VOL. 191, NO. 4 THE AMERICAN NATURALIST APRIL 2018

# Inferring Causalities in Landscape Genetics: An Extension of Wright's Causal Modeling to Distance Matrices

**Lisa Fourtune,[1,*] Jérôme G. Prunier,[1] Ivan Paz-Vinas,[2,3] Géraldine Loot,[1,4] Charlotte Veyssière,[4] and Simon Blanchet[1,3]**

1. Station d'Ecologie Théorique et Expérimentale, Unité Mixte de Recherche (UMR) 5321, Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), 09200 Moulis, France; 2. Institut Mediteranéen de la Biodiversité et d'Ecologie marine et continentale, UMR 7263, CNRS, Université Aix-Marseille, Institut de Recherche pour le Developpement, Université d'Avignon, Centre Saint-Charles, Case 36, Marseille, France; and Laboratoire d'Ecologie des Hydrosystèmes Naturels et Anthropisés, UMR 5023, CNRS, Université de Lyon, École Nationale des Travaux Publics de l'Etat, 6 rue Raphaël Dubois, 69622 Villeurbanne, France; 3. Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174, CNRS, UPS, École Nationale de Formation Agronomique, 31062 Toulouse cedex 4, France; 4. EDB, UMR 5174, Université de Toulouse, UPS, 31062 Toulouse cedex 4, France

*Submitted February 5, 2017; Accepted November 1, 2017; Electronically published February 14, 2018*

*Online enhancements:* appendix figure.

ABSTRACT: Identifying landscape features that affect functional connectivity among populations is a major challenge in fundamental and applied sciences. Landscape genetics combines landscape and genetic data to address this issue, with the main objective of disentangling direct and indirect relationships among an intricate set of variables. Causal modeling has strong potential to address the complex nature of landscape genetic data sets. However, this statistical approach was not initially developed to address the pairwise distance matrices commonly used in landscape genetics. Here, we aimed to extend the applicability of two causal modeling methods—that is, maximum-likelihood path analysis and the directional separation test—by developing statistical approaches aimed at handling distance matrices and improving functional connectivity inference. Using simulations, we showed that these approaches greatly improved the robustness of the absolute (using a frequentist approach) and relative (using an information-theoretic approach) fits of the tested models. We used an empirical data set combining genetic information on a freshwater fish species (*Gobio occitaniae*) and detailed landscape descriptors to demonstrate the usefulness of causal modeling to identify functional connectivity in wild populations. Specifically, we demonstrated how direct and indirect relationships involving altitude, temperature, and oxygen concentration influenced within- and between-population genetic diversity of *G. occitaniae*.

*Keywords:* causal modeling, landscape genetics, path analysis, d-sep test, pairwise data.

* Corresponding author; e-mail: lisa.fourtune@gmail.com.

**ORCID:** Prunier, http://orcid.org/0000-0003-4110-2567; Paz-Vinas, http://orcid.org/0000-0002-0043-9289.

Am. Nat. 2018. Vol. 191, pp. 000–000. 
DOI: 10.1086/696233

## Introduction

Landscape genetics is a discipline aimed at understanding spatial patterns of genetic diversity by exploring the relationships between landscape features and microevolutionary processes such as genetic drift, selection, mutation, and gene flow (Manel et al. 2003; Manel and Holderegger 2013). This discipline builds on the latest advances in molecular biology and landscape data processing and is becoming increasingly important for fundamental and applied sciences (Storfer et al. 2010; Keller et al. 2015). Landscape genetics addresses issues ranging from the identification of barriers to dispersal, to the inference of the spread of nonnative species (Storfer et al. 2010).

The main objectives of landscape genetics are to spatially describe effective dispersal (i.e., gene flow) and to identify landscape features (e.g., roads, dams, urban areas, and rivers) that affect functional connectivity (Manel et al. 2003; Storfer et al. 2010; Manel and Holderegger 2013). To achieve these objectives, landscape geneticists calculate genetic descriptors that are subsequently compared with landscape features and potential dispersal barriers (Balkenhol et al. 2009; Jaquiéry et al. 2011; Bradburd et al. 2013). Analytical tools developed for analyzing landscape genetic data often rely on empirical correlations that allow an assessment of the possible influence of various evolutionary processes. For example, a significant and positive correlation between genetic and geographic distances is generally considered indicative of isolation by distance (IBD; a spatial pattern whereby the homogenizing effect of gene flow decreases and the relative effect of genetic drift increases as the geographic distance between sites increases; Hutchison and Templeton 1999).

However, because correlation does not imply causation, processes can be incorrectly inferred from empirical correla-

tions (Guillot et al. 2009). The likelihood of incorrectly inferring causalities from correlation is exacerbated in landscape genetics because it often implies intricate relationships among landscape variables. In the IBD example described above, the correlation between genetic and geographic distances might be direct, indirect, and/or spurious. The correlation is direct (i.e., the migration rate between two sites decreases because they are far from one another) if no other variable covarying with geographic distances causes the observed pattern of genetic distance. However, if a variable covaries with geographic distances (e.g., the number of barriers between two sites) and causes the observed pattern, then the correlation between genetic and geographic distances is considered indirect. Alternatively, a correlation is spurious when two variables are correlated because they are both influenced by a third (unmeasured) variable (Cushman and Landguth 2010; Prunier et al. 2015). In the two latter cases, processes are incorrectly inferred from simple correlations. Consequently, the relationships linking landscape features to genetic descriptors have to be carefully interpreted, and they sometimes remain unexplained due to our inability to disentangle intricate relationships between variables (Shipley 2000*a*; Grace 2006). Clarifying causal relationships in landscape genetics is thus challenging but important (Guillot et al. 2009).

A solution to improve inferences of causal relationships in landscape genetics may build on methods of causal modeling (e.g., Cushman et al. 2006). Causal modeling procedures, such as path analysis (Grace 2006), rely on the assessment of the validity of a causal graph describing the expected direct and indirect causal relationships among variables. Path analysis was initially developed by one of the founding fathers of population genetics, namely, Sewall Wright (1921). In path analysis, the influence along each path of the causal graph (i.e., the link between two variables) is estimated from correlation/covariance among the involved variables. Almost a century after its introduction by one of the most influential population geneticists, and despite its relevance for analyzing complex observational data, path analysis is still only occasionally used in landscape genetics and in population genetics in general.

Landscape geneticists generally focus on two main types of dependent variables that describe genetic diversity: (i) point summary statistics, which describe the genetic diversity at the sampling-site level (e.g., allelic richness or heterozygosity), and (ii) pairwise summary statistics, which describe the genetic differentiation (or distance) between pairs of sampled populations or individuals (e.g., $F_{ST}$, Jost's $D$). Several well-established methods allow a straightforward analysis of point summary statistics in a path analysis framework (Shipley 2000*a*; Grace 2006). For pairwise statistics, however, the process is more complex since the analysis of pairwise matrices poses a series of analytical issues, notably because of the non-independence of pairwise data (Legendre and Legendre 2012; Graves et al. 2013). Although pairwise data can be handled by reducing multidimensionality, using nonmetric multidimensional scaling or distance-based redundancy analysis, for instance (e.g., Legendre and Fortin 2010), these types of analyses were not developed to tease apart direct and indirect relationships and are more suited to answer questions involving dissimilarity matrices rather than distance matrices (Legendre and Fortin 2010; Legendre et al. 2015). To address this specific data type, Cushman et al. (2006, 2013) proposed a causal modeling procedure based on partial Mantel tests to compare several competing causal models that link a matrix of genetic distances to matrices of explanatory variables. This approach permits an assessment of the goodness-of-fit of each model by independently comparing the observed results of partial Mantel tests (partial correlation coefficients and associated $P$ values) to what is theoretically expected under each model specification. This approach has been proven to be powerful for inferring causalities from relatively simple models (Cushman and Landguth 2010). However, the design of the causal graph is constrained by the number of matrices of explanatory variables that can be handled in partial Mantel tests (only two), which limits the complexity of competing models and prevents the assessment of indirect relationships among variables. We believe that the use of alternative causal modeling procedures, such as maximum-likelihood-based path analysis (hereafter, "path analysis" for the sake of simplicity) and the directional separation test (hereafter, "d-sep test"; Shipley 2000*a*, 2000*b*), can represent an interesting improvement over the approach proposed by Cushman et al. (2006), as they may simultaneously account for all correlations implied in a model and permit the design (and comparison) of more complex models, explicitly addressing both direct and indirect effects.

We propose a simple and integrative framework to study direct and indirect links in the context of the analysis of landscape genetic data (and, more generally, of ecological and evolutionary data involving pairwise matrices). As an introduction, we briefly present the philosophy, advantages, and disadvantages of path analysis and the d-sep test. Then, we extend the applicability of these two methods to pairwise matrices (including distance and dissimilarity matrices) by developing two statistical approaches aimed at analyzing complex causal models (i.e., including several pairwise matrices linked both directly and indirectly) in landscape genetics. We then test the robustness of path analysis and the d-sep test applied to pairwise matrices using simulations. Finally, we use an empirical data set involving patterns of genetic diversity in a freshwater fish species (*Gobio occitaniae*) and landscape descriptors at the river basin scale to demonstrate how these two statistical procedures can be used in landscape genetics to answer important biological questions. This study provides an opportunity to reconcile two important legacies of Sewall Wright's scientific life: population genetics and path analysis.

## A Brief Description of Path Analysis and the D-Sep Test

### *An Introduction to Causal Graphs*

Any causal modeling procedure is based on a causal graph illustrating the a priori hypotheses underlying the potential causal relationships within a set of variables. These relationships are depicted by vertices (i.e., nodes) representing variables that are linked by edges. A causal graph can contain manifest variables, which are directly observed and measured (Shipley 2000*a*); error variables, which represent all of the factors that are not considered in the current graph; and latent variables, which are hypothesized to exist but have not been measured directly (Grace 2006). Causal graphs are an intuitive approach to translate a causal hypothesis into a statistical language. The next step is to statistically test the relevance of the causal model in relation to data. Here, we focused on path analysis and the d-sep test, two methods dedicated to testing causal models without latent variables (Shipley 2000*a*; Grace 2006). These two methods are described below. When the causal graphs contain latent variables, the dedicated method is called structural equation modeling (SEM; Grace 2006), which is a generalization of path analysis. This method will not be presented here.

