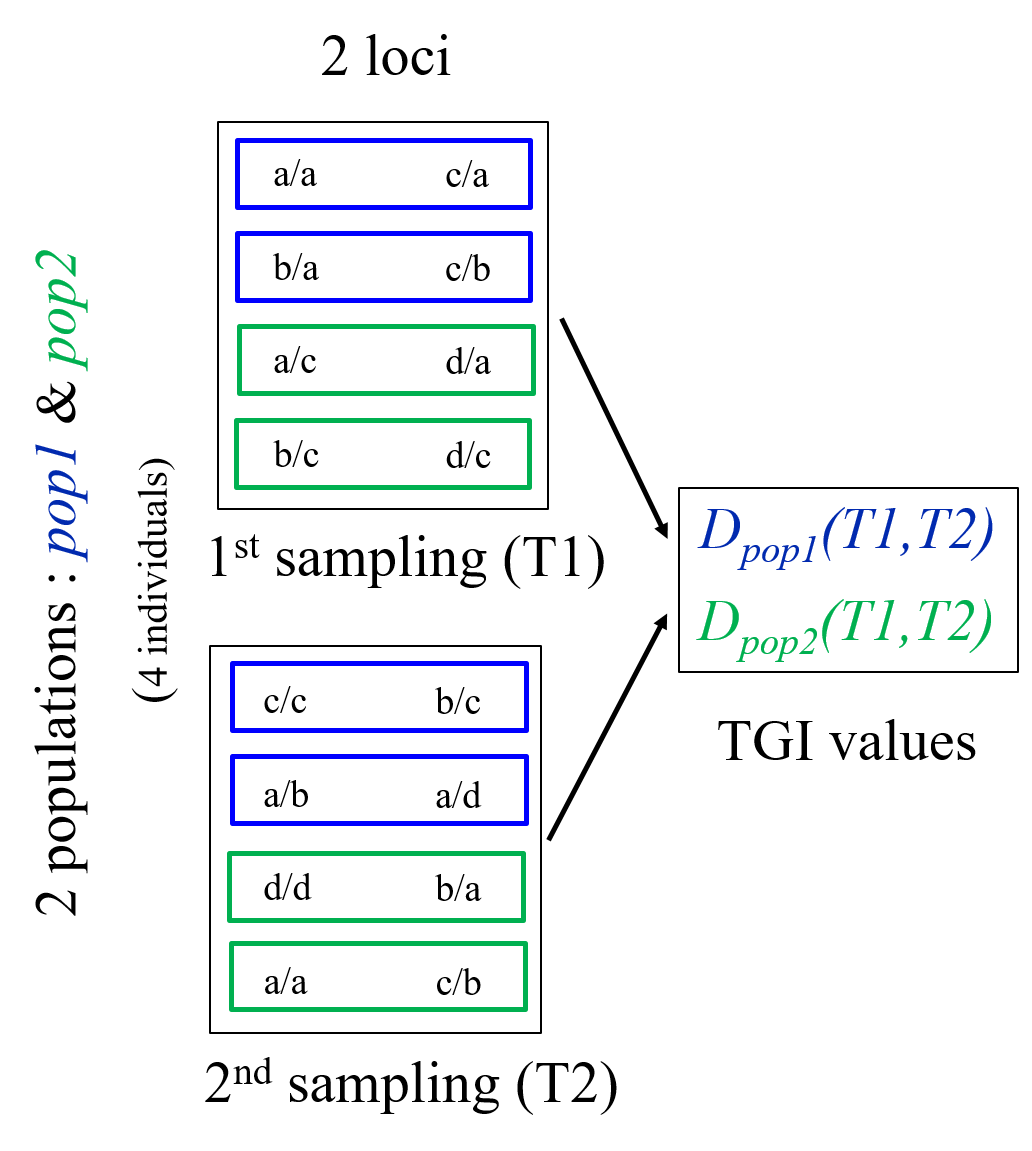
*Conclusions*

Identifying changes in genetic diversity, beyond what one would expect due to background micro-evolutionary processes, can help researchers and conservation managers identify locations that have experienced important past demographic events. These events could be detrimental (*e.g.* loss of diversity, maladaptation) or beneficial (*e.g.* higher effective population size, genetic rescue). Such sites could then be prioritized for increased monitoring and further investigation into the origin of these changes. As shown in our application of TGI on an endangered animal system, our method provides a direct test of the hypothesis of exceptional temporal genetic change. Our approach to detecting temporal genetic differentiation does not require extensive genomic information and can be used to explore the temporal dynamics of genetic diversity changes using relatively small genetic datasets (*e.g.* hundreds of SNPs). As such, our approach holds great promises to facilitate spatio-temporal analysis of wild, non-model organisms for which extensive genomic resources are yet to be developed.

