



**HAL**  
open science

## Spatial graphs highlight how multi-generational dispersal shapes landscape genetic patterns

Emilie Boulanger, Alicia Dalongeville, Marco Andrello, David Mouillot,  
Stephanie Manel

► **To cite this version:**

Emilie Boulanger, Alicia Dalongeville, Marco Andrello, David Mouillot, Stephanie Manel. Spatial graphs highlight how multi-generational dispersal shapes landscape genetic patterns. *Ecography*, 2020, 43, pp.1167-1179. 10.1111/ecog.05024 . hal-03404999

**HAL Id: hal-03404999**

**<https://hal.umontpellier.fr/hal-03404999>**

Submitted on 2 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## Research

### Spatial graphs highlight how multi-generational dispersal shapes landscape genetic patterns

Emilie Boulanger, Alicia Dalongeville, Marco Andrello, David Mouillot and Stéphanie Manel

E. Boulanger (<https://orcid.org/0000-0002-6446-7342>) and S. Manel (<https://orcid.org/0000-0001-8902-6052>) ✉ ([stephanie.manel@cefe.cnrs.fr](mailto:stephanie.manel@cefe.cnrs.fr)), CEFE, Univ. Montpellier, CNRS, EPHE-PSL, IRD, Univ. Paul Valéry Montpellier, Montpellier, France. – A. Dalongeville (<https://orcid.org/0000-0002-8969-8243>), *People and The Sea*, Logon, Daanbantayan, Cebu, Philippines. – M. Andrello (<https://orcid.org/0000-0001-7590-2736>), D. Mouillot (<https://orcid.org/0000-0003-0402-2605>) and EB, MARBEC, Univ. Montpellier, CNRS, Ifremer, IRD, Montpellier, France.

#### Ecography

43: 1167–1179, 2020

doi: 10.1111/ecog.05024

Subject Editor: Pedro Peres-Neto

Editor-in-Chief: Miguel Araújo

Accepted 17 April 2020



Current approaches that compare spatial genetic structure of a given species and the dispersal of its mobile phase can detect a mismatch between both patterns mainly due to processes acting at different temporal scales. Genetic structure result from gene flow and other evolutionary and demographic processes over many generations, while dispersal predicted from the mobile phase often represents solely one generation on a single time-step. In this study, we present a spatial graph approach to landscape genetics that extends connectivity networks with a stepping-stone model to represent dispersal between suitable habitat patches over multiple generations. We illustrate the approach with the case of the striped red mullet *Mullus surmuletus* in the Mediterranean Sea. The genetic connectivity of *M. surmuletus* was not correlate with the estimated dispersal probability over one generation, but with the stepping-stone estimate of larval dispersal, revealing the temporal scale of connectivity across the Mediterranean Sea. Our results highlight the importance of considering multiple generations and different time scales when relating demographic and genetic connectivity. The spatial graph of genetic distances further untangles intra-population genetic structure revealing the Siculo-Tunisian Strait as an important corridor rather than a barrier for gene flow between the Western- and Eastern Mediterranean basins, and identifying Mediterranean islands as important stepping-stones for gene flow between continental populations. Our approach can be easily extended to other systems and environments.

Keywords: genetic connectivity, Mediterranean Sea, *Mullus surmuletus*, seascape genetics, spatial graphs, stepping-stone dispersal

#### Introduction

Connectivity and gene flow influence the evolutionary dynamics of spatially structured populations, maintain genetic diversity and promote population adaptive potential and resilience after disturbance (Hughes and Stachowicz 2004, Vandergast et al. 2008, Bagueette et al. 2013, Donati et al. 2019). Connectivity refers to the movement of individuals in a heterogenous landscape (Taylor et al. 1993).



[www.ecography.org](http://www.ecography.org)

© 2020 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

When individual dispersal is followed by successful reproduction, demographic connectivity results in genetic connectivity, a measure of gene flow and other evolutionary processes (Lowe and Allendorf 2010). Understanding how populations are connected across space and time is essential to assess the impact that habitat change and fragmentation can have on population persistence, and to develop sustainable resource management and appropriate conservation strategies (Fischer and Lindenmayer 2007, Andrello et al. 2015, Magris et al. 2018). It can also help to quantify the potential spreading of adaptive alleles in areas threatened by climate change (Razgour et al. 2019). Landscape genetic studies aim to characterize how habitat and environmental features promote or impede the movement of individuals through landscapes, riverscapes and seascapes and thus influence microevolutionary processes, including gene flow and genetic drift (Manel et al. 2003, Anderson et al. 2010). For species whose connectivity is mainly realised by a dispersing propagule stage (e.g. eggs, larvae, seeds, spores), much interest has been given to understanding the drivers of population connectivity by estimating the dispersal of these propagules through water, wind or animal vectors and by assessing the extent to which propagule dispersal explains the observed genetic structure across space (Nathan and Muller-Landau 2000, Castorani et al. 2017, Escalante et al. 2018).

Comparing genetic connectivity and propagule dispersal requires matching propagule dispersal to the temporal scales of the processes shaping genetic connectivity. While propagule dispersal occurs within one generation, genetic structure results from gene flow and demographic processes over many generations which can, in the absence of strong barriers to dispersal, lead to connectivity over large spatial scales (Hedgecock et al. 2007). Considering only single-generation dispersal events ignores the effect of stepping-stone dispersal through habitat patches leading to genetic connectivity over multiple generations (Saura et al. 2014). We thus need to integrate propagule dispersal across time to better understand genetic connectivity patterns and understand how multi-generational stepping-stone dispersal shapes genetic population structure.

In the terrestrial environment, most landscape genetic studies analyse genetic connectivity using least-cost path estimates of habitat connectivity (Epps et al. 2007, Castillo et al. 2014, Row et al. 2018, Schoville et al. 2018). Some recent advances have been made to combine spatial and temporal scales in landscape connectivity models. Saura et al. (2014) developed a generalized habitat network connectivity model which accounts for the potential role of stepping-stones enhancing species dispersal across generations. Martensen et al. (2017) modelled spatio-temporal networks of connectivity in dynamic landscapes where patch availability changes through time. However, neither of these studies tested how such improved estimates of landscape connectivity correlate with genetic connectivity.

In the marine environment, important advances combined biophysical models and genetic analyses to describe population connectivity. Examples include mobile species such as

reef fishes (e.g. *Plectropomus maculatus* in the southern Great Barrier Reef, Bode et al. 2019) and lobsters (e.g. *Panulirus argus* in the Caribbean sea, Truelove et al. 2016), and sedentary species such as sea cucumbers (e.g. *Parastichopus californicus* in the North-Eastern Pacific, Xuereb et al. 2018), molluscs (e.g. *Kelletia kelletii* in the Santa Barbara Channel, White et al. 2010) and seagrasses (e.g. *Zostera marina* in the North Sea, Jahnke et al. 2018). The two latter studies, White et al. (2010) and Jahnke et al. (2018), integrated dispersal over multiple generations by applying Markov Chain matrix multiplication but missed herein a spatially explicit consideration.

Spatial graphs (Box 1) can help to better understand the dynamics of complex systems composed of many interacting units (Rozenfeld et al. 2008, Dale and Fortin 2010, Peterson et al. 2019). They offer a promising tool to study multi-generational dispersal by integrating stepping-stone dispersal through multiple sites and multiple pathways, and to study the spatial structure of gene flow between populations (Murphy et al. 2015). Applications of spatial graphs to model multi-generational dispersal are scarce and restricted to the marine environment (e.g. modelling the dispersal of rafting brown algae at small spatial scale; Buonomo et al. 2017, and the dispersal of coral larvae at the scale of the Great Barrier Reef; Riginos et al. 2019). To our knowledge, only one recent study (Jahnke et al. 2018) applied spatial graphs to explicitly compare patterns of demographic and genetic connectivity, but this study did not calculate the multi-generational connectivity.

Here, we present a spatial graph approach that extends propagule dispersal networks with a stepping-stone model to represent dispersal between suitable habitat patches over multiple generations. This approach allows to estimate the multi-generation long-distance connectivity that is potentially detected using genetic data, and to compare explicitly the spatial structure of multi-generational genetic connectivity, single-generation propagule dispersal and multi-generation dispersal networks. Our approach also accounts for unsampled stepping-stone sites that contribute to the connectivity patterns between the sampled sites (the so-called 'ghost population' effect; Beerli 2004, Slatkin 2005). We then illustrate our approach by analyzing connectivity of a high gene-flow species, the red mullet *Mullus surmuletus*, in the Mediterranean Sea. Finally, we discuss how to extend our approach to other systems.

## Methods

### General approach: multi-generational spatial graphs for genetic connectivity

The first step of our approach to study multi-generational genetic connectivity is to model propagule dispersal in a given environment or habitat for a single generation. This first step produces dispersal probabilities between each pair of sites used to build a spatial graph (Box 1) of single-generation dispersal (Trembl et al. 2008, Kininmonth et al. 2010, Andrello et al. 2013).