### *Path Analysis*

Path analysis is based on maximum-likelihood estimation (Fisher 1950) of model parameters through the computation of covariance matrices. Each causal model includes a set of parameters, some of which are known (e.g., variances and covariances of variables), whereas others are unknown (e.g., path coefficients that quantify the direct influence of a variable along a given path; Wright 1921). The first step is to infer values for these unknown parameters. This inference is made iteratively by computing a maximum-likelihood fitting function ($F_{ML}$; Bollen 1989) that quantifies the difference between the observed covariance matrix and a covariance matrix computed using the inferred values. The best parameter values are those that minimize this function. The absolute fit of the model can be assessed by computing a $\chi^2$ statistic and an associated $P$ value to determine whether the minimal value of $F_{ML}$ is small enough to conclude that the observed data fit the hypothesized causal model; a high $P$ value indicates a high probability that the observed data fit the hypothesized causal model. Additionally, the relative fits of competing models can be tested using an information-theoretic approach (e.g., using Akaike's information criterion [AIC]; Bollen 1989).

Path analysis requires linear relationships between variables—preferentially, multivariate normal data (Shipley 2000*a*; Grace 2006)—and assumes that observations are independent, which is notoriously not the case when considering pairwise matrices (Legendre and Legendre 2012). Because of this latter limitation, path analysis is not frequently used in landscape genetics.

### *D-Sep Test*

Shipley's d-sep test simultaneously tests for conditional independence relationships that should be true if the causal model is verified. If these conditional independence relationships do not exist in the empirical data, then the causal hypothesis is rejected. These relationships are identified using the directional-separation (d-separation) criterion (Pearl and Verma 1987; Pearl 1988; Shipley 2003). However, there are usually far too many d-separation relationships to test all of them. The principle of the d-sep test is, hence, to identify a basis set of mutually independent d-separation relationships that together imply all others (Pearl 1988; Shipley 2000*b*). Once this basis set is identified, each of these $k$ independence claims has to be tested against the empirical data. This can be achieved through the use of Pearson's partial correlation coefficients or linear regressions, if the variables are normally distributed and are linked by linear relationships, as well as using more complex statistical methods in other cases. The $k$ $P$ values obtained are equivalent to the probability levels of the data, given each of the $k$ d-separation relationships. If all these tests are mutually independent, the $k$ $P$ values can be combined using the equation

$$C = -2\sum_{i=1}^{k} \ln(P_i) \qquad (1)$$

(Fisher 1938).

If all the independence relationships hold in the data, this statistic follows a $\chi^2$ distribution with $2k$ degrees of freedom. The resulting test is called Fisher's $C$ test (Shipley 2000*b*). A large $C$ value, and thus a small resulting $P$ value, implies a poor absolute fit of the data to the model. In path analysis, the relative fits of competing models can also be assessed through the use of AIC adapted to the d-sep test (Cardon et al. 2011; Shipley 2013).

As the d-sep test does not impose any inference of parameters, the conditions for its application are flexible: it can, for instance, be applied to data sets with small sample sizes. Importantly, two nodes in a causal graph that are d-separated will also be conditionally independent in any data set generated by this graph, irrespective of the distribution of the variables (Shipley 2000*a*; Pearl 2009). This means that different modeling approaches (e.g., linear or nonlinear models, Bayesian models, hierarchical models; Shipley 2009; Cardon et al. 2011) can be used for testing the conditional independence relationships provided these tests are appropriate for the type of variables involved in the d-sep claims (Shipley 2009). The d-sep test is, therefore, a flexible method that

cannot be used directly to infer path coefficients, although estimates can be computed from a combination of independent models.

### *A Note about the Use of* P *Values and AIC in the Context of Causal Modeling*

In the context of causal modeling, *P* values and AIC provide different—yet complementary—information. While AIC values allow the identification of the best-fitting model among a set of candidate models (the relative fit), *P* values provide information about the absolute fit of the empirical covariance matrix for a given model. This means that the best-fitting model—with the lowest AIC—may be a poorly fitting model if diagnosed through the inspection of *P* values. We therefore encourage use of both the AIC and *P* values to infer the causal structure of data. On a philosophical side note, and following Goodman (1999), we here chose not to set any significance threshold (e.g., $\alpha = 0.05$): *P* values are hereafter interpreted as the probability of obtaining a result equal to, or more extreme than, what was actually observed under the null hypothesis.

## Extending Path Analysis and the D-Sep Test to Pairwise Matrices

We hereafter present four statistical approaches aimed at applying path analysis and the d-sep test to the analysis of causal models involving pairwise matrices. Fully usable R functions (R Development Core Team 2017) are available online at https://doi.org/10.5281/zenodo.1048975.

### *Path Analysis Applied to Pairwise Matrices*

To take into account the nonindependence of pairwise data in path analysis, we used the maximum-likelihood population effects (MLPE) approach developed by Clarke et al. (2002; see also Van Strien et al. 2012). In MLPE models, identities of the two sites involved in a pairwise comparison are treated as two random factors to take into account the spatial dependency of pairwise data: each site is associated with a random deviation from the intercept, and any pairwise values sharing a common source site thus share a common random deviation. To apply path analysis to pairwise matrices, we used the lavaan.survey R package (Oberski 2014) that was initially developed to use SEM and path analysis with hierarchically structured data. In this approach (hereafter, "clustering-based path analysis"), the identities of sites involved in a pairwise comparison were treated as clusters (i.e., each pairwise value was associated with two clusters), with the second cluster being nested within the first. As a result, all pairwise values originating from the same first cluster will share a common random deviation from the intercept, although each pairwise value will also be attributed a unique random effect associated with the second cluster being nested within the first. To prevent some sites from being more influential than others in the computation of the first random deviation, the identities of the two sites involved in a given pairwise comparison are randomly permuted. This approach allows the assessment of both the relative and absolute fits of a model, as well as the *P* values associated with the path coefficients, while partially taking into account the nonindependence of pairwise data.

As an alternative approach to assess the *P* values associated with path coefficients, we used a permutation procedure aimed at randomly permuting rows and columns of input matrices (Legendre 2000). This procedure provides the theoretical distribution of a given statistic (e.g., a Mantel correlation) under the null hypothesis of the absence of relationships between variables. An unbiased *P* value can then be computed by comparing the observed value of the statistic to its null distribution. This approach (hereafter, "permutation-based path analysis") was not used to quantify the probability that the data fit the model, as the null distribution corresponds to a scenario in which none of the paths are true; rather, we aimed to test whether only the defined paths are true. This approach involves the following four steps. First, all matrices are independently permuted many times. Second, the values of the unknown parameters (i.e., path coefficients linking permuted matrices according to the considered causal model) are inferred by minimizing the difference between the covariance matrix computed from permuted data and the observed (optimized) covariance matrix to create a set of null causal models. The third and fourth steps consist of creating null distributions for the parameters of interest (here, values of the path coefficients) and computing unbiased *P* values, respectively. The one-tailed *P* value of each path coefficient is then computed as the proportion of permuted path coefficient values lower than or equal to (respectively, greater than or equal to) the observed path coefficient value (Legendre 2000).

Finally, we built a parametric bootstrap procedure to quantify the range of values (i.e., confidence intervals) that act as good estimates of each unknown parameter value, while taking into account the presence of pairwise matrices in the path analysis. This procedure is based on sampling, with simultaneous replacement of rows and columns, performed numerous times. Parameter values are estimated each time through path analysis, and 95% confidence intervals are provided as the 2.5 and 97.5 percentiles of these bootstrapped parameters. This method can be applied only to standardized data.

### *The D-Sep Test Applied to Pairwise Matrices*

Building on the flexibility of the d-sep test, we developed an approach that allowed the proper application of this method

to distance matrices. This approach takes the form of a new R function called dsep.test, permitting the use of the d-sep test for both point summary and pairwise data.

In the first step, the basis set of independent d-separation relationships implied by a given causal model is determined using the basiSet function from the R package ggm (Marchetti 2006). In the next step, each conditional relationship of independence is tested using multiple regressions applied to distance matrices (MRM; Smouse et al. 1986), a permutation-based method classically used to infer parameters and *P* values from regressions that can involve more than two pairwise matrices as explicative variables (contrary to the partial Mantel test). We used the MRM function from the R package ecodist (Goslee and Urban 2007). For each tested model, the *P* values can then be obtained for each independence claim of the basis set and used as the $P_i$ values in formula (1) to test the absolute fit of the model. This new approach is hereafter called the permutation-based d-sep test for the sake of clarity.

In addition, we implemented an automatic calculation of the AIC score related to the tested model to test for the relative fits of competing models. AIC was only recently developed for the d-sep test (Cardon et al. 2011; Shipley 2013), and the R function has not been implemented in the ggm package.

## Testing the Reliability of Path Analysis and the D-Sep Test Applied to Pairwise Matrices: A Simulation Test

### *General Approach*

To test the reliability of path analysis and the d-sep test to take into account the nonindependence of pairwise data, we simulated 1,000 data sets consisting of 50 sites each, and each site was associated with five observations that were independently drawn from five normal distributions. These variables were separated into two independent variables ($X_1$ and $X_2$) and three response variables ($X_3$, $X_4$, and $X_5$). The response variables were calculated as linear combinations of one or two variables plus random noise (SD $=$ 2). The causal model structure was held constant across simulations, but linear coefficients were randomly selected from a uniform distribution ranging from 0.8 to 1.6 for each simulation (scripts are available online at https://doi.org/10.5281/zenodo.1048872). From the five variables, we computed five Euclidean distance matrices (1,225 pairwise values). With this procedure, distance matrices within simulated data sets were connected by four predefined causal links (fig. 1*A*). We then used these simulated data sets to test (i) the reliability of the *P* values and AIC scores to detect adequate causal models among different model structures and (ii) the reliability of *P* values and confidence intervals for specific path coefficients in the case of path analysis applied to pairwise matrices only.