**TABLES AND FIGURES**

**Table 1**

|  |  |  |  |
| --- | --- | --- | --- |
| *Dispersal (B)*  *No. populations* | **Low** (720) | **Moderate** (720) | **High** (720) |
| **1** (540) | L1 (180) | M1 (180) | H1 (180) |
| **2** (540) | L2 (180) | M2 (180) | H2 (180) |
| **3** (540) | L3 (180) | M3 (180) | H3 (180) |
| **0: control** (540) | CL (180) | CM (180) | CH (180) |

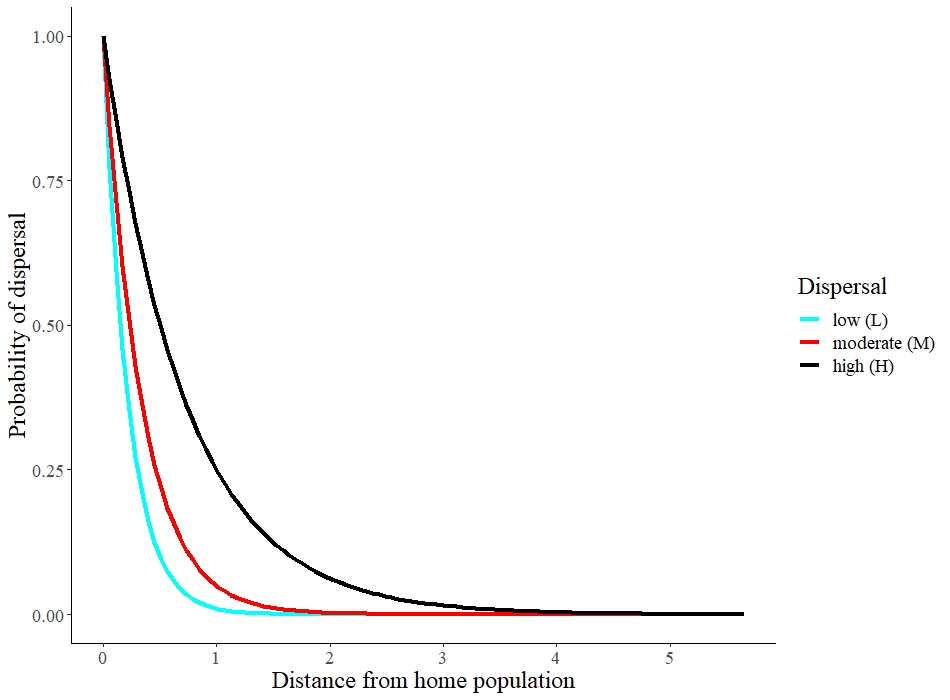
****

A)

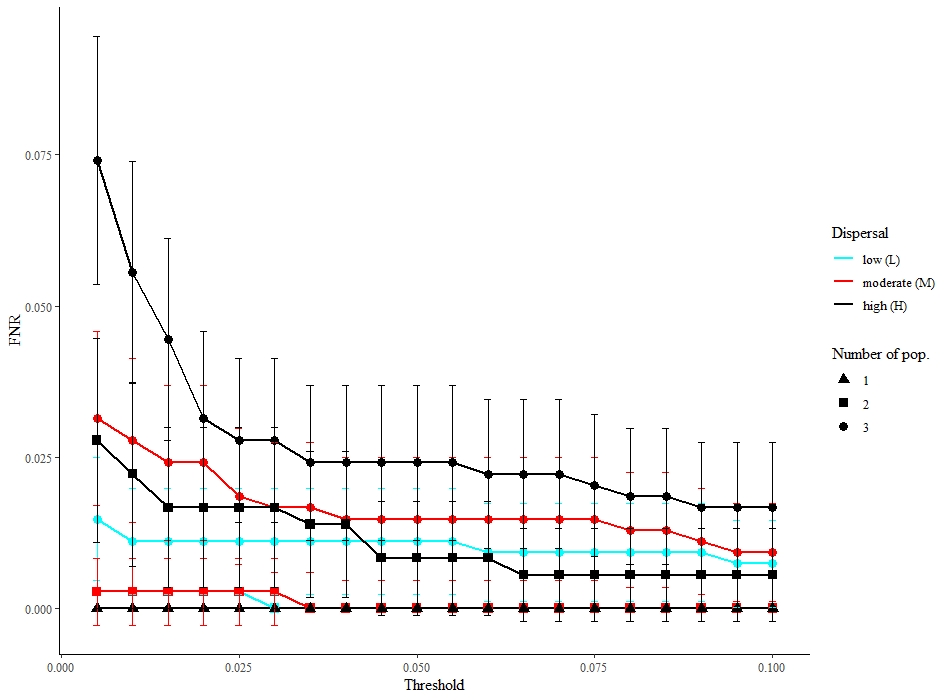
****

B)