### Box 1. How spatial graphs can illuminate connectivity

Spatial graphs stem from graph theory and help to better understand the dynamics of complex systems composed of many interacting units (Rozenfeld et al. 2008, Dale and Fortin 2010, Peterson et al. 2019). They explicitly consider the spatial context of data (Urban et al. 2009) and can be used to study dispersal and connectivity between multiple sites simultaneously (Dale and Fortin 2010, Eros et al. 2011). In graph theory, a graph  $G$  consists of a set of nodes or vertices that are pairwise linked by a set of edges  $E(G)$  (Urban and Keitt 2001). A graph can be built to study connectivity in a network by using populations or sampling localities as georeferenced nodes, rendering a spatial graph, and weighing the edges between them by their modelled dispersal to represent connection probability, or by their pairwise genetic distance to study gene flow (under the hypothesis that genetic differentiation is mainly the outcome of gene flow). In doing so, graph theory allows to visually represent the network of genetic connectivity and analyse topological features to test hypotheses of population structure (Dyer and Nason 2004, Dale and Fortin 2010, Xuereb et al. 2018). At the same time, graph theory allows the quantitative estimation of network- and node-level properties (Rozenfeld et al. 2008, Albert et al. 2013).

#### Interpreting graph metrics

Node metrics allow the identification of nodes, e.g. populations, that are of increased importance to the whole network by simultaneously considering all nodes and edges of the graph. High node degrees indicate populations that exchange propagules with many others (in a dispersal graph) or that have already exchanged genes with many others (in a genetic connectivity graph). Such populations have potentially sent or received many migrants over evolutionary time scales, acting as sources or sinks in the network (Rozenfeld et al. 2008). Node betweenness centrality quantifies the importance of single nodes in passing information through the network by identifying nodes that represent important stepping-stones for dispersers (in propagule dispersal graphs) or genes (in genetic connectivity graphs) (Garroway et al. 2008, Rozenfeld et al. 2008, Andreello et al. 2013). Network-level metrics allow us to quantitatively characterize the topology of a graph. They furthermore allow us to test landscape-genetic hypotheses by comparing the topology and associated characteristics of a genetic connectivity graph to graphs representing explanatory variables relating to structural or functional connectivity (Dale and Fortin 2010).

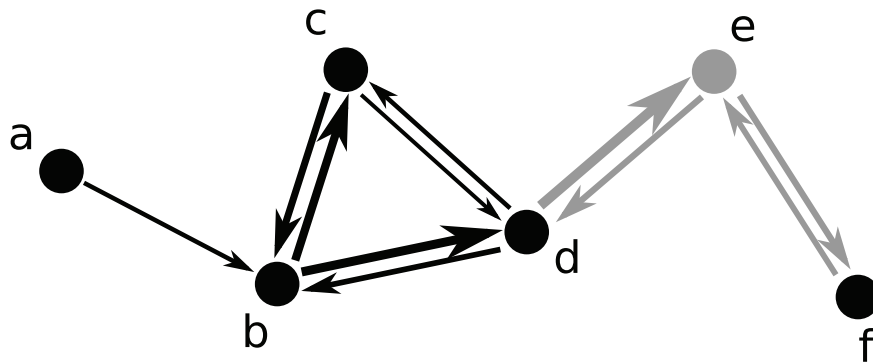
The second step examines nodes that are not directly connected in the single-generation dispersal graph and calculate the number of stepping-stones to connect the shortest paths between all pairs of nodes (Table 1). For example, if no direct connection exists between nodes  $a$  and  $c$  (Fig. 1, single-generation propagule dispersal), the shortest path algorithm identifies node  $b$  as the most suitable stepping-stone and indicates that one stepping-stone is needed to connect nodes  $a$  and  $c$  through  $b$ . In this way, the single-generation dispersal graph

allows to calculate multi-generational connectivity through stepping-stone nodes (Fig. 1, multi-generational propagule dispersal). To calculate the shortest paths, dispersal probabilities between nodes are transformed into distance measures by taking the natural logarithm of their inverse; this transformation yields a distance matrix where pairs of nodes with the highest dispersal probabilities have the smallest distance, thereby ensuring that the shortest path between two nodes is effectively the most probable one (Costa et al. 2017).

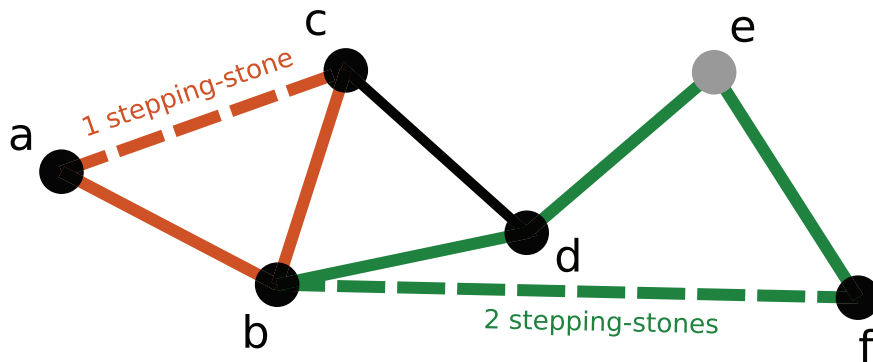
Table 1. Spatial graph metrics used in this study (Urban and Keitt 2001, Newman 2003, Urban et al. 2009).

Network assortativity	Measure of how well nodes in a network show similar characteristics. Typically calculated as degree assortativity, which quantifies whether high-degree nodes tend to attach to other high-degree nodes (assortative mixing) or whether high-degree nodes rather attach to low-degree ones (disassortative mixing). Measured as the Pearson correlation coefficient $r$ of the degrees at either ends of an edge, averaged over edges.
Network clustering	Probability that two nodes connected to a given third node are themselves connected, averaged for all the nodes in a network.
Node betweenness	Fraction of all shortest paths between nodes that passes through a given node. Quantifies the relative importance of a single node in passing information through the network. The betweenness centrality ( $C_B$ ) of node $i$ is calculated by taking the proportion of shortest paths connecting nodes $s$ and $t$ ( $\sigma_{st}$ ) that passes through node $i$ ( $\sigma_{st}(i)$ ), and summing them over all possible node pairs: $C_B(i) = \sum_{s \neq t \neq i} \frac{\sigma_{st}(i)}{\sigma_{st}}$
Node degree	Number of edges connected to a given node. Quantifies the local importance of a node.
Shortest path	Minimum-weight path connecting two nodes. In an unweighted graph, the shortest path represents the path with the smallest number of edges. In a weighted graph, the shortest path represents the path in which the total weight of the sequence of edges is the smallest.

### Single-generation propagule dispersal



### Multi-generational propagule dispersal



### Multi-generational genetic connectivity

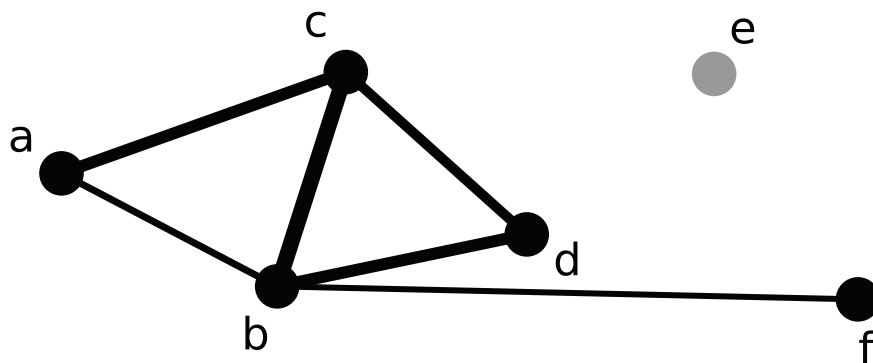


Figure 1. Spatial graphs to analyse single-generation propagule dispersal (top), multi-generational propagule dispersal (middle) and multi-generational genetic connectivity networks (bottom). Dots and arrows represent the nodes and the directional edges of the graph. Black dots are the sampled populations and grey dots are unsampled populations. Black and grey arrows are modelled propagule dispersal probabilities between sampled and unsampled populations, respectively. The width of the arrows is proportional to propagule dispersal probability (in the single-generation propagule dispersal graph) and inversely proportional to genetic distance (in the multi-generational genetic connectivity graph). In the multi-generational propagule dispersal graph, dispersal probabilities are used to estimate the number of stepping-stones needed to connect nodes that are not connected in the single-generation graph. Dashed lines represent multi-generational connectivity through stepping-stones and are highlighted for nodes a and c (brown) and nodes b and f (green). Adapted from Urban et al. (2009).