First, we tested the ability of the new approaches to detect adequate causal models. We tested three types of models, each of which included four paths (fig. 1*A*); the first type of model fitted the four causal relationships predefined in the simulated data sets (hereafter, "adequate model"), the second type of model fitted two of the four causal relationships (hereafter, "intermediate model"), and the third type of model fitted none of the four causal relationships (hereafter, "inadequate model"). For each of the 1,000 simulated data sets, the three types of models were tested using the path analysis/d-sep test applied to point summary statistics, as well as the clustering-based path analysis and the permutation-based d-sep test applied to pairwise matrices. Clustering-based path analysis and the permutation-based d-sep test can be considered robust to assess the relative fit of each model if the AIC score of the adequate model calculated using these approaches is lower than the AIC scores calculated for the intermediate and inadequate models. We used the ΔAIC (the difference between the AIC of the considered model and the AIC of the best-fitting model; Burnham and Anderson 2002) as a measure of the relative support of each model relative to the best-fitting model (with $\Delta\text{AIC} < 2$ and $\Delta\text{AIC} < 4$ as thresholds). Additionally, clustering-based path analysis and the permutation-based d-sep test can be considered more accurate than the approaches classically used for point-summary statistics to assess the absolute fit of the data to the model if the *P* value computed for the adequate model is higher than the *P* values computed using the classical approaches. For the inadequate model, we expected both the classical approaches and the approaches developed for pairwise matrices to generate low *P* values.

Second, we tested the ability of clustering-based path analysis and permutation-based path analysis to compute reliable *P* values for the path coefficients of a given model, as well as the ability of the parametric bootstrap procedure developed for pairwise matrices to provide more reliable confidence intervals than bootstrap procedures not taking into account the nonindependence of pairwise data. To do so, we built a model combining six paths (fig. 2*A*): three of the paths fitted the causal relationships predefined in the simulated data sets (i.e., path coefficients of the adequate model; hereafter, "adequate coefficients"), whereas the other three did not (i.e., path coefficients from the inadequate model; hereafter, "inadequate coefficients"). This model was fitted using classical path analysis, clustering-based path analysis, and permutation-based path analysis, and confidence intervals were assessed using the parametric bootstrap procedure developed for pairwise matrices. The *P* values of each path coefficient were obtained across the 1,000 simulated data sets. In methods designed to account for the nonindependence of pairwise data, the *P* values of the inadequate coefficients should generally be high, whereas the *P* values of the adequate coefficients should generally be low. On the contrary, the classical

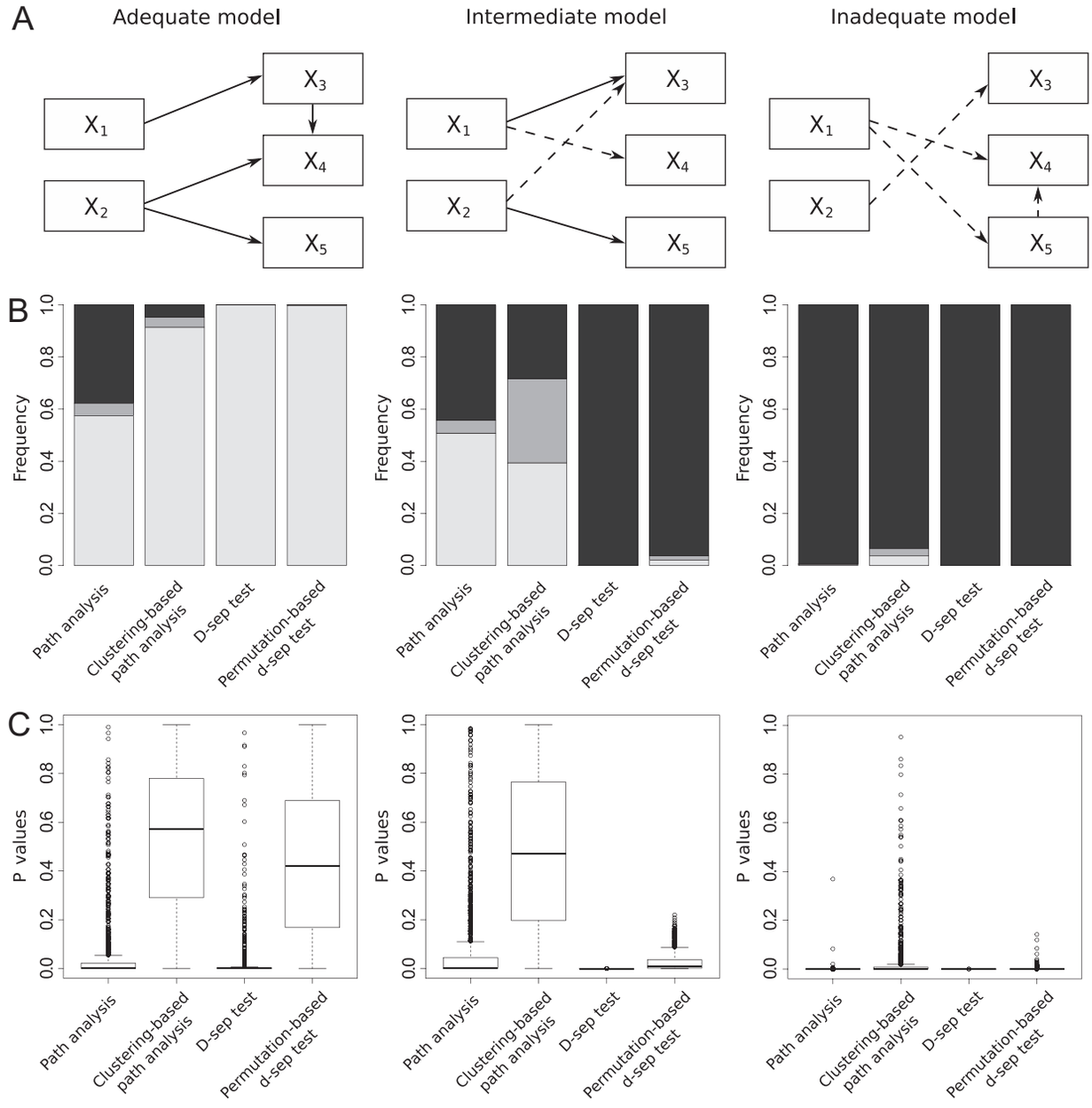


**Figure 1:** Simulated pairwise data were used to test the reliability of the new causal approaches developed for pairwise matrices to detect adequate causal models. *A*, Simulated data were generated according to a given causal scenario (including four paths), and we fitted a first causal model accounting for the four causal relationships predefined in the simulated data sets (adequate model), a second causal model accounting for two of the four causal relationships (intermediate model), and a third causal model accounting for none of the four causal relationships (inadequate model). We expected the adequate model to have better relative and absolute fits to the simulated data than the intermediate and inadequate models when using causal approaches developed for pairwise data. *B*, For each statistical approach (approaches that did not account for pairwise data: path analysis and d-sep test; approaches accounting for pairwise data: clustering-based path analysis and the permutation-based d-sep test), the relative fit of each model to the simulated data is provided as the frequencies of the difference between the Akaike's information criterion (AIC) of the considered model and the AIC of the best-fitting model ($\Delta$AIC) values (dark gray: $\Delta AIC > 4$; medium gray: $2 < \Delta AIC < 4$; light gray: $\Delta AIC < 2$). *C*, For each statistical approach, the absolute fit of each model to the simulated data is provided as boxplots of the *P* values obtained over 1,000 simulations. The solid line within each box indicates the median; the length of the box is the interquartile range (from the first to the third quartile). The lower whisker extends to the first quartile minus 1.5 times the interquartile range; the upper whisker extends to the third quartile plus 1.5 times the interquartile range. Small circles represent data points located beyond whiskers.

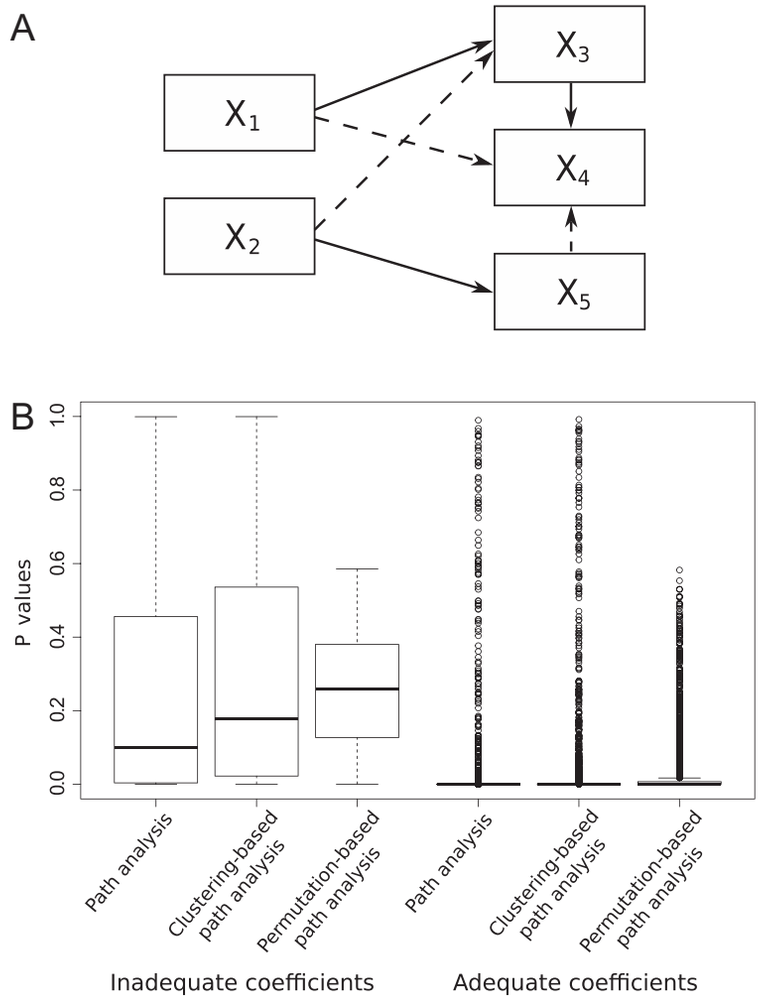


**Figure 2:** Simulated pairwise data were used to test the reliability of the new causal approaches developed for pairwise matrices to properly estimate the fit of the path coefficients. *A*, Simulations were generated according to a predefined scenario, and both path coefficients corresponding to this predefined scenario (adequate paths; black arrows in the graph) and path coefficients not defined in this predefined scenario (inadequate paths; dotted arrows in the graph) were estimated and tested using classical path analysis, clustering-based path analysis and permutation-based path analysis. *B*, Boxplots summarizing the *P* values of the adequate and inadequate coefficients obtained over 1,000 simulations using classical path analysis, clustering-based path analysis, and permutation-based path analysis. For the sake of clarity, the *P* values of the three adequate and three inadequate coefficients are combined. See the figure 1*C* legend for details.

approach should consistently provide low *P* values irrespective of the coefficient. Similarly, the 95% confidence intervals of adequate coefficients calculated using the parametric bootstrap procedure developed for pairwise matrices should not include zero, whereas the confidence intervals of inadequate coefficients should include zero.