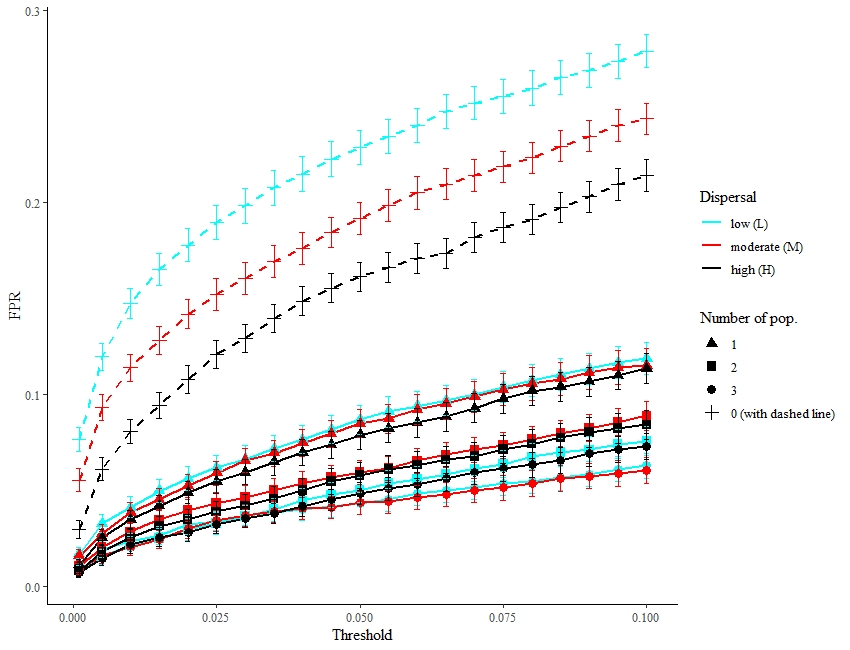
**Figure 1.**



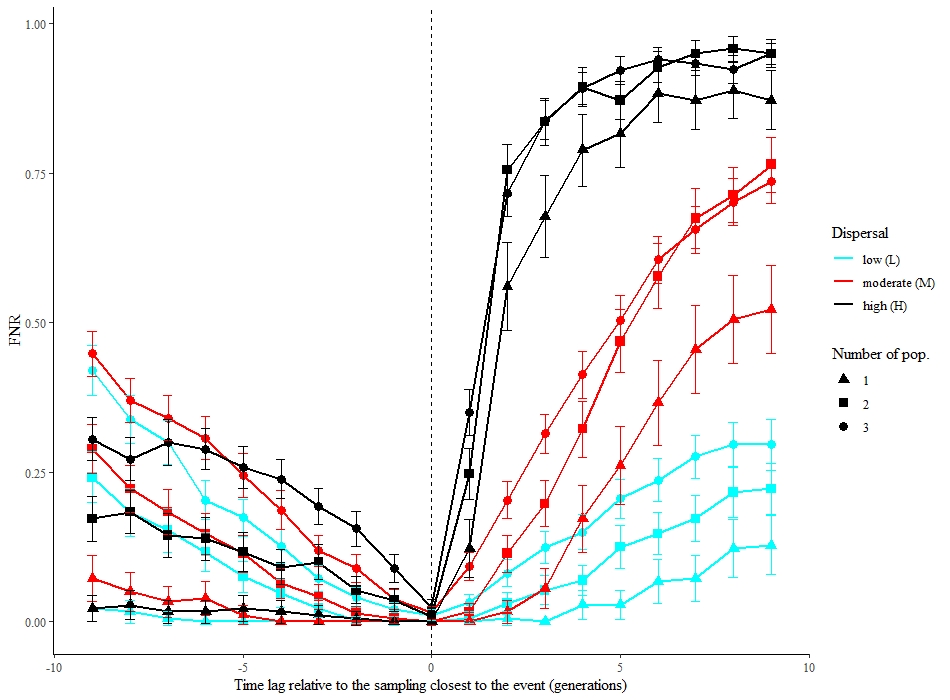
**Figure 2.**



**Figure 3.**

****

**Figure 4.**

****

Influence of the timing of the posterior sampling

(0 represents sampling right after the event)

Influence of the timing of the prior sampling

(0 represents sampling right before the event)

**Figure 5**.

****

Influence of the timing of the posterior sampling

(0 represents sampling right after the event)

Influence of the timing of the prior sampling

(0 represents sampling right before the event)

**Figure 6.**