The shortest paths between all possible node pairs can then be estimated using the Floyd–Warshall algorithm for directed graphs (Floyd 1962, Warshall 1962). Unsampled sites potentially acting as stepping-stones are included in the multi-generational propagule dispersal graph.

The third step of our approach estimates a spatial graph of genetic distances among sites, hereafter referred

to as multi-generational genetic connectivity (Fig. 1) (Rozenfeld et al. 2008, Fortuna et al. 2009, Albert et al. 2013). The apparent discrepancies between genetic connectivity and propagule dispersal can in part be explained by the process of multi-generational dispersal connecting nodes over larger spatial scales than single-generation dispersal.

## Example: multi-generational genetic connectivity in the striped red mullet across the Mediterranean Sea

To illustrate our approach, we re-analysed a published genetic dataset of an economically important fish species with a pelagic larval stage, the striped red mullet *Mullus surmuletus*, in the Mediterranean Sea (Dalongeville et al. 2018b). We hypothesize that building a spatial graph of potential larval dispersal over multiple generations, can improve our understanding of the genetic structure across the Mediterranean Sea and of the temporal scale at which connectivity operates. We build the spatial graph of genetic connectivity to characterize the spatial pattern of genetic structure and to compare it to the graph of multi-generational larval dispersal.

### Species description, study area, sampling and genetic data

The striped red mullet *Mullus surmuletus* is a demersal fish species that is commonly found in the Mediterranean Sea particularly in areas characterised by a narrow continental shelf with rough substrate and at depths ranging from 10 to 100 m (Lombarte et al. 2000). Adults move to deeper sites for spawning (70–150 m, Machias et al. 1998), which occurs from April to May and produces larvae with a pelagic larval duration (PLD) of approximately 30 d (Reñones et al. 1995, Macpherson and Raventos 2006, Arslan and İş 2015). *Mullus surmuletus* is one of the most economically valuable species in commercial landings of coastal Mediterranean demersal fisheries (Reñones et al. 1995, Félix-Hackradt et al. 2013).

The Mediterranean Sea harbours a complex oceanic circulation that shapes larval dispersal patterns and creates putative barriers for gene flow (Pascual et al. 2017). To maximise the detection of these various processes influencing population structure, *M. surmuletus* samples were collected along the entire Mediterranean coastal range including islands (47 sites, Fig. 2a). Fin clips were taken from specimens obtained at local artisanal fisheries landing sites. DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's protocol, and individuals were genotyped using a Genotyping by Sequencing method (Elshire et al. 2011). Individual genotypes were then pooled by their sampling locality to maximize sequencing coverage (see Dalongeville et al. 2018b for a complete description of sampling, molecular and bioinformatics procedure). This resulted in a dataset containing the allele frequencies of 1153 Single Nucleotide Polymorphism (SNP) loci for 47 pools of nine to 18 individuals, hereafter referred to as populations.

Genetic differentiation among these populations was calculated in the 'ade4' package in R v3.2.3 (Dray and Dufour 2007, R Core Team) using Cavalli-Sforza and Edwards' chord distance  $D_c$  (Cavalli-Sforza and Edwards 1967), which is well-suited to distinguish genetically similar populations in high gene flow species (Libiger et al. 2009).

### Step 1. Single-generation propagule dispersal

Larval dispersal was simulated using the biophysical model of Andreollo et al. (2013), which was adapted to fit the life history parameters of *M. surmuletus* (Dalongeville et al. 2018b).

In brief, three-dimensional sea current velocities were obtained from the hydrodynamics model NEMOMED12 (Beuvier et al. 2012) and used to simulate passive larval dispersion using the software Ichthyop 3.1 (Lett et al. 2008). Larvae were released in 1/10th degree cells covering the Mediterranean continental shelf (7703 cells; Andreollo et al. 2015) every three days during the species' spawning season (1–28 May) and allowed to passively disperse during 30 d. Pairwise probabilities of larval dispersal  $c(i,j)$  between cells  $i$  and  $j$  were calculated from the numbers of simulated larvae released in cell  $j$  that arrived in cell  $i$  after 30 d. Probabilities of dispersal were averaged over a grid of 100 cells covering the full coastal range of the Mediterranean Sea, resulting in a  $100 \times 100$  asymmetric connectivity matrix  $C_{100}$ . Not every grid cell could be sampled for the genetic analysis described above. To allow comparison between the larval dispersal probabilities and the genetic connectivity estimates between the 47 sampled sites, a  $47 \times 47$  connectivity matrix  $C_{47}$  was extracted from the  $C_{100}$  matrix with the cell centroids corresponding to the 47 sites sampled for the genetic analysis (Supplementary material Appendix 1 Fig. A1).

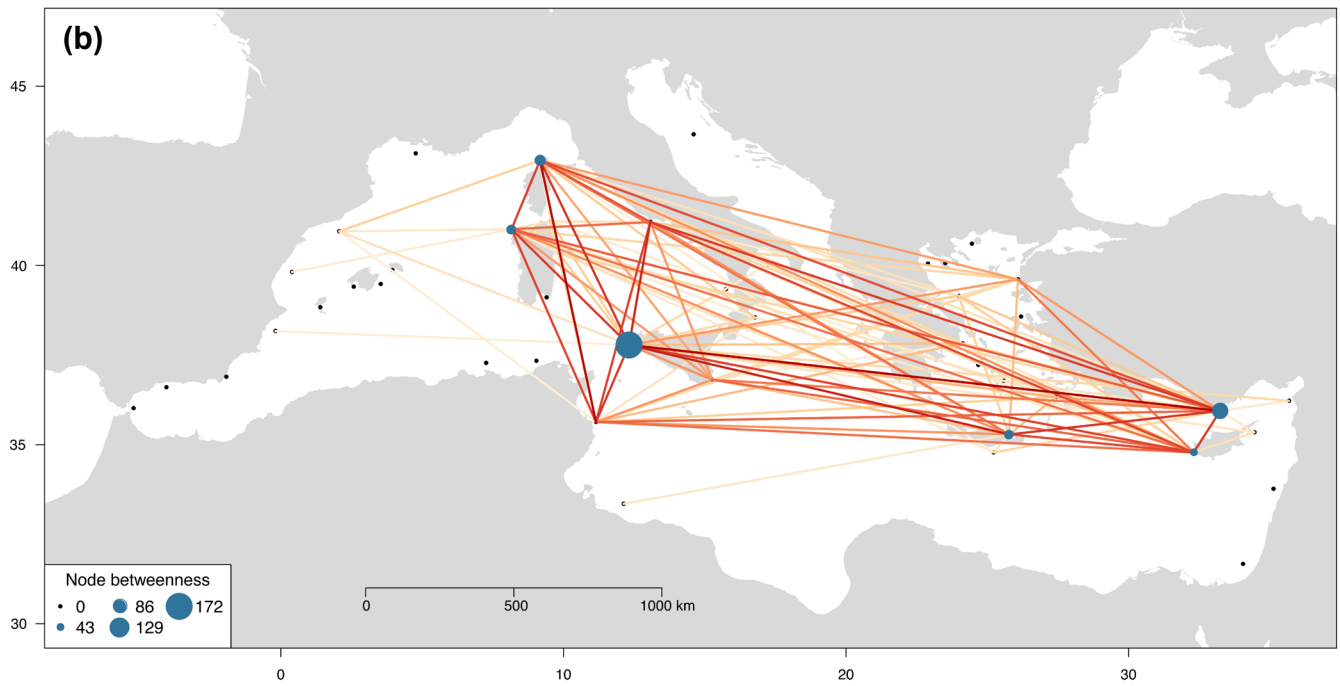
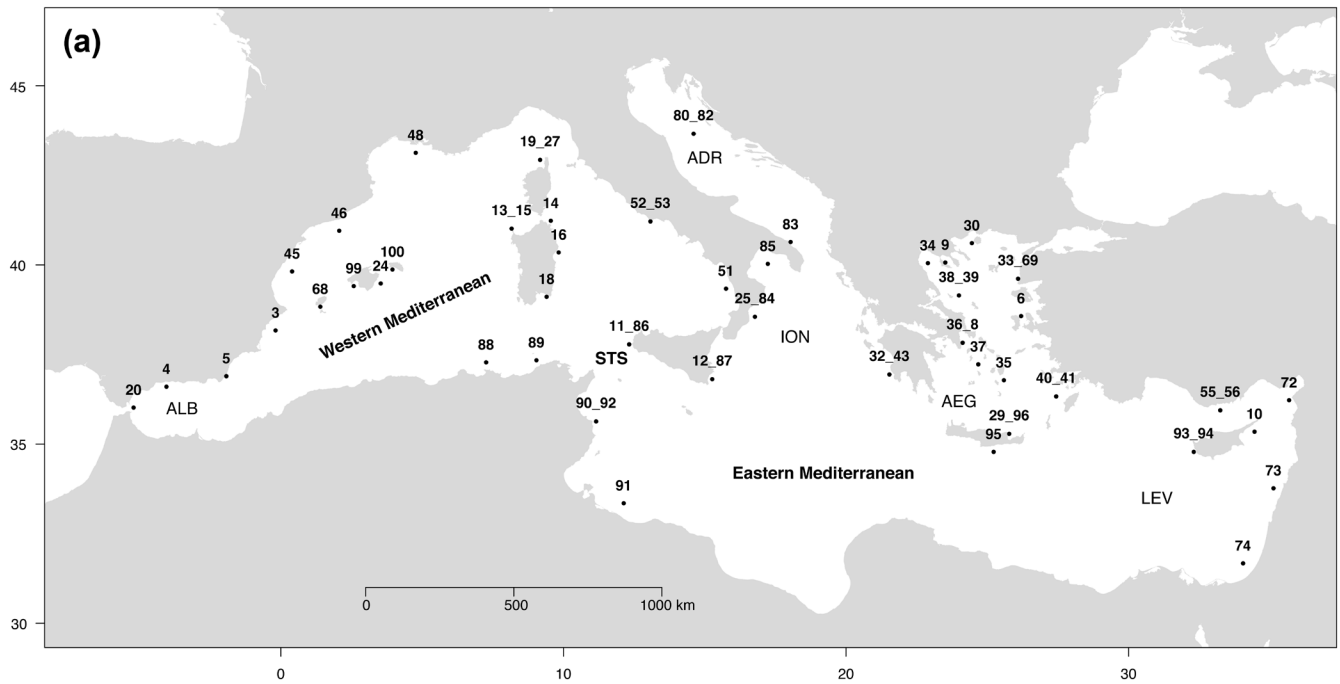
We constructed directed spatial graphs of single-generation larval dispersal both for the  $47 \times 47$  and  $100 \times 100$  larval connectivity matrices. Edges of the spatial graphs were weighted by the modelled larval dispersal probabilities and nodes represented the cell-centroids of the 47 populations and the 100 modelled sites, respectively. The 47 populations are connected through single-generation larval dispersal by 108 edges.