### *Results*

The AIC scores computed using clustering-based path analysis were more reliable than the AIC scores computed using classical path analysis to assess the relative fits of competing models (fig. 1*B*). Notably, the adequate model was identified as one of the best-fitting models in 91.4% of the simulations ($\Delta\text{AIC} < 2$; light gray bars in fig. 1*B*) when using clustering-based path analysis, compared with only 57.5% of the simulations when using classical path analysis. Additionally, the intermediate model was identified as one of the best-fitting models in only 39.4% of the simulations when using clustering-based path analysis, compared with 50.8% when using classical path analysis. However, $\Delta$AIC of the intermediate model ranged from 2 to 4 units in 32.2% of the simulations when using clustering-based path analysis (medium gray bars in fig. 1*B*), indicating that this method penalized the absence

of a part of the causal structure by only a small increase in the AIC score, and thus there was only a small decrease in relative fit. Interestingly, the relative fits of all three models were almost always correctly estimated using both the permutation-based d-sep test and the classical d-sep test; for example, $\Delta AIC < 2$ was observed in 99.2% and 100% of the simulations, respectively, using the adequate model (dark gray bars in fig. 1*B*).

Values of *P* assessing the absolute fit of the adequate model were strikingly higher when clustering-based path analysis and the permutation-based d-sep test (median of the 1,000 simulated *P* values = 0.422 and 0.571, respectively) were used compared with the classical approaches (median < 0.001 for both the classical path analysis and d-sep test; fig. 1*C*). This indicates that when the null hypothesis is true, the *P* value appears much more effective when taking into account the nonindependence of pairwise data. Clustering-based path analysis and the permutation-based d-sep test hence appeared far more reliable than classical approaches to assess the absolute fit of a causal graph. The *P* values of the intermediate model were high when clustering-based path analysis (median = 0.471) was used but were low when the permutation-based d-sep test (median = 0.010) was used, indicating a difference in the sensitivity of these two methods to detect models that do not perfectly reflect the causal structure underlying the data; the permutation-based d-sep test offers higher sensitivity than clustering-based path analysis. As expected, the *P* values of the inadequate model computed using any of the approaches were very low (medians < 0.001 in all cases; fig. 1*C*).

Clustering-based path analysis and permutation-based path analysis were also more reliable in estimating the *P* values of path coefficients than classical path analysis (fig. 2*B*). Indeed, accounting for the nonindependence of pairwise data led to an increase in the *P* values of inadequate coefficients (median of 0.179 and 0.259 with clustering-based path analysis and permutation-based path analysis, respectively; fig. 2*B*), whereas the *P* values of adequate coefficients remained low (median smaller than 0.001 for all methods). Additionally, the parametric bootstrap procedure developed for pairwise matrices greatly improved the reliability of the 95% confidence intervals of inadequate coefficients: only 4% of confidence intervals computed for inadequate coefficients did not include zero when taking into account the pairwise matrix structure of the data. However, the confidence intervals of adequate coefficients computed using this procedure included zero in 28.5% of the simulations.

## Empirical Illustration of Path Analysis and the D-Sep Test

We used a data set involving a freshwater fish species (the gudgeon *Gobio occitaniae*) to illustrate how path analysis and the d-sep test can be used when the causal models include either point summary statistics or pairwise matrices. Our aims were (i) to unravel direct and indirect relationships between landscape features (and associated processes) and genetic diversity in dendritic river networks and (ii) to disentangle the relative effects of natural versus anthropogenic factors on spatial patterns of genetic diversity in gudgeon.

### *Data Collection*

*Genetic Data.* A total of 92 sites scattered across the whole Garonne-Dordogne river catchment (southwestern France; see fig. 3*A*) were sampled, and a maximum of 30 gudgeons per site were caught by electrofishing during spring 2010 and 2011. Nine out of these 92 sites were discarded from the analysis because fewer than 10 samples were available. The final database included 1,928 individuals sampled from 80 sites distributed across 32 rivers. For each individual, a small piece of pelvic fin was collected and preserved in 70% ethanol. DNA was extracted using a salt-extraction protocol (Aljanabi and Martinez 1997), and individuals were genotyped for eight microsatellite loci, as described in Blanchet et al. (2010). Neither departures from Hardy-Weinberg equilibrium nor null alleles were detected for any of these loci (Fourtune et al. 2016). Eight samples were not successfully genotyped and were removed from the database.

Genetic diversity at the sampling-site level was assessed using Fstat 2.9.3 (Goudet 2001) by computing the standardized allelic richness (Ar), that is, the expected mean number of alleles (over all loci) in a random subsample of *N* individuals at each sampling location, where *N* is the smallest sample size across populations ($N = 10$). Genetic differentiation among sampling sites was measured using Jost's *D* (Jost 2008). This metric measures the allelic variation between pairs of populations; it has a null (or slightly negative) value when there is no differentiation between two populations and a value of one when two populations have no alleles in common. Jost's *D* among sites was calculated using the mmod package (Winter 2012) in the R environment.

*River Topography.* We selected three variables describing the topography and network arrangement at each sampling site, as network connectivity and topology are known to affect biodiversity patterns in river networks (Campbell Grant et al. 2007; Carrara et al. 2012; Paz-Vinas and Blanchet 2015). First, the betweenness centrality value (an index of river connectivity quantifying the positional importance of a node within a network; Freeman 1977) of the closest confluence upstream from each site was estimated using NetworkX (Hagberg et al. 2008), with higher values corresponding to nodes of higher importance for network connectivity. Second, local altitude and distance from the river mouth were obtained

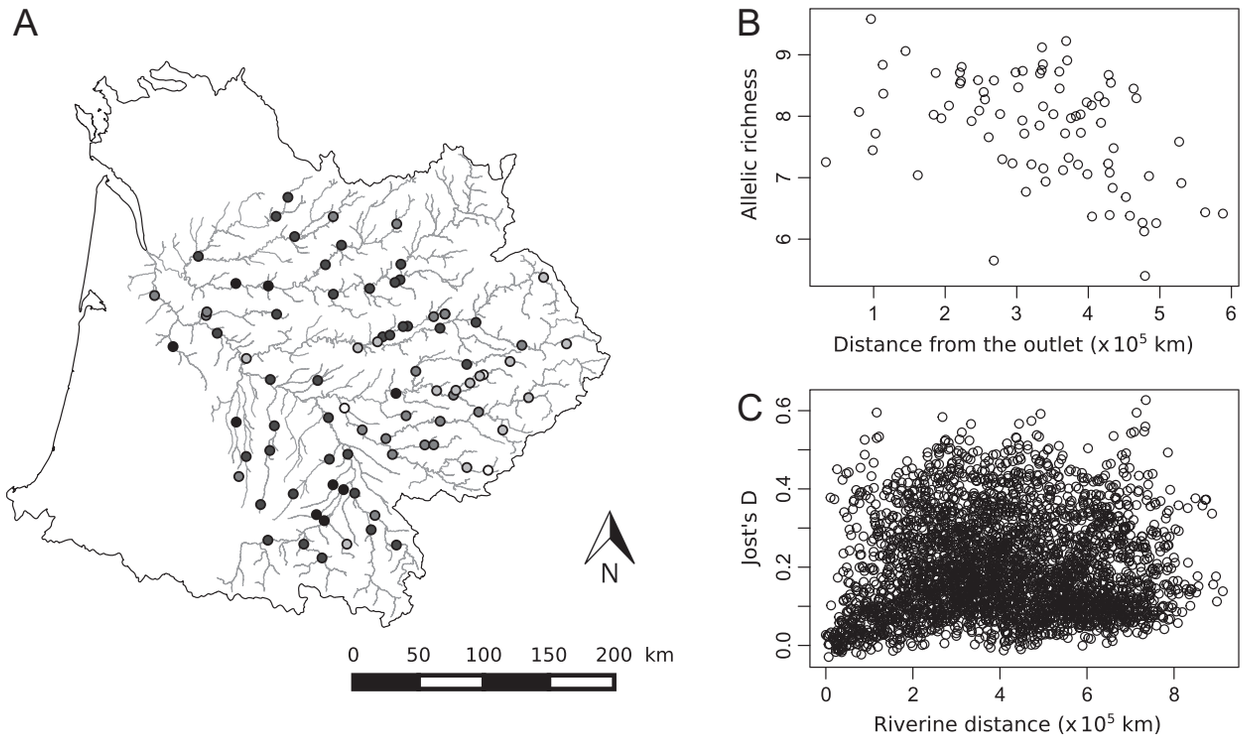


**Figure 3:** *A*, Geographic distribution of the 83 sites in which *Gobio occitaniae* was sampled within the Garonne-Dordogne river basin to characterize the genetic diversity (allelic richness and Jost's *D* measured at microsatellite markers) of each sampling site. The color of each dot is related to the relative allelic richness: white dots = lowest values; black dots = highest values. *B*, Biplot of the allelic richness measured at each sampling site plotted against the distance of each site from the outlet. *C*, Biplot of Jost's *D* measured between each sample pair plotted against pairwise riverine distance.

from the French Theoretical Hydrological Network (Réseau Hydrologique Théorique français; Pella et al. 2012). Third, the geographic distance along the river network (riverine distance) between each pair of sites was computed using QuantumGIS software (Quantum GIS Development Team 2017).

*Physicochemical Quality*. We hypothesized that the physicochemical characteristics of the sampling sites may affect the density of the fish populations (i.e., sampling sites with good water quality and optimal physical properties should sustain high fish densities) and, hence, the effective population size (assuming a positive correlation between abundance and effective population size and, ultimately, genetic diversity). The data thus indirectly reflected the possible influence of genetic drift on the genetic summary statistics. Data were obtained from the database of the Water Information System of the Adour Garonne basin (Système d'Information sur l'Eau du Bassin Adour Garonne [SIEAG]; http://adour-garonne.eaufrance.fr). Among other variables, this database compiles chemical characteristics of surface water (e.g., concentrations of various chemical compounds), measured several times a year at numerous sites in the Garonne-Dordogne river basin. Only sites with data available for March, May, July, September, and November of 2011 were selected from the SIEAG database. Most of our sampling sites overlapped with a SIEAG site, in which case the mean of the five temporal measures was used as a proxy for the chemical quality of our sampling sites. When the overlap was not perfect, each sampling site was assigned to the nearest SIEAG site (on the same river), and the average values of variables at this nearest SIEAG site were used as surrogates for the chemical quality of the sampling site. Three sampling sites had no SIEAG site close enough to obtain reliable information (distance greater than 10 km) and were therefore discarded from the final database.

We specifically obtained water temperature and oxygen concentration data to test the assumption that a site with an optimal temperature and a high oxygen concentration can host larger fish densities. Additionally, we selected five chemical components directly affected by human activities and considered as good indicators of water quality (ammonium, nitrate, nitrite, orthophosphate, and phosphorus concentrations). Principal component analysis of these five chemical components (see fig. A1, available in the online appendix)

was performed using the R package ade4 (Dray and Dufour 2007). The coordinates of each site on the first axis, which accounted for 62.4% of the variance, were used to create a synthetic variable (hereafter, "chemicals"), representing the amount of chemical components at each site. Low values correspond to high ammonium, nitrate, nitrite, orthophosphate, and phosphorus concentrations.

*Habitat Fragmentation.* We selected three variables related to habitat fragmentation, as it has previously been shown to affect genetic diversity in gudgeon (Blanchet et al. 2010). Habitat fragmentation variables were obtained from the Référentiel des Obstacles à l'Écoulement database (Onema 2010) that identifies and georeferences barriers to water flow along French rivers. Two main types of obstacles, weirs (<4 m high) and dams (5–30 m high, in general), were considered here. We measured the home range of each population, that is, the riverine geographic distance a fish can access without being stopped (both upstream, downstream, and in tributaries) by a weir or dam (Prunier et al. 2017). This is thus a direct measure of the impact of weirs and dams on fish habitat availability. Furthermore, the total number of weirs and dams along the river stretch between each pair of sites was calculated. Because these obstacles may have differential influence on genetic diversity, they were considered separately.

## *Statistical Analysis*

*Point Summary Statistics.* Allelic richness (Ar) was analyzed using both classical path analysis and the d-sep test. For the two approaches, a full causal model was first designed using theoretical and a priori knowledge. In this model (fig. 4*A*), allelic richness is expected to be directly linked to two human-related factors (chemicals and home range) and to five natural factors (betweenness centrality, oxygen concentration, temperature, distance from the mouth, and altitude). We constrained altitude and distance from the mouth to be related one to the other, and we assumed direct but also indirect relationships between altitude and allelic richness through temperature and oxygen concentration (fig. 4). All other relationships were direct (fig. 4). Prior to analysis, all variables were centered and scaled to obtain standardized parameter estimates (Schielzeth 2010).

This full model was tested through path analysis using the sem function from the lavaan R package (Rosseel 2012). We thus obtained the $F_{ML}$ value, the corresponding $P$ value, and the AIC of the model. We used the function dsep.test (available online at https://doi.org/10.5281/zenodo.1048975) to test the full model using a d-sep test approach in which conditional dependencies were tested through linear regressions. The full model was then simplified by removing paths one by one until the model with the best relative fit (i.e., the lowest AIC score) was identified. The absolute fit of this model was computed to ensure that the observed data were coherent with the model. This process permitted the identification of variables with a major influence on allelic richness. Finally, path analysis was used to collect the inferred path coefficients and the residual variance of allelic richness (corresponding to the amount of variance that was not explained by the model).

*Pairwise Summary Statistics.* In our data set, four variables were present in the form of pairwise distance matrices: Jost's $D$ (dependent variable), the counts of weirs and dams (separately) between each pair of sites, and the river distance between pairwise sites. Additionally, pairwise dissimilarity matrices for chemicals, altitude, temperature, and oxygen concentration were computed as the absolute differences between sites. These dissimilarity matrices reflected the isolation-by-environment hypothesis (IBE; Rundle and Nosil 2005; Sexton et al. 2014). Using these eight variables, we designed a complete model (fig. 5*A*) in which the number of weirs and dams between sites and pairwise differences in oxygen concentration and chemicals have a direct effect on genetic differentiation, whereas pairwise differences in altitude, temperature, and riverine distance have both direct and indirect effects (fig. 5*A*). As conducted previously, all variables were centered and scaled to facilitate interpretation. The full model was tested through both clustering-based path analysis and the permutation-based d-sep test. The model was simplified as explained previously until the model with the best relative fit was obtained. The coherence of the observed data to the best-fitted model was tested using clustering-based path analysis and the permutation-based d-sep test, whereas the $P$ values and confidence intervals of its path coefficients were computed using clustering-based path analysis, permutation-based path analysis, and the parametric bootstrap procedure developed for pairwise matrices.

## *Results*

*Description of Genetic Data.* Allelic richness ranged from 5.401 to 9.582, with a mean value of 7.793 ($\pm$0.887). No apparent spatial pattern could be visually identified (fig. 3*A*). However, there was a significant decrease in Ar from downstream to upstream sections of the landscape (Pearson correlation between allelic richness and distance from the river mouth, $r = -0.513$, df $= 1$, $P < .001$; fig. 3*B*). Full data sets are deposited in the Zenodo repository: https://doi.org/10.5281/zenodo.1048918 (Fourtune et al. 2017).

Jost's $D$ ranged from $-0.029$ to 0.627, with a mean value of 0.196 ($\pm$0.124). We failed to detect any IBD pattern since Jost's $D$ did not increase with the riverine distance between sites (Mantel test, $r = 0.055$, $P = .172$; fig. 3*C*).

*Causal Modeling Applied to Point Summary Statistics.* Simplification of the full model led to the removal of seven paths

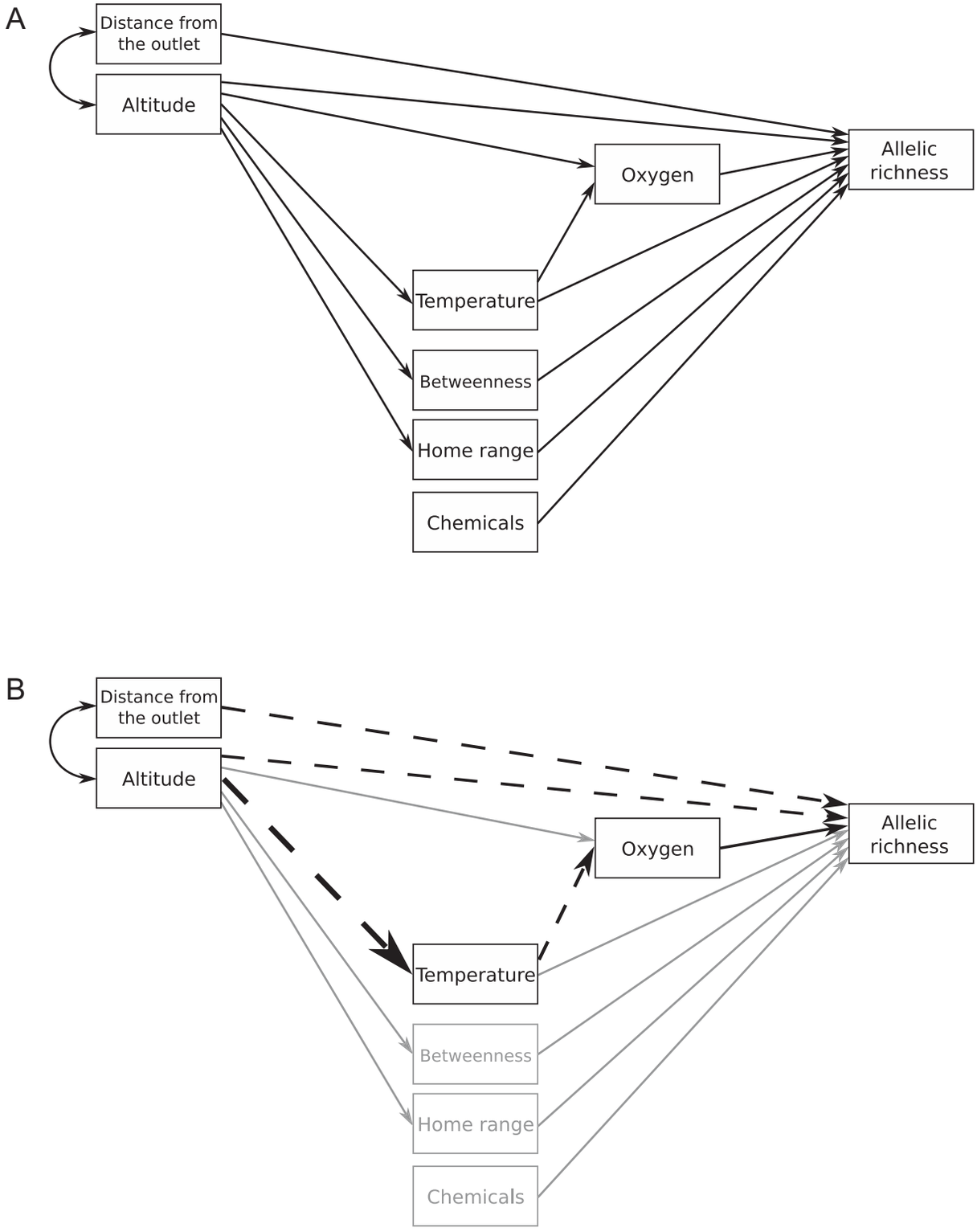


**Figure 4:** Graphical representations of the complete model depicting possible causal relationships between the allelic richness of the freshwater fish *Gobio occitaniae* quantified across 83 sampling sites in the Garonne-Dordogne river basin and anthropogenic and natural factors (*A*) and the best causal model obtained for fitting the data using classical path analysis (results were similar when using the classical d-sep test) and after a simplification procedure (*B*). Single-headed arrows indicate a causal link. Double-headed arrows indicate covariation. Solid and dashed lines represent positive and negative values, respectively; their width is proportional to the absolute value of the corresponding path coefficient. Gray arrows represent paths removed during the simplification procedure.