Marine geographic distances (minimum distances constrained by water) were computed as least-cost path distances between all pairs of populations by assigning infinite resistance to land areas and constant zero resistance to water using the functions 'transition', 'geoCorrection' and 'costDistance' of the R package 'gdistance' v. 1.1-1 (van Etten 2017). The nearest geographical neighbour was additionally determined for each population.

### Step 2. Multi-generational propagule dispersal

Both larval dispersal matrices ( $C_{100}$  and  $C_{47}$ ) reflect dispersal only from the current generation. We calculated larval connections for the matrix  $C_{100}$  between non-connected nodes based on a stepping-stone approach. We hypothesize that larvae that dispersed from site  $j$  to  $i$  successfully settle in their arrival site and, upon reaching maturity, produce larvae that potentially disperse to other sites. As *M. surmuletus* populations occur across the full Mediterranean coastal range, we hypothesize that all the 100 grid cells represent suitable habitat and all sites contribute equally to larval production.

After transforming the single-generation larval dispersal probabilities of  $C_{100}$  into distance measures, we calculated the shortest paths between all possible node pairs using the Floyd-Warshall algorithm embedded in the R package 'Rfast' (Papadakis et al. 2018) (Supplementary material Appendix 1 Fig. A2). We then extracted the multi-generational dispersal measures between the 47 sampled populations from the  $C_{100}$  multi-generational dispersal graph (Supplementary material Appendix 1 Fig. A3). The final spatial graph of



Pairwise  $D_c$

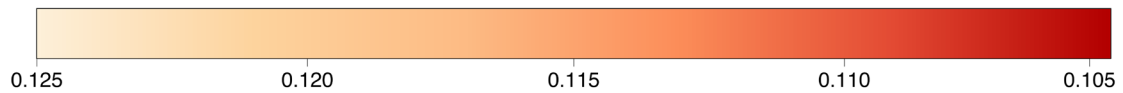


Figure 2. Spatial graph of genetic connectivity. (a) Map of the Mediterranean Sea with the location of the 47 sampling sites, the Siculo-Tunisian Strait (STS), Mediterranean regions and ecoregions (ALB, Alboran Sea; ADR, Adriatic Sea; ION, Ionian Sea; AEG, Aegean Sea; LEV, Levantine Sea). (b) Spatial graph representing the genetic connectivity network of *Mullus surmuletus* in the Mediterranean Sea. Edges between nodes (populations) are weighted by their pairwise genetic distance ( $D_c$ ), represented by a colour gradient (dark red for the smallest distances and thus the highest genetic similarity). The size of the nodes is proportional to their betweenness centrality.

multi-generational dispersal is constructed using the 47 populations as nodes and weighting the edges by the number of stepping stones connecting them.

### **Step 3. Spatial graph of multi-generational genetic connectivity**

The network of genetic connectivity was constructed by designating the 47 sampled populations as nodes, and their pairwise genetic distance ( $D_i$ ) as weighted edges. The construction of a spatial graph from all  $D_i$  values results in a fully saturated network, where every node is connected to every other node. The calculation of most graph metrics, such as node degree, network assortativity and network clustering, depends on the number of edges connected to the graph's nodes and cannot be calculated in a saturated network (Table 1). As the weights of the genetic graph have a small range (see further), the shortest-path algorithm always identified the direct path between two nodes as the most efficient one and impeded the calculation of informative node betweenness. We thus proceeded to prune the graph edges to get a more informative graph topology (Garroway et al. 2008). Edges were pruned by retaining only the 108 edges with the smallest genetic distance (approximately 10% of all edges), thus obtaining a multi-generational genetic connectivity graph with the same number of edges as the single-generation larval dispersal graph. The retained connections are considered as the relevant genetic relationships for further network analyses.

As this pruning parameter is arbitrary, we also implemented an edge removal scenario where edges of decreasing genetic distances were discarded one by one. The resulting genetic connectivity spatial graphs were compared after each removal to test the influence of the pruning parameter choice on the graph topology. Additionally, we ranked each node according to their betweenness values for each node removal scenario and compared the variation in node ranks.

### **Analysis of the spatial graphs**

Network topology was analysed by estimating several graph metrics (Table 1). For the multi-generational genetic connectivity graph, we calculated node degree and node betweenness centrality to identify important stepping-stone nodes in the network (Box 1) (Rozenfeld et al. 2008). Network assortativity and network clustering were calculated to evaluate the structure of the genetic network. To determine whether assortativity and clustering resulted from biological processes (such as habitat patch location and quality and between-habitat migration) versus random configuration, we generated 10 000 random networks with the same number of nodes and edges as our genetic connectivity network following the Erdős–Rényi model (Erdős and Rényi 1959). The mean graph metrics (assortativity and clustering) and their standard deviations were calculated and compared to the values of the genetic graph. Network assortativity and clustering were also calculated for the single-generation propagule dispersal graph.

The pruned genetic graph was constructed with the 'gengraph' function in the R package 'adegenet', and all

subsequent graph-theoretic analyses were conducted with the R package 'igraph' (functions 'degree', 'betweenness', 'assortativity\_degree' and 'transitivity') (Csardi and Nepusz 2006, Jombart 2008). Random networks were generated using the 'sample\_gnm' function.

Correlations between genetic distances and marine geographic distances, single-generation larval dispersal probabilities as well as multi-generational larval dispersal distances were estimated using the Mantel coefficient along with 9999 permutations (R package 'vegan', Oksanen et al. 2016). As the Mantel test is known to have an inflated type I error rate and low statistical power (Legendre and Fortin 2010, Guillot and Rousset 2013), we also computed maximum-likelihood population-effects (MLPE) mixed models (Clarke et al. 2002) which account for the non-independence of pairwise data and provide the greatest probability of identifying the true model from competing alternatives (Shirk et al. 2018). We used the 'lmer' function of the R package 'lme4' to fit the MLPE mixed models with a random-effects term accounting for population-level influence (Bates et al. 2015, Row et al. 2017). Model fit was evaluated using the coefficient of determination  $R^2$  calculated with the 'MuMIn' package (Bartoń 2019).

## **Results**

### **Single-generation larval dispersal**

The larval dispersal matrix  $C_{47}$  is highly sparse, with only 108 realised connections out of 1081 (Supplementary material Appendix 1 Fig. A4). The constructed larval network connects all nodes to at least one other node with connections primarily made to the nearest neighbours. It is fragmented in three main clusters roughly representing the western Mediterranean, the Aegean sea and the Levantine sea (Fig. 3a, Supplementary material Appendix 1 Fig. A5). The graph has a high level of clustering and a positive degree assortativity (Supplementary material Appendix 1 Table A1).