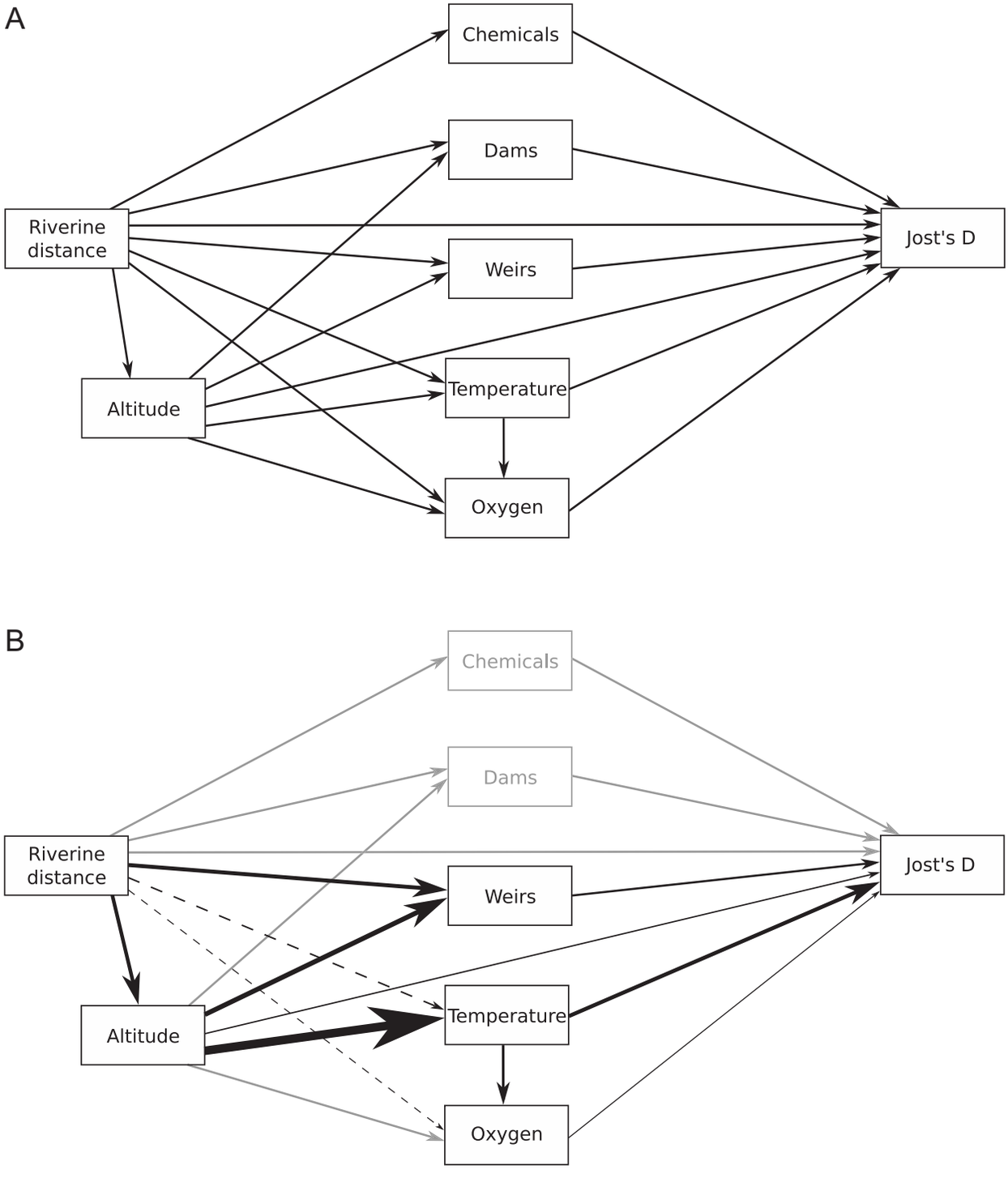


**Figure 5:** Graphical representations of the complete model depicting possible causal relationships between the genetic differentiation (Jost's *D*) of the freshwater fish *Gobio occitaniae* quantified across 83 sampling sites in the Garonne-Dordogne river basin and anthropogenic and natural factors (*A*) and the best causal model obtained for fitting the data using clustering-based path analysis (results were similar when using the permutation-based d-sep test) and after a simplification procedure (*B*). Single-headed arrows indicate a causal link. Double-headed arrows indicate covariation. Solid and dashed lines represent positive and negative values, respectively; their width is proportional to the absolute value of the corresponding path coefficient. Gray arrows represent paths removed during the simplification procedure.

**Table 1:** Path analysis and d-sep test statistics used to disentangle the effects of environmental factors on allelic richness (Ar, point summary statistics) and genetic differentiation (Jost's *D*, pairwise summary statistics) in a freshwater fish species (*Gobio occitaniae*) sampled in a river network

| | Test statistics | df | *P* | AIC |
|---|---|---|---|---|
| Path analysis applied to point summary statistics (Ar): | | | | |
| Full model | 41.455 | 15 | <.001 | 11.445 |
| Best-fitted model | 1.656 | 4 | .799 | −6.344 |
| D-sep test applied to point summary statistics (Ar): | | | | |
| Full model | 59.676 | 30 | .001 | 109.676 |
| Best-fitted model | 4.549 | 8 | .804 | 32.549 |
| Clustering-based path analysis applied to pairwise summary statistics (Jost's *D*): | | | | |
| Full model | 82.273 | 10 | <.001 | 62.273 |
| Best-fitted model | 3.286 | 4 | .512 | −4.714 |
| Permutation-based d-sep test applied to pairwise summary statistics (Jost's *D*): | | | | |
| Full model | 57.509 | 20 | <.001 | 121.509 |
| Best-fitted model | 4.149 | 8 | .843 | 46.149 |

Note: For each metric (Ar and Jost's *D*), we simplified a full model (i.e., a model including all paths described in figs. 4, 5) until a model with the lowest Akaike's information criterion (AIC) score (best-fitted model) was obtained. The model simplification was performed using either path analysis or the d-sep test for Ar and either path analysis (clustering-based path analysis) or the d-sep test (permutation-based d-sep test) taking into account the structure of pairwise matrices. The term "test statistics" represents the maximum-likelihood fitting function in the case of path analysis and Fisher's *C* in the case of the d-sep test.

before the model with the lowest AIC score was obtained (table 1; fig. 4*B*). Both approaches (path analysis and the d-sep test) led to the same best-fitted model. In this best-fitted model, allelic richness was directly correlated with the distance from the outlet, the altitude, and the oxygen concentration. Allelic richness was higher in downstream sites at a low altitude and with a high oxygen concentration (table 2). Altitude was also indirectly correlated with allelic richness

**Table 2:** Estimates of the path coefficients and their associated *P* values and 95% confidence intervals (CIs) of the best-fitted models linking environmental features to allelic richness (Ar, point summary statistics) and genetic differentiation (Jost's *D*, pairwise summary statistics) for a freshwater fish (*Gobio occitaniae*) sampled in a river network

| | Path coefficient | *P* | 95% CI |
|---|---|---|---|
| Point summary statistics (Ar): | | | |
| Distance from the outlet → allelic richness | −.407 | .005 | [−.720;−.127] |
| Altitude → allelic richness | −.392 | .005 | [−.663;−.095] |
| Oxygen → allelic richness | .377 | <.001 | [.227;.554] |
| Temperature → oxygen | −.453 | <.001 | [−.648;−.266] |
| Altitude → temperature | −.745 | <.001 | [−.899;−.599] |
| Pairwise summary statistics (Jost's *D*): | | | |
| Altitude → Jost's *D* | .094 | <.001 | [.052;.139] |
| Oxygen → Jost's *D* | .071 | <.05 | [.039;.105] |
| Temperature → Jost's *D* | .265 | <.001 | [.226;.304] |
| Weirs → Jost's *D* | .144 | <.001 | [.110;.182] |
| Riverine distance → weirs | .266 | <.001 | [.234;.299] |
| Altitude → weirs | .329 | <.001 | [.289;.369] |
| Riverine distance → oxygen | −.069 | <.001 | [−.101;−.038] |
| Temperature → oxygen | .177 | <.001 | [.142;.213] |
| Altitude → temperature | .577 | <.001 | [.544;.607] |
| Riverine distance → temperature | −.104 | <.001 | [−.135;−.072] |
| Riverine distance → altitude | .247 | <.001 | [.214;.283] |

Note: For Ar, path coefficients, *P* values, and 95% CIs were obtained using classical path analysis. For genetic differentiation, path coefficients and *P* values were obtained using clustering-based path analysis (results were similar when using permutation-based path analysis), whereas 95% CIs were obtained using the parametric bootstrap procedure developed for pairwise matrices.

through a pathway that sequentially involved temperature and the oxygen concentration (fig. 4*B*). The residual variance of Ar was 0.508 ± 0.080, which indicated that almost 50% of its total variance was explained by this best-fitted model.

*Causal Modeling Applied to Pairwise Matrices.* Irrespective of the approach, the simplification procedure led to the gradual removal of six paths before the model with the lowest AIC score was obtained (table 1; fig. 5*B*). In this best-fitted model, Jost's *D* was directly correlated with the differences in altitude, temperature, oxygen concentration, and the number of weirs between sites. All these explanatory variables were also correlated with riverine distance through direct pathways and, in the case of weirs, temperature and oxygen, through indirect pathways involving other environmental variables (table 2; fig. 5*B*).

## Discussion

Within a landscape, environmental variables can have compounding and contrasting impacts on spatial patterns of genetic diversity, and properly inferring these impacts is a key challenge in landscape genetics (Storfer et al. 2010). Here, we built on the framework of path analysis and the d-sep test (which we extended to the analysis of pairwise matrices) to provide landscape geneticists with a reliable statistical tool to improve their ability to unravel direct and indirect relationships between landscape features and the spatial distribution of genetic diversity.

### *Validation of Path Analysis and the D-Sep Test Applied to Pairwise Matrices*

We improved the two commonly used causal modeling approaches (path analysis and the d-sep test; Shipley 2000*b*; Grace 2006) by extending their validity to the analysis of causal models comprising pairwise matrices and by using various procedures such as permutations, bootstrapping, and specification of random effects. We provide operational R functions for these four improved approaches online at https://doi.org/10.5281/zenodo.1048975, making them directly transferable to other biological systems. The simulations demonstrated that, as expected, our improved procedures were robust in identifying the best causal model compared to path analysis or d-sep tests that do not explicitly account for nonindependence of pairwise data. Although clustering-based path analysis does not account for total nonindependence between pairwise data because of the nested structure of random effects, it is noteworthy that the approach is a major improvement over classical path analysis. However, clustering-based path analysis imperfectly assessed the relative and absolute fits of models that did not perfectly reflect the causal structure underlying the data (intermediate model). We therefore suggest that clustering-based path analysis be used as a secondary approach to ensure/refine results obtained from the permutation-based d-sep test, if needed.