The  $C_{100}$  network of modelled larval dispersal however fully connects the Mediterranean Sea, i.e. the network consists of one large cluster connecting all nodes to at least three other nodes (Supplementary material Appendix 1 Fig. A6).

### **Multi-generational larval dispersal**

The stepping-stone approach to larval connectivity converted our sparse connectivity matrix  $C_{47}$  into a fully saturated connectivity matrix, although a substructuring separating the western and eastern basin was still visible (Fig. 3a). The spatial graph of the multi-generational larval dispersal network shows that the Aegean and Levantine Sea are connected when including one stepping-stone (Fig. 3b), and that the west-east barrier is breached when including two stepping-stones (Fig. 3c). Allowing for more stepping-stones increased graph connectivity (Fig. 3d–e), saturating the network with 24 stepping-stones (suggesting 25 generations).



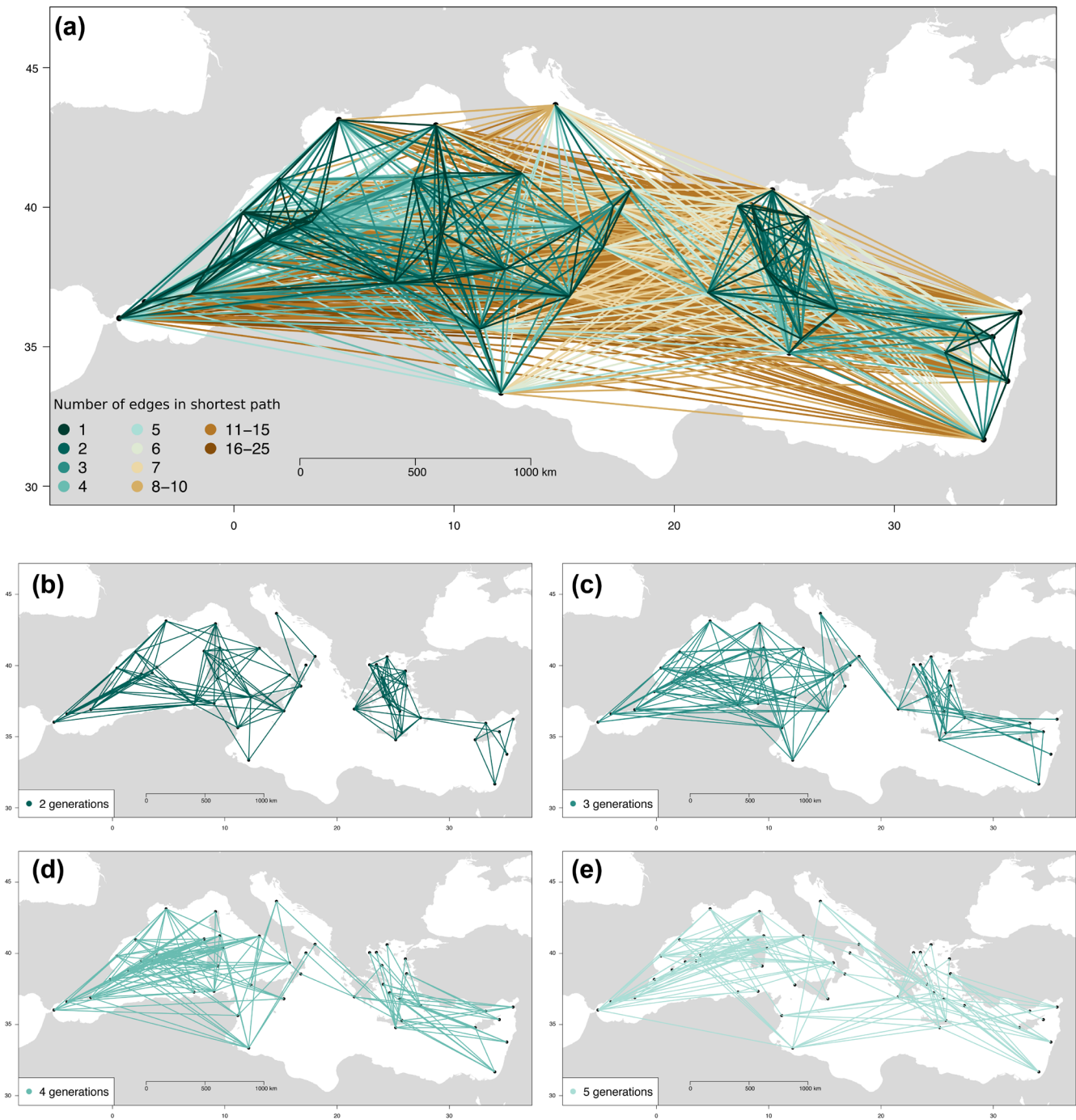


Figure 3. Temporal perspective of marine connectivity: spatial graphs representing the amount of generations needed to connect our 47 populations through stepping-stone larval dispersal. (a) Graph with all pairwise edges, the colours represent the number of edges in the shortest path (= number of generations) connecting each node pair. The darkest green colour (1 edge) represents the direct larval dispersal over one generation. The Mediterranean Sea is fully connected through larval dispersal when allowing up to 25 edges (or 24 stepping stones) in the shortest paths (a). (b–e) Graph built using edges representing two, three, four and five generations respectively.

### Multi-generational genetic connectivity

Pairwise genetic distance  $D_c$  varied between 0.10 and 0.20. The pruned spatial graph built with these genetic distances connected 29 out of 47 nodes (Fig. 2b). Topological

analysis of the graph reveals no substructuring into clusters but instead more and stronger connections between the western and eastern Mediterranean basin than within each basin. Network metrics showed a high level of disassortative mixing with  $r = -0.57$ , which was considerably

lower than that of the randomized networks ( $r_0 = -0.05$ ) (Supplementary material Appendix 1 Table A1), which indicates a centralized network. This means that the genetic network has a central core of a few highly connected populations (i.e. with a high degree), that are in turn linked to lesser connected populations (i.e. with a low degree) in the periphery (Supplementary material Appendix 1 Fig. A7). The level of clustering  $c = 0.58$  was considerably higher than the mean of the randomized graphs ( $c_0 = 0.10$ ). These stronger levels of disassortativity and clustering demonstrate that the genetic connectivity network is more structured than would be expected from a random process alone. The centralised structure was robust under varying levels of pruning (Supplementary material Appendix 1 Fig. A8). The sequential removal of edges also shows that the unconnected nodes in the pruned graph are not completely isolated but rather increasingly differentiated from the central structure.

The values of betweenness centrality also suggest a centralized structure, which attribute an importance to only six out of the 29 connected nodes. In particular, populations 11\_86 in Sicily (present in 50.0% of all shortest paths) and 55\_56 in Turkey (23.3%) appear to be the main stepping stones for the gene flow through the network (Fig. 1b). The remaining populations with a nonzero betweenness are 19\_27 in Corsica (9.9%), 13\_15 in Sardinia (8.1%), 29\_96 in Crete (7.8%) and 93\_94 in Cyprus (0.9%). These six nodes consistently ranked highest out of all nodes by their betweenness values under varying pruning criteria (Supplementary material Appendix 1 Fig. A9).

### Comparison between single-generation larval dispersal, multi-generational larval dispersal and multi-generational genetic connectivity

The mean larval dispersal was estimated to 266 km (SD = 137) for the single-generation dispersal graph and to 1139 km (SD = 694) for the genetic connectivity graph, revealing the different spatial scales of both components of connectivity. Sixty-six of the 72 nearest neighbour pairs were connected in the single-generation larval dispersal graph, whereas only 16 of the nearest neighbour pairs were connected in the genetic connectivity graph.

The Mantel test and maximum-likelihood population-effects (MLPE) mixed model between the genetic distances (pairwise  $D_c$ ) and the geographic distances for the 47 populations were significant (Mantel  $r = 0.31$ , MLPE  $R^2_m = 0.17$ ,

both  $p < 0.001$ ) (Table 2) indicating a pattern of isolation by distance (IBD). The Mantel test and MLPE model between genetic distances and single-generation larval dispersal were not significant (Table 2). Conversely, the Mantel test and MLPE model between genetic distances and multi-generational larval dispersal revealed a significant positive correlation (Mantel  $r = 0.18$ , MLPE  $R^2_m = 0.12$ , both  $p < 0.001$ ) (Table 2), indicating an isolation by larval dispersal when considering multiple generations. The multi-generational estimates of larval dispersal were also significantly correlated to the geographic distances (Mantel  $r = 0.70$ ,  $p = 0.001$ ).

## Discussion

Here we demonstrate how stepping-stone dispersal graphs can be used as a proxies of multi-generational connectivity and increase our understanding of genetic connectivity. We present a spatial graph approach applied to red mullet in the Mediterranean Sea that calculates shortest paths between populations unconnected by direct modelled larval dispersal but showing strong genetic connectivity. The increased correlation between the resulting multi-generational connectivity and genetic connectivity, compared to single-generation dispersal, highlights the underlying discrepancy in spatio-temporal scales in our landscape genetic study.