This simulation study was also used to test whether clustering-based path analysis and permutation-based path analysis were reliable in inferring the *P* values of path coefficients connecting pairwise matrices and, hence, model parameters. We showed that both methods greatly outperformed traditional path analysis. Additionally, the parametric bootstrap procedure developed for pairwise matrices provided reliable confidence intervals for path coefficients, despite a tendency to underestimate these intervals in the case of adequate coefficients.

Although robust, the results of our simulation study should be considered with caution as we explored only a small set of variables. We therefore call for additional simulations (Landguth et al. 2015) to further assess the reliability of clustering-based path analysis, permutation-based path analysis, and the permutation-based d-sep test, especially when compared to other traditional statistical procedures used in landscape genetics. Furthermore, we encourage further methodological developments for the implementation of the MLPE procedure (Clarke et al. 2002) into the framework of clustering-based path analysis, as it is currently based on a hierarchical structure of random effects.

### *Empirical Application of Path Analysis and the D-Sep Test Applied to Pairwise Matrices*

When applied to empirical genetic data for *Gobio occitaniae* obtained from a whole river basin, both path analysis and the d-sep test identified the best-fitted causal models depicting both direct and indirect relationships between genetic summary statistics and landscape predictors. Our results strongly suggest that both allelic richness and pairwise measures of genetic differentiation were mainly related to natural landscape features (altitude, temperature, and oxygen concentration) and that at such a large spatial scale, anthropogenic factors (related to habitat fragmentation and water pollution) were negligible drivers of genetic diversity in this species (with the exception of weirs).

Interestingly, some landscape features such as altitude were identified as direct drivers of both allelic richness and genetic differentiation. There was indeed a strong direct negative relationship between allelic richness and altitude, indicating that allelic richness was higher in sampling sites located at lower altitudes. Similarly, we found a direct positive relationship between the difference in altitude and pairwise measures of genetic differentiation between sites: the higher the difference in altitude, the higher the genetic differentiation. This type of direct relationship between altitude and genetic summary statistics for a freshwater fish is, to our knowledge, rarely presented in the literature (but see Faulks et al. 2011) and

could reflect two nonexclusive processes: the past colonization history of *G. occitaniae* and the contemporary influence of asymmetric gene flow toward downstream sites. First, altitude may reflect historical contingencies whereby glacial refugia during the last glaciation event (~10,000 years ago) were mainly found in lowlands. Glacial refugia (in downstream sites) should indeed be more genetically diverse than recently colonized areas (in upstream sites; Paz-Vinas et al. 2015), whereas larger distances from the refugia should be associated with higher genetic differentiation (Costedoat and Gilles 2009). Second, altitude could be a good surrogate for the unidirectional water flow of rivers that favors gene flow from upstream (sites at high altitude) to downstream (sites at low altitude; Paz-Vinas et al. 2015). The direct negative relationship between distance from the outlet and allelic richness could similarly stem from these two mechanisms.

For both allelic richness and pairwise measures of genetic differentiation, we additionally found indirect relationships between altitude and genetic diversity. These indirect relationships were mediated through water temperature and oxygen concentration. These two indirect relationships were expected to underline the effect of genetic drift. Oxygen is an important driver of fish species distribution in river networks (Crispo and Chapman 2008), and we hypothesized that higher oxygen concentrations may sustain higher fish densities (or at least sites with extremely low oxygen availability may have higher fish mortality). Assuming that density is positively related to the effective population size in fish (Belmar-Lucero et al. 2012), oxygen limitation may directly alter allelic richness and, ultimately, genetic differentiation through genetic drift (Hutchison and Templeton 1999). In the same way, because water temperature and oxygen availability are negatively correlated, higher water temperatures may also be related to lower effective population sizes, leading to an increase in genetic drift and, consequently, genetic differentiation. This unmeasured effect of genetic drift may also explain the direct effects of differences in altitude and water temperature on genetic differentiation. The direct links between differences in water temperature, differences in oxygen concentration, and genetic differentiation may stem from the additional effect of IBE, a process that occurs when populations inhabiting different environments experience divergent patterns of selection (Rundle and Nosil 2005; Sexton et al. 2014). As a consequence, dispersing individuals may be maladapted to new environments, with reduced fitness and reproductive success, thereby decreasing gene flow between environmentally different areas (Crispo et al. 2006). Water temperature and oxygen concentration may thus act as important selective pressures in *G. occitaniae*—a hypothesis that deserves further investigation. Nonetheless, our study presents, to our knowledge, one of the first demonstrations of a direct relationship between oxygen availability, water temperature, and genetic diversity in a freshwater fish species.

Regarding anthropogenic factors, weirs were found to have an impact on genetic differentiation, whereas the link between dams and genetic differentiation was discarded. This is surprising since weirs are not as high (1–4 m) as dams (5–30 m) and are generally expected to be more permeable to dispersal than dams (Blanchet et al. 2010). However, weirs are also generally older (they can be as old as 400 years, whereas dams are generally no more than 60 years old): considering the possible delay between anthropogenic impacts and the ensuing genetic response (Smith and Bernatchez 2008), it is possible that dams are too recent to have left a significant genetic imprint on the spatial patterns of genetic differentiation. A nonexclusive hypothesis may be that the relatively low influence of dams on genetic differentiation can be explained by their small numbers in the network (there are, on average, fourfold fewer dams than weirs between sites). The influence of weirs as factors limiting dispersal and increasing genetic differentiation in freshwater fish has been shown in previous studies (Raeymaekers et al. 2009; Blanchet et al. 2010; Faulks et al. 2011), and our study therefore confirms these findings at a larger spatial scale while taking other covariables into account.

It is noteworthy that we also found indirect relationships between geographic isolation and genetic differentiation (through the number of weirs and the differences in altitude, water temperature, and oxygen concentration between sites), despite the absence of any direct relationship between riverine distance and Jost's *D*. IBD patterns are generally interpreted as imprints of gene flow and genetic drift (Hutchison and Templeton 1999), although the exact mechanisms underlying these patterns are rarely unraveled. Here, using path analyses, we were able to highlight potential causal pathways linking geographic distance to genetic differentiation: this relationship most probably arose from the spatial covariation between geographic distances, number of weirs, and differences in altitude, water temperature, and oxygen concentration. The use of causal modeling allowed the unraveling of multiple and complex relationships between geographic distance and genetic differentiation, which is essential for fundamental knowledge and applied perspectives. Of course, the relationship between geographic isolation and genetic differentiation may also result from alternative processes that we failed to model properly (Ewers and Didham 2006), and this should be investigated in future studies. For instance, the complexity of the river network is expected to play a major role in genetic differentiation of aquatic organisms (Paz-Vinas and Blanchet 2015), with isolated upstream populations acting as reservoirs for unique and rare alleles, hence triggering high genetic differentiation between upstream and other populations.

Causal modeling taking both direct and indirect effects into account provides a better appraisal of factors driving the spatial distribution of alleles in river networks. Most previous

studies on spatial patterns of genetic diversity in aquatic organisms focused on the relationships between allelic richness and distance from the river mouth and generally found an increase in genetic diversity from the source to the mouth of river networks (reviewed in Paz-Vinas et al. 2015). We also uncovered this general pattern, although the use of path analysis procedures suggested that several other environmental variables were linked to allelic richness, even in a model in which the relationship between distance from the outlet and allelic richness was taken into account. This result illustrates the strength of causal modeling to unravel complex processes shaping spatial patterns of genetic diversity that simultaneously involve several key environmental variables, notably in landscapes with high spatial covariation such as river networks.

### *General Conclusion*

When adapted to pairwise matrices, causal modeling allows the assessment of complex competing causal models, depicting the a priori hypotheses concerning causal relationships among explanatory variables. The proposed framework constitutes a promising alternative to the causal modeling procedure proposed by Cushman et al. (2006), as it allows the assessment of both direct and indirect causal relationships among numerous predictors.

Nevertheless, caution must be taken when using causal modeling. Causal modeling procedures rely directly on the formally stated a priori causal hypotheses depicted in the initial causal model. As a first consequence, inferred relationships among variables cannot be considered as absolute causal links; rather, these relationships can be considered as only possible causal links because some important but unknown (or unmeasured) variables may have been overlooked. Although investigation of the interplay between direct and indirect relationships in the optimal model may reveal hidden pathways, thus shedding light on the biological processes acting on the dependent variable, researchers should always keep in mind that a causal model cannot provide information beyond stated a priori hypotheses. Our empirical data set exemplifies this observation appropriately, with genetic drift identified as a possible driver of spatial genetic variation in *G. occitaniae* only under the hypothesis of a direct relationship between oxygen concentration and the effective population size. As a second consequence, causal modeling cannot be confidently considered as a data mining procedure, as investigation of correlation coefficients in the absence of any implicit a priori hypothesis (and thus in the absence of any formal causal model) may produce spurious conclusions (Legendre and Legendre 2012; Prunier et al. 2015). Note also that the interpretation of causal modeling becomes more complex as the number of predictors increases. The keys toward the successful use of causal modeling are thus (i) a well-thought-out initial set of possible causal models, (ii) a cautious interpretation of the combination of AIC, *P* values and confidence intervals, together providing a body of evidence as to the relevance of considered models, and (iii) proper biological interpretation of inferred direct and indirect relationships in the light of formally stated a priori hypotheses. Keeping these prerequisites in mind, we advocate the use of causal modeling as a powerful explanatory tool in landscape genetics.

## Acknowledgments

We warmly thank the Onema staff as well as the many colleagues who contributed to the field sampling. Charlotte Veyssière, Roselyne Etienne, and Vincent Dubut are also thanked for their help with molecular analyses. We are grateful to Yves Rosseel for constructive discussions in early stages of this work. We also warmly thank Bill Shipley, Mark Taper, and two anonymous reviewers for their constructive comments on earlier versions of the manuscript. L.F. was financially supported by a Ministère de l'Enseignement Supérieur et de la Recherche PhD scholarship during this study. This work was supported by the Laboratoire d'Excellence TULIP (ANR-10-LABX-41).

## Literature Cited

Aljanabi, S. M., and I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25:4692–4693.

Balkenhol, N., L. P. Waits, and R. J. Dezzani. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32:818–830.

Belmar-Lucero, S., J. L. A. Wood, S. Scott, A. B. Harbicht, J. A. Hutchings, and D. J. Fraser. 2012. Concurrent habitat and life history influences on effective/census population size ratios in stream-dwelling trout. Ecology and Evolution 2:562–573.

Blanchet, S., O. Rey, R. Etienne, S. Lek, and G. Loot. 2010. Species-specific responses to landscape fragmentation: implications for management strategies. Evolutionary Applications 3:291–304.

Bollen, K. A. 1989. Structural equations with latent variables. Wiley, New York.

Bradburd, G. S., P. L. Ralph, and G. M. Coop. 2013. Disentangling the effects of geographic and ecological isolation on genetic differentiation. Evolution 67:3258–3273.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. Springer, New York.

Campbell Grant, E. H., W. H. Lowe, and W. F. Fagan. 2007. Living in the branches: population dynamics and ecological processes in dendritic networks. Ecology Letters 10:165–175.

Cardon, M., G. Loot, G. Grenouillet, and S. Blanchet. 2011. Host characteristics and environmental factors differentially drive the burden and pathogenicity of an ectoparasite: a multilevel causal analysis. Journal of Animal Ecology 80:657–667.

Carrara, F., F. Altermatt, I. Rodriguez-Iturbe, and A. Rinaldo. 2012. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. Proceedings of the National Academy of Sciences of the USA 109:5761–5766.

Clarke, R. T., P. Rothery, and A. F. Raybould. 2002. Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. Journal of Agricultural, Biological, and Environmental Statistics 7:361–372.

Costedoat, C., and A. Gilles. 2009. Quaternary pattern of freshwater fishes in Europe: comparative phylogeography and conservation perspective. Open Conservation Biology Journal 3:36–48.

Crispo, E., P. Bentzen, D. N. Reznick, M. T. Kinnison, and A. P. Hendry. 2006. The relative influence of natural selection and geography on gene flow in guppies. Molecular Ecology 15:49–62.

Crispo, E., and L. J. Chapman. 2008. Population genetic structure across dissolved oxygen regimes in an African cichlid fish. Molecular Ecology 17:2134–2148.

Cushman, S. A., and E. L. Landguth. 2010. Spurious correlations and inference in landscape genetics. Molecular Ecology 19:3592–3602.

Cushman, S. A., K. S. McKelvey, J. Hayden, and M. K. Schwartz. 2006. Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. American Naturalist 168:486–499.

Cushman, S. A., T. N. Wasserman, E. L. Landguth, and A. J. Shirk. 2013. Re-evaluating causal modeling with Mantel tests in landscape genetics. Diversity 5:51–72.

Dray, S., and A.-B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22: 1–20.

Ewers, R. M., and R. K. Didham. 2006. Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews 81:117–142.

Faulks, L. K., D. M. Gilligan, and L. B. Beheregaray. 2011. The role of anthropogenic vs. natural in-stream structures in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). Evolutionary Applications 4: 589–601.

Fisher, R. A. 1938. Statistical methods for research workers. 7th ed. Oliver & Boyd, Edinburgh.

———. 1950. Contributions to mathematical statistics. Wiley, New York.

Fourtune, L., I. Paz-Vinas, G. Loot, J. G. Prunier, and S. Blanchet. 2016. Lessons from the fish: a multi-species analysis reveals common processes underlying similar species-genetic diversity correlations. Freshwater Biology 61:1830–1845.

Fourtune, L., J. G. Prunier, I. Paz-Vinas, G. Loot, C. Veyssière, and S. Blanchet. 2017. Data from: Inferring causalities in landscape genetics: an extension of Wright's causal modeling to distance matrices. American Naturalist, Zenodo, https://doi.org/10.5281/zenodo.1048918.

Freeman, L. 1977. A set of measures of centrality based on betweenness. Sociometry 40:35–41.

Goodman, S. N. 1999. Toward evidence-based medical statistics. 1: the *P* value fallacy. Annals of Internal Medicine 130:995–1004.

Goslee, S. C., and D. L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22: 1–19.

Goudet, J. 2001. Fstat: a program to estimate and test gene diversities and fixation indices. Version 2.9.3. http://www2.unil.ch/popgen/softwares/fstat.htm.

Grace, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press, Cambridge.

Graves, T. A., P. Beier, and J. A. Royle. 2013. Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. Molecular Ecology 22:3888–3903.

Guillot, G., R. Leblois, A. Coulon, and A. C. Frantz. 2009. Statistical methods in spatial genetics. Molecular Ecology 18:4734–4756.

Hagberg, A. A., D. A. Schult, and P. J. Swart. 2008. Exploring network structure, dynamics, and function using NetworkX. Pages 11–15 *in* G. Varoquaux, T. Vaught, and J. Millman, eds. Proceedings of the 7th Python in Science Conference (ScyPy2008), Pasadena, CA.

Hutchison, D., and A. Templeton. 1999. Correlation of pairwise genetic and geographic distance measure: inferring the relative influences of gene flow and drift on distribution of genetic variability. Evolution 53:1898–1914.

Jaquiéry, J., T. Broquet, A. H. Hirzel, J. Yearsley, and N. Perrin. 2011. Inferring landscape effects on dispersal from genetic distances: how far can we go? Molecular Ecology 20:692–705.

Jost, L. 2008. $G_{ST}$ and its relatives do not measure differentiation. Molecular Ecology 17:4015–4026.

Keller, D., R. Holderegger, M. J. van Strien, and J. Bolliger. 2015. How to make landscape genetics beneficial for conservation management? Conservation Genetics 16:503–512.

Landguth, E., S. A. Cushman, and N. Balkenhol. 2015. Simulation modeling in landscape genetics. Pages 99–113 *in* N. Balkenhol, S. A. Cushman, A. T. Storfer, and L. P. Waits, eds. Landscape genetics: concepts, methods, applications. Wiley-Blackwell, Oxford.

Legendre, P. 2000. Comparison of permutation methods for the partial correlation and partial Mantel tests. Journal of Statistical Computation and Simulation 67:37–73.

Legendre, P., and M.-J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10:831–844.

Legendre, P., M.-J. Fortin, and D. Borcard. 2015. Should the Mantel test be used in spatial analysis? Methods in Ecology and Evolution 6: 1239–1247.

Legendre, P., and L. F. J. Legendre. 2012. Numerical ecology. Elsevier, Amsterdam.

Manel, S., and R. Holderegger. 2013. Ten years of landscape genetics. Trends in Ecology and Evolution 28:614–621.

Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18:189–197.

Marchetti, G. M. 2006. Independencies induced from a graphical Markov model after marginalization and conditioning: the R package ggm. Journal of Statistical Software 15:1–15.

Oberski, D. L. 2014. lavaan.survey: an R package for complex survey analysis of structural equation models. Journal of Statistical Software 57:1–27.

Onema. 2010. Référentiel national des obstacles à l'écoulement: une cartographie nationale des obstacles sur les cours d'eau. Les Fiches de l'ONEMA. http://www.onema.fr.

Paz-Vinas, I., and S. Blanchet. 2015. Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. Journal of Evolutionary Biology 28:986–994.

Paz-Vinas, I., G. Loot, V. M. Stevens, and S. Blanchet. 2015. Evolutionary processes driving spatial patterns of intra-specific genetic diversity in river ecosystems. Molecular Ecology 24:4586–4604.

Pearl, J. 1988. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, San Mateo, CA.

———. 2009. Causality: models, reasoning, and inference. 2nd ed. Cambridge University Press, Cambridge.
Pearl, J., and T. Verma. 1987. The logic of representing dependencies by directed graphs. Pages 374–379 *in* Proceedings of the 6th National Conference on Artificial Intelligence. Vol. 1. AAAI, Seattle.
Pella, H., J. Lejot, N. Lamouroux, and T. Snelder. 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. Géomorphologie: Relief, Processus, Environnement 3:317–336.
Prunier, J. G., M. Colyn, X. Legendre, K. F. Nimon, and M. C. Flamand. 2015. Multicollinearity in spatial genetics: separating the wheat from the chaff using commonality analyses. Molecular Ecology 24:263–283.
Prunier, J. G., V. Dubut, L. Chikhi, and S. Blanchet. 2017. Contribution of spatial heterogeneity in effective population sizes to the variance in pairwise measures of genetic differentiation. Methods in Ecology and Evolution 8:1866–1877.
Quantum GIS Development Team. 2017. Quantum GIS geographic information system. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.
Raeymaekers, J. A. M., D. Raeymaekers, I. Koizumi, S. Geldof, and F. A. M. Volckaert. 2009. Guidelines for restoring connectivity around water mills: a population genetic approach to the management of riverine fish. Journal of Applied Ecology 46:562–571.
R Development Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
Rosseel, Y. 2012. lavaan: an R package for structural equation modeling. Journal of Statistical Software 48:1–36.
Rundle, H. D., and P. Nosil. 2005. Ecological speciation. Ecology Letters 8:336–352.
Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution 1:103–113.
Sexton, J. P., S. B. Hangartner, and A. A. Hoffmann. 2014. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution 68:1–15.
Shipley, B. 2000*a*. Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. Cambridge University Press, Cambridge.
———. 2000*b*. A new inferential test for path models based on directed acyclic graphs. Structural Equation Modeling 7:206–218.
———. 2003. Testing recursive path models with correlated errors using d-separation. Structural Equation Modeling 10:214–221.
———. 2009. Confirmatory path analysis in a generalized multilevel context. Ecology 90:363–368.
———. 2013. The AIC model selection method applied to path analytic models compared using a d-separation test. Ecology 94:560–564.
Smith, T. B., and L. Bernatchez. 2008. Evolutionary change in human-altered environments. Molecular Ecology 17:1–8.
Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology 35:627–632.
Storfer, A., M. A. Murphy, S. F. Spear, R. Holderegger, and L. P. Waits. 2010. Landscape genetics: where are we now? Molecular Ecology 19:3496–3514.
Van Strien, M. J., D. Keller, and R. Holderegger. 2012. A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. Molecular Ecology 21:4010–4023.
Winter, D. J. 2012. MMOD: an R library for the calculation of population differentiation statistics. Molecular Ecology Resources 12:1158–1160.
Wright, S. 1921. Correlation and causation. Journal of Agricultural Research 20:557–585.

Associate Editor: Charles F. Baer
Editor: Alice A. Winn

Sampling location on the Vère river, near Cahuzac-sur-Vère in southern France. Photo credit: Simon Blanchet.