The spatial graph of single-generation larval dispersal was mostly realized between nearest neighbours, resulting in a highly clustered structure. The spatial graph of genetic connectivity revealed an absence of subdivision between the western and eastern Mediterranean sub basins and a stronger between-basin than within-basin connectivity. No significant correlation could be found between the genetic distances and the direct larval dispersal probabilities at the scale of the Mediterranean Sea.

Visualisation of the larval dispersal and gene flow network showed that genetic connectivity is realised at a much larger spatial scale than larval dispersal, as already found by Dalongeville et al. (2018b), reflecting different temporal scales. Biophysical connectivity is usually modelled for a single generation of larvae (White et al. 2010), whereas genetic connectivity integrates gene flows over multi-generational time scales (Hedgecock et al. 2007). Intermediate stepping-stones for dispersal are crucial in maintaining genetic connectivity across large spatial scales (Crandall et al. 2012). The shortest-path approach to larval dispersal explores the temporal

Table 2. Mantel test and maximum-likelihood population-effects (MLPE) mixed model results comparing the multi-generational genetic connectivity (Cavalli-Sforza and Edwards' chord distance  $D_c$ ) to the geographic distance, single-generation propagule dispersal (larval dispersal probability  $c_{ij}$ ) and the multi-generational propagule dispersal (stepping-stone larval distance). Only the geographic distance and stepping-stone larval distance were significantly correlated to the genetic chord distance ( $p < 0.001$ ).

Predictor variable	Genetic chord distance $D_c$			
	Mantel correlation	p value	MLPE $R^2_m$	p value
Geographic distance	$r = 0.31$	<b>&lt; 0.001</b>	0.17	<b>&lt; 2.2e-16</b>
Larval dispersal probability ( $c_{ij}$ )	$r = 0.02$	0.24	0.0007	0.37
Stepping-stone larval distance (Floyd-Warshall shortest path on $\ln(1/c_{ij})$ )	$r = 0.18$	<b>&lt; 0.001</b>	0.12	<b>&lt; 2.2e-16</b>

aspect of marine connectivity by suggesting how sites would be connected if *M. surmuletus* larvae were successfully spread through the Mediterranean Sea over multiple generations following a stepping-stone model. Multigenerational dispersal can thus increase the explanatory power of dispersal models for population genetic analyses, and our understanding of the drivers of population genetic structure (White et al. 2010, Buonomo et al. 2017).

Connectivity can promote the spread of adaptive alleles from populations locally adapted to presently extreme environmental conditions that could become more common in the future. Such spreading of adaptive alleles can strengthen the adaptation of populations to future environmental conditions, and in some cases even result in genetic rescue of declining populations (Whiteley et al. 2015, Xuereb et al. 2019). In the marine environment, populations can be locally adapted to climate-related variables such as salinity (Dalongeville et al. 2018a). Climate projections under a business-as-usual scenario (RCP8.5) predict that sea surface salinity (SSS) in the Mediterranean Sea will increase, with marked regional differences, by 0.13 practical salinity unit ( $\pm 0.13$  PSU) for the 2021–2050 period and to come back to its current global climate value ( $\pm 0.01$  PSU) for the 2071–2100 period (Moullec et al. 2019). In some areas, the increase in SSS might result in extreme environmental conditions that cannot be tolerated by populations of *M. surmuletus*, because the future conditions will be too different from the present conditions to which the populations are adapted (Rellstab et al. 2016). In order to persist, these populations will have to move to more suitable areas or adapt to the new local conditions from standing genetic variation and/or immigration of adaptive genotypes. Analysis of the association between SNP and salinity have suggested a signal of adaptation to local water salinity in Mediterranean populations of *M. surmuletus* (Dalongeville et al. 2018a). The results of our spatial graph analysis indicate that about five generations will be necessary to spread beneficial alleles from populations adapted to current high salinity in the east, to populations with similar projected future physiochemical conditions in the central Mediterranean (Fig. 3e). Given an average age at first reproduction of 1.5 yr for *M. surmuletus* (Reñones et al. 1995), this corresponds to one decade which has been judged as a small enough timescale to keep pace with the temporal scale of climatic changes (Jönsson and Watson 2016).

The strong basin-wide genetic connectivity highlighted by the genetic spatial graph identified the Siculo-Tunisian Strait (STS) as an important corridor for gene flow. Beyond analysing the topology of a spatial network, graph theory allows to quantitatively analyse the properties of the nodes and edges in the spatial graph (Rozenfeld et al. 2008). The calculation of network- and node-metrics revealed a disassortative, centralised network. This means that the network consists of a central core with a few high-degree hubs, which are in turn linked to lesser connected nodes in the periphery (Rozenfeld et al. 2008). These highly connected hubs could act as sources and/or sinks in the gene flow network (Peery et al. 2008). Betweenness centrality measures the

importance of a single node in relaying information through the network (Rozenfeld et al. 2008). The populations with high genetic betweenness could thus be interpreted as prominent stepping-stones for gene flow. These are populations where individuals from multiple locations seem to have aggregated and reproduced, to subsequently further spread their genes to multiple other locations. Six populations, located in Sicily (11\_86), Turkey (55\_56), Corsica (19\_27), Sardinia (13\_15), Crete (29\_96) and Cyprus (93\_94) were identified as important stepping-stones in our network. These populations lay along the longitudinal axis and seem to form a path connecting the western and eastern Mediterranean. Remarkably, five out of these six populations are located along islands and coincide with the location of the high degree nodes. These results detect islands as important stepping-stones in the network-wide genetic connectivity of *M. surmuletus*, and suggest that islands can be important sites for gene flow between continental populations by providing intermediate suitable habitats for larval settlement in the deeper parts of the Mediterranean basin.

The construction of a genetic spatial graph requires defining a criterion to prune the initially saturated graph. The appropriate strategy depends largely on the questions asked. For example Dyer and Nason (2004) developed the population graph approach where graph edges are trimmed based on the genetic covariance among populations. This approach requires individual genotypes and is not applicable to our population-level data. Because one of our hypotheses was that genetic connectivity is in part the result of larval dispersal, edges were pruned by retaining only the 108 edges with the smallest genetic distance (approximately 10% of all edges), thus obtaining a multi-generational genetic connectivity graph with the same number of edges as the single-generation larval dispersal graph. The disassortative structure with a strong cross-basin connectivity in the gene flow graph remained robust under the edge removal scenario. The nodes with the higher betweenness values in the pruned gene flow graph also consistently ranked higher under varying pruning values. The consistency of these results validates that our pruned graph correctly represents the gene flow patterns of *M. surmuletus* in the Mediterranean Sea.

A large portion of the genetic differences still remains unexplained. Bottom-trawl surveys suggest that *M. surmuletus* migrates to deeper water during maturation but realises rather short-distance than long-distance movement (Machias et al. 1998). Together with the small home ranges of sympatric demersal species (reviewed by Calò et al. 2013) this suggests that adult dispersal might be rather limited. Other processes that could explain the remaining genetic structure are larval behaviour and higher-resolution oceanographic circulation patterns not resolved by the hydrodynamic model used to simulate larval dispersal (Briton et al. 2018, Faillettaz et al. 2018). Larvae at the settlement stage can use oriented swimming driven by e.g. odour or sound to detect favourable settlement habitats (Simpson 2005, Paris et al. 2013). Shoreward swimming at the end of the pelagic stage can strongly influence the rate of recruitment

of larvae, resulting in longer dispersal distances than predicted with purely passive dispersal (Faillietaz et al. 2017). In contrast, vertical migration of larvae at night as well as homing behaviour of larvae returning to their natal reef result in greater retention than predicted by advection models and smaller dispersal distances (Gerlach et al. 2007, Andrello et al. 2013, Bottesch et al. 2016). This would decrease the amount of connectivity in the larval dispersal system and would highlight even more the importance of stepping-stone migration over multiple generations to realize the observed long-distance genetic connectivity. Predicting how larval behaviour would affect larval dispersal patterns thus requires including larval swimming and orientation mechanisms in the larval dispersal model, but the parametrization of these processes would be difficult because of the lack of biological knowledge on larval dispersal in *M. surmuletus*.

Although applied here to the larval dispersal of a demersal fish species, the stepping-stone approach to dispersal can be generalized to other environments or living systems. Connectivity estimates can be applied to any organisms with a dispersing propagule stage, by modelling wind dispersal (Nathan et al. 2011), animal-mediated seed dispersal (García-Fernández et al. 2019) or riverine resistance (Schick and Lindley 2007, Oliveira et al. 2019). The approach can be adapted by modelling dispersal between species-specific suitable habitats such as coral reefs (Williamson et al. 2016), forests (Wang et al. 2008) or breeding ponds (Fortuna et al. 2006, Decout et al. 2012). Additionally, the model could be improved by taking into account species-specific settlement success rates, population-specific productivity rates and behavioural responses.

## Data availability statement

Data for this study are available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.31tk592>> (Dalongeville et al. 2018c). The R script for the shortest-path analysis and to produce Fig. 3 is available at <<https://github.com/eboulanger/seaConnect--spatialGraph>>.

**Acknowledgements** – We thank Raquel Torres, Laure Velez, Jennifer Tournois, Eva Maire, Sebastien Villeger, Johanna Brun-Vitelli, Manuel Muntoni, Claire Secch, Sofène Ben Abdelhamid, Achraf Barbena, Itai Granot, Ori Frid and Shahar Malamud for collecting *M. surmuletus* samples, as well as all the fishermen that provided the fishes. We are grateful to the SMGE platform, and especially Veronique Arnal and Marie-Pierre Dubois for technical support during molecular analyses. We also thank Stéphane Lobreaux for the bioinformatic analyses. We thank the anonymous referees as well as the editor for their insightful comments.

**Funding** – This study was funded by the ‘ANR’ through the RESERVEBENEFIT project, PSL Environment through the project RESMOD, and EB was partially funded by the Region Occitanie.

**Author contributions** – All authors generated hypotheses and designed research. AD conducted the sampling and molecular analyses; EB analysed the data and wrote the first draft. All authors helped to improve manuscript writing.

## References

- Albert, E. M. et al. 2013. Assessing the robustness of networks of spatial genetic variation. – *Ecol. Lett.* 16: 86–93.
- Anderson, C. D. et al. 2010. Considering spatial and temporal scale in landscape-genetic studies of gene flow. – *Mol. Ecol.* 19: 3565–3575.
- Andrello, M. et al. 2013. Low connectivity between Mediterranean marine protected areas: a biophysical modeling approach for the dusky grouper *Epinephelus marginatus*. – *PLoS One* 8: e68564.
- Andrello, M. et al. 2015. Extending networks of protected areas to optimize connectivity and population growth rate. – *Ecography* 38: 273–282.
- Arslan, M. and İş, A. 2015. Age, growth and reproduction of *Mullus surmuletus* (Linnaeus, 1758) in Saros Bay (northern Aegean Sea). – *J. Black Sea/Mediterr. Environ.* 19: 217–233.
- Baguette, M. et al. 2013. Individual dispersal, landscape connectivity and ecological networks. – *Biol. Rev.* 88: 310–326.
- Bartoń, K. 2019. MuMIn: multi-model inference. – R package ver. 1.43.6.
- Bates, D. et al. 2015. Fitting linear mixed-effects models using lme4. – *J. Stat. Softw.* 67: 1–48.
- Beerli, P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. – *Mol. Ecol.* 13: 827–836.
- Beuvier, J. et al. 2012. MED12, oceanic component for the modeling of the regional Mediterranean Earth system. – *Mercat. Ocean Q. Newsl.* 46: 60–66.
- Bode, M. et al. 2019. Successful validation of a larval dispersal model using genetic parentage data. – *PLoS Biol.* 17: e3000380.
- Bottesch, M. et al. 2016. A magnetic compass that might help coral reef fish larvae return to their natal reef. – *Curr. Biol.* 26: R1266–R1267.
- Briton, F. et al. 2018. High-resolution modelling of ocean circulation can reveal retention spots important for biodiversity conservation. – *Aquat. Conserv. Mar. Freshwater Ecosyst.* 28: 882–893.
- Buonomo, R. et al. 2017. Habitat continuity and stepping-stone oceanographic distances explain population genetic connectivity of the brown alga *Cystoseira amentacea*. – *Mol. Ecol.* 26: 766–780.
- Calò, A. et al. 2013. A review of methods to assess connectivity and dispersal between fish populations in the Mediterranean Sea. – *Adv. Oceanogr. Limnol.* 4: 150–175.
- Castillo, J. A. et al. 2014. Landscape effects on gene flow for a climate-sensitive montane species, the American pika. – *Mol. Ecol.* 23: 843–856.
- Castorani, M. C. N. et al. 2017. Fluctuations in population fecundity drive variation in demographic connectivity and metapopulation dynamics. – *Proc. R. Soc. B* 284: 20162086.
- Cavalli-Sforza, L. L. and Edwards, A. W. 1967. Phylogenetic analysis. Models and estimation procedures. – *Evolution* 21: 550–570.
- Clarke, R. T. et al. 2002. Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. – *J. Agric. Biol. Environ. Stat.* 7: 361–372.
- Costa, A. et al. 2017. On the calculation of betweenness centrality in marine connectivity studies using transfer probabilities. – *PLoS One* 12: e0189021.
- Crandall, E. D. et al. 2012. Coalescent and biophysical models of stepping-stone gene flow in neritid snails. – *Mol. Ecol.* 21: 5579–5598.

- Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. – *InterJounal Complex Syst.* 1695: 1–9.
- Dale, M. R. T. and Fortin, M.-J. 2010. From graphs to spatial graphs. – *Annu. Rev. Ecol. Evol. Syst.* 41: 21–38.
- Dalongeville, A. et al. 2018a. Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (*Mullus surmuletus*). – *BMC Genomics* 19: 1–13.
- Dalongeville, A. et al. 2018b. Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. – *Evol. Appl.* 11: 1437–1447.
- Dalongeville, A. et al. 2018c. Data from: Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. – Dryad Digital Repository, <<https://doi.org/10.5061/dryad.31tk592>>.
- Decout, S. et al. 2012. Integrative approach for landscape-based graph connectivity analysis: a case study with the common frog (*Rana temporaria*) in human-dominated landscapes. – *Landscape Ecol.* 27: 267–279.
- Donati, G. F. A. et al. 2019. A process-based model supports an association between dispersal and the prevalence of species traits in tropical reef fish assemblages. – *Ecography* 42: 1–12.
- Dray, S. and Dufour, A. B. 2007. The ade4 package: implementing the duality diagram for ecologists. – *J. Stat. Softw.* 22: 1–20.
- Dyer, R. J. and Nason, J. D. 2004. Population graphs: the graph theoretic shape of genetic structure. – *Mol. Ecol.* 13: 1713–1727.
- Elshire, R. J. et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. – *PLoS One* 6: e19379.
- Epps, C. W. et al. 2007. Optimizing dispersal and corridor models using landscape genetics. – *J. Appl. Ecol.* 44: 714–724.
- Erdős, P. and Rényi, A. 1959. On random graphs. – *Publ. Math.* 6: 290–297.
- Eros, T. et al. 2011. Network thinking in riverscape conservation – a graph-based approach. – *Biol. Conserv.* 144: 184–192.
- Escalante, M. A. et al. 2018. The interplay of riverscape features and exotic introgression on the genetic structure of the Mexican golden trout (*Oncorhynchus chrysogaster*), a simulation approach. – *J. Biogeogr.* 45: 1500–1514.
- Faillietaz, R. et al. 2017. Swimming speeds of Mediterranean settlement-stage fish larvae nuance Hjort's aberrant drift hypothesis. – *Limnol. Oceanogr.* 63: 509–523.
- Faillietaz, R. et al. 2018. Larval fish swimming behavior alters dispersal patterns from marine protected areas in the north-western Mediterranean Sea. – *Front. Mar. Sci.* 5: 1–12.
- Félix-Hackradt, F. C. et al. 2013. Discordant patterns of genetic connectivity between two sympatric species, *Mullus barbatus* (Linnaeus, 1758) and *Mullus surmuletus* (Linnaeus, 1758), in south-western Mediterranean Sea. – *Mar. Environ. Res.* 92: 23–34.
- Fischer, J. and Lindenmayer, D. B. 2007. Landscape modification and habitat fragmentation: a synthesis. – *Global Ecol. Biogeogr.* 16: 265–280.
- Floyd, R. W. 1962. Algorithms 97: shortest path. – *Commun. ACM* 5: 345.
- Fortuna, M. A. et al. 2006. Spatial network structure and amphibian persistence in stochastic environments. – *Proc. R. Soc. B* 273: 1429–1434.
- Fortuna, M. A. et al. 2009. Networks of spatial genetic variation across species. – *Proc. Natl Acad. Syst. USA* 106: 19044–19049.
- García-Fernández, A. et al. 2019. Herbivore corridors sustain genetic footprint in plant populations: a case for Spanish drove roads. – *PeerJ* 7: e7311.
- Garroway, C. J. et al. 2008. Applications of graph theory to landscape genetics. – *Evol. Appl.* 1: 620–630.
- Gerlach, G. et al. 2007. Smelling home can prevent dispersal of reef fish larvae. – *Proc. Natl Acad. Sci. USA* 104: 858–863.
- Guillot, G. and Rousset, F. 2013. Dismantling the Mantel tests. – *Methods Ecol. Evol.* 4: 336–344.
- Hedgecock, D. et al. 2007. Genetic approaches to measuring connectivity. – *Oceanography* 20: 70–79.
- Hughes, A. R. and Stachowicz, J. J. 2004. Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. – *Proc. Natl Acad. Sci. USA* 101: 8998–9002.
- Jahnke, M. et al. 2018. Seascape genetics and biophysical connectivity modelling support conservation of the seagrass *Zostera marina* in the Skagerrak-Kattegat region of the eastern North Sea. – *Evol. Appl.* 11: 645–661.
- Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. – *Bioinformatics* 24: 1403–1405.
- Jönsson, B. F. and Watson, J. R. 2016. The timescales of global surface-ocean connectivity. – *Nat. Commun.* 7: 1–6.
- Kininmonth, S. J. et al. 2010. Graph theoretic topology of the Great but small Barrier Reef world. – *Theor. Ecol.* 3: 75–88.
- Legendre, P. and Fortin, M. J. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. – *Mol. Ecol. Resour.* 10: 831–844.
- Lett, C. et al. 2008. A Lagrangian tool for modelling ichthyoplankton dynamics. – *Environ. Model. Softw.* 23: 1210–1214.
- Libiger, O. et al. 2009. Comparison of genetic distance measures using human SNP genotype data. – *Hum. Biol. Int. Rec. Res.* 81: 389–406.
- Lombarte, A. et al. 2000. Spatial segregation of two species of Mullidae in relation to habitat. – *Mar. Ecol. Prog. Ser.* 206: 239–249.
- Lowe, W. H. and Allendorf, F. W. 2010. What can genetics tell us about population connectivity? – *Mol. Ecol.* 19: 3038–3051.
- Machias, A. et al. 1998. Bathymetric distribution and movements of red mullet *Mullus surmuletus*. – *Mar. Ecol. Prog. Ser.* 166: 247–257.
- Macpherson, E. and Raventos, N. 2006. Relationship between pelagic larval duration and geographic distribution of Mediterranean littoral fishes. – *Mar. Ecol. Prog. Ser.* 327: 257–265.
- Magris, R. A. et al. 2018. Biologically representative and well-connected marine reserves enhance biodiversity persistence in conservation planning. – *Conserv. Lett.* 11: e12439.
- Manel, S. et al. 2003. Landscape genetics: combining landscape ecology and population genetics. – *Trends Ecol. Evol.* 18: 189–197.
- Martensen, A. C. et al. 2017. Spatio-temporal connectivity: assessing the amount of reachable habitat in dynamic landscapes. – *Methods Ecol. Evol.* 8: 1253–1264.
- Moullec, F. et al. 2019. An end-to-end model reveals losers and winners in a warming Mediterranean Sea. – *Front. Mar. Sci.* 6: 1–19.
- Murphy, M. et al. 2015. Graph theory and network models in landscape genetics. – In: Balkenhol, N. et al. (eds), *Landscape genetics: concepts, methods, applications*. Wiley, pp. 165–180.
- Nathan, R. and Muller-Landau, H. C. 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. – *Trends Ecol. Evol.* 15: 278–285.

- Nathan, R. et al. 2011. Mechanistic models of seed dispersal by wind. – *Theor. Ecol.* 4: 113–132.
- Newman, M. E. J. 2003. The structure and function of complex networks. – *SIAM Rev.* 45: 167–256.
- Oksanen, J. et al. 2016. *vegan*: community ecology package. – R package ver. 2.4-1.
- Oliveira, J. dos A. et al. 2019. Model-based riverscape genetics: disentangling the roles of local and connectivity factors in shaping spatial genetic patterns of two Amazonian turtles with different dispersal abilities. – *Evol. Ecol.* 33: 273–298.
- Papadakis, M. et al. 2018. Rfast: a collection of efficient and extremely fast R functions. – R package ver. 1.99.1.
- Paris, C. B. et al. 2013. Reef odor: a wake up call for navigation in reef fish larvae. – *PLoS One* 8: e72808.
- Pascual, M. et al. 2017. Impact of life history traits on gene flow: a multispecies systematic review across oceanographic barriers in the Mediterranean Sea. – *PLoS One* 12: e0176419.
- Peery, M. Z. et al. 2008. Characterizing source–sink dynamics with genetic parentage assignments. – *Ecology* 89: 2746–59.
- Peterson, E. E. et al. 2019. Spatially structured statistical network models for landscape genetics. – *Ecol. Monogr.* 82: e01355.
- Razgour, O. et al. 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. – *Proc. Natl Acad. Sci. USA* 116: 10418–10423.
- Rellstab, C. et al. 2016. Signatures of local adaptation in candidate genes of oaks (*Quercus* spp.) with respect to present and future climatic conditions. – *Mol. Ecol.* 25: 5907–5924.
- Reñones, O. et al. 1995. Life history of the red mullet *Mullus surmuletus* from the bottom-trawl fishery of the Island of Majorca (north-west Mediterranean). – *Mar. Biol.* 123: 411–419.
- Riginos, C. et al. 2019. Asymmetric dispersal is a critical element of concordance between biophysical dispersal models and spatial genetic structure in Great Barrier Reef corals. – *Divers. Distrib.* 25: 1684–1696.
- Row, J. R. et al. 2017. Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. – *Ecol. Evol.* 7: 3751–3761.
- Row, J. R. et al. 2018. Quantifying functional connectivity: the role of breeding habitat, abundance and landscape features on range-wide gene flow in sage-grouse. – *Evol. Appl.* 11: 1305–1321.
- Rozenfeld, A. F. et al. 2008. Network analysis identifies weak and strong links in a metapopulation system. – *Proc. Natl Acad. Sci. USA* 105: 18824–18829.
- Saura, S. et al. 2014. Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks. – *J. Appl. Ecol.* 51: 171–182.
- Schick, R. S. and Lindley, S. T. 2007. Directed connectivity among fish populations in a riverine network. – *J. Appl. Ecol.* 44: 1116–1126.
- Schoville, S. D. et al. 2018. Preserving genetic connectivity in the European Alps protected area network. – *Biol. Conserv.* 218: 99–109.
- Shirk, A. J. et al. 2018. A comparison of regression methods for model selection in individual-based landscape genetic analysis. – *Mol. Ecol. Resour.* 18: 55–67.
- Simpson, S. D. 2005. Homeward Bound. – *Science* 308: 221–221.
- Slatkin, M. 2005. Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. – *Mol. Ecol.* 14: 67–73.
- Taylor, P. D. et al. 1993. Connectivity is a vital element of landscape structure. – *Oikos* 68: 571–573.
- Treml, E. A. et al. 2008. Modeling population connectivity by ocean currents, a graph–theoretic approach for marine conservation. – *Landscape Ecol.* 23: 19–36.
- Truelove, N. K. et al. 2016. Biophysical connectivity explains population genetic structure in a highly dispersive marine species. – *Coral Reefs* 36: 233–244.
- Urban, D. and Keitt, T. 2001. Landscape connectivity: a graph-theoretic perspective. – *Ecology* 82: 1205–1218.
- Urban, D. L. et al. 2009. Graph models of habitat mosaics. – *Ecol. Lett.* 12: 260–273.
- van Etten, J. 2017. R Package gdistance: distances and routes on geographical grids. – *J. Stat. Softw.* 76: 1–21.
- Vandergast, A. G. et al. 2008. Are hotspots of evolutionary potential adequately protected in southern California? – *Biol. Conserv.* 141: 1648–1664.
- Wang, Y. H. et al. 2008. Habitat suitability modelling to correlate gene flow with landscape connectivity. – *Landscape Ecol.* 23: 989–1000.
- Warshall, S. 1962. A theorem on Boolean matrices. – *J. ACM* 9: 11–12.
- White, C. et al. 2010. Ocean currents help explain population genetic structure. – *Proc. R. Soc. B* 277: 1685–1694.
- Whiteley, A. R. et al. 2015. Genetic rescue to the rescue. – *Trends Ecol. Evol.* 30: 42–49.
- Williamson, D. H. et al. 2016. Large-scale, multidirectional larval connectivity among coral reef fish populations in the Great Barrier Reef Marine Park. – *Mol. Ecol.* 25: 6039–6054.
- Xuereb, A. et al. 2018. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). – *Mol. Ecol.* 27: 2347–2364.
- Xuereb, A. et al. 2019. Marine conservation and marine protected areas. – In: Oleksiak, M. and Rajora, O. (eds), *Population genomics: marine organisms*. Springer, pp. 423–446.

Supplementary material (available online as Appendix ecog-05024 at <[www.ecography.org/appendix/ecog-05024](http://www.ecography.org/appendix/ecog-05024)>). Appendix 1.